



Uniwersytet
ŁÓDZKI

UNIWERSYTET ŁÓDZKI
WYDZIAŁ EKONOMICZNO-SOJOLOGICZNY

PRACA DOKTORSKA

EKONOMIA

**Zastosowanie modeli zmienności
stochastycznej typu
Ornsteina-Uhlenbecka
w analizie finansowych szeregów
czasowych**

mgr Piotr Szczepocki

Promotor: prof. dr hab. Czesław Domański

Łódź, maj 2019

.....

podpis promotora

.....

podpis autora

Spis treści

Wykaz skrótów, oznaczeń i symboli	2
Wstęp	3
1 Pojęcia podstawowe	9
1.1 Uwagi wstępne	9
1.2 Procesy stochastyczne	9
1.3 Procesy Lévy’ego	22
1.4 Podklasy procesów Lévy’ego	39
1.5 Proces Ornsteina-Uhlenbecka	42
2 Niegaussowskie modele zmienności stochastycznej typu Ornsteina-Uhlenbecka	47
2.1 Uwagi wstępne	47
2.2 Model stochastycznej zmienności Bandorffa-Nielsena i Shepharda . . .	51
2.3 Złożenie procesów zmienności	66
2.4 Efekt dźwigni	71
2.5 Rozkłady stacjonarne procesu wariancji chwilowej	74
2.6 Wariancja zrealizowana	81
3 Estymacja niegaussowskich modeli zmienności stochastycznej typu Ornsteina-Uhlenbecka	90
3.1 Uwagi wstępne	90

3.2	Modele przestrzeni stanów	92
3.3	Filtr Kalmana	96
3.4	Liniowe modele przestrzeni dla nieguassowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka	103
3.5	Filtry cząsteczkowe	116
3.6	Nieliniowe modele przestrzeni dla nieguassowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbeckaa	125
3.7	Implementacja filtrów cząsteczkowych	133
3.8	Iterowana filtracja	139
4	Modelowanie zmienności finansowych szeregów czasowych	158
4.1	Uwagi wstępne	158
4.2	Charakterystyka danych empirycznych	160
4.3	Wybór rozkładu stacjonarnego procesu wariancji chwilowej	170
4.4	Wyniki estymacji dla indeksu WIG	176
4.5	Wyniki estymacji dla kursu USD/PLN	186
4.6	Wyniki estymacji dla indeksu WIG20	195
	Zakończenie	203
	Dodatek A Rozkłady bezwarunkowe logarytmicznych stóp zwrotu	207
	Dodatek B Symulacja niegaussowskich procesów Ornsteina-Uhlenbecka	211
	Dodatek C Symulacja procesu logarytmicznych cen	215
	Bibliografia	217
	Spis rysunków	238
	Spis tablic	245

Wykaz skrótów, oznaczeń i symboli

$\mathbb{N}$	zbiór liczb naturalnych (z zerem)
$\mathbb{N}_+$	zbiór liczb naturalnych dodatnich
$\mathbb{Z}$	zbiór liczb całkowitych
$\mathbb{R}$	zbiór liczb rzeczywistych
$\mathbb{R}_+$	zbiór liczb rzeczywistych dodatnich
$\mathbb{R}^k$	k -wymiarowa przestrzeń Euklidesowa ($\mathbb{R}^k = \mathbb{R} \times \dots \times \mathbb{R}$)
Ω	przestrzeń prób
Σ	σ -ciało
$(\Sigma_t)_{t \in \mathbb{T}}$	filtracja
$(\Sigma_t^X)_{t \in \mathbb{T}}$	filtracja generowane przez proces stochastyczny
$(X(t))_{t \in \mathbb{T}}$	
$\mathbb{P}$	miara probabilistyczna
$(X(t))_{t \in \mathbb{T}}$	proces stochastyczny
$(L(t))_{t \geq 0}$	proces Lévy'ego
$(Z(t))_{t \geq 0}$	proces podporządkowany
$(\bar{Z}(t))_{t \geq 0}$	skompensowany proces podporządkowany
$(W(t))_{t \geq 0}$	proces Wienera
$(\tilde{W}(t))_{t \geq 0}$	proces Wienera z dryfem

$(N(t))_{t \geq 0}$	proces Poissona
$(\bar{N}(t))_{t \geq 0}$	skompensowany proces Poissona
$(\tilde{N}(t))_{t \geq 0}$	złożony proces Poissona
$(\bar{\tilde{N}}(t))_{t \geq 0}$	skompensowany złożony proces Poissona
$(\sigma^2(t))_{t \geq 0}$	proces wariancji chwilowej
$(\sigma^{2*}(t))_{t \geq 0}$	proces wariancji scałkowanej
$(\sigma_n^2)_{n \in \mathbb{N}}$	proces wariancji aktualnej
$(S(t))_{t \geq 0}$	proces ceny aktywa finansowego
$(Y(t))_{t \geq 0}$	proces logarytmu ceny aktywa finansowego (proces logarytmicznych cen)
$(y_n)_{n \in \mathbb{N}}$	proces zwrotów logarytmicznych
$\mathcal{D}$ -OU	proces Ornsteina-Uhlenbecka o rozkładzie stacjonarnym $\mathcal{D}$
OU- $\mathcal{D}$	proces Ornsteina-Uhlenbecka o procesie prowadzącym Lévy'ego ukrytym w tle o rozkładzie $\mathcal{D}$ dla zmiennej losowej $Z(1)$
BDPL	proces prowadzący Lévy'ego ukryty w tle
$N(\mu, \sigma^2)$	rozkład normalny o średniej μ i wariancji σ^2
$\text{GIG}(\nu, \delta, \gamma)$	uogólniony rozkład Gaussa
$\text{IG}(\delta, \gamma)$	rozkład Gaussa
$\text{Ga}(\nu, \alpha)$	rozkład gamma
$\text{PH}(\delta, \gamma)$	dodatni rozkład hiperboliczny
$\text{IGa}(\nu, \alpha)$	rozkład odwrotny gamma
$\text{TS}(\kappa, \nu, \alpha)$	temperowany rozkład stabilny
$\text{LN}(\mu, \sigma^2)$	rozkład log-normalny
$\text{GH}(\mu, \alpha, \beta, \delta, \nu)$	uogólniony rozkład hiperboliczny
$\text{NIG}(\mu, \alpha, \beta, \delta)$	rozkład normalny odwrotny Gaussa
$\text{VG}(\mu, \sigma, \nu, \theta)$	rozkład <i>variance gamma</i>

Wstęp

W centrum wielu zagadnień finansów jest zmienność aktywów finansowych. O fundamentalnym znaczeniu zmienności w teorii finansów świadczy ich rola w klasycznej teorii portfela zaproponowanej przez Markowitza, modelach wyceny opcji, czy pomiarze ryzyka metodą wartości zagrożonej (*Value at Risk*). Zmienność jest obecną również przedmiotem handlu na rynkach finansowych m.in. poprzez kontrakty terminowe na zmienność implikowaną oraz kontrakty typu variance swap. Prawdopodobnie najlepiej znaczenie zmienności finansach zostało podsumowane przez Andersen i Bollerslev (1998), którzy stwierdzili krótko: „Zmienność przenika finanse” (*“Volatility permeates finance”*).

Zmienność na rynkach finansowych rozumiana jest jako miara niepewności co do zmian cen badanego aktywa. Zmienność nie jest bezpośrednio obserwowana, ale może być aproksymowana przez różne mierniki, z których najbardziej popularnymi są odchylenie standardowe i wariancja stóp zwrotu z aktywów. Jednak już wcześnie prace Mandelbrot (1963) i Fama (1965) wskazywały, że tak mierzona zmienność nie jest stała w czasie. Również dla Fishera Blacka i Myrona Scholes twórców modelu wyceny opcji europejskich założenie jednorodnej w czasie zmienności było nierealistycznym, ale wygodnym założeniem upraszczającym. W artykule Cohen i in. (1972) pisali: „...istnieje dowód niestacjonarności wariancji. Więcej pracy powinno zostać poświęcone, aby przewidywać wariancje na podstawie dostępnej informacji” (*“...there is evidence of non-stationarity in the variance. More work must be done to predict variances using the information available”*). Co więcej jak zauważył Mandelbrot (1963): „większe zmiany mają tendencje do następowania po dużych

zmianach, a małe zmiany po małych” (*“large changes tend to be followed by large changes (...) and small changes tend to be followed by small changes”*). Zjawisko to jest teraz powszechnie nazywane grupowaniem zmienności. W celu uchwycenia tego zjawiska zostały sformułowane dwie klasy modeli: uogólnione autoregresyjne z warunkową heteroskedastycznością (*generalized autoregressive conditional heteroskedasticity*, GARCH) oraz stochastycznej zmienności (*stochastic volatility*, SV).

Obie te klasy umożliwiają odwierciedlać obserwowane własności finansowych szeregów czasowych, ale w sposób odmienny modelują zmienność. W przypadku model GARCH oraz ich uogólnień, zmienność definiuje się jako wariancję warunkową, którą wyznacza się jako deterministyczną funkcję poprzednich wariancji warunkowych i kwadratów poprzednich obserwacji. Natomiast cechą charakterystyczną modeli SV jest to, że zmienność ceny instrumentu finansowego traktowana jest jako oddzielny proces stochastyczny, nazywany wariancją chwilową lub natychmiastową (*instantaneous volatility*). Oznacza to, że zmienność ma osobne źródło losowości niż ceny instrumentu. Dzięki temu modele SV są bardziej elastyczne w zakresie opisu ewolucji zmienności w czasie niż modele GARCH, jednak przy tym są dużo trudniejsze w estymacji. Modele SV mają również tę przewagę, że stanowią naturalne uogólnienie modelu Blacka-Scholesa, w którym zmienność nie jest stałym parametrem, ale oddzielnym procesem stochastycznym. Podejście to zostało zapoczątkowane przez pracę Hull i White (1987). Niemniej powszechnie uważa się, że historia modeli SV jest dłuższa (Shephard i Andersen, 2009). Po raz pierwszy tego typu modele zostały przedstawione pracy Clark (1973), w której cena instrumentu jest procesem stochastycznym, w którym czas jest oddzielnym procesem modelującym losowy napływ informacji.

Literatura dotycząca modeli SV jest bardzo bogata. Do najbardziej znanych należą modele, w których logarytm procesu wariancji chwilowej jest gaussowskim procesem Ornsteina-Uhlenbecka i klasa modeli stałej elastyczności wariancji (*constant elasticity of variance*, CEV), której najbardziej znanym przedstawicielem jest model Hestona (Heston, 1993). Niegaussowskie modele stochastycznej zmienności zaproponowane przez Barndorff-Nielsen i Shephard (2001b) są pod wieloma względami wyjątkowe. Nie jest to jeden model, ale raczej cała klasa modeli oparta na

całkowicie nowym podejściu do opisywania ewolucji zmienności w czasie. Procesem wariancji chwilowej jest niegaussowski proces Ornsteina-Uhlenbecka, dla którego stochastyczne równanie różniczkowe ma różniczkę względem niemalejącego procesu Lévy'ego. Ta nietypowa postać procesu wariancji chwilowej ma wiele istotnych dla modelowania zmienności właściwości: proces wariancji chwilowej jest dodatni bez konieczności dodatkowych transformacji, ma ściśle stacjonarny rozkład oraz niezależną od wyboru rozkładu stacjonarnego funkcję autokorelacji. Ponadto można dokonać dyskretyzacji procesu wariancji chwilowej bez konieczności stosowania schemtu dyskretyzacji Eulera lub Milsteina co pozwala na uniknięciu błędu dyskretyzacji. Na podstawie tak skonstruowanego procesu wariancji chwilowej Barndorff-Nielsen i Shephard (2001b) pokazali również jak można budować bardziej zaawansowane modele pozwalające odtwarzać bardziej subtelne własności finansowych szeregów czasowych między innymi długą pamięć, czy skorelowanie pomiędzy zmiennością a stopami zwrotu, zwane „efektem dźwigni”. Niegaussowskie modele stochastycznej zmienności typu Ornsteina-Uhlenbecka zostały z powodzeniem zastosowane w matematyce finansowej, m.in. do wyceny opcji w stylu europejskim (Nicolato i Venardos, 2003), opcji azjatyckich (Shi i Yang, 2014), opcji barierowych (Li i Linetsky, 2015), swapów (Benth i in., 2007; Habtemicael i SenGupta, 2016; Issaka i SenGupta, 2017), wyceny opcji za pomocą transformacji Essechera (Hubalek i Sgarra, 2009), optymalizacji portfeli inwestycyjnych (Benth i in., 2003).

Zastosowanie niegaussowskich modeli SV typu Ornsteina-Uhlenbecka ogranicza problem wspólny wszystkim modelom stochastycznej zmienności – są trudne w estymacji. W modelach SV, w tym także niegaussowskim typu Ornsteina-Uhlenbecka, zazwyczaj nie można wyznaczyć funkcji wiarygodności w postaci analitycznej, ponieważ wymagałoby to policzenia skomplikowanej wielowymiarowej całki. W konsekwencji wnioskowanie w oparciu o metodą największej wiarygodności nie jest możliwe. Powstała bardzo obszerna literatura na temat estymacji niegaussowskich modeli stochastycznej zmienności Ornsteina-Uhlenbecka, ale w dużej mierze opiera się na wnioskowaniu bayesowskim. Zastosowanie metod opartych na próbkowaniu Monte Carlo łańcuchami Markowa (*Markov Chain Monte Carlo*, MCMC) pozwala na ominięcie problemu wyznaczania wartości funkcji wiarygodności. Tylko nieliczne

prace podejmują problem estymacji od strony klasycznego wnioskowania (Lindberg, 2008; Hubalek i Posedel, 2011; Taufer i in., 2011). Na gruncie klasycznego wnioskowania również można zastosować metody estymacji, które opierają się na twierdzeniu Bayesa: filtry Kalmana i filtry cząsteczkowe. Już w pierwszej pracy poświęconej niegaussowskim modelom stochastycznej zmienności Ornsteina-Uhlenbecka (Barndorff-Nielsen i Shephard, 2001b) wskazali, że można użyć filtrów Kalmana i cząsteczkowych do estymacji stanów procesu wariancji aktualnej (dyskretyzacji procesu wariancji chwilowej). Oba podejścia wiążą się z pewnymi trudnościami. W przypadku filtru Kalmana należy przedstawić model w postaci liniowej co nie zawsze jest możliwe, a estymacja parametrów odbywa się poprzez przyjęcie założenia o normalności składników losowych co w przypadku niegaussowskich modeli stochastycznej jest założeniem upraszczającym. Filtry cząsteczkowe choć pozwalają na estymację stanu w przypadku modeli nieliniowych i niegaussowskich nie można zastosować bezpośrednio do estymacji parametrów.

Głównym celem rozprawy jest przedstawienie niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka oraz opracowanie i zastosowanie metod opartych na filtrach Kalmana i filtrach cząsteczkowych do estymacji wariancji aktualnej i parametrów tych modeli.

W ramach wymienionego celu głównego rozważane są następujące cele szczegółowe:

- przegląd i omówienie stosowanych niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka
- propozycja estymacji parametrów niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka za pomocą metody iterowanej filtracji
- przedstawienie wyników badań empirycznych dotyczących modelowania zmienności szeregów czasowych pochodzących z polskiego rynku finansowego.

Metoda iterowanej filtracji została zainicjowana w Ionides i in. (2006), a jej teoretyczne uzasadnienie zostało przedstawione w Ionides i in. (2011). W iterowanej filtracji estymatory parametrów uzyskuje się wykonując sekwencję filtracji na przestrzeni stanów poszerzonej o wektor zmiennych w czasie parametrów, zmniejszając

z kroku na krok zmienność parametrów, tak aby w rezultacie otrzymać oszacowanie parametru maksymalizujące logarytm funkcji wiarygodności modelu ze stałymi parametrami. Metoda iterowanej filtracji została wykorzystana w szeregu prac głównie w kontekście modeli ekologicznych i epidemiologicznych (Bhadra i in., 2011; Roy i in., 2013; Blackwood i in., 2013), ale także do estymacji parametrów w modelu stochastycznej zmienności ze stochastycznym parametrem mierzącym efekt dźwigni Bretó (2014). Zastosowanie metody iterowanej filtracji w modelach niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka zostało zainicjowane w pracach Szczepocki (2019a) oraz Szczepocki (2019b). W tej pracy zostanie zastosowane do bardziej ogólnych przypadków m.in. złożenia procesów zmienności i modeli uwzględniających efekt dźwigni.

W oparciu o powyższe cele pracy sformułowano następujące hipotezy badawcze rozprawy. Pierwsza hipoteza postuluje, że niegaussowskie modele stochastycznej zmienności typu Ornsteina-Uhlenbecka pozwalają na odtworzenie obserwowanych własności finansowych szeregów czasowych. Hipoteza druga określa, że można dobrać rozkłady stacjonarne procesów wariancji chwilowej dla szeregów stóp zwrotu z wybranych instrumentów finansowych pochodzących z polskiego rynku finansowego i rynku walutowego. Hipoteza trzecia zakłada, że metoda iterowanej filtracji pozwala na skuteczną estymację parametrów dla niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka.

Pracę ujęto w czterech rozdziałach. W pierwszym rozdziale przedstawione zostały definicje i własności rozważanych w pracy procesów stochastycznych, głównie procesów Lévy’ego i Ornsteina-Uhlenbecka. Przybliżono twierdzenia o reprezentacji procesów Lévy’ego, wykorzystywane w dalszej części podklasy procesów Lévy’ego oraz podstawowe własności procesów Ornsteina-Uhlenbecka. Rozdział drugi poświęcony został modelom stochastycznej zmienności opartym na niegaussowskich procesach Ornsteina-Uhlenbecka. Omówiono szczegółowo własności podstawowego modelu, a następnie przedstawiono rozszerzenia tego modelu pozwalające na uchwycenie zależności długookresowych w procesie zmienności oraz korelacji pomiędzy procesem obserwacji i procesem zmienności, co pozwala na modelowanie efektu dźwigni. W rozdziale trzecim omówiono zagadanie estymacji niegaussowskich modeli stocha-

stycznej zmienności typu Ornsteina-Uhlenbecka. Punktem wyjścia, a zarazem także odniesienia są filtry Kalmana i estymacja metodą *quasi*-największej wiarygodności. Następnie przedstawione zostanie estymacja procesu zmienności za pomocą filtrów cząsteczkowych i estymacja parametrów niegaussowskich modeli stochastycznej zmienności za pomocą metody iterowanej filtracji. Czwarty rozdział pracy obejmuje przykłady modelowania zmienności na podstawie szeregów czasowych pochodzących z polskiego rynku finansowego.

Prace uzupełniają trzy dodatki oznaczone kolejno literami A, B i C. W pierwszym przedstawiono zastosowane w pracy rozkłady stóp zwrotu. W dodatku B przedstawiono kody w języku programowania R umożliwiające symulację niegaussowskich modeli stochastycznej zmienności oraz kolejne generacje cząsteczek w filtrze cząsteczkowym. Dodatek C zawiera metody symulacji procesów logarytmu cen w niegaussowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka.

Pojęcia podstawowe

1.1 Uwagi wstępne

W poniższym rozdziale przedstawione zostaną podstawy matematyczne nieguassowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka. Pierwszy podrozdział stanowi przegląd podstawowych terminów związanych z procesami stochastycznymi wykorzystywanymi w dalszej części pracy. Drugi podrozdział poświęcony jest procesom Lévy’ego, które leżą u podstawy rozważanych procesów stochastycznej zmienności. Omówione zostaną twierdzenia o reprezentacji procesów Lévy’ego oraz najważniejsze przykłady procesów Lévy’ego. W podrozdziale 1.4 omówione zostaną dwie ważne podklasy procesów Lévy’ego: procesy podporządkowane i procesy o rozkładach samorozkładalnych. Procesy podporządkowane, to niemalejące procesy Lévy’ego, które mogą stanowić model „czasu operacyjnego”, czyli opisywać ewolucję czasu dla pewnego innego procesu stochastycznego. Rozkłady samorozkładalne będą w dalszej części pracy stanowiły rozkłady stacjonarne rozważanych procesów wariancji chwilowej. Ostatni podrozdział poświęcony został procesom Ornsteina-Uhlenbecka.

1.2 Procesy stochastyczne

Dla potrzeb niniejszej pracy będziemy przez $(\Omega, \Sigma, \mathbb{P})$ oznaczać przestrzeń probabilistyczną zupełną. Zbiór Ω , to zbiór wszystkich zdarzeń elementarnych zwany także

przestrzenią prób. Zbiór ten zawiera wszystkie możliwe wyniki obserwacji. Rodzina podzbiorów Σ zbioru Ω jest σ -ciałem (co oznacza, że do Σ należy zbiór pusty, dopełnienie dowolnego zbioru należące do Σ również należy do zbioru Σ oraz przeliczalna suma zbiorów należących do Σ również należy do Σ). Elementy tej rodziny nazywać będziemy zdarzeniami. Natomiast $\mathbb{P}$ jest miarą probabilistyczną określona na Σ wyznaczającą prawdopodobieństwo zajścia zdarzeń należących do rodziny Σ . Przestrzeń probabilistyczna $(\Omega, \Sigma, \mathbb{P})$ jest zupełna, czyli dla dowolnych $B \subset A$, warunek $\mathbb{P}(B) = 0$ implikuje $\mathbb{P}(A) = 0$.

Ponadto przestrzeń probabilistyczną będziemy zazwyczaj rozważać z pewną filtracją, czyli niemalejącą rodziną σ -ciał określającą jak zmienia się nasza wiedza o zajściu zdarzeń z rodziny Σ w czasie.

Definicja 1.1.

Filtracją przestrzeni $(\Omega, \Sigma, \mathbb{P})$ nazywamy rodzinę σ -ciał $(\Sigma_s)_{s \in \mathbb{T}} \subset \Sigma$ spełniającą następujący warunek: $\Sigma_s \subset \Sigma_t \subset \Sigma$ dla $0 \leq s \leq t, s, t \in \mathbb{T}$.

W przypadku filtracji zbiór indeksów $\mathbb{T}$ jest nieujemnym podzbiorem liczb rzeczywistych $\mathbb{R}$. Może natomiast być dyskretny lub ciągły, ograniczony lub nie, w zależności od kontekstu. W dalszej części pracy będziemy zakładać, że rozważane przestrzenie probabilistyczne z filtracją $(\Omega, \Sigma, \mathbb{P}, (\Sigma_s)_{s \in \mathbb{T}})$ będą spełniać warunki zwykłe (*usual conditions*, por. Schountens (2003, str. 12) oraz Latała (2011, str. 18)):

- 1) Σ_0 zawiera wszystkie zbiory miary 0 rodziny Σ ,
- 2) $(\Sigma_s)_{s \in \mathbb{T}}$ filtracja jest prawostronnie ciągła, tzn. $\Sigma_t = \bigcap_{s < t} \Sigma_s$.

Definicja 1.2.

Niech $(\Omega, \Sigma, \mathbb{P})$ będzie przestrzenią probabilistyczną, $(E, \mathcal{E})$ przestrzenią mierzalną. Procesem stochastycznym nazywamy rodzinę zmiennych losowych $(X(t))_{t \in \mathbb{T}}$ indeksowanych parametrem $t \in \mathbb{T}$ określonych na przestrzeni probabilistycznej $(\Omega, \Sigma, \mathbb{P})$ o wartościach w zbiorze E .

W dalszej części będziemy zakładać, że zbiór wartości E jest zbiorem liczb rzeczywistych $\mathbb{R}$. Natomiast za $\mathcal{E}$ przyjmować będziemy σ -ciała zbiorów borelowskich $\mathcal{B}(\mathbb{R})$.

Zauważmy, że proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ można również traktować jako odwzorowanie dwóch zmiennych $X : \Omega \times \mathbb{T} \rightarrow \mathbb{R}$. Dla ustalonego $t \in \mathbb{T}$ funkcja $t \mapsto X(\cdot, t)$ jest zmienną losową. Natomiast przy określeniu zdarzenia elementarnego $\omega \in \Omega$ funkcję $\omega \mapsto X(\omega, \cdot)$ nazywamy trajektorią (lub realizacją) procesu X .

Zbiór indeksów $\mathbb{T}$ interpretuje się jako czas. W przypadku, gdy $\mathbb{T}$ jest zbiorem przeliczalnym (np. zbiorem liczb naturalnych) to proces stochastyczny nazywa się procesem z czasem dyskretnym. Dla procesów z czasem dyskretnym będziemy przyjmować zapis: $X_t = X(t)$. W przypadku, gdy $\mathbb{T}$ jest przedziałem zawartym w zbiorze liczb rzeczywistych $\mathbb{R}$ (np. $\mathbb{T} = [0, +\infty)$), to proces stochastyczny nazywany jest procesem z czasem ciągłym. W dalszej części pracy zakładamy, że $\mathbb{T} = \mathbb{N}$ lub $\mathbb{T} = [0, +\infty)$.

Definicja 1.3.

Dwa procesy stochastyczne $(X(t))_{t \in \mathbb{T}}$ i $(Y(t))_{t \in \mathbb{T}}$ nazywamy:

- 1) *wariantami (modyfikacjami)*, jeżeli $\forall_{t \in \mathbb{T}} \mathbb{P}(X(t) = Y(t)) = 1$,
- 2) *nierozróżnialnymi*, jeżeli $\mathbb{P}(\forall_{t \in \mathbb{T}} X(t) = Y(t)) = 1$.

Warunek nierozróżnialności jest silniejszy niż modyfikacji: jeżeli dwa procesy są nierozróżnialne to są również modyfikacjami, natomiast implikacja odwrotna nie jest zawsze prawdziwa (Latała, 2011, str. 14). Szczególną rolę w dalszej części pracy będą odgrywały modyfikacje procesu mające ciągłe trajektorie¹.

Definicja 1.4.

Filtrację $(\Sigma_t^X)_{t \in \mathbb{T}}$ nazywamy generowaną przez proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ jeżeli dla dowolnego $t \in \mathbb{T}$ zachodzi warunek $\Sigma_t^X = \sigma(X_s : s \leq t)$.

Filtrację generowaną przez proces stochastyczny nazywa się również filtracją naturalną procesu stochastycznego. Można ją interpretować w następujący sposób: Σ_t^X jest zbiorem informacji dostępnej na podstawie obserwacji procesu $(X(t))_{t \in \mathbb{T}}$ do momentu $t \in \mathbb{T}$ (Jakubowski i in., 2003, str. 73).

¹Twierdzenie o istnieniu takich wariantów można znaleźć w Latała (2011, str. 14).

Definicja 1.5.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ nazywamy zgodnym z filtracją $(\Sigma_t)_{t \in \mathbb{T}}$ (lub adaptowanym do filtracji $(\Sigma_t)_{t \in \mathbb{T}}$) jeśli zmienna losowa $X(t)$ jest mierzalna względem σ -ciała Σ_t dla dowolnego $t \in \mathbb{T}$.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ jest zgodny z filtracją $(\Sigma_t)_{t \in \mathbb{T}}$ wtedy i tylko wtedy, gdy $\Sigma_t^X \subset \Sigma_t$ dla $t \in \mathbb{T}$ (Latała, 2011, str. 17). Oznacza to również, że każdy proces $(X(t))_{t \in \mathbb{T}}$ jest zgodny z filtracją naturalną. Jest to również najmniejsza filtracja, do której proces jest adoptowany (Jakubowski i in., 2003, str. 73).

Procesy stochastyczne, podobnie jak w przypadku zmiennych losowych, można także charakteryzować za pomocą momentów, z tą różnicą, że w przypadku zmiennych losowych momenty są liczbami rzeczywistymi, a w przypadku procesów stochastycznych funkcjami.

Definicja 1.6.

Wartością oczekiwaną procesu $(X(t))_{t \in \mathbb{T}}$ nazywamy funkcję m_X przyporządkowującą, każdemu $t \in \mathbb{T}$ wartość oczekiwaną zmiennej losowej $X(t)$, czyli funkcję

$$m_X(t) = \mathbb{E}(X(t)), \quad t \in \mathbb{T}.$$

Wariancję procesu $(X(t))_{t \in \mathbb{T}}$ nazywamy funkcję postaci

$$v_X(t) = \text{Var}(X(t)) = \mathbb{E}(X(t) - m_X(t))^2, \quad t \in \mathbb{T}.$$

Funkcją korelacyjną procesu $(X(t))_{t \in \mathbb{T}}$ nazywamy funkcję postaci

$$R_X(t_1, t_2) = \mathbb{E}(X(t_1)X(t_2)), \quad t_1, t_2 \in \mathbb{T}.$$

Funkcją kowariancji procesu $(X(t))_{t \in \mathbb{T}}$ nazywamy funkcję postaci

$$k_X(t_1, t_2) = \mathbb{E}[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))], \quad t_1, t_2 \in \mathbb{T}.$$

Dla dowolnych $t_1, t_2 \in \mathbb{T}$, dla funkcji kowariancji zachodzi równość

$$k_X(t_1, t_2) = R_X(t_1, t_2) - m_X(t_1)m_X(t_2).$$

Wartość oczekiwana i wariancja procesu stochastycznego zależą od jednowymiarowych rozkładów prawdopodobieństwa, natomiast funkcje korelacji i kowariancji zależą od dwuwymiarowych rozkładów prawdopodobieństwa procesu stochastycznego $(X(t))_{t \in \mathbb{T}}$.

W dalszej części pracy ważną rolę odgrywać będą procesy stacjonarne. Są to procesy niezmiennicze względem przesunięć czasu. Wyróżniamy dwa typy stacjonarności: ścisłą (węższą) i słabą (szerszą).

Definicja 1.7.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ nazywamy ściśle stacjonarnym (stacjonarnym w wąszym sensie), jeżeli dla dowolnego $\{t_1, t_2, \dots, t_n\} \subset \mathbb{T}$ i dla każdego h takiego że $t_{i+h} \in \mathbb{T}$ ($i = 1, \dots, n$) n -wymiarowe rozkłady prawdopodobieństwa wektorów losowych $[X(t_1) \ X(t_2) \ \dots \ X(t_n)]^T$ oraz $[X(t_{1+h}) \ X(t_{2+h}) \ \dots \ X(t_{n+h})]^T$ są takie same.

Dla ściśle stacjonarnego procesu stochastycznego rozkłady jednowymiarowe mają identyczną postać dla dowolnego $t \in \mathbb{T}$, natomiast rozkłady dwuwymiarowe zależą tylko od różnicy momentów $t_2 - t_1$, gdzie $t_1, t_2 \in \mathbb{T}$. Jeżeli $(X(t))_{t \in \mathbb{T}}$ jest procesem stacjonarnym, to jednowymiarowy rozkład $X(t)$ dla dowolnego $t \in \mathbb{T}$ nazywany jest rozkładem stacjonarnym tego procesu.

Definicja 1.8.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ nazywamy słabo stacjonarnym (stacjonarnym w szerszym sensie), jeżeli $\mathbb{E}|X(t)|^2 < +\infty$ dla dowolnego $t \in \mathbb{T}$ oraz dla dowolnych $t_1, t_2 \in \mathbb{T}$ zachodzą warunki

$$1) \ m_X(t_1) = m_X(t_2),$$

$$2) \ k_X(t_1, t_2) = k_X(t_1 - t_2, t_2 - t_2) = k_X(t_1 - t_2, 0).$$

Procesy stacjonarne w węższym sensie są również procesami stacjonarnymi w szerszym sensie. Natomiast odwrotna implikacja nie jest prawdziwa.

W przypadku procesów stacjonarnych funkcje kowariancji i korelacji procesu upraszczają się do funkcji jednej zmiennej:

$$\tilde{k}_X(\tau) = k_X(t_1 - t_2, 0) \quad \tilde{R}_X(\tau) = R_X(t_1 - t_2, 0), \quad (1.1)$$

gdzie $\tau = t_1 - t_2$. Zmienna τ przyjmuje wartości całkowite, gdy proces jest z czasem dyskretnym lub z rzeczywiste, gdy jest rozważany z czasem ciągłym. Zmienną τ można utożsamić z opóźnieniem w czasie dwóch zmiennych losowych $X(t + \tau)$ oraz $X(t)$.

W przypadku procesów słabo stacjonarnych zamiast funkcji korelacji częściej korzysta się z jej unormowanej wersji zwanej funkcją autokorelacji lub krótko autokorelacją procesu.

Definicja 1.9.

Funkcją autokorelacji procesu słabo stacjonarnego $(X(t))_{t \in \mathbb{T}}$ o funkcji korelacyjnej $\tilde{R}_X(\tau)$ nazywamy funkcję

$$r(\tau) = \frac{\tilde{R}_X(\tau)}{\text{Var}(X(0))}, \quad (1.2)$$

gdzie $\tau \in \mathbb{Z}$, gdy $\mathbb{T} = \mathbb{N}$ lub $\tau \in \mathbb{R}$, gdy $\mathbb{T} = [0, +\infty)$.

Ważną koncepcją w analizie procesów stochastycznych jest „pamięć procesu” stochastycznego. Termin „pamięć procesu” odnosi się do związku pomiędzy zmiennymi $(X(t))$ i $(X(t + \tau))$, które oddziela pewien długi okres czasu τ . Wyróżnia się dwa przypadki: procesy z krótką (*short memory, short range dependency*) lub z długą pamięcią (*long memory, long range dependency*).

Definicja 1.10.

Niech $(X(t))_{t \in \mathbb{T}}$ będzie procesem słabo stacjonarnym. Wówczas $(X(t))_{t \in \mathbb{T}}$ ma krótką pamięć, jeżeli

$$\int_{-\infty}^{+\infty} |r(\tau)| d\tau < +\infty,$$

w przypadku, gdy $\mathbb{T} = \mathbb{N}$ lub

$$\sum_{\tau=-\infty}^{+\infty} |r(\tau)| < +\infty,$$

gdy $\mathbb{T} = [0, +\infty)$. Natomiast $(X(t))_{t \in \mathbb{T}}$ ma długą pamięć, jeżeli

$$\int_{-\infty}^{+\infty} |r(\tau)| d\tau = +\infty,$$

w przypadku, gdy $\mathbb{T} = \mathbb{N}$ lub

$$\sum_{\tau=-\infty}^{+\infty} |r(\tau)| = +\infty,$$

gdy $\mathbb{T} = [0, +\infty)$.

Pojęcie pamięci procesu nie jest jednoznaczne w literaturze².

Kolejną ważną klasą są procesy Markowa, dla których przyszła wartość zależy tylko od wartości teraźniejszej, a nie od całej przeszłości procesu. Formalnie określa to następująca definicja.

Definicja 1.11.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ nazywamy procesem Markowa, jeżeli dla dowolnego ciągu indeksów $t_1, t_2, \dots, t_n, t_{n+1} \in \mathbb{T}$, $0 \leq t_1 < t_2 < \dots < t_n < t_{n+1}$, dowolnych wartości $x_1, x_2, \dots, x_n \in \mathbb{R}$ oraz dla dowolnego zbioru borelowskiego $B \in \mathcal{B}(\mathbb{R})$ zachodzi warunek

$$\begin{aligned} \mathbb{P}(X(t_{n+1}) \in B | X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \dots, X(t_1) = x_1) = \\ = \mathbb{P}(X(t_{n+1}) \in B | X(t_n) = x_n). \end{aligned}$$

Ważnym pojęciem w teorii procesów Markowa jest prawdopodobieństwo przejścia.

Definicja 1.12.

Prawdopodobieństwem przejścia procesu Markowa $(X(t))_{t \in \mathbb{T}}$ ze stanu x w momencie s do zbioru stanu $B \in \mathcal{B}(\mathbb{R})$ w chwili $t \geq s$ nazywamy funkcję przypisującą czwórce (x, s, B, t) wartość

$$P(x, s, B, t) = \begin{cases} \mathbb{P}(X(t) \in B | X(s) = x) & \text{dla } t > s \\ \mathbb{1}_B(x) & \text{dla } t = s \end{cases}.$$

Funkcja $\mathbb{1}_B(\cdot)$ jest indykátorem zbioru B , czyli funkcją:

$$\mathbb{1}_B(x) = \begin{cases} 1 & \text{dla } x \in B \\ 0 & \text{dla } x \notin B \end{cases}.$$

Dla ustalonych x, s, t , ($t \geq s$) funkcja $P(x, s, \cdot, t)$ jest rozkładem prawdopodobieństwa. Jeżeli rozkład ten jest absolutnie ciągły względem miary Lebesgue'a na $\mathbb{R}$, to posiada gęstość, która nazywana jest gęstością przejścia (*transition density*).

²Praca Kaarakka (2015) zawiera omówienie różnych definicji pamięci procesu stochastycznego. Definicja 1.10 jest zgodna z m.in. Piontek (2004), Gieriej (2008), Kaarakka (2015).

Prawdopodobieństwo przejścia pozwala na rozróżnienie wśród procesów Markowa procesów jednorodnych i niejednorodnych.

Definicja 1.13.

Proces Markowa $(X(t))_{t \in \mathbb{T}}$ nazywamy jednorodnym (o stacjonarnych przyrostach), jeżeli dla dowolnych momentów $t > s$ ($t, s \in \mathbb{T}$) prawdopodobieństwo przejścia zależy jedynie od różnicy $\tau = t - s$

$$P(x, s, B, t) = P(x, 0, B, t - s) = P(x, B, \tau).$$

Proces Markowa, który nie jest procesem jednorodnym nazywamy niejednorodnym.

Proces Markowa o stacjonarnych przyrostach nie musi być procesem stacjonarnym. Aby tak było musi spełniać dodatkowe warunki, które określa następujące twierdzenie.

Twierdzenie 1.1. *(Sobczyk, 1996, str. 55) Warunkiem koniecznym i dostatecznym, aby proces Markowa $(X(t))_{t \in \mathbb{T}}$ był procesem ściśle stacjonarnym jest, aby*

- 1) *proces X był jednorodnym procesem Markowa*
- 2) *istniał rozkład niezmienniczy P_0 dla całej przestrzeni stanów, tzn. taki rozkład prawdopodobieństwa P_0 , dla którego dla dowolnego $B \in \mathcal{B}(\mathbb{R})$ i dla dowolnego $\tau > 0$*

$$P_0(B) = \int_{\mathbb{R}} P(x, B, \tau) P_0(dx), \quad (1.3)$$

- 3) *proces niezmienniczy P_0 był rozkładem początkowym procesu X .*

W modelowaniu finansowych szeregów czasowych ważną rolę pełnią procesy mające punkty nieciągłości w trajektoriach, potocznie zwane skokami. Do modelowania ich wykorzystuje się procesy typu *càdlàg* i *càglàg*.

Definicja 1.14.

Proces stochastyczny $(X(t))_{t \geq 0}$ nazywamy prawostronnie ciągłym z lewostronnymi granicami (right continuous with left limits, w skrócie RCLL, fr. continue à droite, limite à gauche, w skrócie càdlàg) jeżeli istnieje zbiór $\Omega_0 \in \Sigma$, $\mathbb{P}(\Omega_0) = 1$ taki

że dla każdego $\omega \in \Omega_0$ trajektoria $\omega \mapsto X(\cdot, \omega)$ jest ciągła prawostronnie oraz ma skończone granice lewostronne.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ nazywamy lewostronnie ciągłym z prawostronnymi granicami (*left continuous with right limits*, *LCRL*, fr. *continue à gauche, limite à droit, càglàd*) jeżeli istnieje zbiór $\Omega_0 \in \Sigma$, $\mathbb{P}(\Omega_0) = 1$, taki że dla każdego $\omega \in \Omega_0$ trajektoria $\omega \mapsto X(\cdot, \omega)$ jest ciągła lewostronnie oraz ma skończone granice prawostronne.

Przestrzenie wszystkich procesów adoptowanych typu *càdlàg* oraz *càglàd* będziemy oznaczać odpowiednio przez $\mathbb{D}$ i $\mathbb{L}$.

Rozróżnienie pomiędzy procesami typu RCLL i LCRL ma istotne znaczenie, ponieważ zmienna t jest interpretowana jako czas. Zatem granica z lewej strony jest „przed”, a granica z prawej strony „po” danym zdarzeniu. Jeżeli proces RCLL ma skok w momencie t_0 , to skok nie jest możliwy do przewidzenia poprzez obserwację procesu do momentu t_0 – skok trajektorii jest zdarzeniem nieprzewidywalnym. Natomiast, gdyby proces miał trajektorię ciągłą z lewej strony, to obserwując trajektorię procesu w kierunku t_0 mogłby przewidzieć wartość procesu w momencie t_0 .

Niech proces stochastyczny $(X(t))_{t \geq 0}$ będzie typu *càdlàg*. Oznaczmy przez $\Delta X(t) = X(t) - X(t-)$, gdzie $X(t-) = \lim_{s \rightarrow t-} X(s)$ to granica lewostronna. Granica jest dobrze określona z definicji procesu *càdlàg*. Proces $(\Delta X(t))_{t \geq 0}$ nazywany jest procesem skoków związanym z procesem $(X(t))_{t \geq 0}$.

W modelowaniu zmienności cen instrumentów finansowych z czasem ciągłym konieczne jest określenie co jest zmiennością procesu stochastycznego w przedziale czasu. Jako miarę zmienności przyjmuje się zazwyczaj wariację kwadratową (Doman i Doman, 2009, str. 155).

Definicja 1.15.

Wariacją kwadratową procesu $(X(t))_{t \geq 0}$ na przedziale $[0, t]$ względem ciągu podziałów $(\Pi_n)_{n \in \mathbb{N}}$ nazywamy granicę (zbieżność wg prawdopodobieństwa)

$$\text{plim}_{n \rightarrow +\infty} [X]_t^{\Pi_n} = [X]_t, \quad (1.4)$$

gdzie

$$[X]_t^{\Pi} = \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2,$$

$(\Pi_n)_{n \in \mathbb{N}}$ jest ciągiem podziałów $[0, t]$ takim, że średnice podziałów zbiegają do zera:
 $\lim_{n \rightarrow +\infty} \text{diam}(\Pi_n) = 0$.

Jeżeli, granica 1.4 istnieje dla dowolnego ciągu podziałów, to można wykazać (por. Karandikar i Rao (2014)), że wartość tej granicy jest niezależna od wyboru ciągu podziału. Wówczas granicę tę nazywa się wariacją kwadratową procesu $(X(t))_{t \geq 0}$ na przedziale $[0, t]$. Jeżeli wariacja kwadratowa procesu stochastycznego $(X(t))_{t \geq 0}$ jest określona dla dowolnego $t > 0$, to może być traktowana również jako proces stochastyczny $([X](t))_{t \geq 0}$. Wystarczy przyjąć, że dla dowolnego $t > 0$: $[X](t) = [X]_t$ oraz $[X](0) = 0$ (Karandikar i Rao, 2014). Potrzebne jest jednak dodatkowo założenie, aby określić trajektorie procesu, ponieważ granice w definicji 1.15 są wyznaczone prawie na pewno dla dowolnego t , co nie pozwala określić ich wartości łącznie dla nieskończonej liczby punktów. Zazwyczaj wybiera się wariant procesu $([X](t))_{t \geq 0}$ prawostronnie ciągły z lewostronnymi granicami (RCLL) (Karandikar i Rao, 2014) i tak będziemy też przyjmować w dalszej części pracy.

Kolejnymi rozważanymi w tej pracy klasami procesów stochastycznych są martyngały i semimartynały.

Definicja 1.16.

Proces $X = (X(t))_{t \in \mathbb{T}}$ adaptowany do filtracji $(\Sigma_t)_{t \in \mathbb{T}}$ nazywamy:

- *martyngałem*, jeżeli $\mathbb{E}[X(t) | \Sigma_s] = X(s)$ prawie na pewno, dla wszystkich $s < t$; $s, t \in \mathbb{T}$,
- *podmartyngałem*, jeżeli $\mathbb{E}[X(t) | \Sigma_s] \geq X(s)$ prawie na pewno, dla wszystkich $s < t$; $s, t \in \mathbb{T}$,
- *nadmartyngałem*, jeżeli $\mathbb{E}[X(t) | \Sigma_s] \leq X(s)$ prawie na pewno, dla wszystkich $s < t$; $s, t \in \mathbb{T}$.

Definicje martyngałów, podmartyngałów i nadmartyngałów zakładają, że proces stochastyczny $X = (X(t))_{t \in \mathbb{T}}$ jest rozważany wraz z pewną filtracją $\Sigma = (\Sigma_t)_{t \in \mathbb{T}}$ i miarą probabilistyczną $\mathbb{P}$. Powinno się zatem mówić o $(\Sigma, \mathbb{P})$ -martyngałach. W matematyce finansowej zazwyczaj jednak rozważa się proces stochastyczny X z pewną

filtracją $\Sigma = (\Sigma_t)_{t \in \mathbb{T}}$ oraz miarą probabilistyczną $\mathbb{P}$ i szuka takiej miary probabilistycznej $\mathbb{Q}$, dla której proces X z filtracją Σ jest martyngałem. Wówczas dla podkreślenia roli miary $\mathbb{Q}$ proces X nazywa się $\mathbb{Q}$ -martyngałem, a miarę $\mathbb{Q}$ miarą martyngałową procesu X . Metoda ta jest stosowana w martyngałowej wycenie instrumentów pochodnych (Jakubowski i in., 2003, str. 74).

Martyngały są powszechnie wykorzystywane do modelowania zjawisk finansowych. Inwestowanie w instrumenty finansowe w matematyce finansowej rozważa się jako specyficzne „gry losowe”. Wiąże się, to z interpretacją martyngałów jako modelu „gry sprawiedliwej”. Jeśli przyjmiemy proces $(X(t))_{t \in \mathbb{T}}$ jako wyniki pewnej gry, to martyngały można traktować jako model „gry sprawiedliwej”, gdyż średnia wartość procesu w chwili t , gdy znany jest cały przebieg procesu do chwili s ($s < t$), jest równy $X(t)$, czyli wyniku w grze do momentu s . Miarę martyngałową $\mathbb{Q}$ jest w literaturze finansowej nazywana prawdopodobieństwem neutralizującym ryzyko (*risk-neutral probability*) (Jakubowski i in., 2003, str. 46).

Ważną rolę pełnią również semimartyngały, ponieważ jest to najszersza klasa procesów stochastycznych, dla których można zdefiniować całkę stochastyczną (Protter, 2005, roz. 4). Do ich określenia potrzebne są najpierw pewne definicje pomocnicze.

Definicja 1.17.

Zmienną losową $\tau : \Omega \rightarrow \mathbb{T} \cup \{+\infty\}$ nazywamy momentem stopu względem filtracji $(\Sigma_t)_{t \in \mathbb{T}}$ jeśli dla dowolnego $t \in \mathbb{T}$ spełnia następujący warunek $\{\omega \in \Omega : \tau(t) \leq t\} \subset \Sigma_t$.

Z momentami stopu związane są procesy stochastyczne zwane procesami zatrzymanymi.

Definicja 1.18.

Niech $\tau : \Omega \rightarrow \mathbb{T} \cup \{+\infty\}$ będzie momentem stopu względem filtracji $(\Sigma_t)_{t \in \mathbb{T}}$. Niech $(X(t))_{t \in \mathbb{T}}$ będzie procesem stochastycznym adoptowanym do filtracji $(\Sigma_t)_{t \in \mathbb{T}}$. Wówczas proces stochastyczny $(X^\tau(t))_{t \in \mathbb{T}}$ określony wzorem $X^\tau(t) = X(\min\{t, \tau\})$ nazywamy procesem zatrzymanym w τ .

Momenty zatrzymania są często wykorzystywane do uogólniania pewnych wła-

sności procesów stochastycznych na przypadek, w którym żądana własność jest spełniona jedynie lokalnie. Jeżeli mamy pewną klasę procesów $\mathcal{K}$, to klasa procesów lokalnych $\mathcal{K}_{loc}$ jest taką rodziną procesów, dla których istnieje ciąg momentów stopu $(\tau_n)_{n \in \mathbb{N}}$ rosnący i rozbieżny do $+\infty$ taki, że procesy $(X^\tau(t))_{t \in \mathbb{T}}$ należą do $\mathcal{K}$ (Jakubowski i in., 2003, str. 88). Jedną z najważniejszych klas lokalnych jest klasa martyngałów lokalnych.

Definicja 1.19.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ nazywamy martyngałem lokalnym względem filtracji $(\Sigma_t)_{t \in \mathbb{T}}$ jeżeli jest adaptowany do filtracji $(\Sigma_t)_{t \in \mathbb{T}}$ oraz istnieje ciąg momentów stopu $(\tau_n)_{n \in \mathbb{N}}$ rosnący i rozbieżny do $+\infty$ taki, że $1_{\tau_n} X^\tau(t)$ jest martyngałem dla dowolnego $n \in \mathbb{N}$.

Każdy martyngał jest lokalnym martyngałem, ponieważ jako ciąg momentów stopów wystarczy przyjąć $(\tau_n)_{n \in \mathbb{N}} = (n)_{n \in \mathbb{N}}$. Natomiast twierdzenie odwrotne nie jest prawdziwe – można wskazać martyngały lokalne, które nie są martyngałami (Jakubowski i in., 2003, str. 88).

Definicja 1.20.

Proces stochastyczny $(X(t))_{t \in \mathbb{T}}$ adaptowany do filtracji $(\Sigma_t)_{t \in \mathbb{T}}$, który można zdekomponować na sumę

$$X(t) = M(t) + A(t), \quad t \in \mathbb{T}$$

gdzie $(M(t))_{t \in \mathbb{T}}$ i $(A(t))_{t \in \mathbb{T}}$ są odpowiednio: martyngałem lokalnym i procesem adaptowanym względem filtracji $(\Sigma_t)_{t \in \mathbb{T}}$ o lokalnie ograniczonym wahaniiu, przy czym trajektorie procesu $(A(t))_{t \in \mathbb{T}}$ są prawostronnie ciągłe z lewostronnymi granicami (càdlàg), nazywamy semimartyngałem względem filtracji $(\Sigma_t)_{t \in \mathbb{T}}$.

Semimartyngały są szczególnie ważne w zagadnieniach analizy stochastycznej. Po pierwsze, stanowią najszerszą klasę procesów, dla których określona jest całka stochastyczna Ito. Po drugie, każdy martyngał, podmartyngał i nadmartyngał typu càdlàg jest również semimartyngałem. Ponadto, wariacja kwadratowa jest określona dla dowolnego semimartyngału. Odpowiednie twierdzenia można znaleźć w Protter (2005, roz. 4).

W dalszej części pracy rozważane będą stochastyczne równania różniczkowe względem procesów Lévy'ego. W tym celu zdefiniowana zostanie całka stochastyczna względem semimartyngału³. Jest to szeroka klasa procesów stochastycznych obejmująca także procesy Lévy'ego⁴.

Definicja 1.21.

Proces stochastyczny $(H(t))_{t \geq 0}$ nazywamy prosto przewidywalnym (simply predictable), jeżeli można go przedstawić w postaci

$$H(t) = H(0)\mathbb{1}_{\{0\}}(t) + \sum_{i=1}^n H(i)\mathbb{1}_{(\tau_i, \tau_{i+1}]}(t),$$

gdzie $0 = \tau_1 \leq \tau_1 \leq \dots \leq \tau_n < \infty$ jest ciągiem momentów zatrzymania i H_i jest mierzalna względem Σ^{τ_i} . Przestrzeń wszystkich procesów prosto przewidywalnych oznaczymy przez $\mathcal{S}$.

Możemy teraz określić całkę z procesu prosto przewidywalnego względem procesu *càdlàd*

$$I_X(H) = H(0)X(0) + \sum_{i=1}^n H(i)(X^{\tau_{i+1}} - X^{\tau_i}),$$

gdzie $(X^\tau)_{t \geq 0}$ jest procesem zatrzymanym w τ .

Przekształcenie $I_X(H) : \mathcal{S} \rightarrow \mathbb{D}$ jest liniowe i nie zależy od wyboru reprezentacji procesu H . Można pokazać, że zbiór $\mathcal{S}$ jest gęsty w zbiorze $\mathbb{L}$ z topologią *ucp* (jednostajnie zbieżnie według prawdopodobieństwa na zbiorach zwartych, *uniform convergence on compacts in probability*) określoną następująco: ciąg procesów $(H_n(t))_{t \geq 0}$, $n = 1, 2, \dots$ zbiega w topologii *ucp* do procesu $(H(t))_{t \geq 0}$, jeżeli dla dowolnego $t > 0$ $\sup_{0 \leq s \leq t} |H_n(s) - H(s)|$ zbiega według prawdopodobieństwa do 0. Oznacza, to że każdy proces typu *càglàd* może być jednostajnie aproksymowany poprzez ciąg procesów prosto przewidywalnych.

Zakładając dodatkowo, że proces $(X(t))_{t \geq 0}$ jest semimartyngałem można pokazać, że przekształcenie $I_X(H) : \mathcal{S}_{ucp} \rightarrow \mathbb{D}_{ucp}$ jest ciągłe.

³Przedstawiony zarys teorii całki stochastycznej oparty jest na pracy Behme (2011). Precyzyjne opracowanie stochastycznych równań różniczkowych w oparciu o semimartyngały zawierają monografie Cont i Tankov (2004), Protter (2005), Applebaum (2009).

⁴Twierdzenie 1.4.

Definicja 1.22.

Dla dowolnego semimartynału $(X(t))_{t \geq 0}$ ciągle przekształcenie $I_X(H) : \mathcal{L}_{ucp} \rightarrow \mathbb{D}_{ucp}$ będące rozszerzeniem przekształcenia $I_X(H) : \mathcal{S}_{ucp} \rightarrow \mathbb{D}_{ucp}$ nazywamy całką stochastyczną.

Dla ustalonego procesu $(H(t))_{t \geq 0}$ typu *càglàd* całkę stochastyczną $I_X(H)$ będziemy zapisywać

$$I_X(H) = \int H dX. \quad (1.5)$$

W przypadku, gdy chcemy wyznaczyć wartość procesu $I_X(H)$ w pewnym momencie $t \geq 0$ będziemy używać notacji

$$\int_0^t H(s) dX(s).$$

1.3 Procesy Lévy'ego

Procesy Lévy'ego stanowią jedną z najważniejszych klas procesów stochastycznych. Procesy te zostały nazwane od nazwiska wybitnego matematyka francuskiego Paula Pierre'a Lévy'ego (1886-1971), który pierwszy studiował ich własności w latach trzydziestych XX w. Wiele powszechnie stosowanych procesów stochastycznych jest specjalnymi przypadkami procesów Lévy'ego, na przykład procesy Wienera, Poissona, złożony Poissona, α -stabilne. Procesy Lévy'ego są tak wygodnym narzędziem do modelowania zjawisk zarówno ekonomicznych jak i przyrodniczych, ponieważ ich trajektorie mogą być interpretowane jako ciągły ruch przerywany skokami wartości (punktami nieciągłości). Szczegółowe informacje o procesach Lévy'ego można znaleźć w monografiach Janicki i Izydorczyk (2001), Schountens (2003), Cont i Tankov (2004), Kliber (2013).

Definicja 1.23.

Proces stochastyczny $(L(t))_{t \geq 0}$ nazywamy procesem Lévy'ego, jeżeli

- 1) $\mathbb{P}(L(0) = 0) = 1$,
- 2) dla dowolnego skończonego podzbioru zbioru indeksów $\{t_0, t_1, t_2, \dots, t_n\}$ takiego, że $0 \leq t_1 < t_2 < \dots < t_n$ zmienne losowe $L(t_0), L(t_1) - L(t_0), \dots, L(t_n) - L(t_{n-1})$

są niezależne,

3) dla dowolnych $0 \leq t_1 < t_2$ rozkłady prawdopodobieństwa $L(t_2) - L(t_1)$ oraz $L(t_2 - t_1)$ są jednakowe,

4) proces stochastyczny $(L(t))_{t \geq 0}$ jest ciągły według prawdopodobieństwa, tzn. dla dowolnego $\epsilon > 0$ i $t \geq 0$ zachodzi $\lim_{s \rightarrow t} \mathbb{P}(|L(s) - L(t)| > \epsilon) = 0$.

Pierwszy warunek definicji oznacza, że proces Lévy'ego z prawdopodobieństwem 1 ma początkową wartość równą zero. Drugi warunek definicji to niezależność przyrostów, co oznacza, że dla dowolnych dwóch przedziałów czasowych $[t_0, t_1], [t_2, t_3]$, które się nie nakładają (choć mogą mieć wspólny brzeg) przyrosty $L(t_1) - L(t_0)$ oraz $L(t_3) - L(t_2)$ są niezależne. Własność tą można uogólnić na dowolną skończoną liczbę przedziałów. Trzeci warunek oznacza, że rozkład przyrostu zależy jedynie od długości przedziału. Przyrosty o tej samej długości mają te same rozkłady (proces ma stacjonarne przyrosty). Ostatni warunek oznacza, że momenty nieciągłości procesu Lévy'ego są niedeterministyczne, to znaczy pojawiają się w losowych momentach czasu. Prawdopodobieństwo, że w ustalonym punkcie t pojawi się skok jest równe 0. Nie oznacza to jednak, że trajektorie procesu Lévy'ego są ciągłe. Dla dowolnego procesu Lévy'ego $(L(t))_{t \geq 0}$ można skonstruować wersję procesu z prawostronnie ciągłą z lewostronnymi granicami, która jest także procesem Lévy'ego (Schountens, 2003). W dalszej części pracy będziemy zakładać o procesach Lévy'ego że są właśnie tak dobranymi wariantami.

Zauważmy, że dowolny proces Lévy'ego $(L(t))_{t \geq 0}$ jest określony poprzez pewien rozkład jednowymiarowy zmiennej losowej $L(t)$, gdzie $t > 0$. Istotnie, rozważmy $L(t)$ i ustalmy liczbę naturalną $n > 1$. Wówczas

$$L(t) = \left[L(t) - L\left(t \frac{n-1}{n}\right) \right] + \left[L\left(t \frac{n-1}{n}\right) - L\left(t \frac{n-2}{n}\right) \right] + \dots + \left[L\left(\frac{t}{n}\right) - L(0) \right]. \quad (1.6)$$

Z niezależności i stacjonarności przyrostów wynika, że rozkład zmiennej losowej $L(t)$ jest identyczny z rozkładem sumy n zmiennych losowych o tym samym rozkładzie prawdopodobieństwa. Rozkłady prawdopodobieństwa, które można przedstawić jako sumę niezależnych zmiennych o takim samym rozkładzie nazywane są nieskończenie podzielными.

Definicja 1.24.

Rozkład prawdopodobieństwa F_X zmiennej losowej X nazywamy nieskończenie podzielny, jeżeli dla dowolnej liczby naturalnej $n > 1$ istnieje n niezależnych zmiennych losowych $X_1, \dots, X_n$ o identycznych rozkładach takich, że $X \stackrel{d}{=} X_1 + \dots + X_n$, gdzie $\stackrel{d}{=}$ oznacza równość według rozkładów.

Przykładem rozkładu nieskończenie podzielnego jest rozkład normalny. Wystarczy przyjąć, że zmienne losowe $X_1, \dots, X_n$ mają rozkład normalny $N(\mu/n, \sigma^2/n)$ wówczas zmienna losowa $X \stackrel{d}{=} X_1 + \dots + X_n$ ma rozkład normalny $N(\mu, \sigma^2)$. Przykładem rozkładu, który nie jest nieskończenie podzielny jest rozkład jednostajny Cont i Tankov (2004), str. 69.

Ze wzoru 1.6 widać zatem, że dla każdego procesu Lévy'ego rozkład jednowymiarowy $L(t)$ dla dowolnego $t > 0$ jest rozkładem nieskończenie podzielny. Oznacza to również, że rozkład zmiennej losowej $L(t)$ dla dowolnego $t > 0$ można wyznaczyć ze znajomości rozkładu w dowolnym innym momencie czasu, w szczególności z rozkładem dla $L(1)$. Zachodzi także związek odwrotny: z każdym rozkładem nieskończenie podzielny można związać proces Lévy'ego. Opisuje to następujące twierdzenie.

Twierdzenie 1.2. (Sato, 2014, tw. 1.1) Niech $(L(t))_{t \geq 0}$ będzie procesem Lévy'ego. Wówczas dla każdego $t \geq 0$ rozkład zmiennej losowej $L(t)$ jest nieskończenie podzielny. Odwrotnie, dla dowolnego rozkładu nieskończenie podzielnego F_L istnieje proces Lévy'ego $(L(t))_{t \geq 0}$, taki że $L(1) \sim F_L$.

Z twierdzenia 1.2 wynika ściśle powiązanie rozkładów nieskończenie podzielnych i procesów Lévy'ego. W dalszej części pracy pisząc o rozkładzie brzegowym procesu Lévy'ego $(L(t))_{t \geq 0}$ będziemy mieli na myśli rozkład prawdopodobieństwa dla $L(1)$.

Omówimy teraz kilka najważniejszych przykładów procesów Lévy'ego. Pierwszym i najczęściej wykorzystywanym w modelowaniu zjawisk ekonomicznych jest proces Wienera.

Definicja 1.25.

Proces Lévy'ego $(W(t))_{t \geq 0}$, dla którego przyrosty $W(t_2) - W(t_1)$ (dla dowolnych $0 \leq t_1 < t_2$) mają rozkład normalny z wartością oczekiwaną 0 i wariancją równą $\sigma^2(t_2 - t_1)$ nazywamy procesem Wienera z parametrem dyfuzji σ . W przypadku, gdy

parametr σ jest równy 1, proces Wienera nazywamy standardowym.

W literaturze można się również z innymi określeniami tego procesu: ruch Browna, proces Gaussa. Z definicji procesu Wienera wynika rozkład procesu dla dowolnego $t > 0$: $W(t) \sim N(0, \sigma^2 t)$. Jednowymiarowa funkcja charakterystyczna procesu Wienera dana jest wzorem

$$\varphi_t(\zeta) = \exp\left(\frac{-\sigma^2 \zeta^2 t}{2}\right).$$

Proces Wienera ma prawie wszystkie trajektorie ciągłe i dla każdego procesu Wienera można wybrać modyfikację procesu z ciągłą trajektorią. W dalszej części będziemy posługiwać się wariantami procesu Wienera z ciągłymi trajektoriami. Można również wykazać, że prawie każda trajektoria procesu Wienera ma nieograniczone wahania na dowolnym przedziale oraz prawie wszystkie trajektorie procesu Wienera są nigdzie nieróżniczkowalne (Jakubowski i Sztencel, 2001, str. 312). Wariacja kwadratowa procesu Wienera jest deterministycznie równa $[W](t) = \sigma^2 t$ (Cont i Tankov, 2004, str. 270). Proces Wienera rozważany wraz z filtracją generowaną przez ten proces jest martyngałem (Jakubowski i Sztencel, 2001, str. 313), zatem w długim okresie proces Wienera nie ma tendencji ani do wzrostu ani do spadku. Na rysunku 1.1 przedstawiono wykres przykładowej trajektorii standardowego procesu Wienera.

Kolejnym przykładem procesu Lévy'ego jest proces Wienera z dryfem, zwany także arytmetycznym ruchem Browna.

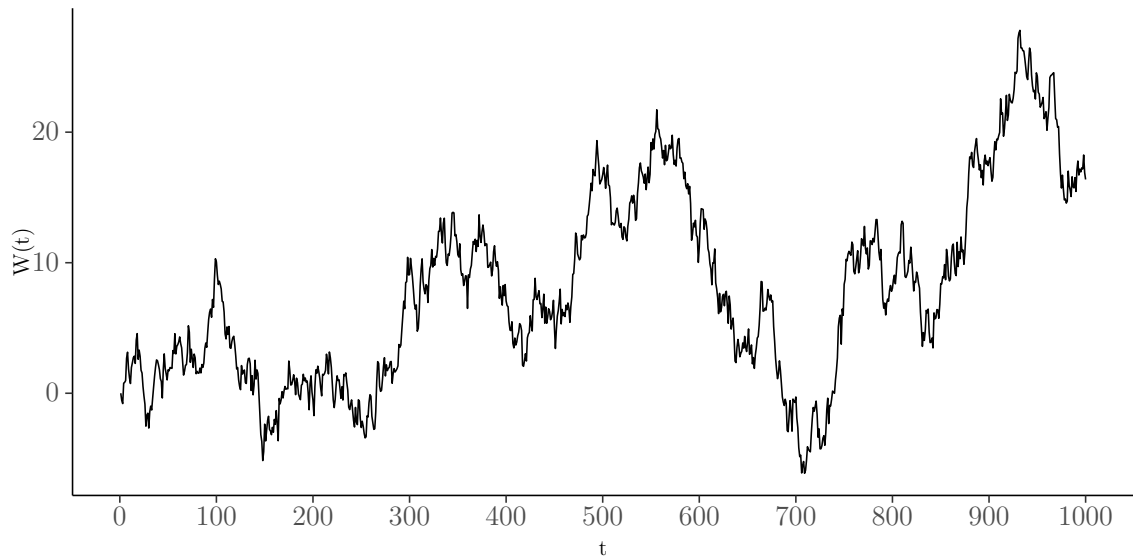
Definicja 1.26.

Procesem Wienera z dryfem (z parametrami dryfu μ i dyfuzji σ) nazywamy proces Lévy'ego $(\tilde{W}(t))_{t \geq 0}$, dla którego przyrosty $\tilde{W}(t_2) - \tilde{W}(t_1)$ (dla dowolnych $0 \leq t_1 < t_2$) mają rozkład normalny z wartością oczekiwaną $\mu(t_2 - t_1)$ i wariancją równą $\sigma^2(t_2 - t_1)$.

Każdy proces Wienera z dryfem $(\tilde{W}(t))_{t \geq 0}$ można przedstawić w postaci

$$\tilde{W}(t) = \mu t + \sigma W(t), \quad t \geq 0,$$

gdzie $(W(t))_{t \geq 0}$ jest standardowym procesem Wienera. Widać zatem, że arytmetyczny ruch Browna ma również trajektorie (prawie wszystkie) ciągłe, ale nieróż-



Rysunek 1.1: Trajektoria standardowego procesu Wienera.

Źródło: opracowanie własne.

niczkiwalne. Przykładami procesu o nieciągłych trajektoriach jest proces Poissona i złożony proces Poissona.

Definicja 1.27.

Proces Lévy'ego $(N(t))_{t \geq 0}$, dla którego przyrosty $N(t_2) - N(t_1)$ (dla dowolnych $0 \leq t_1 < t_2$) mają rozkład Poissona z wartością oczekiwaną $\lambda(t_2 - t_1)$ nazywamy procesem Poissona. Parametr $\lambda > 0$ jest ustaloną stałą zwaną intensywnością procesu Poissona.

Proces Poissona w momencie t ma funkcję rozkładu prawdopodobieństwa

$$p(k) = \mathbb{P}(N(t) = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad \text{dla } k = 0, 1, 2, \dots$$

Rozkład ten przyjmuje z dodatnim prawdopodobieństwem wartości naturalne, a trajektorie procesu Poissona są niemalejące i przedziałami stałe. Skoki procesu Poissona mają wartość 1, a moment pojawienia się kolejnego punktu nieciągłości jest zmienną losową o rozkładzie wykładniczym z parametrem λ (co oznacza średni czas pomiędzy skokami równy $1/\lambda$ jednostki czasu). Proces Poissona można zatem zapisać w

postaci

$$N(t) = \sum_{i=1}^{N(t)} 1. \quad (1.7)$$

Pamiętając, że $\Delta N(t) = 1$ (gdy pojawia się skok) lub $\Delta N(t) = 0$ (w przeciwnym przypadku), można zauważyć, że

$$\Delta N(t) = |\Delta N(t)| = (\Delta N(t))^2.$$

W konsekwencji proces Poissona można zapisać w postaci

$$N(t) = \sum_{s=1}^t \Delta N(s). \quad (1.8)$$

Choć suma we wzorze przebiega wszystkie liczby rzeczywiste z przedziału od $[0, t]$, to składa się tylko ze skończonej ilości składników różnych od zera. Ponadto zachodzi warunek

$$[N](t) = \sum_{i=1}^t (\Delta N(i))^2 = \sum_{i=1}^t \Delta N(i) = N(t).$$

Zatem, wariancja kwadratowa procesu Poissona $(N(t))_{t \geq 0}$ jest również procesem stochastycznym tożsamym z procesem $(N(t))_{t \geq 0}$ (Kliber, 2013, str. 20).

Wartość oczekiwana procesu Poissona $(N(t))_{t \geq 0}$ z parametrem intensywności λ jest równa $\mathbb{E}(N(t)) = \lambda t$ dla $t \geq 0$. Zatem, proces Poissona wraz z filtracją generowaną przez ten proces nie jest martyngałem, ponieważ ma tendencję do wzrostu w czasie. Rozważa się jednak skompensowany proces Poissona dany wzorem

$$\bar{N}(t) = N(t) - \lambda t,$$

który wraz z generowaną przez ten proces filtracją jest martyngałem (Cont i Tankov, 2004, str. 65). Na rysunku 1.2 przedstawiono przykładowe trajektorie procesu Poissona i skompensowanego procesu Poissona. Wybrana wartość parametru intensywności $\lambda = 5$, oznacza, że na jednostkę czasu przypada średnio 5 skoków.

Definicja 1.28.

Złożonym procesem Poissona o parametrze intensywności λ nazywamy proces stochastyczny postaci

$$\tilde{N}(t) = \sum_{n=1}^{N(t)} Y_n, \quad (1.9)$$

gdzie $(N(t))_{t \geq 0}$ jest procesem Poissona, $Y_1, Y_2, \dots$ są niezależnymi zmiennymi losowymi o takich samych rozkładach (*independent and identically distributed, i.i.d.*):

$$Y_1, Y_2, \dots \sim i.i.d. F_Y.$$

Porównując wzory (1.7) oraz (1.9) łatwo zauważyć, że proces Poissona jest specjalnym przypadkiem złożonego procesu Poissona, dla którego skoki są deterministycznie równe 1. Wzór (1.8) można analogicznie uogólnić dla złożonego procesu Poissona do postaci

$$\tilde{N}(t) = \sum_{s=1}^t \Delta \tilde{N}(s). \quad (1.10)$$

Podobnie jak w przypadku wzoru (1.8) suma we wzorze (1.10) choć przebiega wszystkie liczby rzeczywiste $s \in [0, t]$, to składa się tylko ze skończonej ilości elementów, dla których $\Delta \tilde{N}(s) \neq 0$. Skoki te mają rozkład

$$\Delta \tilde{N}(s) \sim F_Y.$$

Czas oczekiwania na kolejne skoki złożonego procesu Poissona jest również zmienną losową o rozkładzie wykładniczym z parametrem λ . Oznacza to, że na jednostkę czasu średnio przypada λ skoków.

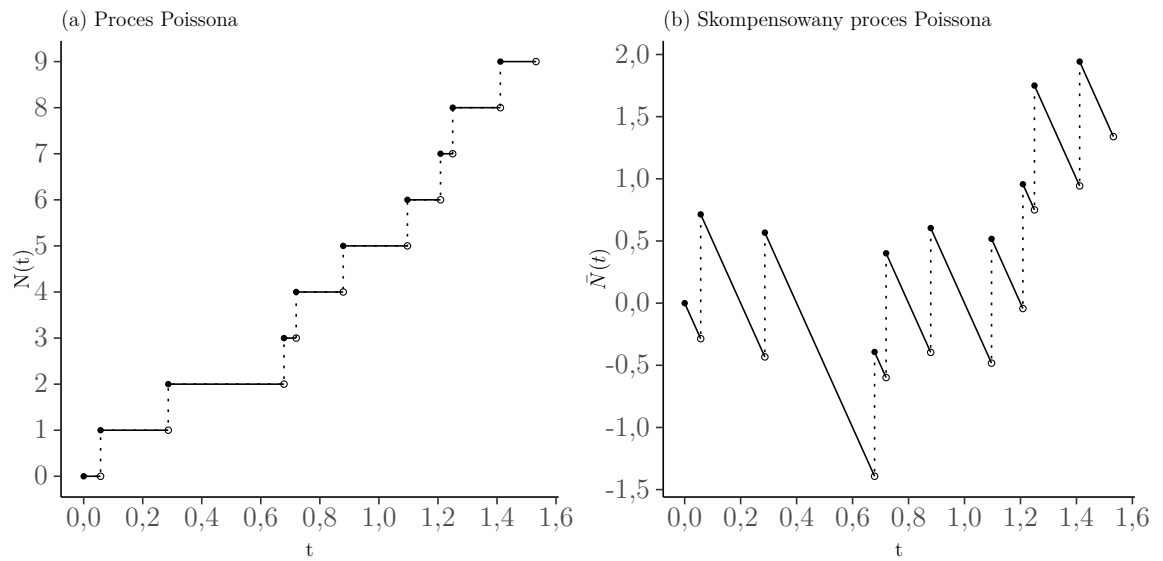
Jeżeli rozkład F_Y ma skończoną wartość oczekiwaną i wariancję, to można wyznaczyć wartość oczekiwaną i wariancję złożonego procesu Poissona $(N(t))_{t \geq 0}$ i wynoszą one odpowiednio $\mathbb{E}(\tilde{N}(t)) = \lambda t \mathbb{E}Y_1$ oraz $\text{Var}(\tilde{N}(t)) = \lambda t \text{Var}Y_1$. Można wówczas także wyznaczyć skompensowany złożony proces Poissona

$$\bar{\tilde{N}}(t) = \tilde{N}(t) - t\lambda \mathbb{E}Y_1,$$

który wraz z filtracją generowaną przez ten proces jest martyngałem.

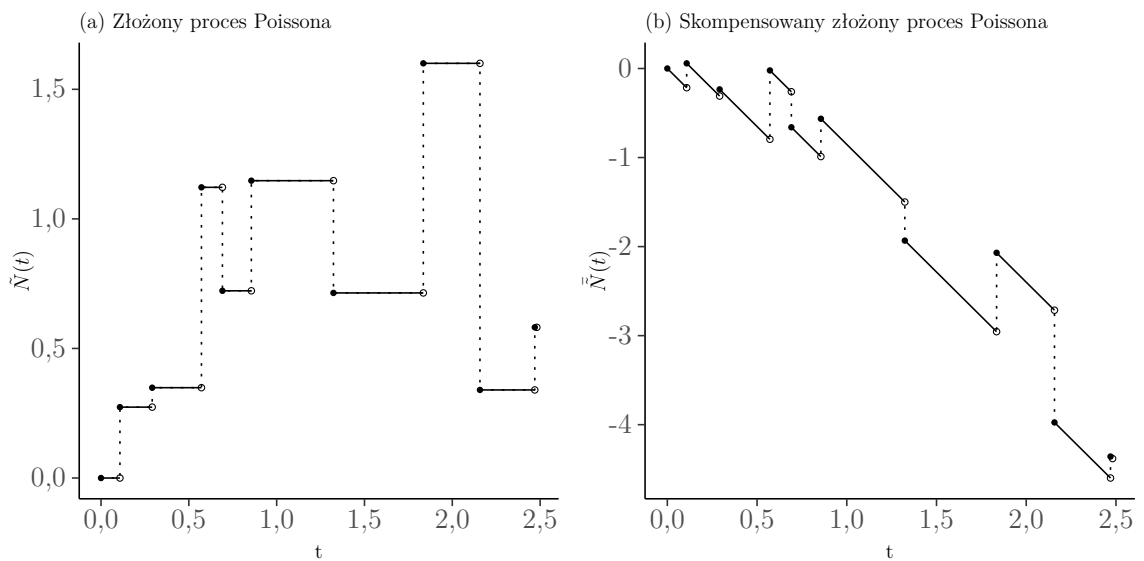
Na rysunku 1.2 przedstawiono trajektorie złożonego procesu Poissona i skompensowanego złożonego procesu Poissona, dla których rozkład skoków F_Y jest rozkładem normalnym standardowym.

Procesy Lévy'ego mają punkty nieciągłości zwane skokami. Nie zawsze jednak ilość skoków w jednostce czasu jest skończona jak w przypadku procesu Poissona czy złożonego procesu Poissona, ale zawsze jest co najwyżej przeliczalna. Z ciągłości według prawdopodobieństwa procesu Lévy'ego (definicja 1.23 [p. 4]) wynika że



Rysunek 1.2: Trajektoria procesu Poissona z parametrem intensywności $\lambda = 5$ (a) oraz odpowiadająca temu procesowi trajektoria skompensowanego procesu Poissona (b).

Źródło: opracowanie własne.



Rysunek 1.3: Trajektoria złożonego procesu Poissona z parametrem intensywności $\lambda = 2$ ze skokami o rozkładzie normalnym standardowym (a) oraz odpowiadająca temu procesowi trajektoria skompensowanego złożonego procesu Poissona (b).

Źródło: opracowanie własne.

dla ustalonego $t \geq 0$ zachodzi warunek $\Delta L(t) = 0$ (prawie na pewno). Oznacza to jednak tylko, że proces Lévy'ego nie ma ustalonych deterministycznie momentów nieciągłości. Dla dowolnego $t > 0$ suma

$$\sum_{s \leq t} \Delta L(s)$$

jest dobrze określona (zawiera co najwyżej skończoną ilość elementów), ale może nie być zbieżna (Iacus, 2011, str. 135). Można określić miarę, która będzie mierzyła intensywność pojawiania się skoków procesów Lévy'ego.

Definicja 1.29.

Funkcję $\mu : \Omega \times \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ nazywamy miarą losową, jeżeli:

- 1) dla każdego $\omega \in \Omega$, $\mu(\omega, \cdot)$ jest miarą na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,
- 2) dla każdego $A \in \mathcal{B}(\mathbb{R})$, $\mu(\cdot, A)$ jest funkcją mierzalną.

Dla ustalonego $t \geq 0$ można określić następującą miarę losową określającą liczącą liczbę skoków procesu $(L(t))_{t \geq 0}$ o wielkości A występujących do czasu t :

$$\mu^L(\omega; t, A) = \# \{0 \leq s \leq t : \Delta L(\omega, s) \neq 0 \text{ i } \Delta L(\omega, s) \in A\} = \sum_{s \leq t} \mathbb{1}_A(\Delta L(\omega, s)), \quad (1.11)$$

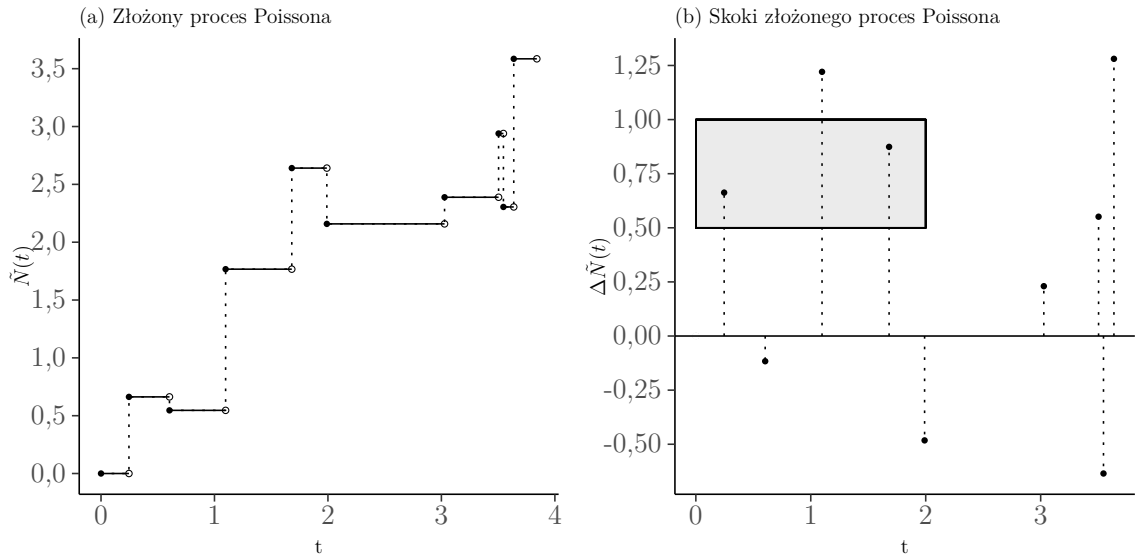
gdzie funkcja $\mathbb{1}_A$ jest indykátorem zbioru A . Na rysunku 1.4 przedstawioną przykładową trajektorię złożonego procesu Poissona oraz odpowiadający mu proces skoków. Miara losowa dla tego procesu dla zbioru $A = [0, 5, 1]$ i $t = 2$ przyjmuje wartość 2.

Miara zadana wzorem (1.11) dla ustalonego $t \geq 0$ jest zmienną losową (definicja 1.29 [p. 2]), zatem można wyznaczyć jej wartość oczekiwaną dla $t = 1$. W ten sposób powstaje funkcja przypisująca zbiorom $A \in \mathcal{B}(\mathbb{R})$ pewną nieujemną liczbę rzeczywistą. Funkcję tą nazywamy miarą Lévy'ego. Precyzyjnie określa to następująca definicja.

Definicja 1.30.

(Cont i Tankov, 2004, str. 88) Miarą Lévy'ego procesu $(L(t))_{t \geq 0}$ nazywamy miarę ν określoną na przestrzeni mierzalnej $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ następująco

$$\nu(A) = \mathbb{E}(\mu^L(1, A)) = \mathbb{E}\left(\sum_{s \leq 1} \mathbb{1}_A(\Delta L(\omega, s))\right),$$



Rysunek 1.4: Trajektoria złożonego procesu Poissona z parametrem intensywności $\lambda = 0,1$ ze skokami o rozkładzie normalnym standardowym (a) oraz odpowiadająca temu procesowi trajektoria procesu skoków wraz zaznaczonym zbiorem $[0, 2] \times [0, 5, 1]$ (b).

Źródło: opracowanie własne.

przyjmując dodatkowo $\nu(\{0\}) = 0$.

Miara Lévy'ego określa oczekiwaną liczbę skoków o określonej wielkości A w przedziale czasowym o długości 1. Miara ta jest σ -skończona, ale nie jest miarą probabilistyczną, ponieważ nie spełnia warunku unormowania (Papapantoleon, 2008). Spełnia natomiast następujący warunek (Cont i Tankov, 2004, str. 91):

$$\int_{\mathbb{R}} \min \{1, |x|^2\} \nu(dx) < +\infty.$$

Na podstawie własności miary Lévy'ego można rozróżnić dwa rodzaje procesów Lévy'ego, co przedstawia poniższe twierdzenie.

Twierdzenie 1.3. (Iacus, 2011, tw. 3.18.7) Dla dowolnego procesu Lévy'ego $(L(t))_{t \geq 0}$ o mierze Lévy'ego ν zachodzi jeden z dwóch warunków:

- 1) jeżeli $\nu(\mathbb{R}) < +\infty$, to niemal wszystkie trajektorie procesu Lévy'ego mają skończoną ilość skoków na każdym przedziale zwartym (proces ma skończoną aktywność (finite activity)),

2) jeżeli $\nu(\mathbb{R}) = +\infty$, to niemal wszystkie trajektorie procesu Lévy'ego mają nieskończoną ilość skoków na każdym przedziale zwartym (proces ma nieskończoną aktywność (*infinite activity*)).

Przykładem procesu Lévy'ego o skończonej aktywności jest proces Poissona. Oczekiwana wartość pojawienia się skoku o wielkości 1 jest równa parametrowi intensywności λ . Miara Lévy'ego procesu Poissona przyjmuje zatem postać $\nu(A) = \lambda \mathbb{1}_{\{1\}}(A)$. Przykładami procesów o nieskończonej aktywności są procesy gamma i odwrotny Gaussa przedstawione w dalszej części rozdziału.

W przypadku, gdy proces Lévy'ego ma nieskończoną aktywność, należy zwrócić uwagę na małe skoki: suma wszystkich skoków co do wartości bezwzględnej mniejszych od dowolnego $\epsilon > 0$ nie jest zbieżna. W kontekście miary Lévy'ego oznacza to, że dla dowolnego zbioru otwartego A zawierającego 0, miara Lévy'ego przyjmuje wartość $\nu(A) = +\infty$. Można jednak w takim przypadku „podzielić” miarę Lévy'ego na dwie części $\nu = \nu_1 + \nu_2$. Pierwsza część wydziela obszar wokół 0, na przykład przedział $(-1, 1)$, tj. $\nu_1(A) = \nu(A \cap (-1, 1))$. Druga część mierzy skoki co do wartości bezwzględnej równe co najmniej 1: $\nu_2(A) = \nu(A \cap ((-\infty, -1] \cup [1, +\infty)))$. Pierwszy składnik miary jest granicą ciągu skompensowanych złożonych procesów Poissona, a drugi składnik określa pewien złożony proces Poissona. Dokładniej określa to następujące twierdzenie.

Twierdzenie 1.4 (Dekompozycja Lévy'ego-Ito). (*Cont i Tankov, 2004, tw. 3.7*)
Dowolny proces Lévy'ego $(L(t))_{t \geq 0}$ można zdekomponować na trzy składniki:

$$L(t) = L^1(t) + L^2(t) + L^3(t)$$

gdzie:

$L^1(t)$, to proces Wienera z dryfem,

$L^2(t)$ jest złożonym procesem Poissona odpowiadającym za skoki o wielkości większej niż 1,

$L^3(t)$ jest martyngałem będącym granicą (prawie na pewno) skompensowanych złożonych procesów Poissona

$$L^3(t) = \lim_{\epsilon \rightarrow 0^+} L^\epsilon(t),$$

o mierze $\nu^\epsilon(A) = \nu(A \cap ((-1, -\epsilon) \cup (\epsilon, 1)))$ odpowiadających za skoki o wielkości $(-1, -\epsilon) \cup (\epsilon, 1)$.

Z dekompozycji Lévy'ego-Ito wynika, że każdy proces Lévy'ego $(L(t))_{t \geq 0}$ jest semimartyngałem. Porównując twierdzenie 1.4 z definicją 1.20 widać, że jako martyngał lokalny $M(t)$ należy przyjąć proces $L^3(t) + L^2(t)$, a jako $A(t)$ proces $L^1(t)$. Kolejne twierdzenie charakteryzuje dowolny proces Lévy'ego za pomocą funkcji charakterystycznych.

Twierdzenie 1.5 (Reprezentacja Lévy'ego-Chinczyna). (*Iacus, 2011, tw. 3.18.7*)
Dla dowolnego procesu Lévy'ego $L = (L(t))_{t \geq 0}$ jednowymiarowa funkcja charakterystyczna tego procesu przyjmuje postać

$$\varphi(\zeta) = \exp(t\psi(\zeta)) = \exp \left[t \left(\mu\zeta - \frac{1}{2}\sigma^2\zeta^2 + \int_{\mathbb{R}} (e^{i\zeta x} - 1 - i\zeta x \mathbb{1}_{|x| < 1}) \nu(dx) \right) \right], \quad (1.12)$$

gdzie $\mu \in \mathbb{R}$, $\sigma \geq 0$, natomiast ν jest σ -skończoną miarą na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ spełniającą

$$\int_{\mathbb{R}} \min\{1, x^2\} \nu(dx) < +\infty \quad \text{oraz} \quad \nu(\{0\}) = 0.$$

Odwrotnie, każda funkcja jednowymiarowa funkcja charakterystyczna postaci (1.12) wyznacza pewien proces Lévy'ego.

Funkcja $\psi(\zeta)$ ze wzoru (1.12) jest nazywana eksponentą charakterystyczną procesu Lévy'ego $(L(t))_{t \geq 0}$. Ponadto, zgodnie z twierdzeniem 1.2 $\tilde{L} \stackrel{d}{=} L(1)$ jest zmienną losową o rozkładzie nieskończenie podzielny. Funkcja charakterystyczna tej zmiennej losowej jest zatem postaci

$$\varphi_{\tilde{L}}(\zeta) = \exp(\psi(\zeta)).$$

Z reprezentacji Lévy'ego-Chinczyna wynika, że każdy proces Lévy'ego może składać się z trzech komponentów: części liniowej deterministycznej (dryfu), procesu Wienera z parametrem dyfuzji σ (te dwa składniki łącznie określają proces Wienera z dryfem) oraz części czysto skokowej opisanej przez miarę Lévy'ego $\nu(dx)$. Zatem każdy proces Lévy'ego można scharakteryzować poprzez trójkę Lévy'ego: $(\mu, \sigma, \nu(dx))$, gdzie μ jest współczynnikiem dryfu, σ odchyleniem standardowym

procesu Wienera (parametrem dyfuzji), $\nu(dx)$ jest miarą Lévy'ego mierzącą intensywność pojawiania się skoków. Z twierdzenia 1.5 wynika także odwrotna zależność: każda trójka charakterystyczna Lévy'ego $(\mu, \sigma, \nu(dx))$ określa pewien proces Lévy'ego $(L(t))_{t \geq 0}$.

Proces Wienera ma w reprezentacji Lévy'ego-Chinczyna postać $(0, 1, 0)$, proces Wienera z dryfem $(\mu, \sigma, 0)$, proces Poissona $(0, 0, \lambda \delta_{\{1\}})$, gdzie $\delta_{\{1\}}$ jest deltą Diraca skoncentrowaną w 1. Natomiast dla złożonego procesu Poissona z intensywnością λ i rozkładem skoków F_Y trójka charakterystyczna przyjmuje postać $(0, 0, \lambda F_Y(dx))$.

Definicja 1.31.

Jeżeli miara Lévy'ego ν jest absolutnie ciągła względem miary Lebesgue'a, to jej pochodną Radona–Nikodyma względem miary Lebesgue'a, to jest funkcję

$$u(x) = \frac{\nu(dx)}{dx}$$

nazywamy gęstością Lévy'ego miary ν .

Kolejne dwa przykłady procesów Lévy'ego które zostaną wykorzystane w dalszej części pracy mają nieskończoną aktywność i przyrosty odpowiednio o rozkładzie gamma oraz odwrotnym Gaussa (*inverse Gaussian*)⁵.

Definicja 1.32.

Procesem gamma nazywamy proces Lévy'ego $(L(t))_{t \geq 0}$, taki że dla każdego $t > 0$ zmienna losowa $L(t)$ ma rozkład gamma $\text{Ga}(\nu t, \alpha)$.

Trójka charakterystyczna procesu gamma przyjmuje postać (Schoutens, 2003)[s. 52]

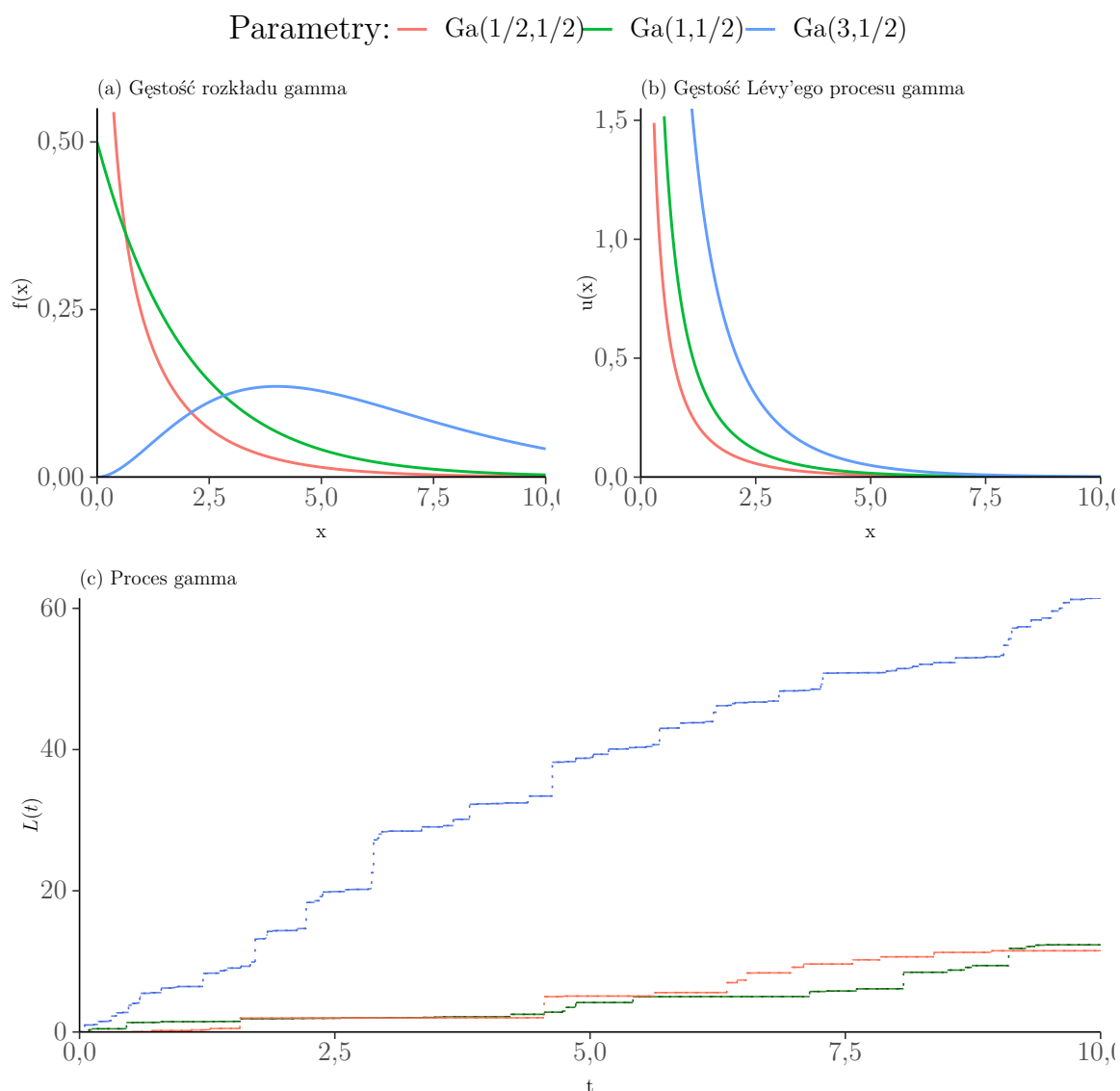
$$\left(\frac{\nu}{\alpha} (1 - \exp(-\alpha)), 0, \nu \exp(-\alpha x) x^{-1} \mathbb{1}_{(0, +\infty)}(x) dx \right) \quad (1.13)$$

Proces gamma jest czystko skokowy (tzn. bez składnika gaussowskiego), o nieskończonej aktywności. Ze wzoru (1.13) wynika, że gęstość Lévy'ego procesu gamma jest postaci

$$u(x) = \frac{\nu}{x} e^{-\alpha x} \mathbb{1}_{(0, +\infty)}(x).$$

⁵Wzory na gęstość, funkcje charakterystyczne oraz momenty dla rozkładów gamma, odwrotnego Gaussa oraz uogólnionego odwrotnego rozkładu Gaussa znajdują się w podrozdziale 2.5.

Miara Lévy'ego procesu gamma jest skoncentrowana na dodatniej półosi, zatem skoki procesu gamma są dodatnie. Ponieważ parametr dryfu jest nieujemny, to trajektorie procesu gamma są nieujemne i niemalejące (prawie na pewno).



Rysunek 1.5: Gęstości rozkładu gamma (a), gęstości Lévy'ego procesu gamma (b) oraz symulacje trajektorii procesu gamma (c) odpowiadające trzem rozkładom gamma

$$Ga(\nu, \alpha).$$

Źródło: opracowanie własne.

Definicja 1.33.

Procesem odwrotnym Gaussa nazywamy proces Lévy'ego $(L(t))_{t \geq 0}$, taki że dla każdego $t > 0$ zmienna losowa $L(t)$ ma rozkład odwrotny Gaussa $IG(\delta t, \gamma t^2)$.

Proces odwrotny Gaussa ma w reprezentacji Lévy'ego-Chinczyna następującą trójkę charakterystyczną

$$\left(\frac{\delta}{\gamma} (2\Phi_N(\gamma)), 0, \frac{\delta x^{-3/2}}{\sqrt{2\pi}} \exp(-\gamma^2 x) \mathbb{1}_{(0, +\infty)}(x) dx \right),$$

gdzie $\Phi_N(\cdot)$ jest gęstością rozkładu normalnego standardowego. Proces odwrotny Gaussa podobnie jak proces gamma jest niemalejącym, nieujemnym, czystko skokowym procesem.

Rozkładem, który uogólnia zarówno rozkład gamma jak i odwrotny Gaussa jest uogólniony odwrotny rozkład Gaussa $GIG(\nu, \delta, \gamma)$. Uogólniony odwrotny rozkład Gaussa jest również rozkładem nieskończenie podzielnym. Można z nim związać proces Lévy'ego.

Definicja 1.34.

Proces Lévy'ego $(L(t))_{t \geq 0}$, dla którego przyrosty $L(t+s) - L(t)$ (dla dowolnych $t \geq 0$ oraz $s > 0$) mają rozkład o funkcji charakterystycznej $(\varphi_{GIG}(\zeta; \nu, \delta, \gamma))^s$ nazywamy uogólnionym odwrotnym procesem Gaussa.

Uogólniony odwrotny proces Gaussa ma gęstość Lévy'ego postaci (Barndorff-Nielsen i Shephard, 2001b)

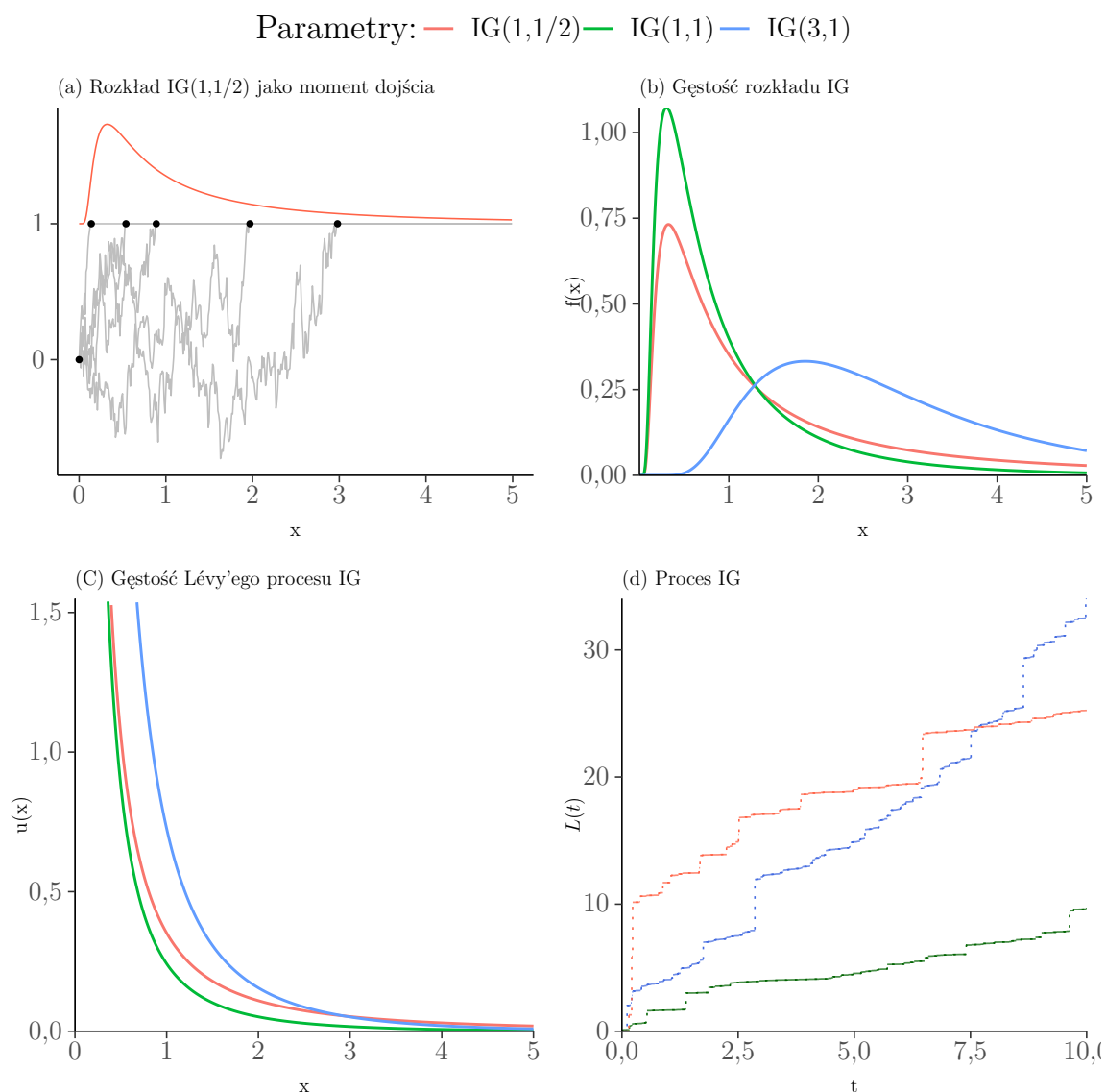
$$u(x) = \frac{1}{x} \exp(-\gamma^2 x/2) \left(\frac{1}{2} \int_0^{+\infty} \exp(-\frac{1}{2} \delta^{-2} x z) g(z) dz + \max\{0, \nu\} \right), \quad (1.14)$$

gdzie

$$g(x) = \frac{2}{\pi^2 x} \left(J_{|\gamma|}^2(\sqrt{x}) + N_{|\gamma|}^2(\sqrt{2x}) \right)^{-1},$$

gdzie J_ν i N_ν są funkcjami Besella odpowiednio pierwszego i drugiego rodzaju. Uogólniony odwrotny proces Gaussa jest procesem o nieskończonej aktywności.

Uogólniony odwrotny rozkład Gaussa pełni bardzo ważną rolę w modelowaniu stóp zwrotu jako rozkład miksujący (*mixture distribution*) w normalnych mieszaninach średnio-wariacyjnych (*normal mean-variance mixture*) dla uogólnionego rozkładu hiperbolicznego. Przedstawia to poniższe twierdzenie.



Rysunek 1.6: Rozkład IG(1,1/2) jako moment dojścia do poziomu $\delta = 1$ procesu Wienera z dryfem $\gamma = 1/2$ (por. podrozdział 2.5) oraz gęstości rozkładu IG (b), gęstości Lévy'ego procesu IG (c) oraz trajektorie procesu IG (d) odpowiadające trzem rozkładom odwrotnym Gaussa $IG(\delta, \gamma)$.

Źródło: opracowanie własne.

Twierdzenie 1.6. (*Breymann i Lüthi, 2013*) Jeżeli zmienną losową X można zapisać w postaci

$$X = \mu + W\varepsilon + \sqrt{W}Z, \quad (1.15)$$

gdzie: $Z \sim N(0, 1)$, $W \sim GIG(\nu, \delta, \gamma)$, to zmienna X ma rozkład uogólniony hiperboliczny⁶. W szczególności, gdy $W \sim IG(\delta, \gamma)$ to X ma rozkład normalny odwrotny Gaussa NIG (*Normal Inverse Gaussian*), a gdy $W \sim Ga(\nu, \alpha)$ to X ma rozkład VG (*Variance Gamma*). Ponadto, rozkład warunkowy $X|W = w$ jest rozkładem normalnym postaci

$$X|W \sim N(\mu + w\varepsilon, w). \quad (1.16)$$

Uogólnionego rozkład hiperboliczny ma półciężkie (*semi-heavy*) ogony (Schoutens, 2003, str. 65), w szczególności

$$f_{GH}(x, \alpha, \beta, \delta, \nu) \sim |x|^{\nu-1} \exp((\mp)\alpha + \beta)x, \quad x \rightarrow \pm\infty,$$

gdzie $\sim$ oznacza proporcjonalność z dokładnością do multiplikatywnej stałej. Właśność ta oraz możliwość uwzględnienia asymetrii rozkładu powoduje bardzo dobre dopasowanie do obserwowanych empirycznie stóp zwrotu, m.in. Eberlein i Keller (1995) oraz Prause (1999) pokazali doskonałe dopasowanie do dziennych logarytmicznych stóp zwrotów zarówno dla pojedynczych spółek jak i portfeli spółek na podstawie danych pochodzących z wiodących niemieckich spółek. Z twierdzenia 1.6 wynika zatem, że uogólniony odwrotny rozkład Gaussa może pełnić rolę rozkładu stacjonarnego dla procesu wariancji pozwalając uzyskać w rezultacie modele o bardzo dobrym dopasowaniu do empirycznie obserwowanych rozkładów stóp zwrotu. Również rozkłady VG oraz NIG cechują się bardzo dobrym dopasowaniem do dziennych logarytmicznych stóp zwrotów i były wielokrotnie stosowane do modelowania zjawisk finansowych (por. Madan i Seneta (1990); Madan i in. (1998) dla rozkładu VG oraz Barndorff-Nielsen (1997); Albrecher i Predota (2004) dla NIG).

⁶Wzory na gęstość, funkcje charakterystyczną oraz momenty dla uogólnionego odwrotnego rozkładu hiperbolicznego, NIG oraz VG znajdują się w dodatku A.

1.4 Podklasy procesów Lévy’ego

Ważną podklasę procesów Lévy’ego stanowią procesy podporządkowane (*subordinator*, por. Feller (1977, str. 310) oraz Kliber (2013, str. 60)). Są to procesy stosowane do modelowania czasu dla pewnego innego procesu stochastycznego (tzw. zrandomizowane czasy operacyjne Feller (1977), str. 310). Proces stochastyczny modelujący ewolucję czasu musi być niemalejący, aby czas nie mógł się „cofać”. Warunek ten spełniają procesy podporządkowane.

Definicja 1.35.

Procesem podporządkowanym nazywamy proces Lévy’ego $(Z(t))_{t \geq 0}$ o niemalejących (prawie na pewno) trajektoriach.

Przykładem procesu podporządkowanego jest proces Poissona, ponieważ proces ten ma trajektorie przedziałami stałe, które skokowo rosną o jednostkę. Procesem podporządkowanym może być także złożony proces Poissona, pod warunkiem, że skoki mają miarę Lévy’ego skoncentrowaną na dodatniej półprostej. Przykładami procesów podporządkowanych o nieskończonej aktywności są procesy gamma i odwrotny Gaussa. Natomiast proces Wienera nie jest procesem podporządkowanym, ponieważ dla tego procesu prawdopodobieństwo wzrostu w dowolnej jednostce czasu jest takie same jak spadku i wynosi $1/2$. Kolejne twierdzenie opisuje postać funkcji charakterystycznej dla procesu podporządkowanego.

Twierdzenie 1.7. (*Applebaum, 2009, tw. 1.3.15*) *Jednowymiarowa funkcja charakterystyczna procesu podporządkowanego $(Z(t))_{t \geq 0}$ dana jest wzorem*

$$\varphi(\zeta) = \exp \left[t \left(\mu i \zeta + \int_0^{+\infty} (e^{i \zeta x} - 1) \nu(dx) \right) \right], \quad (1.17)$$

gdzie $\mu \geq 0$. Miara Lévy’ego ν procesu podporządkowanego $(Z(t))_{t \geq 0}$ spełnia ponadto warunek

$$\nu((-\infty, 0)) = 0. \quad (1.18)$$

Odwrotnie, każda funkcja charakterystyczna postaci (1.17) z miarą Lévy’ego spełniającą warunek (1.18) wyznacza pewien proces podporządkowany $(Z(t))_{t \geq 0}$.

Z twierdzenia 1.7 wynika, że procesy podporządkowane nie mają składnika gaussowskiego (w reprezentacji Lévy’ego-Chinczyna drugi składnik ma wartość 0). Dodatkowe warunki dla miary Lévy’ego zapewnia nieujemność skoków trajektorii.

Zastosowanie procesów podporządkowanych do modelowania czasu dla pewnego innego procesu stochastycznego przedstawia następujące twierdzenie.

Twierdzenie 1.8. *(Cont i Tankov, 2004, tw. 4.3) Dla dowolnej miary Lévy’ego ν i dowolnego $\mu \in \mathbb{R}$ istnieje proces Lévy’ego $(L(t))_{t \geq 0}$ z miarą Lévy’ego ν , taki że dla dowolnego $t \geq 0$*

$$L(t) = \mu Z(t) + W(Z(t)), \quad (1.19)$$

dla pewnego procesu podporządkowanego $(Z(t))_{t \geq 0}$ oraz niezależnego od tego procesu porządkowanego procesu Wienera $(W(t))_{t \geq 0}$ wtedy i tylko wtedy, gdy

- 1) *miara Lévy’ego ν ma gęstość $u(x)$*
- 2) *$u(x)e^{-\mu x} = u(-x)e^{\mu x}$*
- 3) *funkcja $f(x) = u(\sqrt{x})e^{-\mu\sqrt{x}}$ jest ściśle monotoniczna na przedziale $(0, +\infty)$.*

W matematyce finansowej twierdzenie 1.8 wykorzystuje się do modelowania ewolucji w czasie cen przekształconych przez funkcję logarytmiczną (wówczas przyrosty są logarytmicznymi stopami zwrotu). W tym przypadku, proces podporządkowany interpretuje się jako „rynkową skalę czasu” (Kliber, 2013, s. 60). Proces ewolucji czasu $(Z(t))_{t \geq 0}$ jest losowy, ponieważ rynek reaguje na napływ informacji raz szybciej, raz wolniej. Nie można jednak z góry przewidzieć szybkości reakcji rynku. Zauważmy ponadto, że specjalnym przypadkiem wzoru (1.19) dla $Z(t) = t$ jest standardowy model Blacka-Scholesa, w którym cena podlega geometrycznemu ruchowi Browna, a logarytmiczne stopy zwrotu mają rozkład normalny (Kliber, 2013, s. 60).

Twierdzenie 1.8 jest odpowiednikiem twierdzenia 1.6 dla procesów stochastycznych. Zgodnie z twierdzeniem 1.8, gdy jako procesy czasu operacyjnego wykorzystuje się procesy gamma oraz odwrotny Gaussa, to jako model logarytmu cen postaci (1.19) otrzymuje się odpowiednio proces VG oraz normalny odwrotny Gaussa NIG.

Kolejną ważną podklasą procesów Lévy’ego są procesy o rozkładach samorozkładalnych (*self decomposable*).

Definicja 1.36.

Rozkład jednowymiarowej zmiennej losowej X nazywamy samorozkładalnym jeżeli dla dowolnego $c \in (0, 1)$ istnieje zmienna losowa $X^{(c)}$ niezależna od X , taka że

$$X \stackrel{d}{=} cX + X^{(c)}. \quad (1.20)$$

Proces Lévy'ego $(L(t))_{t \geq 0}$, dla którego rozkład brzegowy $L(1)$ jest rozkładem samorozkładalnym nazywamy procesem samorozkładalnym.

Zgodnie ze wzorem (1.20) zmienna losowa ma rozkład samorozkładalny, jeżeli da się ją przedstawić w postaci sumy przeskalowanej w dół wersji tej zmiennej oraz pewnej niezależnej zmiennej rezydualnej. Warunek z definicji 1.36 można równoważnie zapisać w postaci funkcji charakterystycznych. Zmienna losowa X ma rozkład samorozkładalny dokładnie wtedy, gdy jej funkcję charakterystyczną można dla dowolnego $c \in (0, 1)$ zapisać w postaci

$$\varphi_X(\zeta) = \varphi_X(c\zeta)\varphi_c(\zeta), \quad (1.21)$$

gdzie φ_c należy do pewnej rodziny funkcji charakterystycznych $\{\varphi_c : c \in (0, 1)\}$.

Każdy rozkład samorozkładalny jest rozkładem nieskończenie podzielny, ale nie każdy rozkład nieskończenie podzielny jest samorozkładalny. Opisuje to następujące twierdzenie.

Twierdzenie 1.9. (*Sato, 1999, wniosek 15.11*) *Miara Lévy'ego ν procesu samorozkładalnego $(L(t))_{t \geq 0}$ ma gęstość postaci*

$$u(x) = \frac{k(x)}{|x|},$$

gdzie: $k : \mathbb{R} \rightarrow \mathbb{R}^+$ jest funkcją o ograniczonym wahanu, malejącą na $(0, +\infty)$ i rosnącą na $(-\infty, 0)$.

Przykładami rozkładów samorozkładalnych są m.in rozkład normalny, t-Studenta, gamma, odwrotny Gaussa (Barndorff-Nielsen i Shephard, 2001b).

Rozkłady samorozkładalne są interesujące w zastosowaniach ekonomicznych, ponieważ są rozkładami granicznymi wystandaryzowanych sum pewnych (niekoniecznie identycznych) niezależnych zmiennych losowych. Niech $X_1, X_2, \dots$ będzie ciągiem

niezależnych zmiennych losowych. Oznaczmy przez $S_n = \sum_{k=1}^n X_k$. Załóżmy, że istnieją ciągi: centrujący $(c_n)_{n \in \mathbb{N}}$ oraz skalujący $(b_n)_{n \in \mathbb{N}}$, takie że $b_n S_n + c_n$ zbiega (według rozkładu) do pewnej zmiennej losowej X . Wówczas zmienna X ma rozkład samorozkładalny (Carr i in., 2007). Początkowo rozkłady będące granicą pewnych wystandaryzowanych sum badali Lévy (1937) oraz Chinczyn (1938) (za: Carr i in. (2007)). Ten drugi nazwał tę klasę rozkładów jako klasę L (Carr i in., 2007). Klasa L jest identyczna z klasą rozkładów samorozkładalnych (twierdzenie wraz z dowodem przedstawia Sato (1999, twr. 15.3)).

Rozkłady samorozkładalne stanowią istotne uogólnienie centralnego twierdzenia granicznego, w którym czynnik skalujący może być różny od $1/\sqrt{n}$. Jest to istotne, ponieważ ceny aktywów finansowych interpretuje się zazwyczaj jako zagregowane sumy niezależnych transakcji i uwolnienie od czynnika skalującego $1/\sqrt{n}$ (prowadzącego do modeli gaussowskich) doprowadza do ogólniejszej klasy modeli samorozkładalnych (Trabs, 2014). Ma to szczególne zastosowanie w przypadku wykładniczo ważonych średnich ruchomych, często stosowanych w przypadku analizy finansowych szeregów czasowych (Carr i in., 2007). W rozdziale drugim rozkłady samorozkładalne zostaną wykorzystane jako rozkłady stacjonarne procesów wariancji chwilowej.

1.5 Proces Ornsteina-Uhlenbecka

Proces Ornsteina-Uhlenbecka został zaproponowany przez dwóch holenderskich fizyków Leonarda Ornsteina (1880-1941) i George’a Uhlenbecka (1900-1988) w pracy dotyczącej ruchu cząsteczki w płynie (Uhlenbeck i Ornstein, 1930). Proces Ornsteina-Uhlenbecka jest procesem łączącym wiele istotnych z punktu widzenia zastosowań własności: ma stacjonarne przyrosty, jest gaussowski i markowski. Ponadto jest uważany za ciągły odpowiednik dyskretnego procesu autoregresyjnego rzędu pierwszego AR(1). Proces Ornsteina-Uhlenbecka znalazł liczne zastosowanie w fizyce (Blomberg, 2017), biologii (Beaulieu i in., 2012), oraz ekonomii. W tej ostatniej najbardziej znane jest zastosowanie procesu Ornsteina-Uhlenbecka do modelowania stóp procentowych (Vasicek, 1977). Ważną rolę proces Ornsteina-Uhlenbecka pełni również w modelowaniu zmienności instrumentów finansowych, jako proces modelujący

zmiennosc (Stein i Stein, 1991) lub logarytm zmiennosci (Wiggins, 1987).

Definicja 1.37.

Proces stochastyczny $X = (X(t))_{t \geq 0}$ nazywany jest procesem Ornsteina-Uhlenbecka, jeżeli jest rozwiązaniem następującego stochastycznego równania różniczkowego

$$dX(t) = \lambda(\mu - X(t))dt + \sigma dW(t). \quad (1.22)$$

gdzie $\mu, \theta \in \mathbb{R}, \sigma > 0$ są parametrami, natomiast $W = (W(t))_{t \geq 0}$ jest standardowym procesem Wienera.

Wartość początkowa X_0 procesu X jest zmienną losową (możliwie stałą), niezależną od procesu Wienera W .

Rozwiązanie równania (1.22) jest postaci

$$X(t) = X(0)e^{-\lambda t} + \mu(1 - e^{-\lambda t}) + \sigma \int_0^t e^{-\lambda(t-s)} dW(s) \quad (1.23)$$

Warunkowa wartość oczekiwana procesu Ornsteina-Uhlenbecka jest równa $t > s$

$$\mathbb{E}[X(t)|X(s) = x_0] = x_0 e^{-\lambda(t-s)} + \mu(1 - e^{-\lambda(t-s)}) \quad (1.24)$$

natomiast warunkowa wariancja przyjmuje postać

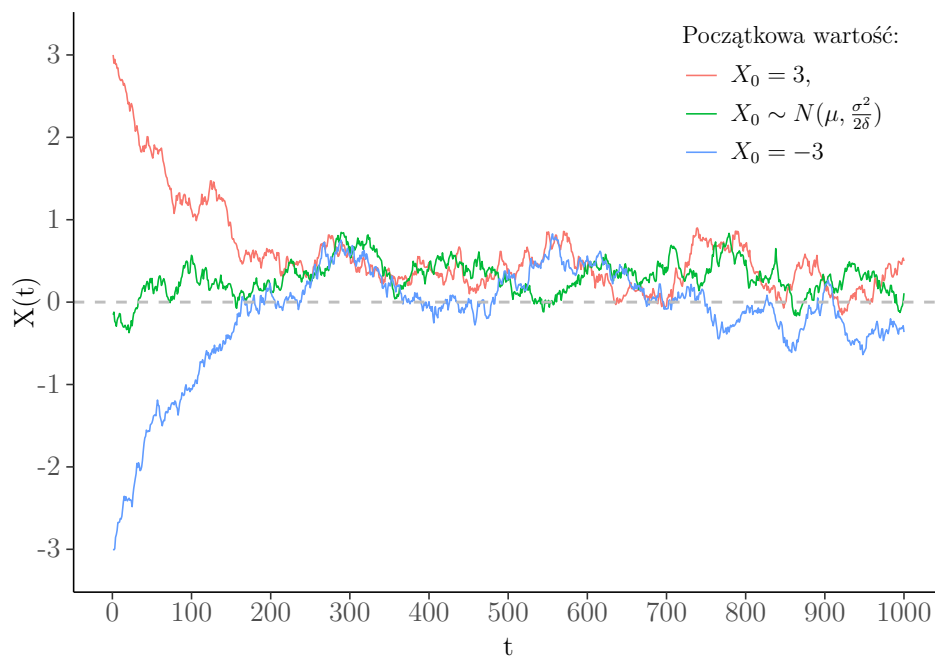
$$\text{Var}[X(t)|X(s) = x_0] = \frac{\sigma^2(1 - e^{-2\lambda(t-s)})}{2\lambda} \quad (1.25)$$

Parametr μ można interpretować jako długookresową wartość średnią procesu. Ze wzoru (1.23) wynika, że wraz z czasem (gdy $t \rightarrow +\infty$) proces „zapomina” o początkowej wartości i dąży do μ . Parametr λ mierzy szybkość „zbiegania” do długookresowej średniej (i jednocześnie „zapominania” wartości początkowej). Natomiast parametr σ odpowiada za zmienność procesu.

Proces Ornsteina-Uhlenbecka jest jednorodnym procesem Markowa. Ma stacjonarne i niezależne przyrosty o rozkładzie normalnym z warunkową wartością oczekiwaną daną wzorem (1.24) i warunkową wariancją daną wzorem (1.25). Rozkładem niezmienniczym (spełniającym warunek (2) twierdzenia 1.1) procesu Ornsteina-Uhlenbecka jest rozkład normalny z wartością oczekiwaną μ i wariancją $\frac{\sigma^2}{2\lambda}$ (graniczne wartości przy $t \rightarrow +\infty$ we wzorach odpowiednio (1.24) i (1.25)). Przyjmując

jako wartość początkowa procesu Ornsteina-Uhlenbecka X_0 zmienną losową o rozkładzie niezmienniczym $X_0 \sim N(\mu, \frac{\sigma^2}{2\lambda})$ otrzymujemy stacjonarny proces Ornsteina-Uhlenbecka (por. twierdzenie 1.1). Natomiast przyjmując założenie, że prawie na pewno $X(0) = 0$, to otrzymujemy proces Ornsteina-Uhlenbecka będący szczególnym przypadkiem procesu Lévy'ego.

Na rysunku 1.7 przedstawiono trzy przykładowe trajektorie procesu Ornsteina-Uhlenbecka. Widać wyraźnie, że trajektorie powracają do długoterminowej wartości średniej. Kolorem zielonym zaznaczono trajektorię stacjonarnego procesu Ornsteina-Uhlenbecka, dla której wartość początkową procesu wylosowano z rozkładu niezmienniczego. Przerywaną linią zaznaczono długookresową wartość średnią procesu.



Rysunek 1.7: Trajektorie trzech różnych procesów Ornsteina-Uhlenbecka o tych samych parametrach $\mu = 0, \lambda = 0,01, \sigma = 0,05$, ale różnych wartościach początkowych.

Źródło: opracowanie własne.

Zauważmy, że korzystając ze wzoru (1.23) można rozszerzyć na różniczkę względem dowolnego procesu Lévy'ego. Przedstawia to poniższa definicja.

Definicja 1.38.

(Sato, 1999, Definicja 17.2) Procesem Ornsteina-Uhlenbecka względem procesu Lévy'ego

$L = (L(t))_{t \geq 0}$ nazywamy proces $X = (X(t))_{t \geq 0}$ określony następująco

$$X(t) = X(0)e^{-\lambda t} + \mu(1 - e^{-\lambda t}) + \int_0^t e^{-\lambda(t-s)} dL(s). \quad (1.26)$$

Określony przez definicję 1.37 proces Ornsteina-Uhlenbecka jest specjalnym przypadkiem definicji 1.38, w którym procesem Lévy'ego jest proces Wienera ze współczynnikiem dyfuzji σ . W dalszej części proces Ornsteina-Uhlenbecka z definicji 1.37 nazywany będzie gaussowskim procesem Ornsteina-Uhlenbecka, a proces zdefiniowany przez definicję 1.38 uogólnionym procesem Ornsteina-Uhlenbecka.

Każdy uogólniony proces Ornsteina-Uhlenbecka ma modyfikację będącą procesem z prawostronnie ciągłym z lewostronnymi granicami, będącą również uogólnionym procesem Ornsteina-Uhlenbecka (Sato, 1999, uwaga 17.3).

Uogólniony proces Ornsteina-Uhlenbecka może przyjmować zarówno dodatnie jak i ujemne wartości. Natomiast w modelowaniu zmienności istotne jest, aby proces zmienności był nieujemny. Konieczną nieujemność można uzyskać poprzez nieliniową transformację np. wykładniczą (modele stochastycznej zmienności typu exp-OU) lub podnoszenie do drugiej potęgi. Barndorff-Nielsen i Shephard (2001b) zaproponowali odmienne rozwiązanie: użycie dodatnich niegaussowskich procesów Ornsteina-Uhlenbecka. Pozwala to na uniknięcie nieliniowych przekształceń co znacznie ułatwia wykorzystanie do wyceny instrumentów pochodnych.

Definicja 1.39.

Dodatnim niegaussowskim procesem Ornsteina-Uhlenbecka nazywamy proces $X = (X(t))_{t \geq 0}$ określony następująco

$$X(t) = X(0)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dZ(s), \quad (1.27)$$

gdzie $\lambda > 0$. Wartość początkowa $X(0)$ jest zmienną losową o rozkładzie skoncentrowanym na dodatniej półosi, niezależną od procesu podporządkowanego $Z = (Z(t))_{t \geq 0}$.

Dodatni niegaussowski proces Ornsteina-Uhlenbecka można równoważnie zapisać w postaci stochastycznego równania różniczkowego

$$dX(t) = -\lambda X(t)dt + dZ(t), \quad X(0) > 0.$$

Własności dodatniego niegaussowskiego procesem Ornsteina-Uhlenbecka oraz ich zastosowanie do modelowania zmienności instrumentów finansowych zostaną przedstawione w rozdziale 2.

Niegaussowskie modele zmienności stochastycznej typu Ornsteina-Uhlenbecka

2.1 Uwagi wstępne

Modele stochastycznej zmienności (*Stochastic Volatility*, SV) są jednym z podstawowych narzędzi stosowanych do uchwycenia zmienności instrumentów finansowych. Są używane zarówno w matematyce finansowej jak i ekonometrii finansowej. Modele SV wykorzystuje się głównie do wyceny instrumentów pochodnych i wyznaczania ekspozycji instrumentów finansowych na ryzyko rynkowe.

Cechą wspólną modeli SV jest to, że zmienność ceny instrumentu finansowego jest oddzielnym procesem stochastycznym. Oznacza to, że zmienność ma osobne źródło losowości niż ceny instrumentu. Jest to główna różnica pomiędzy modelami SV, a klasą modeli autoregresywnej warunkowej heteroskedastyczności (*Autoregressive Conditional Heteroscedasticity*, ARCH), dla których wariancja błędu losowego w danym okresie jest funkcją kwadratów wartości cen w okresach poprzednich. Oznacza to, że w modelach ARCH ich uogólnieniach (klasa modeli GARCH) wariancja procesu ceny instrumentu w okresie t przy danym zbiorze informacji Σ_{t-1} generowanym przez historię procesu do momentu $t - 1$ jest funkcją deterministyczną. Natomiast w przypadku modeli SV oddzielne źródło losowości powoduje, że warian-

cja procesu ceny instrumentu w okresie t przy danym zbiorze informacji Σ_{t-1} jest zmienną losową Doman i Doman (2009), str. 176.

Dodatkowe źródło losowości istotnie zwiększa elastyczność modeli SV do opisu ewolucji zmienności w czasie, ale z drugiej strony bardzo utrudnia estymację parametrów i wyznaczenie prognoz zmienności. Dlatego modele SV choć historycznie starsze (za pierwszą pracę uznaje się Clark (1973)) i rozwijane w latach osiemdziesiątych XX wieku równolegle z modelami ARCH/GARCH, długo pozostawały w cieniu tych drugich. Dopiero szybki rozwój metod estymacji modeli SV w latach 90. XX wieku spowodował wzrost zainteresowania modelami stochastycznej zmienności. Obecnie modele SV stanowią standardowe narzędzie matematyki finansowej i ekonometrii.

Z punktu widzenia matematyki finansowej modele SV są ważne, ponieważ umożliwiają przezwyciężyć jedno z istotnych niedociągnięć powszechnie stosowanego przez praktyków do wyceny opcji europejskich modelu Blacka-Scholesa. Model ten zakłada stałość parametru zmienności, co jest założeniem nierealistycznym, sprzecznym z obserwowanymi własnościami finansowych szeregów czasowych. Ponadto, powoduje to obciążenie modelu: wartości teoretyczne wyznaczone w modelu różnią się istotnie od obserwowanych rzeczywiście. Zgodnie z modelem Blacka-Scholesa zmienność implikowana powinna być wielkością stałą, niezależną od terminu wygaśnięcia opcji. W rzeczywistości obserwuje się natomiast efekt „uśmiechu zmienności”, czyli struktury czasowej zmienności implikowanej (Piontek, 2003). Modele SV stanowią naturalne uogólnienie modelu Blacka-Scholesa, w którym zmienność jest oddzielnym procesem stochastycznym. Podejście to zostało zapoczątkowane przez prace Hull i White (1987) i weszło do kanonu wyceny instrumentów pochodnych.

W literaturze rozważa się wiele modeli stochastycznej zmienności, ale większość propozycji jest rozwinięciem dwóch modeli. Pierwszy z nich, jest to model, w którym logarytm procesu wariancji chwilowej jest gaussowskim procesem Ornsteina-Uhlenbecka

$$d \ln \sigma^2(t) = -\lambda (\ln \sigma^2(t) - \mu) dt + \varsigma dW(t), \quad (2.1)$$

gdzie $W = (W(t))_{t \geq 0}$ jest procesem Wienera. W tym modelu stochastycznej zmienności stochastyczne równanie różniczkowe opisuje dynamikę w czasie logarytmu pro-

cesu wariancji chwilowej, ponieważ gaussowski proces Ornsteina-Uhlenbecka może przyjmować wartości ujemne, a zmienność z natury rzeczy musi być nieujemna¹. Równanie (2.1) posiada rozwiązanie (por. wzór (1.23)). Model stochastycznej zmienności określony przez równanie (2.1) ma liczne zalety, m.in. jest markowski, gaussowski, ma stacjonarne przyrosty. Ma jednak istotną wadę ograniczającą jego wykorzystanie: nie można wyznaczyć całki stochastycznej z wariancją chwilowej w postaci analitycznej (bez wprowadzenia błędu dyskretyzacji np. poprzez schemat Eulera, czy Milsteina²), co utrudnia zastosowanie do wyceny opcji egzotycznych.

Model ten rozważany był w kontekście wyceny opcji eurposjskich przez Wiggins (1987), Chesney i Scott (1989), Melino i Turnbull (1990). Zastosowanie dyskretyzacji Eulera prowadzi do modelu stochastycznej zmienności z czasem dyskretnym rozważanym przez Taylor (1982). Model Taylora zyskał popularność i stał się przedmiotem licznych prac ekonometrycznych: m.in. Harvey i in. (1994) zaproponowali estymację metodą quasi największej wiarygodności (*Quasi-maximum likelihood estimate*, QMLE) za pomocą filtrów Kalmana, Jacquier i in. (1999) oraz Kim i in. (1998) podejście bayesowskie z wykorzystaniem algorytmów MCMC (*Markov Chain Monte Carlo*). Zaproponowano też liczne modyfikacje modelu Taylora, m.in. korelacje składników losowych (Jacquier i in., 1999), grube ogony rozkładu warunkowego cen (np. t-Studenta) (Jacquier i in., 1999), skokami procesu cen (Eraker i in., 2003).

Drugim modelem jest model stałej elastyczności wariancji (*constant elasticity of variance*, CEV)

$$d\sigma^2(t) = -\lambda(\sigma^2(t) - \mu)dt + \varsigma(\sigma^2(t))^d dW(t), \quad d \geq 1/2. \quad (2.2)$$

Proces CEV znany wcześniej z zastosowania do modelowania cen instrumentów finansowych (Cox, 1975) został zaproponowany do modelowania stochastycznej zmien-

¹Stein i Stein (1991) oraz Schöbel i Zhu (1999) rozważali model stochastycznej zmienności, w którym wariancja chwilowa jest gaussowskim procesem Ornsteina-Uhlenbecka, argumentując, że przy odpowiednio dobranych parametrach prawdopodobieństwo przyjęcia wartości ujemnych jest bardzo małe.

²O wpływie dyskretyzacji na obciążenie estymatorów można znaleźć w Iacus (2009, str. 124-126)). Obciążenie to utrudnia poprawne oszacowanie cen instrumentów oraz wrażliwości opcji na zmianę czynników, tzw. greckie współczynniki (Broadie i Kaya, 2006).

ności przez Meddahi i Renault (1998). Dla dowolnej wartości parametru d nie istnieje rozwiązanie równania (2.2). Nie można także symulować trajektorii procesu wariancji chwilowej i jej całki stochastycznej bez wprowadzania błędu dyskretyzacji. Dwa specjalne przypadki modelu CEV rozważane były już wcześniej jako modele stochastycznej zmienności: dyfuzyjny model ARCH (*ARCH diffusion*) zaproponowany przez Nelson (1990) ($d=1$) oraz model Hestona (Heston, 1993) ($d=1/2$). Ten pierwszy może być interpretowany jako graniczny przypadek modeli GARCH, gdy odstęp pomiędzy kolejnymi obserwacjami dąży do 0. Rozkład stacjonarny procesu wariancji chwilowej $(\sigma^2(t))_{t \geq 0}$, to odwrotny gamma (*inverse gamma*) o $(4\lambda/\zeta^2) + 2$ stopniach swobody. Model ten mimo atrakcyjnej interpretacji ma słabe dopasowanie do danych empirycznych. Wynika to z faktu, że model ten zależy jedynie od trzech parametrów, które dodatkowo nie mogą być dobierane swobodnie, bo łączy je zależność określająca stopnie swobody rozkładu (Barndorff-Nielsen i Shephard, 1998).

W przypadku modelu Hestona rozwiązaniem równania (2.2) jest proces pierwiastka kwadratowego (*square root process*), znany z zastosowania do modelowania krótkoterminowych stóp procentowych jako model CIR od nazwiska autorów Coxa, Ingersolla i Rossa (Cox i in., 1985). Dla procesu pierwiastka kwadratowego warunkowy rozkład $\sigma^2(\Delta(n+1)) | \sigma^2(\Delta n)$ jest niecentralnym rozkładem χ^2 . W modelu Hestona proces wariancji chwilowej może być skorelowany z procesem cen, co umożliwia modelowanie obserwowanego efektu dźwigni, czyli ujemnej korelacji cen instrumentów finansowych i ich mierników zmienności. Dokładna (bez wprowadzania błędu dyskretyzacji) symulacja procesu pierwiastka kwadratowego jest trudna i była przedmiotem licznych prac, m.in. Broadie i Kaya (2006), Zhu (2008), Glasserman i Kim (2011). Tym niemniej ze względu na możliwość wyceny opcji europejskich z poziomu funkcji charakterystycznej (np. poprzez zastosowanie szybkiej transformaty Fouriera Carr i Madan (1999)) model Hestona stał się jednym z najbardziej popularnych modeli stochastycznej zmienności.

W rozdziale tym przedstawione zostaną modele stochastycznej zmienności oparte na dodatnich niegaussowskich procesach Ornsteina-Uhlenbecka, które w dalszej części pracy posłużą do modelowania zmienności dla danych pochodzących z polskiego rynku finansowego i rynku walutowego. Przedstawiona teoria opiera się głównie na

trzech artykułach Bandorffa-Nielsena i Shepharda (Barndorff-Nielsen i Shephard (2001b), Barndorff-Nielsen i Shephard (2002), Barndorff-Nielsen i Shephard (2003)) w których zostały zainicjowane wykorzystanie dodatnich niegaussowskich procesów Ornsteina-Uhlenbecka do modelowania zmienności cen instrumentów finansowych. Ponadto przedstawione zostaną zaproponowane w literaturze modyfikacje i uogólnienia oraz powiązania z estymatorami wariancji zrealizowanej opartymi na danych śróddziennych (*intraday*).

2.2 Model stochastycznej zmienności Bandorffa-Nielsena i Shepharda

Modele stochastycznej zmienności oparte na dodatnich niegaussowskich procesach Ornsteina-Uhlenbecka zostały zaproponowane przez Bandorffa-Nielsena i Shepharda w artykule Barndorff-Nielsen i Shephard (2001b), ale częściowe wyniki były publikowane wcześniej (np. w Barndorff-Nielsen i Shephard (1998) oraz Barndorff-Nielsen i Shephard (1999)). Od nazwiska autorów w dalszej części model będzie oznaczany w skrócie jako BNS.

Definicja 2.1.

Podstawowym modelem stochastycznej zmienności Bandorffa-Nielsena i Shepharda (BNS) nazywamy model stochastycznej zmienności określony przez następujące stochastyczne równania różniczkowe

$$d \ln S(t) = (\mu + \beta \sigma^2(t)) dt + \sigma(t) dW(t), \quad \ln S(0) = 0, \quad (2.3)$$

$$d\sigma^2(t) = -\lambda \sigma^2(t) dt + dZ(\lambda t), \quad \sigma^2(0) > 0, \quad (2.4)$$

gdzie $\mu, \beta \in \mathbb{R}, \lambda \in \mathbb{R}_+$. Proces Wienera $W = (W(t))_{t \geq 0}$ jest niezależny od procesu podporządkowanego $Z = (Z(\lambda t))_{t \geq 0}$ ³.

Równanie (2.3) opisuje dynamikę logarytmu procesu cen $(S(t))_{t \geq 0}$. W dalszej części pracy będziemy przyjmować zapis $Y(t) = \ln S(t)$ dla logarytmu procesu cen

³Rozważane procesy stochastyczne są określone na przestrzeni probabilistycznej $(\Omega, \Sigma, \mathbb{P})$ i adaptowane względem pewnej filtracji $(\Sigma_t)_{t \in \mathbb{T}}$.

i proces $(Y(t))_{t \geq 0}$ nazywać logarytmicznym procesem cen. Proces ten jest obserwowany w równych odstępach czasu o długości $\Delta = t_n - t_{n-1}$, dla $n = 1, \dots$. Wówczas przyrosty logarytmicznego procesu cen to zwroty logarytmiczne z okresu n :

$$y_n = \int_{(n-1)\Delta}^{n\Delta} dY(t) = Y(n\Delta) - Y((n-1)\Delta). \quad (2.5)$$

Pojedynczy okres n o długości Δ może oznaczać np. godzinę, dzień, tydzień lub miesiąc.

W równaniu (2.3) parametr μ jest dryfem, a parametr β jest interpretowany jako premia za ryzyko. Ze względu na rolę jaką proces $(Z(\lambda t))_{t \geq 0}$ pełni w równaniu (2.4) nazywany jest prowadzącym procesem Lévy'ego ukrytym w tle (*background driven Lévy'ego proces*, BDLP). Jest to proces podporządkowanym zatem zgodnie z definicją 1.35 jest nieujemny i niemalejący.

Rozwiązanie równania (2.4) dane jest wzorem

$$\sigma^2(t) = \sigma^2(0)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dZ(\lambda s) \quad (2.6)$$

Proces wariancji chwilowej jest ściśle dodatni i ograniczony z dołu przez deterministyczną funkcję $\sigma^2(0)e^{-\lambda t}$. Wynika to z faktu, że proces BDLP jest niemalejący i $\sigma^2(0) > 0$. Ujemny znak przed parametrem λ zapewnia, że proces wariancji chwilowej jest procesem powracającym (*mean-reverting*) do średniej 0. Dodatniość nieguassowskiego procesu Ornsteina-Uhlenbecka pozwala na modelowanie wariancji bezpośrednio, bez uciekania się do nieliniowych transformacji, np. logarytmicznej (jak w przypadku modelu SV danego wzorem 2.1).

Trajektorie wariancji chwilowej ewoluują w czasie tylko na dwa sposoby: albo maleją wykładniczo według stopy λ albo pojawia się punkt nieciągłości (skok trajektorii do góry). Pierwsza możliwość związana jest z sytuacją, gdy proces BDLP jest w pewnym okresie stały ($dZ(\lambda t) = 0$). Wówczas równanie (2.6) sprowadza się tylko do części deterministycznej, której rozwiązaniem jest funkcja eksponencjalna. Tym większa wartość parametru λ , tym szybciej proces wariancji chwilowej maleje pomiędzy skokami. Punkty nieciągłości procesu zmienności chwilowej odpowiadają punktom nieciągłości procesu $(Z(\lambda t))_{t \geq 0}$. Zauważmy, że przyjęta postać wariancji chwilowej jest zgodna z obserwowanymi faktami dotyczącym zjawisk finansowych:

pojawienie się nowej informacji na rynku powoduje nagły wzrost zmienności, która potem maleje stopniowo wraz z upływem czasu. Przeskalowanie czasu w procesie BDLP powoduje uniezależnienie rozkładu brzegowego procesu zmienności chwilowej od parametru λ i ścisłą stacjonarność procesu wariancji chwilowej. Szczegóły, tej zależności zostaną opisane w twierdzeniu 2.1. Własności procesu wariancji chwilowej przedstawia rysunek 2.1, gdzie wybrano proces BDLP o skończonej aktywności (proces ma skończoną ilość skoków w skończonym przedziale czasu).

Twierdzenie 2.1. (*Wolfe, 1982*) *Jeśli $\mathcal{D}$ jest rozkładem samorozkładalnym, to istnieje ściśle stacjonarny proces $(\sigma^2(t))_{t \geq 0}$ i proces Lévy'ego $(Z(t))_{t \geq 0}$ takie, że dla każdego $t \geq 0$ zmienna losowa $\sigma^2(t)$ ma rozkład $\mathcal{D}$ oraz*

$$\sigma^2(t) = \sigma^2(0)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dZ(\lambda s),$$

dla dowolnej $\lambda > 0$.

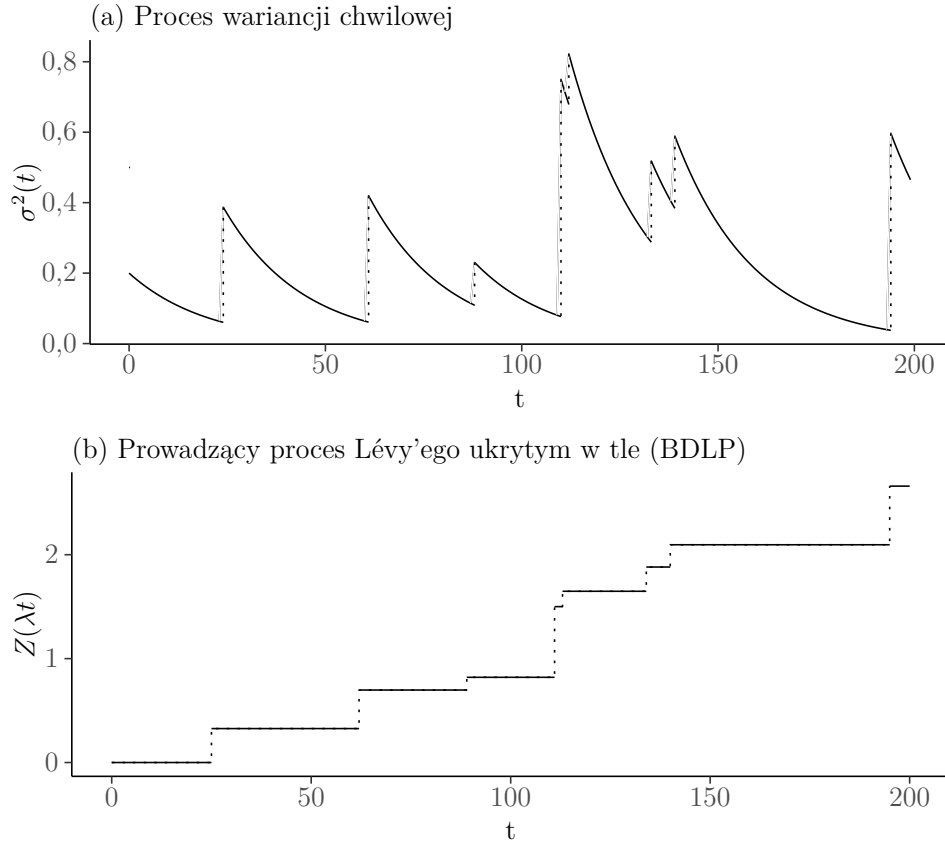
Odrotnie, jeżeli ściśle stacjonarny proces $(\sigma^2(t))_{t \geq 0}$ o rozkładzie stacjonarnym $\mathcal{D}$ i proces Lévy'ego $(Z(t))_{t \geq 0}$ dla dowolnej $\lambda > 0$ spełniają stochastyczne równanie różniczkowe

$$d\sigma^2(t) = -\lambda\sigma^2(t)dt + dZ(\lambda t),$$

to rozkład $\mathcal{D}$ jest samorozkładalny.

W konsekwencji proces wariancji chwilowej określony przez równanie (2.4) ma ściśle stacjonarny rozkład samorozkładalny $\mathcal{D}$, jeśli tylko $\sigma^2(0)$ ma rozkład $\mathcal{D}$. W dalszej części pracy będziemy zakładać, że $\sigma^2(0)$ ma rozkład $\mathcal{D}$, a proces wariancji chwilowej jest ściśle stacjonarny. Twierdzenie 2.1 pozwala poprzez użyci dodatnich niegaussowskich procesów Ornsteina-Uhlenbecka na zastosowanie szerokiej klasy rozkładów samorozkładalnych do modelowania wariancji chwilowej, m.in. rozkładów gamma, odwrotnego Gaussa, log-normalnego, temperowanego rozkładu stabilnego⁴. Twierdzenie 2.1 nie jest ograniczone wyłącznie do rozkładów samorozkładalnych skoncentrowanych na dodatniej półosi. Dodatniość procesu wariancji chwilowej

⁴Przegląd rozkładów samorozkładalnych, które mogą pełnić rolę rozkładu stacjonarnego zostanie przedstawiony w podrozdziale 2.5.



Rysunek 2.1: Trajektoria procesu wariancji chwilowej o rozkładzie brzegowym gamma z parametrem kształtu $\nu = 1$ oraz parametrem skali $1/\alpha = 1/3$ (a) oraz odpowiadający mu proces Lévy'ego ukryty w tle (BDLP) z parametrem $\lambda = 0,01$ (b).

Źródło: opracowanie własne.

wynika z dodatkowego założenia przyjętego przez Bandorffa-Nielsena i Shepharda: przyjęcia jako procesu prowadzącego w tle procesu podporządkowanego.

Z twierdzenia 2.1 wynika także możliwość modelowania wariancji chwilowej albo poprzez wybór rozkładu stacjonarnego $\mathcal{D}$ i dobranie odpowiedniego procesu BDLP albo poprzez wybranie procesu BDLP i dopasowanie odpowiedniego procesu wariancji chwilowej. W pierwszym przypadku dla procesu $(\sigma^2(t))_{t \geq 0}$ stosuje się notację $\mathcal{D}$ -OU, a w tym drugim OU- $\mathcal{D}$. Rysunek 2.1 przedstawia proces wariancji chwilowej o rozkładzie stacjonarnym gamma, co w skrócie można zapisać jako Ga-OU. Konstrukcja procesu wariancji chwilowej poprzez określenie procesu BDLP nakłada pewne ograniczenia na wybór procesu podporządkowanego $(Z(t))_{t \geq 0}$. Zgodnie ze

wcześniejszą uwagę proces BDLP musi być procesem podporządkowanym, aby zapewnić nieujemność procesu wariancji. Poniższy warunek zapewnia stacjonarność procesu Ornsteina-Uhlenbecka.

Twierdzenie 2.2. (Wolfe, 1982) Proces $(\sigma^2(t))_{t \geq 0}$ będący rozwiązaniem równania

$$d\sigma^2(t) = -\lambda\sigma^2(t)dt + dZ(\lambda t),$$

jest procesem stacjonarnym wtedy i tylko wtedy, gdy proces $(Z(t))_{t \geq 0}$ spełnia warunek

$$\mathbb{E} [\ln (1 + |Z(1)|)] < +\infty. \quad (2.7)$$

Sprawdzając, czy dany proces podporządkowany może być BDLP dla pewnego procesu wariancji chwilowej można także zweryfikować warunek na miarę Lévy'ego procesu $(Z(t))_{t \geq 0}$.

Twierdzenie 2.3. (Barndorff-Nielsen i Shephard, 2001b) Załóżmy, że $(Z(t))_{t \geq 0}$ jest procesem podporządkowanym z miarą Lévy'ego $W(dx)$ oraz gęstością Lévy'ego $w(x)$. Określmy funkcję $u : \mathbb{R}^+ \rightarrow \mathbb{R}$ daną wzorem

$$u(x) = \int_1^{+\infty} w(\tau x) d\tau. \quad (2.8)$$

Jeżeli

$$\int_1^{+\infty} \ln x W(dx) < +\infty \quad (2.9)$$

to równanie

$$\sigma^2(t) = \sigma^2(0)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dZ(\lambda s),$$

określa stacjonarny proces, dla którego gęstość Lévy'ego jest dana przez funkcję u zdefiniowaną wzorem (2.8).

Przykładem procesu, który może pełnić rolę BDLP jest proces gamma. Proces ten jest procesem podporządkowanym oraz spełnia warunek z twierdzenia 2.3. Aby to pokazać wystarczy skorzystać ze wzoru na gęstość Lévy'ego procesu gamma (por. wzór (1.13)) oraz prostej nierówności $\ln x < x$ (prawdziwej dla $x > 0$):

$$\int_1^{+\infty} \ln x W(dx) = \int_1^{+\infty} \frac{\nu \ln x}{x} e^{-\alpha x} dx \leq \int_1^{+\infty} \nu e^{-\alpha x} dx \leq \frac{\nu e^{-1}}{\alpha} \leq \frac{\nu}{\alpha} < +\infty.$$

Niegaussowski proces Ornsteina-Uhlenbecka, dla którego procesem BDLP jest procesem gamma można zapisać jako OU-Ga. Podobnie można wykazać, że rozkład odwrotny Gaussa może zarówno pełnić rolę rozkładu stacjonarnego procesu wariancji chwilowej (IG-OU), ponieważ jest rozkładem samorozkładalnym jak i rozkładu zmiennej losowej $Z(1)$ dla procesu BDLP (OU-IG), ponieważ proces odwrotny Gaussa spełnia twierdzenie 2.3.

Do prześledzenia związków pomiędzy procesem wariancji chwilowej i procesem BDLP można posłużyć się funkcją generującą kumulanty. Dla dowolnej zmiennej losowej X funkcją generującą kumulanty nazywamy logarytm naturalny funkcji generującej momenty

$$C_X(s) = \ln \mathbb{E} \left(e^{sX} \right) = \ln M_X(s), \quad \text{dla } s \in \mathbb{R},$$

gdzie M_X jest funkcją generującą momenty⁵.

Funkcje generujące kumulanty zmiennych losowych $Z(1)$ i $\sigma^2(t)$ (dla dowolnego $t \geq 0$) łączy następująca zależność (Barndorff-Nielsen i Shephard, 2001b):

$$C_{Z(1)}(s) = s \frac{dC_{\sigma^2(t)}(s)}{ds}.$$

W konsekwencji

$$C_{Z(1)}^n = nC_{\sigma^2(t)}^n, \quad \text{dla } t \geq 0.$$

w szczególności zachodzą równości

$$\mathbb{E} [Z(1)] = \mathbb{E} [\sigma^2(t)] \quad \text{oraz} \quad \text{Var} [Z(1)] = 2 \text{Var} [\sigma^2(t)].$$

Oznacza to, że wartość oczekiwaną i wariancję rozkładu stacjonarnego procesu wariancji chwilowej można wyznaczyć z rozkładu procesu BDLP i na odwrót.

⁵Kumulanty C_X^n otrzymuje się jako rozwinięcie funkcji generującej kumulanty w szereg Mac-laurina

$$C_X(s) = \sum_{n=1}^{\infty} C_X^n \frac{s^n}{n!}.$$

Kumulanty można wyznaczyć także bezpośrednio ze wzoru

$$C_X^n = \left. \frac{d^n C_X(s)}{d s^n} \right|_{s=0}.$$

Kumulanty są ściśle powiązane z momentami zmiennej losowej X . W szczególności pierwsze dwie kumulanty odpowiadają wartości oczekiwanej i wariancji zmiennej losowej X .

Kolejne powiązania pomiędzy procesem wariancji chwilowej i procesem BDLP można znaleźć na gruncie gęstości Lévy'ego. Oznaczmy przez w i u gęstości Lévy'ego (o ile istnieją) odpowiednio dla procesów $(Z(t))_{t \geq 0}$ i $(\sigma^2(t))_{t \geq 0}$. Jeżeli dodatkowo funkcja u jest różniczkowalna, to zachodzi równość

$$w(x) = -u(x) - xu'(x). \quad (2.10)$$

Relację pomiędzy gęstościami Lévy'ego można wykorzystać do wyznaczenia procesu BDPL w przypadku procesu wariancji chwilowej o rozkładzie gamma (Ga-OU). Gęstość Lévy'ego procesu gamma dana jest wzorem

$$u(x) = \frac{\nu}{x} e^{-\alpha x} 1_{(0, \infty)}(x).$$

Zatem zgodnie ze wzorem (2.10) gęstość Lévy'ego rozkładu procesu BDPL jest równa

$$\begin{aligned} w(x) &= -u(x) - xu'(x) = \\ &= -\frac{\nu}{x} e^{-\alpha x} 1_{(0, \infty)}(x) - x \left(-\frac{\nu}{x^2} e^{-\alpha x} 1_{(0, \infty)}(x) - \frac{\nu}{x} e^{-\alpha x} (-\alpha) 1_{(0, \infty)}(x) \right) = \\ &= \nu \alpha e^{-\alpha x} 1_{(0, \infty)}(x) = \nu g(x), \end{aligned}$$

gdzie $g(x)$ jest gęstością rozkładu wykładniczego z parametrem α . Oznacza to, że BDLP dla modelu Ga(ν, α)-OU jest złożonym procesem Poissona z parametrem intensywności ν oraz wykładniczym rozkładem skoków z parametrem α . Po przeskalowaniu czasu parametrem λ otrzymujemy

$$Z(\lambda t) = \sum_{n=1}^{N(\lambda t)} Y_n, \quad (2.11)$$

gdzie $(N(\lambda t))_{t \geq 0}$ jest procesem Poissona o intensywności $\lambda \nu$, natomiast $Y_1, Y_2, \dots$ jest ciągiem niezależnych zmiennych o rozkładzie wykładniczym z parametrem α .

Wróćmy do równania (2.3). Proces logarytmiczny cen można zapisać w postaci

$$Y(t) = \int_0^t dY(s) = \mu t + \beta \int_0^t \sigma^2(u) du + \int_0^t \sigma^2(u) dW(u). \quad (2.12)$$

Parametr μ jest dryfem, czyli stałą tendencją, która może pojawiać się w logarytmicznym procesie cen. Parametr β , który interpretuje się jako premia za ryzyko,

powoduje asymetrię $Y(t)$ od dryfu μt , który jest symetrycznie skoncentrowany względem μt , gdy $\beta = 0$.

Proces stochastyczny $(\sigma^{2*}(t))_{t \geq 0}$ określony jako

$$\sigma^{2*}(t) = \int_0^t \sigma^2(u) du \quad (2.13)$$

nazywany jest wariancją scałkowaną (*integrated variance*).

W przypadku, gdy $\mu = \beta = 0$, to wariancja kwadratowa (definicja 1.15) logarytmicznego procesu cen jest równa wariancji scałkowanej:

$$[Y](t) = \sigma^{2*}(t), \quad (2.14)$$

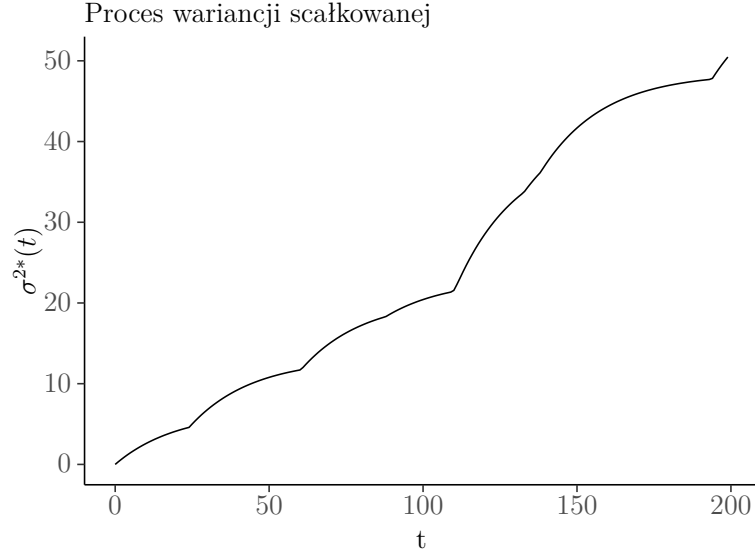
bez względu na wybór procesu wariancji chwilowej - równość jest prawdziwa dla dowolnego modelu stochastycznej zmienności, niekoniecznie typu Ornsteina-Uhlenbecka. Jako miarę zmienności cen rozważanych z czasem ciągłym wariancja kwadratowa (lub w niektórych pracach zwrotów cen) została wprowadzonego w pracach Andersena i in. (1999) i Andersen i in. (2001). Estymatorem wariancji kwadratowej są estymatory wariancji zrealizowanej. Szerzej zostanie to przedstawione w podrozdziale 2.6.

Ważną własnością w modelu BNS jest możliwość wyznaczenia wariancji scałkowanej w postaci jawnego wzoru

$$\begin{aligned} \sigma^{2*}(t) &= \int_0^t \sigma^2(u) du = \int_0^t \left(\sigma^2(0)e^{-\lambda u} + \int_0^u e^{-\lambda(u-s)} dZ(\lambda s) \right) du = \\ &= \int_0^t \left(\sigma^2(0)e^{-\lambda u} \right) du + \int_0^t \left(e^{-\lambda u} \int_0^u e^{-\lambda s} dZ(\lambda s) \right) du = \\ &= \frac{1}{\lambda} (1 - e^{-\lambda t}) \sigma^2(0) + \int_0^t \left(1 - e^{-\lambda(t-s)} \right) dZ(\lambda s) = \\ &= \frac{1}{\lambda} \left(\sigma^2(0) - \sigma^2(0)e^{-\lambda t} \right) + \frac{1}{\lambda} \left(Z(\lambda t) - \int_0^t e^{-\lambda(t-s)} dZ(\lambda s) \right) = \\ &= \frac{1}{\lambda} \left(\sigma^2(0) + Z(\lambda t) - \sigma^2(0)e^{-\lambda t} - \int_0^t e^{-\lambda(t-s)} dZ(\lambda s) \right) = \\ &= \frac{1}{\lambda} \left(Z(\lambda t) - \sigma^2(t) + \sigma^2(0) \right). \end{aligned} \quad (2.15)$$

Choć procesy BDLP i wariancji chwilowej mają punkty nieciągłości, to ich kombinacja liniowa, czyli wariancja scałkowana jest ciągła. W literaturze taka sytuacja

nazywana jest współ-skokami (*co-breaks*, por. Hendry i Massmann (2007)). Na rysunku 2.2 przedstawiono przykład trajektorii wariancji scałkowanej odpowiadającej procesowi wariancji chwilowej przedstawionemu na rysunku 2.1.



Rysunek 2.2: Trajektoria procesu wariancji scałkowanej odpowiadającej procesowi wariancji chwilowej przedstawionemu na rysunku 2.1.

Źródło: opracowanie własne.

Wariancja scałkowana pełni w modelu BNS rolę „czasu operacyjnego” dla logarytmu procesu cen. Przedstawia to następujące twierdzenie.

Twierdzenie 2.4. (*Barndorff-Nielsen i Shephard, 1998*) Dla procesu wariancji scałkowanej w modelu BNS (definicja 2.1) istnieje proces Wienera $(\dot{W}(t))_{t \geq 0}$ na przestrzeni probabilistycznej z filtracją $(\Omega, \Sigma, \mathbb{P}, (\Sigma_s)_{s \in \mathbb{T}})$ taki że

$$\int_0^t \sigma^2(s) dW(s) = \dot{W}(\sigma^{2*}(t)) \quad (2.16)$$

Twierdzenie to wynika z twierdzenia Dubins-Schwarza o reprezentacji ciągłego martyngału lokalnego jako „czasu operacyjnego” dla pewnego procesu Wienera (por. Rogers i Williams (1996), str. 96). Stosując twierdzenie 2.4 do wzoru (2.12) otrzymujemy

$$Y(t) = \int_0^t dY(s) = \mu t + \beta \sigma^{2*}(t) + \dot{W}(\sigma^{2*}(t)) \quad (2.17)$$

Wzór (2.17) ma wiele konsekwencji. Po pierwsze, wariancja scałkowana pełni w modelu BNS rolę „czasu operacyjnego” dla procesu logarytmu cen. Po drugie, umożliwia wyznaczenie rozkładu warunkowego i bezwarunkowego powszechnie stosowanych w matematyce finansowej stóp zwrotów. Logarytmiczna stopa zwrotu w n -tym okresie trwającym $\Delta > 0$ dana jest wzorem

$$\begin{aligned} y_n &= \int_{(n-1)\Delta}^{n\Delta} dY(t) = Y(n\Delta) - Y((n-1)\Delta) = \\ &= \mu\Delta + \beta \int_{(n-1)\Delta}^{n\Delta} \sigma^2(t)dt + \int_{(n-1)\Delta}^{n\Delta} \sigma^2(t)dW(t) = \\ &= \mu\Delta + \beta \left(\sigma^{2*}(n\Delta) - \sigma^{2*}((n-1)\Delta) \right) + \sqrt{(\sigma^{2*}(n\Delta) - \sigma^{2*}((n-1)\Delta))}U, \end{aligned} \quad (2.18)$$

gdzie U jest zmienną losową o rozkładzie normalnym standaryzowanym. W ostatniej równości, we wzorze 2.18, wykorzystano, że proces standardowy Wienera ma przyrosty o rozkładzie normalnym o zerowej wartości oczekiwanej i odchyleniu standardowym równym pierwiastku przyrostu czasu.

Przyrosty wariancji scałkowanej w przedziale czasu n długości Δ postaci

$$\sigma_n^2 = \sigma^{2*}(n\Delta) - \sigma^{2*}((n-1)\Delta) \quad (2.19)$$

Barndorff-Nielsen i Shephard nazwali wariancją aktualną w okresie n . Ze wzorów (2.15) oraz (2.19) wynika, że wariancja aktualna w okresie n ma prostą formę liniowej kombinacji przyrostów procesu BDPL i procesu wariancji chwilowej

$$\sigma_n^2 = \frac{1}{\lambda} \left[Z(\lambda n\Delta) - Z(\lambda(n-1)\Delta) - \left(\sigma^2(\lambda n\Delta) - \sigma^2(\lambda(n-1)\Delta) \right) \right]. \quad (2.20)$$

Równoważnie wzór (2.20) można zapisać jako

$$\sigma_n^2 = \frac{1}{\lambda} \left[(1 - e^{-\lambda\Delta})\sigma^2(\lambda(n-1)\Delta) + \eta_{2n} - \eta_{1n} \right], \quad (2.21)$$

gdzie składowe wektora losowego $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ wyznacza się ze wzorów:

$$\eta_{1n} = e^{-\lambda\Delta} \int_0^{\lambda\Delta} e^s dZ(s). \quad (2.22)$$

oraz

$$\eta_{2n} = \int_0^{\lambda\Delta} dZ(s), \quad (2.23)$$

Podstawiając (2.19) do wzoru (2.18) otrzymujemy

$$y_n = \mu\Delta + \beta\sigma_n^2 + \sqrt{\sigma_n^2}U.$$

Rozkład warunkowy logarytmicznych stóp zwrotu y_n pod warunkiem wariancji aktualnej σ_n^2 przyjmuje zatem postać

$$y_n | \sigma_n^2 \sim N(\mu\Delta + \beta\sigma_n^2, \sigma_n^2). \quad (2.24)$$

W rozdziale trzecim równanie (2.24) zostanie zastosowane do zapisania modelu BNS w reprezentacji przestrzeni stanu, co pozwala na wykorzystanie filtrów Kalmana i cząsteczkowych do estymacji parametrów w tym modelu. Natomiast wzory (2.20), (2.22) oraz (2.23) umożliwiają generowanie kolejnych generacji cząsteczek w filtrze cząsteczkowym.

Bezwarunkowy rozkład logarytmicznych stóp zwrotu nie jest dokładnie znany. Ze wzoru (2.24) wynika, że y_n jest normalną mieszaniną średnio-wariacyjną, co powinno pozwolić uzyskać rozkłady logarytmicznych stóp zwrotu o własnościach obserwowanych empirycznie: asymetrycznych i o grubych ogonach (por. Piontek (2002, str. 108)). Pewną przesłanką co do rozkładu logarytmicznych stóp zwrotu jest twierdzenie 1.6 zastosowane dla wzoru (2.24): gdyby wariancja aktualna σ_n^2 miała uogólniony odwrotny rozkład Gaussa, to bezwarunkowy rozkład logarytmicznych stóp zwrotu byłby uogólnionym rozkładem hyperbolicznym. Jest to rozkład, który cechuje bardzo dobre dopasowanie do obserwowanych empirycznie logarytmicznych stóp zwrotu. Natomiast do klasy uogólnionych rozkładów odwrotnych Gaussa należy m.in. rozkład gamma i odwrotny Gaussa. Należy jednak zwrócić uwagę, że rozkład σ_n^2 nie jest dokładnie znany. Można przypuszczać, że jako dyskretyzacja wariancji chwilowej dla okresu o długości Δ , powinien mieć rozkład zbliżony do rozkładu stacjonarnego wariancji chwilowej⁶. Drugą przesłanką stanowi porównanie wzoru (2.17) z twierdzeniem 1.8. Procesem podporządkowanym jest wtedy wariancja scałkowana⁷. Barndorff-Nielsen i Shephard (2003) zbadali rozkłady wariancji scałkowanej w zależności od wyboru rozkładu stacjonarnego wariancji chwilowej i wykazali, że w

⁶Taką tezę, ale bez dowodu stawia Benth (2011). W pracy tej autor jako rozkład σ_n^2 przyjmuje rozkład stacjonarny procesu wariancji chwilowej.

⁷Dodatkowy wyraz wolny $\mu\Delta$ nie wpływa znacząco na tezę twierdzenia 1.8 – powoduje jedynie przesunięcie w prawo rozkładów o $\mu\Delta$ (Schoutens, 2003, str. 67).

przypadku, gdy rozkładem stacjonarnym jest rozkład gamma lub odwrotny Gaussa, to choć wariancja scałkowana nie ma wówczas dokładnie takiego samego rozkładu jak wariancja chwilowa, to ogony rozkładów mają dokładnie tą samą postać. Sugeruje to zgodnie z twierdzeniem 2.17, że bezwarunkowy rozkład stóp zwrotu będzie miał podobne ogony jak rozkład VG lub NIG w przypadku, gdy rozkładem stacjonarnym wariancji chwilowej jest odpowiednio rozkład gamma lub odwrotny Gaussa. Na podstawie powyższych rozważań można przypuszczać, że choć rozkład bezwarunkowy stóp zwrotu w modelu BNS nie jest dokładnie znany, to powinien stanowić dobre przybliżenie obserwowanych empirycznie stóp zwrotu.

Na rysunku 2.3 przedstawiono rozkład empiryczny wariancji aktualnej wyznaczony za pomocą estymatora jądrowego⁸ (kolor czerwony) oraz gęstość rozkładu gamma (kolor niebieski). Przykład ilustruje dobre dopasowanie rozkładu wariancji aktualnej do rozkładu teoretycznego wariancji chwilowej uzyskany przy dużej liczbie obserwacji.

Następnym faktem empirycznym obserwowanym dla logarytmicznych zwrotów jest zjawisko dążenia ich rozkładów bezwarunkowych do rozkładu normalnego przy zwiększaniu skali jednostki czasu (por. Kliber (2013, str. 51)). Okazuje się, że model BNS spełnia tę własność, przy dodatkowym założeniu, że proces wariancji chwilowej jest ergodyczny⁹. Wówczas

$$\frac{1}{t}\sigma^{2*}(t) = \frac{1}{t} \int_0^t \sigma^2(u)du \xrightarrow[t \rightarrow +\infty]{} \xi.$$

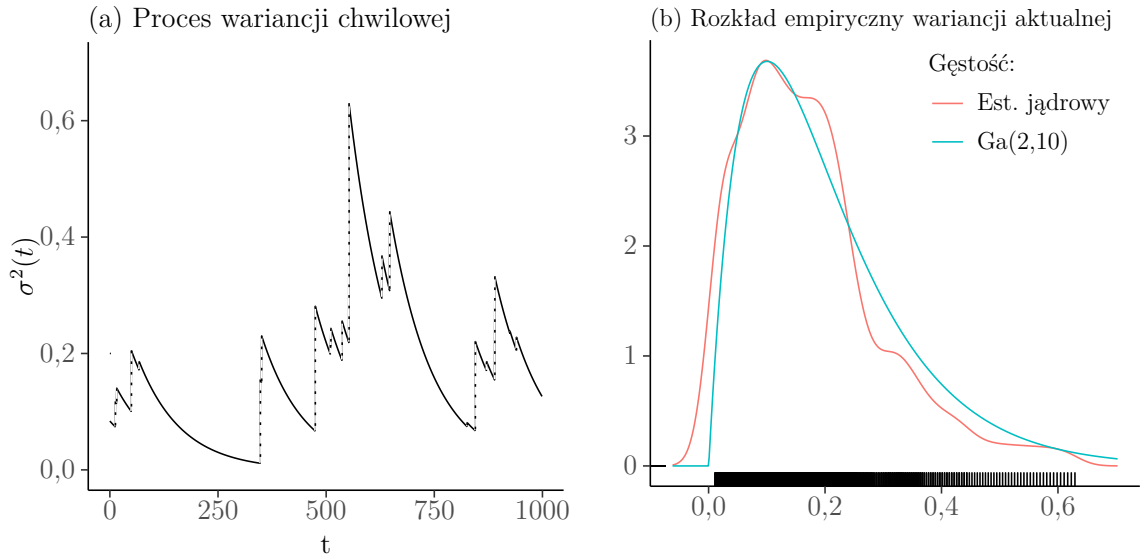
Zachodzi wówczas następująca zbieżność (według rozkładu):

$$t^{-1/2} \left(Y(t) - \mu t - \beta \sigma^{2*}(t) \right) \xrightarrow[t \rightarrow +\infty]{d} N(0, \xi). \quad (2.25)$$

Co oznacza, że bezwarunkowe rozkłady stóp zwrotu wraz ze wzrostem czasu pomiędzy obserwacjami zbiegają do rozkładu normalnego. Podobną własność udowodniono dla modeli klasy ARCH (Diebold, 1988, str. 12-16).

⁸Zastosowano jądro gaussowskie i parametr wygładzany wyznaczony zgodnie z regułą Silvermana (Silverman, 1986).

⁹Procesem ergodycznym nazywamy proces słabo stacjonarny, dla którego wartości parametrów statystycznych po zbiorze realizacji (czyli wartość średnia, wariancja i funkcja autokorelacji) są równe wartościom tych parametrów z jego dowolnej realizacji czasowej Wong (1976, str. 66).



Rysunek 2.3: Trajektoria procesu variancji chwilowej o rozkładzie brzegowym gamma z parametrem kształtu $\nu = 2$, parametrem skali $1/\alpha = 1/10$, parametrem $\lambda = 0,01$ (a) oraz rozkład empiryczny variancji aktualnej: estymator jądrowy gęstości (kolor czerwony) oraz gęstość rozkładu gamma (kolor niebieski) z parametrem kształtu $\nu = 2$ oraz parametrem skali $1/\alpha = 1/10$ (b).

Źródło: opracowanie własne.

W modelu BNS wartość oczekiwana i wariancja variancji scałkowanej, aktualnej oraz logarytmicznych stóp zwrotu zależą od parametrów rozkładu variancji chwilowej. W celu uzależnienia się od wyboru rozkładu stacjonarnego można przejść na parametry związane z wartością oczekiwaną, wariancją i autokorelacją procesu variancji chwilowej.

Założmy, że proces variancji chwilowej ma wartość oczekiwaną ξ oraz wariancję ω^2 . Wówczas istnieje także funkcja autokorelacji procesu variancji chwilowej r i przyjmuje postać

$$r(t) = e^{-\lambda|t|}. \quad (2.26)$$

Wartość oczekiwaną i wariancję procesu variancji scałkowanej można wyznaczyć ze wzorów

$$\mathbb{E}(\sigma^{2*}(t)) = \xi t, \quad \text{Var}(\sigma^{2*}(t)) = 2\omega^2 r^{**}(t), \quad (2.27)$$

gdzie

$$r^*(t) = \int_0^t r(s) ds, \quad r^{**}(t) = \int_0^t r(s) ds.$$

Wówczas

$$r^{**}(t) = \frac{1}{\lambda^2} \left(e^{-\lambda|t|-1+\lambda t} \right). \quad (2.28)$$

Zatem

$$\text{Var} \left(\sigma^{2*}(t) \right) = 2\omega^2 \frac{1}{\lambda^2} \left(e^{-\lambda|t|-1+\lambda t} \right).$$

Korzystając ze wzorów (2.27) można wyznaczyć także wartość oczekiwaną i wariancję procesu wariancji aktualnej

$$\mathbb{E} \left(\sigma_n^2 \right) = \xi \Delta, \quad \text{Var} \left(\sigma_n^2 \right) = 2\omega^2 r^{**}(\Delta) = \frac{2\omega^2}{\lambda^2} \left(e^{-\lambda\Delta-1+\lambda\Delta} \right). \quad (2.29)$$

oraz kowariancję

$$\text{cov}(\sigma_n^2, \sigma_{n+s}^2) = \omega^2 (r^{**}(s + \Delta) - 2r^{**}(s) + r^{**}(s - \Delta)).$$

Wartość oczekiwana i wariancja wariancji scałkowanej i aktualnej zależą od wartości parametrów ξ i ω^2 oraz parametru persystencji λ , a nie zależą od wyboru rozkładu stacjonarnego. Podobna zależność przechodzi na logarytm procesu cen oraz logarytmiczne stopy zwrotu:

$$\begin{aligned} \mathbb{E}(Y(t)) &= (\mu + \beta \xi) t, \\ \text{Var}(Y(t)) &= t \xi + 2\beta^2 \omega^2 r^{**}(t) = t \xi + \frac{\beta^2 \omega^2}{\lambda^2} \left(e^{-\lambda|t|} - 1 + \lambda t \right). \end{aligned}$$

oraz

$$\begin{aligned} \mathbb{E}(y_n) &= \beta \Delta \xi + \mu \Delta, \\ \text{Var}(y_n) &= \Delta \xi + 2\beta^2 \omega^2 r^{**}(\Delta) = \Delta \xi + \frac{2\beta^2 \omega^2}{\lambda^2} \left(e^{-\lambda|\Delta|} - 1 + \lambda \Delta \right). \end{aligned}$$

Kolejną własnością istotną z punktu widzenia zgodności modelu BNS z obserwowanymi faktami empirycznymi dotyczącymi cen są autokorelacje wariancji aktualnej oraz kwadratów logarytmicznych stóp zwrotu. W praktyce obserwuje się, że stopy zwrotu z instrumentów bazowych są nieskorelowane, natomiast kwadraty i wartości

absolutne stóp zwrotu mają bardzo powoli malejącą autokorelację Kliber (2013, str. 51)). W modelu BNS stopy zwrotu są nieskorelowane:

$$\text{cor}(y_n, y_{n+s}) = 0,$$

dla $s = 1, 2, 3, \dots$. Natomiast dla kwadratów stóp zwrotu zachodzi

$$\text{cor}(y_n^2, y_{n+s}^2) = ce^{-\lambda\Delta(s-1)} = ce^{-\lambda\Delta(s-2)}e^{-\lambda\Delta} = \text{cor}(y_n^2, y_{n+(s-1)}^2) e^{-\lambda\Delta}, \quad (2.30)$$

dla $s = 1, 2, 3, \dots$, gdzie:

$$c = \frac{(1 - e^{-\lambda\Delta})^2}{6(e^{-\lambda\Delta} - 1 + \lambda\Delta) + 2\Delta^2(\xi/\omega)^2} \geq 0.$$

Podobnie można wyznaczyć funkcję autokorelacji dla wariancji aktualnej

$$\text{cor}(\sigma_n^2, \sigma_{n+s}^2) = de^{-\lambda\Delta(s-1)} = de^{-\lambda\Delta(s-2)}e^{-\lambda\Delta} = \text{cor}(\sigma_n^2, \sigma_{n+(s-1)}^2) e^{-\lambda\Delta}, \quad (2.31)$$

dla $s = 1, 2, 3, \dots$, gdzie

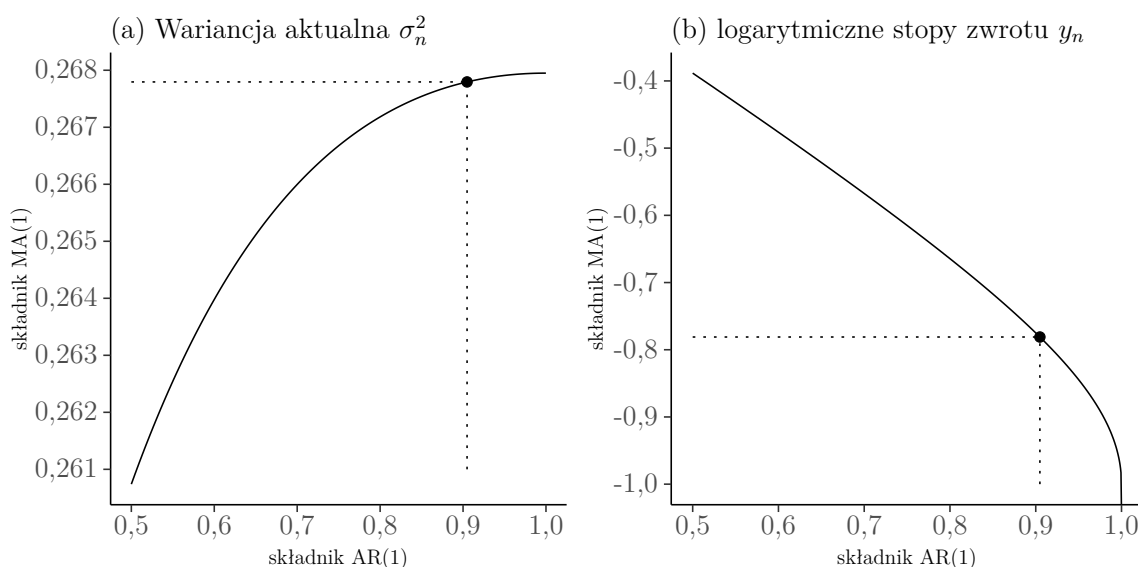
$$d = \frac{(1 - e^{-\lambda\Delta})^2}{2(e^{-\lambda\Delta} - 1 + \lambda\Delta)}.$$

Pomiędzy stałymi c i d ze wzorów (2.30) i (2.31) zachodzi nierówność $1 \geq d \geq c$. Funkcje autokorelacji dla wariancji aktualnej i kwadratów logarytmicznych stóp zwrotu implikują, że σ_n^2 oraz y_n^2 mają postać procesu ARMA(1,1), przy czym mają wspólny parametr autoregresji $e^{-\lambda\Delta}$, który dla typowych wartości parametru ($\lambda \in (0, 1)$) i długości przedziału ($0 < \Delta \leq 1$) przyjmuje wartość bliską 1. Parametr średniej ruchomej wyznacza się numerycznie¹⁰. Parametr średniej ruchomej co do wartości bezwzględnej jest większy dla y_n^2 niż dla σ_n^2 . W konsekwencji korelogram dla y_n^2 znacznie gorzej niż w przypadku σ_n^2 pokazywałby (zmienność aktualna jest nieobserwowana, możemy tylko wyznaczyć jej oszacowanie), że proces generujący dane jest postaci ARMA(1,1). Na rysunku 2.4 zależności pomiędzy składnikami AR(1) a parametrem średniej ruchomej MA(1) odpowiednio dla σ_n^2 (a) i y_n^2 (b). Na obu wykresach przedstawiono punkt dla $\Delta = 1$ i $\lambda = 0, 1$, co oznacza składnik AR(1) o wartości w przybliżeniu równym 0,905. Odpowiada to w przypadku zmienności

¹⁰Model ARMA(1,1) ma funkcję korelacji postaci $r(s) = \alpha r(s-1)$, $s \geq 2$, gdzie α jest parametrem składnika autokorelacji, natomiast parametr β odpowiadający średniej ruchomej wyznacza się numerycznie rozwiązując równanie $r(1) = (\alpha + \beta)/(1 + \beta^2 + 2\alpha\beta)$ (por. Maddala (2008, str. 588)).

aktualnej wartości składnika MA(1) równego w przybliżeniu 0,268, a dla kwadratów stóp zwrotu $-0,782$.

Malejąca wykładniczo funkcja autokorelacji dla kwadratów stóp zwrotu zazwyczaj nie jest zgodna z obserwowanymi własnościami finansowych szeregów czasowych. Wielu autorów sugeruje, że funkcja autokorelacji dla kwadratów stóp zwrotu powinna przyjmować raczej postać funkcji potęgowej o ujemnym wykładniku (por. Gopikrishnan i in. (2000), Cont (2007), Kliber (2013, str. 52)).



Rysunek 2.4: Zależność pomiędzy parametrem autoregresyjnym AR(1) a parametrem średniej ruchomej MA(1) odpowiednio dla σ_n^2 (a) i y_n^2 (b), przy czym zaznaczono punkt dla $\Delta = 1$ i $\lambda = 0,1$.

Źródło: opracowanie własne.

2.3 Złożenie procesów zmienności

Podstawowy model stochastycznej zmienności BNS mimo wielu własności zgodnych z obserwowanymi empirycznie szeregami czasowymi, posiada pewne ograniczenia. Jednym z nich jest struktura typu ARMA(1,1) dla autokorelacji kwadratów logarytmicznych stóp zwrotu opisana wzorem (2.30). Choć taką samą strukturę korelacji ma powszechnie używany model GARCH(1,1), a także model stochastycznej zmien-

ności CEV¹¹, to jednak uważa się to za wadę, a nie zaletę. Proces ARMA(1,1) jest procesem z krótką pamięcią. W szczególności dla funkcji autokorelacji kwadratów stóp zwrotu w modelu BNS zachodzi

$$\sum_{s=-\infty}^{+\infty} |r(s)| = \sum_{s=-\infty}^{+\infty} ce^{-\lambda|\Delta s|} = c \frac{1 + e^{\lambda\Delta}}{1 - e^{\lambda\Delta}} < +\infty.$$

Natomiast obserwacje finansowych szeregów czasowych sugerują, że kwadraty stóp zwrotu mogą charakteryzować się długą pamięcią (por. Kliber (2013, str. 52)). Autokorelacja kwadratów logarytmicznych stóp zwrotu maleje początkowo szybko, a potem od pewnego momentu dużo wolniej, zachowując wartości istotnie różne od zera nawet dla bardzo długich opóźnień. Wyklucza to funkcję autokorelacji w postaci przeskalowanej funkcji wykładniczej. Lepsze dopasowanie uzyskuje się wykorzystując przeskalowane funkcje potęgowe. Są to funkcje autokorelacji odpowiadające procesom o długiej pamięci. Część autorów kwestionuje istnienie długiej pamięci i występujące anomalie wiążą z niestacjonarnością szeregów czasowych (Mikosch i Starica, 2000). Inni autorzy wskazują na niedoskonałości procedur testowania długiej pamięci i uważają istnienie długiej pamięci za kwestię otwartą (Cont, 2005). Nie zawsze także badania empiryczne potwierdzają istnienie długiej pamięci (Gie-rej, 2008).

W celu uzyskania bardziej zbliżonej z obserwacjami funkcji autokorelacji dla kwadratów stóp zwrotów (Barndorff-Nielsen i Shephard, 2001b) zaproponowali złożenie (*superposition*) kilku procesów zmienności w postaci sumy.

Definicja 2.2.

Złożeniem procesów Ornsteina-Uhlenbecka o rozkładzie stacjonarnym $\mathcal{D}$ nazywamy proces wariancji chwilowej $(\sigma^2(t))_{t \geq 0}$ powstały jako suma

$$\sigma^2(t) = \sum_{p=1}^P \sigma_p^2(t) \tag{2.32}$$

niezależnych dostatnich niegaussowskich procesów Ornsteina-Uhlenbecka o stacjonarnych rozkładach należących do jednej rodziny rozkładów samorozkładalnych $\mathcal{D}$

¹¹Model stochastycznej zmienności CEV, ma taką samą postać funkcji autokorelacji dla procesu wariancji chwilowej co model BNS. Jeżeli w modelu CEV składniki losowe procesu cen instrumentu finansowego i zmienności są nieskorelowane, to postać autokorelacji kwadratów logarytmicznych stóp zwrotu jest dokładnie taka sama jak w modelu BNS (Barndorff-Nielsen i Shephard, 2001b).

zamkniętej ze względu na splot i spełniającymi dla $p = 1, \dots, P$ stochastyczne równania różniczkowe

$$d\sigma_p^2(t) = -\lambda_p \sigma_p^2(t)dt + dZ_p(\lambda_p t), \quad \sigma^2(0) > 0, \quad (2.33)$$

gdzie $(Z_1(\lambda_1 t))_{t \geq 0}, \dots, (Z_P(\lambda_P t))_{t \geq 0}$ są niezależnymi procesami podporządkowanymi.

Złożenie procesów Ornsteina-Uhlenbecka o rozkładzie stacjonarnym $\mathcal{D}$ można w skrócie zapisać jako proces $\mathcal{D}$ -supOU $_P$. Identyfikacja modelu wymaga, aby przyjąć dodatkowe założenie dla parametrów $\lambda_1, \dots, \lambda_P$ np. $\lambda_1 < \lambda_2 < \dots < \lambda_P$.

Liniowa postać złożenia (2.32) oraz założenie o niezależności procesów BDLP powoduje że wariancje scałkowane i aktualne są również sumami

$$\sigma^{2*}(t) = \sum_{p=1}^P \sigma_p^{2*}(t), \quad \sigma_n^2 = \sum_{p=1}^P \sigma_{p,n}^2.$$

Przykładami rodzin rozkładów samorozkładalnych zamkniętych ze względu na splot są gamma i odwrotny Gaussa. Natomiast rodzina uogólnionych rozkładów odwrotnych Gaussa nie jest zamknięta ze względu na splot poza tymi dwoma szczególnymi przypadkami. Jeżeli $X_p \sim \text{Ga}(\nu_p, \alpha)$, dla $p = 1, \dots, n$, to suma $X \stackrel{d}{=} X_1 + X_2, \dots, X_P$ ma rozkład $\text{Ga}(\sum_{p=1}^P \nu_p, \alpha)$. Podobnie można pokazać w przypadku rozkładu odwrotnego Gaussa: jeżeli $X_p \sim \text{IG}(\delta_p, \gamma)$, dla $p = 1, \dots, n$, to suma $X \stackrel{d}{=} X_1 + X_2, \dots, X_P$ ma rozkład $\text{IG}(\sum_{p=1}^P \delta_p, \gamma)$. Dodatkowo przyjmując dla rozkładu gamma $\nu = \sum_{p=1}^P \nu_p$ oraz $w_p = \nu_p / \nu$ oraz dla rozkładu odwrotnego Gaussa $\delta = \sum_{p=1}^P \delta_p$ oraz $w_p = \delta_p / \delta$ otrzymujemy

$$\mathbb{E}X_p = w_p \xi \quad \text{oraz} \quad \text{Var}(X_p) = w_p \omega^2, \quad (2.34)$$

gdzie $\mathbb{E}X = \xi$ oraz $\text{Var}(X) = \omega^2$. Pozwala to na skonstruowanie procesu $\mathcal{D}$ -supOU $_P$ o rozkładzie stacjonarnym gamma lub odwrotny Gaussa, dla których każda ze składowych procesów złożenia $(\sigma_p^2(t))_{t \geq 0}$ ma taki sam stosunek wartości oczekiwanej do wariancji rozkładu stacjonarnego ξ_p / ω_p^2 jak suma $(\sigma_p^2(t))_{t \geq 0}$ dana wzorem (2.32):

$$\frac{\xi_p}{\omega_p^2} = \frac{w_p \xi}{w_p \omega^2} = \frac{\xi}{\omega^2}.$$

W konsekwencji można otrzymać proces $\mathcal{D}$ -supOU $_P$ o takim samym o rozkładzie co $\mathcal{D}$ -OU, ale o bogatszej strukturze autokorelacji procesu wariancji chwilowej.

Funkcja ta przyjmuje dla złożenia (2.32) postać

$$r(t) = \sum_{p=1}^P w_p r_p(t) = \sum_{p=1}^P w_p e^{-\lambda_p |t|} \quad (2.35)$$

gdzie wagi $w_p, \dots, w_P$ są nieujemne i sumują się do 1. W konsekwencji funkcja autokorelacji kwadratów stóp zwrotów przyjmuje postać

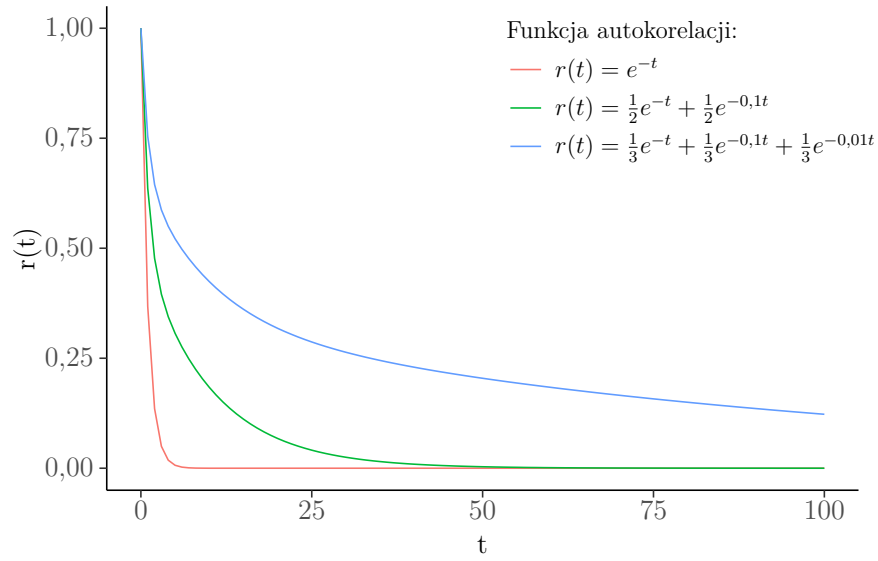
$$\text{cor}(y_n^2, y_{n+s}^2) = \sum_{p=1}^P c_p e^{-\lambda \Delta (s-1)}, \quad (2.36)$$

gdzie

$$c_p = \frac{w_p / \lambda_p^2 (1 - e^{-\lambda_p \Delta})^2}{6 \sum_{p=1}^P w_p / \lambda_p^2 (e^{-\lambda_p \Delta} - 1 + \lambda_p \Delta) + 2 \Delta^2 (\sum_{p=1}^P \xi)^2 / (\sum_{p=1}^P \omega_p)^2}.$$

Należy zwrócić uwagę, że kwadraty stóp zwrotu dla proces wariacji chwilowej supOU mają nadal krótką pamięć, ponieważ ich funkcja autokorelacji jest skończoną sumą funkcji autokorelacji o skończonych sumach. Natomiast funkcja autokorelacji dana wzorem (2.36) ma dużo lepsze dopasowanie do empirycznej funkcji autokorelacji niż model bez złożenia. Badania empiryczne z wykorzystaniem modelu BNS wskazują, że z reguły złożenie dwóch procesów zmienności dostarcza odpowiednie dopasowanie do danych empirycznych (por. Griffin i Steel (2010) oraz Taufer i in. (2011)). Procesy te można następująco interpretować: pierwszy o dużej wartości parametru λ odpowiada za duże, ale rzadkie skoki zmienności, natomiast drugi (o mniejszej wartości parametru λ) za małe, ale częstsze skoki zmienności. Jest to zgodne z obserwacjami m.in. Alizadeh i in. (2002) oraz Lebaron (2001), którzy zaproponowali użycie modeli stochastycznej zmienności o dwóch składnikach (nie będących niegaussowskimi procesami Ornsteina-Uhlenbecka): jednym odpowiadającym za długookresową, a drugi za krótkookresową dynamikę zmienności. Podobnie, na istnienie dwóch czynników w zmienności stóp zwrotu, jednego o bardzo krótkiej, a drugiego o długiej persistencji wskazali Andersen i in. (2002).

Na rysunku 2.5 przedstawiono wykresy funkcji autokorelacji dla pojedynczej funkcji wariacji chwilowej oraz dla złożenia dwóch i trzech funkcji zmienności. Dla uproszczenia przyjęto $\Delta = 1$ oraz równe wagi.



Rysunek 2.5: Funkcje autokorelacji dla jednego procesu zmienności o parametrze $\lambda = 1$ (kolor czerwony), dwóch procesów o parametrach $\lambda = 1$ oraz $\lambda = 0,1$ (kolor zielony) i trzech procesów zmienności o parametrach $\lambda = 1$, $\lambda = 0,1$ oraz $\lambda = 0,01$ (kolor niebieski). Dla uproszczenia zapisu przyjęto $\Delta = 1$.

Źródło: opracowanie własne.

Barndorff-Nielsen i Shephard (2001b) wskazali również, że można uogólnić złożenie skończonej ilości procesów zmienności na przypadek złożenia nieskończonej ilości, co pozwala uzyskać funkcję autokorelacji procesu wariancji chwilowej z długą pamięcią. Zwiększając liczbę składników we wzorze 2.32 można w granicy przy $P \rightarrow +\infty$ otrzymać proces ciągłego złożenia procesów zmienności, dla którego funkcja autokorelacji przyjmuje postać

$$r(t) = (1 + \lambda|t|)^{-2(1-H)}, \quad (2.37)$$

gdzie $H \in (\frac{1}{2}, 1)$ jest parametrem długiej pamięci. Griffin i Steel (2010) zaproponowali, aby złożenie procesów zmienności w niegaussowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka zapisać w postaci

$$\sigma^2(t) = \int \sigma_\lambda^2(t) dF(\lambda), \quad (2.38)$$

gdzie F jest rozkładem prawdopodobieństwa, który Griffin i Steel (2010) nazwali rozkładem miksującym. Funkcja autokorelacji procesu $(\sigma^2(t))_{t \geq 0}$ przyjmuje wówczas

postać

$$r(t) = \int \exp(-\lambda|t|) dF(\lambda). \quad (2.39)$$

Wybór rozkładu determinuje postać funkcji autokorelacji. W szczególności rozkład dyskretny o skończonej liczbie punktów z dodatnim prawdopodobieństwem prowadzi do funkcji autokorelacji postaci 2.35. Natomiast rozkład gamma o parametrze kształtu α i skali ϕ prowadzi do funkcji autokorelacji postaci

$$r(t) = \left(1 + \frac{|t|}{\phi}\right)^{-\alpha}, \quad (2.40)$$

co odpowiada wzorowi (2.37) dla $\lambda = 1/\phi$ oraz $H = 1 - \frac{\alpha}{2}$.

2.4 Efekt dźwigni

Kolejne rozszerzenie podstawowego modelu stochastycznej zmienności BNS polega na uwzględnieniu w modelu „efektu dźwigni”, czyli ujemnej korelacji pomiędzy wartościami stóp zwrotu i miernikami ich zmienności. Obserwacja szeregów finansowych, szczególnie giełdowych, wskazuje, że zachowanie inwestorów jest niesymetryczne, to znaczy inaczej zachowują się w przypadku pojawienia się korzystnej i niekorzystnej informacji. Okazuje się, że w odpowiedzi na „złe informacje” wzrost zmienności jest większy niż w przypadku „dobrych informacji”. Większa zmienność w przypadku niekorzystnej informacji implikuje także większe prawdopodobieństwo wystąpienia skrajnie ujemnych wartości stóp zwrotu niż skrajnie dodatnich wartości w przypadku korzystnej informacji.

Efekt dźwigni po raz pierwszy pojawił się w badaniach ekonomicznych za sprawą pracy Black (1976). Black zauważył, że zmienność rośnie, gdy cena aktywa spada. Zaproponował następujące wytłumaczenie: spadek cen akcji danej spółki powoduje identyczny spadek kapitału własnego spółki; jednocześnie rynkowa cena długu pozostaje niezmienną, co powoduje wzrost współczynnika zadłużenia do kapitału własnego (debt-to-equity ratio) tzw. „dźwigni finansowej”. Teoretyczne uzasadnienie „efektu dźwigni” w obrębie twierdzeń Modigliani-Millera (fundamentalnych twierdzeń dotyczących finansów przedsiębiorstw) zostało przedstawione w Christie (1982). Efekt dźwigni został uwzględniony w uogólnieniach modelu GARCH m.in EGARCH

(Nelson, 1991) i GJR-GARCH (Glosten i in., 1993) oraz w modelach stochastycznej zmienności opartych o transformację logarytmiczną Ornsteina-Uhlenbecka (Jacquier i in., 1999), czy modelu Hestona (Heston, 1993).

Barndorff-Nielsen i Shephard (2001b) zaproponowali uwzględnienie efektu dźwigni w modelu BNS zarówno w modelu bez złożenia jak i ze złożenia procesów wariancji chwilowej.

Definicja 2.3.

Modelem stochastycznej zmienności Bandorffa-Nielsena i Shepharda z efektem dźwigni nazywamy model stochastycznej zmienności, w którym proces logarytmicznych cen $(Y(t))_{t \geq 0}$ spełnia następujące równanie

$$dY(t) = (\mu + \beta\sigma^2(t))dt + \sigma(t)dW(t) + \rho d\bar{Z}(\lambda t), \quad Y(0) = 0, \quad (2.41)$$

gdzie $(\bar{Z}(\lambda t))_{t \geq 0}$ jest skompensowanym procesem podporządkowanym postaci

$$\bar{Z}(\lambda t) = Z(\lambda t) - \mathbb{E}(Z(\lambda t)),$$

gdzie $\mu, \beta, \rho \in \mathbb{R}, \lambda \in \mathbb{R}_+$, natomiast proces wariancji chwilowej $(\sigma^2(t))_{t \geq 0}$ jest zdefiniowany zgodnie z definicją 2.1 modelu BNS.

Parametr ρ mierzy siłę korelacji pomiędzy stopami stóp zwrotu a ich zmiennością i zgodnie z teorią powinien przyjmować wartości ujemne. Kolejną konsekwencją uwzględnienia efektu dźwigni jest nieciągłość trajektorii procesu logarytmów cen. Za punkty nieciągłości odpowiada wprowadzony we wzorze (2.41) składnik $dZ(\lambda t)$.

Efekt dźwigni wpływa na własności logarytmicznych stóp zwrotu. Warunkowy rozkład logarytmicznych stóp zwrotu przyjmuje postać

$$y_n \sim N(\mu\Delta + \beta\sigma_n^2 + \rho\bar{z}_n, \sigma_n^2), \quad (2.42)$$

gdzie

$$\bar{z}_n = \int_{\Delta(n-1)}^{\Delta n} d\bar{Z} - \lambda\Delta\mathbb{E}(Z(1)) = Z(\lambda\Delta n) - Z(\lambda\Delta(n-1)) - \lambda\Delta\xi,$$

$\xi = \mathbb{E}(\sigma^2(t)) = \mathbb{E}(Z(1))$. Ponadto efekt dźwigni przejawia się w skorelowaniu stóp zwrotu z kwadratami stóp zwrotu i autokorelacji kwadratów stóp zwrotu

$$\text{cov}(y_n, y_{n+s}^2) = \mathbb{E}(y_n y_{n+s}^2) = \rho\kappa_2 e^{-\lambda\Delta(s-1)}, \quad (2.43)$$

$$\text{cov}(y_n^2, y_{n+s}^2) = \left(\frac{\kappa_2}{2\lambda^2} + \rho^2\mu_3\right) e^{-\lambda\Delta(s-1)}, \quad (2.44)$$

gdzie κ_2 oraz μ_3 są odpowiednio kumulantą drugiego rzędu i momentem trzeciego rzędu zmiennej losowej $Z(1)$.

Uwzględnienie efektu dźwigni zgodnie ze wzorem (2.41) przejawia się m.in. w ujemnej zależności pomiędzy zwrotami i kwadratami zwrotów, przy czym kwadraty stóp zwrotu uważa się za najprostszą miarę zmienności procesu logarytmu cen.

Wykładnicza względem opóźnienia postać zależności (2.43) i (2.44) jest taka sama jak modelu QARCH (*quadratic ARCH*) zaproponowanym przez Sentana (1995). Ponadto ze względu na występowania we wzorze (2.44) parametru μ_3 wynika, że asymetria prawostronna (lewostronna) procesu BDPL powoduje wzrost (spadek) autokorelacji kwadratów stóp zwrotu. Zazwyczaj procesy BDPL mają prawostronną asymetrię (np. procesy BDPL dla procesów wariancji chwilowej Ga-OU oraz IG-OU).

Efekt dźwigni można także uogólnić na przypadek złożenia p procesów wariancji chwilowej pozwalając, aby z każdym skompensowanym procesem BDPL związany był osobny parametr ρ_j .

Definicja 2.4.

Modelem stochastycznej zmienności Bandorffa-Nielsena i Shepharda ze złożeniem procesów zmienności i efektem dźwigni nazywamy model stochastycznej zmienności, w którym proces logarytmu cen $(Y(t))_{t \geq 0}$ spełnia następujące równanie

$$dY(t) = (\mu + \beta \sigma^2(t)) dt + \sigma(t) dW(t) + \sum_{p=1}^P \rho_p d\bar{Z}_p(\lambda_p t), \quad \ln S(0) = 0, \quad (2.45)$$

gdzie proces wariancji chwilowej $(\sigma^2(t))_{t \geq 0}$ jest złożeniem procesów wariancji chwilowej $(\sigma_1^2(t))_{t \geq 0}, \dots, (\sigma_P^2(t))_{t \geq 0}$ zgodnie z definicją 2.2 oraz $(\bar{Z}_p(\lambda_p t))_{t \geq 0}$ dla $p=1, \dots, P$ są skompensowanymi procesami podporządkowanymi postaci

$$\bar{Z}_p(\lambda t) = Z_p(\lambda_p t) - \mathbb{E}(Z_p(\lambda_p t)),$$

gdzie $\mu, \beta, \rho_1, \dots, \rho_P \in \mathbb{R}, \lambda_1, \dots, \lambda_P \in \mathbb{R}_+$.

Model stochastycznej zmienności ze złożeniem procesów zmienności i efektem dźwigni jest najbardziej ogólnym i najbardziej elastycznym modelem stochastycznej zmienności z zaproponowanych przez Bandorffa-Nielsena i Shepharda w początkowej

pracy Barndorff-Nielsen i Shephard (2001b). Model ten pozwala zarówno skonstruować proces wariacji mający bardziej realistyczną strukturę korelacji i uchwycić efekt dźwigni. Postać równania (2.45) pozwala przypisać każdemu z skompensowanych procesów BDPL oddzielnego parametru ρ_p .

2.5 Rozkłady stacjonarne procesu wariacji chwilowej

Niegaussowskie procesy Ornsteina-Uhlenbecka pozwalają na zastosowanie szerokiej klasy rozkładów samorozkładalnych do modelowania rozkładu stacjonarnego procesu wariacji chwilowej. Jednak nie wszystkie rozkłady samorozkładalne mogą zostać użyte jako rozkład stacjonarny procesu wariacji chwilowej, ponieważ klasa ta obejmuje również rozkłady o nośniku funkcji gęstości obejmującym cały zbiór liczb rzeczywistych. Z natury zjawiska zmienności wynika, że jej rozkład stacjonarny powinien być nieujemny. Przykładami rozkładów samorozkładalnych, które nie mogą być użyte są rozkłady normalny oraz normalny odwrotny Gaussa NIG. Poniżej zostanie przedstawione przegląd rozkładów jakie mogą być użyte jako rozkład stacjonarny w niegaussowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka. Niegaussowskie modele stochastycznej zmienności typu Ornsteina-Uhlenbecka można również tworzyć przez wybór prowadzącego procesu Lévy'ego ukrytego w tle (BDLP) (por. twierdzenie 2.2). W pracach James i in. (2013) oraz James i in. (2018) rozważany był przypadek, w którym procesem BDLP był proces gamma, a w pracy Zhang (2011) proces temperowany stabilny.

1. Uogólniony odwrotny rozkład Gaussa GIG

Uogólniony odwrotny rozkład Gaussa $GIG(\nu, \delta, \gamma)$ (*Generalized Inverse Gaussian*) ma funkcję gęstości daną wzorem

$$f_{GIG}(x) = \frac{(\gamma/\delta)^\nu}{2K_\nu(\delta\gamma)} x^{\nu-1} \exp\left(-\frac{\delta^2 x^{-1} + \gamma^2 x}{2}\right),$$

gdzie $\delta > 0$, $\gamma > 0$, $\nu \in \mathbb{R}$, $K_\nu(x)$ jest zmodyfikowaną funkcją Bessela trzeciego rodzaju.

Funkcja charakterystyczna uogólnionego odwrotnego rozkładu Gaussa przyjmuje postać

$$\varphi_{GIG}(\zeta; \nu, \delta, \gamma) = \frac{1}{K_\nu(\delta\gamma)} \left(1 - \frac{2i\zeta}{\gamma^2}\right)^{\frac{\nu}{2}} K_\nu\left(\delta\gamma\sqrt{1 - \frac{2i\zeta}{\gamma^2}}\right).$$

Specjalnym przypadkami rozkładu uogólnionego odwrotnego Gaussa są rozkład gamma ($GIG(\nu, 0, \sqrt{2\alpha}) = Ga(\nu, \alpha)$), rozkład odwrotny Gaussa ($GIG(-1/2, \delta, \gamma) = IG(\delta, \gamma)$), dodatni rozkład hiperboliczny $GIG(1, \delta, \gamma) = PH(\delta, \gamma)$ oraz odwrotny gamma $GIG(-\nu, \sqrt{2\alpha}, 0) = IGa(\nu, \alpha)$.

Momenty rozkładu $GIG(\nu, \delta, \gamma)$ można wyznaczać ze wzoru

$$\mathbb{E}(X^k) = \left(\frac{\delta}{\gamma}\right)^k \frac{K_{\nu+k}(\delta\gamma)}{K_\nu(\delta\gamma)}, \quad \text{dla } k \in \mathbb{Z}_+.$$

Rozkład GIG został po raz pierwszy użyty przez Good (1953). Blæsild (1978) podał wzory na momenty i kumulanty rozkładu GIG. Samorozkładalność rozkładu GIG udowodnił Halgreen (1979). Barndorff-Nielsen (1978) jako pierwszy przedstawił uogólniony odwrotny rozkład Gaussa jako rozkład miksujący w normalnych mieszaninach średnio-wariacyjnych dla uogólnionego rozkładu hiperbolicznego.

Użycie tego rozkładu w niegaussowskich modelach stochastycznej zmienności ogranicza to, że rozkład ten poza specjalnymi przypadkami (rozkłady odwrotny Gaussa i gamma) nie jest zamknięty ze względu na spłot i nie można procesów supOU o stacjonarnym rozkładzie uogólnionym odwrotnym Gaussa zgodnie z definicją 2.2.

Uogólniony odwrotny rozkład Gaussa został wykorzystany jako rozkład stacjonarny procesu wariancji chwilowej w pracy Gander i Stephens (2007b) w estymacji za pomocą metod Monte Carlo opartych na łańcuchach Markowa w badaniu dotyczącym spółek pochodzących z amerykańskiego rynku finansowego. Ponadto w pracy Taufer i in. (2011) przedstawiono jak wykorzystać uogólniony odwrotny rozkład Gaussa jako rozkład stacjonarny procesu wariancji chwilowej w metodzie empirycznych funkcji charakterystycznych, ale w badaniu empirycznym (kurs indeksu S&P500¹²) zastosowano tylko specjalny przypadek: rozkład odwrotny Gaussa.

¹²S&P500 to indeks giełdowy, w skład którego wchodzi 500 przedsiębiorstw o największej kapitalizacji, notowanych na New York Stock Exchange i NASDAQ.

2. Rozkład gamma

Rozkład gamma $\text{Ga}(\nu, \alpha)$ ma gęstość daną wzorem

$$f_{\text{Ga}}(x; \nu, \alpha) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} \exp(-\alpha x) \mathbb{1}_{(0, +\infty)}(x),$$

gdzie $\Gamma(\cdot)$ oznacza funkcję gamma.

Parametr ν jest nazywany parametrem kształtu, natomiast $1/\alpha$ parametrem skali. Specjalnymi przypadkami rozkładu gamma są rozkłady wykładniczy ($\nu = 1$), Erlanga ($\nu \in \mathbb{N}$) oraz Chi-kwadrat ($\alpha = 1/2, \nu = k/2, k \in \mathbb{N}$). Rozkład odwrotny gamma jest specjalnym przypadkiem rozkładu Pearsona III typu.

Funkcja charakterystyczna zmiennej losowej X o rozkładzie gamma przyjmuje postać

$$\varphi_{\text{Ga}}(\zeta; \nu, \alpha) = \left(1 - \frac{i\zeta}{\alpha}\right)^{-\nu}.$$

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie gamma $\text{Ga}(\nu, \alpha)$ są równe odpowiednio

$$\mathbb{E}X = \frac{\nu}{\alpha} \quad \text{oraz} \quad \text{Var}(X) = \frac{\nu}{\alpha^2}.$$

Rozkład gamma jest zamknięty ze względu na sploty: jeżeli $X_p \sim \text{Ga}(\nu_p, \alpha)$, dla $p = 1, \dots, n$, to suma $X \stackrel{d}{=} X_1 + X_2 + \dots + X_P$ ma rozkład $\text{Ga}(\sum_{p=1}^P \nu_p, \alpha)$. Może być wykorzystany do tworzenia procesów supOU zgodnie z definicją 2.2.

Nicolato i Venardos (2003) podali analityczne wzory na wycenę opcji europejskich w niegaußowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka z rozkładem stacjonarnym procesu wariancji chwilowej gamma. Nicolato i Venardos (2003) przedstawili również wyniki kalibracji modelu do europejskich opcji kupna dla kursu indeksu S&P500.

Niegaußowskie procesy Ornsteina-Uhlenbecka o rozkładzie stacjonarnym gamma są jedynymi rozpatrywanymi w literaturze procesami tego typu o skończonej ilości skoków w jednostce czasu (skończonej aktywności).

Rozkład gamma jest najczęściej wykorzystywanym rozkładem stacjonarnym procesu wariancji chwilowej w niegaußowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka. Został użyty w między innymi następujących pracach:

Roberts i in. (2004), Gander i Stephens (2007b) (jako jeden ze specjalnych przypadków rozkładu GIG), Frühwirth-Schnatter i Sögner (2009), Griffin i Steel (2010) (w modelu z ciągłym złożeniem procesów zmienności) w estymacji za pomocą metod Monte Carlo opartych na łańcuchach Markowa, Hubalek i Posedel (2011) w estymacji za pomocą funkcji estymujących, Benth i in. (2017) w kalibracji modelu kontraktów na certyfikatach do emisji dwutlenku węgla (UE ETS) notowanymi na giełdzie Nord Pool¹³.

3. Rozkład odwrotny Gaussa

Rozkład odwrotny Gaussa $IG(\delta, \gamma)$ ma gęstość daną wzorem

$$f_{IG}(x; \delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} \exp(\delta\gamma) x^{-3/2} \exp\left(-\frac{\delta^2 x^{-1} + \gamma^2 x}{2}\right) \mathbb{1}_{(0, +\infty)}(x).$$

Funkcja charakterystyczna zmiennej losowej X o rozkładzie odwrotnym Gaussa przyjmuje postać

$$\varphi_{IG}(\zeta; \delta, \gamma) = \exp\left[-\delta\left(\sqrt{-2i\zeta + \gamma^2} - \gamma\right)\right].$$

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie odwrotnym Gaussa $IG(\delta, \gamma)$ są równe odpowiednio

$$\mathbb{E}X = \frac{\delta}{\gamma} \quad \text{oraz} \quad \text{Var}(X) = \frac{\delta}{\gamma^3}.$$

Nazwa rozkładu pochodzi od procesu Wienera, zwanego również procesem Gaussa. Niech $(\tilde{W}(t))_{t \geq 0}$ będzie standardowym procesem Wienera z dryfem γ i parametrem dyfuzji $\sigma = 1$. Wówczas czas osiągnięcia przez ten proces wartości $\delta > 0$ jest zmienną losową o rozkładzie $IG(\delta, \gamma)$.

Rozkład odwrotny Gaussa jest zamknięty ze względu na sploty: jeżeli $X_p \sim IG(\delta_p, \gamma)$, dla $p = 1, \dots, n$, to suma $X \stackrel{d}{=} X_1 + X_2 + \dots + X_P$ ma rozkład $IG(\sum_{p=1}^P \delta_p, \gamma)$. Może być, podobnie jak rozkład gamma, wykorzystany do tworzenia procesów supOU zgodnie z definicją 2.2.

Nicolato i Venardos (2003) podali analityczne wzory na wycenę opcji europejskich w niegaussowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka

¹³Giełda Nord Pool (*Nordic Energy Pool*) jest zaliczana do najbardziej rozwiniętych giełd energii elektrycznej na świecie.

z rozkładem stacjonarnym procesu wariancji chwilowej odwrotnym Gaussa oraz przedstawili wyniki kalibracji modelu do europejskich opcji kupna dla kursu indeksu S&P500.

Proces odwrotny Gaussa został wykorzystany jako rozkład stacjonarny procesu wariancji chwilowej w następujących pracach: Gander i Stephens (2007b) w estymacji za pomocą metod MCMC w badaniu dotyczącym spółek pochodzących z amerykańskiego rynku finansowego, Lindberg (2008) w aproksymacji zmienności za pomocą wielkości obrotu i estymacji metodą największej wiarygodności w badaniu dotyczącym stóp zwrotu z akcji spółki Ericsson B notowanej na sztokholmskiej giełdzie OMX Stockholmsbörsen, Taufer i in. (2011) do estymacji metodą empirycznych funkcji empirycznych dla modelowania zmienności indeksu S&P500, Benth i in. (2017) w kalibracji modelu kontraktów na certyfikatach do emisji dwutlenku węgla (UE ETS) notowanymi na giełdzie Nord Pool.

4. Dodatni rozkład hiperboliczny

Dodatni rozkład hiperboliczny $PH(\delta, \gamma)$ (*Positive Hyperbolic*) ma funkcję gęstości daną wzorem

$$f_{PH}(x) = \frac{(\gamma/\delta)}{2K_1(\delta\gamma)} \exp\left(-\frac{\delta^2 x^{-1} + \gamma^2 x}{2}\right),$$

gdzie $\delta > 0$, $\gamma > 0$, $K_1(x)$ jest zmodyfikowaną funkcją Bessela trzeciego rodzaju.

Funkcja charakterystyczna dodatniego rozkładu hiperbolicznego $PH(\delta, \gamma)$ przyjmuje postać

$$\varphi_{PH}(\zeta; \delta, \gamma) = \frac{1}{K_1(\delta\gamma)} \left(1 - \frac{2i\zeta}{\gamma^2}\right)^{\frac{1}{2}} K_1\left(\delta\gamma\sqrt{1 - \frac{2i\zeta}{\gamma^2}}\right).$$

Momenty rozkładu $PH(\delta, \gamma)$ można wyznaczać ze wzoru

$$\mathbb{E}(X^k) = \left(\frac{\delta}{\gamma}\right)^k \frac{K_{k+1}(\delta\gamma)}{K_1(\delta\gamma)}, \quad \text{dla } k \in \mathbb{Z}_+.$$

Dodatni rozkład hiperboliczny został zastosowany jako rozkład stacjonarny procesu wariancji chwilowej jedynie w pracy Gander i Stephens (2007b) jako specjalny przypadek uogólnionego odwrotnego rozkładu Gaussa.

5. Rozkład odwrotny gamma

Rozkład odwrotny gamma $IGa(\nu, \alpha)$ ma funkcję gęstości daną wzorem

$$f_{IGa}(x; \nu, \alpha) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{-\nu-1} \exp(-\alpha x^{-1}) \mathbb{1}_{(0,+\infty)}(x),$$

gdzie $\Gamma(\cdot)$ oznacza funkcję gamma. Parametr ν jest nazywany parametrem kształtu, natomiast α parametrem skali.

Funkcja charakterystyczna rozkładu odwrotny gamma $IGa(\nu, \alpha)$ przyjmuje postać

$$\varphi_{IGa}(\zeta; \nu, \alpha) = \frac{2(-i\alpha\zeta)^{\nu/2}}{\Gamma(\nu)} K_1\left(\sqrt{-4i\alpha\zeta}\right).$$

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie odwrotnym gamma $IG(\delta, \gamma)$ są równe odpowiednio

$$\mathbb{E}X = \frac{\alpha}{\nu - 1}, \quad \text{dla } \nu > 1$$

oraz

$$\text{Var}(X) = \frac{\alpha}{(\nu - 1)^2(\nu - 2)}, \quad \text{dla } \nu > 2.$$

Rozkład odwrotny gamma jest ściśle związany z rozkładem gamma: jeżeli zmienna losowa X ma rozkład gamma $Ga(\nu, \alpha)$, to zmienna losowa $\frac{1}{X}$ ma rozkład $IGa(\nu, \alpha)$. Rozkład odwrotny gamma jest rozkładem Pearsona V typu.

Rozkład odwrotny gamma podobnie jak dodatni rozkład hiperboliczny został wykorzystany jako rozkład stacjonarny procesu wariancji chwilowej jedynie w pracy Gander i Stephens (2007b).

6. Temperowany rozkład stabilny

Temperowany rozkład stabilny o parametrach $0 < \kappa < 1$, $\nu, \alpha > 0$ ma funkcję gęstości daną wzorem

$$f_{TS}(x; \kappa, \nu, \alpha) = e^{\nu\alpha} f_{Y|\kappa, \nu}(x) \exp(-\alpha^{1/\kappa} x/2) \mathbb{1}_{(0,+\infty)}(x),$$

gdzie

$$f_{Y|\kappa, \nu}(x) = \frac{\nu^{-1/\kappa}}{2\pi} \sum_{j=1}^{+\infty} \frac{(-1)^{j-1}}{j!} \sin(j\kappa\pi) \Gamma(j\kappa + 1) 2^{j\kappa+1} \left(\frac{x}{\nu^{1/\kappa}}\right)^{-k\kappa-1}$$

jest gęstością dodatniego κ -stabilnego rozkładu¹⁴. Nazwa rozkładu wynika z tego, że rozkłady te powstają przez „wygaszanie” („temperowanie”) ogonów rozkładu κ -stabilnego (Kliber, 2013, str. 159). Temperowane rozkłady stabilne zaproponował Tweedie (1984). Hougaard (1986) wykorzystał je w analizie przeżycia.

Specjalnym przypadkiem rozkładu stabilnego jest rozkład odwrotny Gaussa $TS(1/2, \gamma, 1/\delta) = IG(\delta, \gamma)$, a granicznym przypadkiem przy $\kappa \rightarrow 0$ jest rozkład gamma.

Funkcja charakterystyczna temperowanego rozkładu stabilnego $TS(\kappa, \nu, \alpha)$ dana jest wzorem

$$\varphi_{TS}(\zeta; \kappa, \nu, \alpha) = \exp \left[\nu \alpha - \nu \left(\alpha^{1/\kappa} - 2i\zeta \right)^\kappa \right].$$

Wartość oczekiwana i wariancja zmiennej losowej X o temperowanym rozkładzie stabilnym $TS(\kappa, \nu, \alpha)$ są równe odpowiednio

$$\mathbb{E}X = 2\nu\kappa\alpha^{\frac{\kappa-1}{\kappa}} \quad \text{oraz} \quad \text{Var}(X) = 4\nu\kappa(1-\kappa)\alpha^{\frac{\kappa-2}{\kappa}}.$$

Temperowany rozkład stabilny jest zamknięty ze względu na sploty: jeżeli $X_p \sim TS(\kappa, \nu_p, \alpha)$, dla $p = 1, \dots, P$, to suma $X \stackrel{d}{=} X_1 + X_2 + \dots + X_P$ ma rozkład $TS(\kappa, \sum_{p=1}^P \nu_p, \alpha)$. Może zatem być wykorzystany do tworzenia procesów supOU zgodnie z definicją 2.2.

Temperowany rozkład stabilny został zastosowany jako rozkład stacjonarny procesu wariancji chwilowej w pracach Gander i Stephens (2007b), Andrieu i in. (2010) (estymacja za pomocą cząsteczkowych Markowskich Łańcuchów Monte Carlo, *Particle Markov Chain Monte Carlo*) oraz Taufer i in. (2011).

7. Rozkład log-normalny

Rozkład log-normalny $LN(\mu, \sigma^2)$ ma funkcję gęstości daną wzorem

$$f_{LN}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma x} \exp \left(-\frac{(\ln x - \mu)^2}{2\sigma^2} \right) \mathbb{1}_{(0, +\infty)}(x),$$

gdzie $\mu \in \mathbb{R}, \sigma > 0$. Jeżeli zmienna losowa X ma rozkład log-normalny $LN(\mu, \sigma^2)$, to zmienna losowa $\ln X$ ma rozkład normalny $N(\mu, \sigma^2)$.

¹⁴W literaturze rozkłady te są nazywane α -stabilnymi. Tu została użyta inna notacja dla ujednolicenia zapisu.

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie $LN(\mu, \sigma^2)$ są równe odpowiednio

$$\mathbb{E}X = e^{\mu + \sigma^2/2}, \quad \text{Var}(X) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}.$$

Rozkład log-normalny jest rozkładem sanorozkładalnym (dowód przedstawił Bondesson (1982)) i nieskończenie podzielny, ale nie jest znana postać gęstości Lévy’ego dla procesu Ornsteina-Uhlenbecka o takim rozkładzie (Barndorff-Nielsen i Shephard, 2003).

2.6 Wariancja zrealizowana

Jedną z najważniejszych nowości w badaniach nad zmiennością finansowych szeregów czasowych ostatnich lat jest wprowadzenie zagadnienia wariancji zrealizowanej.

Główną powodem pojawienia się tej tematyki jest wzrost dostępności danych wysokiej częstotliwości np. zwrotów obserwowanych co 10, 5, 1 minutę, a nawet danych transakcyjnych. To naturalnie przekierowało zainteresowanie od modeli z czasem dyskretnym do modeli z czasem ciągłym, w których można zmieniać częstotliwość obserwacji szeregów czasowych, a wartości obserwowane według mniejszej częstotliwości można uzyskać poprzez sumowanie obserwacji według większej częstotliwości. Miarą zmienności powstałą zgodnie z tym podejściem jest wariancja zrealizowana. Zastosowanie tego miernika zostało zainicjowane przez prace Andersena i in. (1999) oraz Andersen i in. (2001) odnoszące się odpowiednio do rynku walutowego i rynku papierów wartościowych. Otrzymane wyniki wskazywały, że jest to dość dobry miernik zmienności. Barndorff-Nielsen i Shephard (2002) zbadali własności wariancji zrealizowanej przy dość ogólnych założeniach procesu zmienności. Dalsze badania doprowadziły do powstania kolejnych mierników zmienności takich jak wariancja dwupotęgowa i wielopotęgowa (Barndorff-Nielsen i Shephard, 2004), wariancja dwuskałowa (Zhang i in., 2005), czy wariancja jądrowa (Barndorff-Nielsen i in., 2008). Mają one na celu wyeliminowanie pojawiających się dla bardzo wysokich częstotliwości (odstęp pomiędzy kolejnymi obserwacjami poniżej 1 minuty) efektów „mikrostruktury rynku”. Do najważniejszych efektów mikrostruktury rynku zalicza się m.in. dyskretność cen i nieregularne odstęp czasu pomiędzy transakcjami, różnice pomiędzy ceną kupna a ceną sprzedaży (*spread*), niesynchroniczność handlu,

niedostateczną płynność (por. Tsay (2005), roz. 5). Szeroko przeprowadzone¹⁵ przez Liu i in. (2015) badania porównawcze wskazują, że estymator wariancji zrealizowanej wyznaczony na podstawie 5-minutowych zwrotów jest jednym z najlepszych mierników zmienności.

W tym podrozdziale w oparciu o (Barndorff-Nielsen i Shephard, 2002; Barndorff-Nielsen i Shephard, 2004) przedstawione zostanie jak te nowoczesne mierniki zmienności instrumentów finansowych można połączyć z modelami stochastycznej zmienności Ornsteina-Uhlenbecka.

Dla modelu bez efektu dźwigni i przy założeniu $\mu = \beta = 0$ proces wariacji kwadratowej procesu logarytmicznych cen $(Y(t))_{t \geq 0}$ jest równy wariancji scałkowanej (por. wzór (2.14)). Dla ustalonego $t > 0$ wariację proces procesu logarytmicznych cen wyznacza się ze wzoru (por. definicję 1.15):

$$[Y]_t = \text{plim}_{M \rightarrow +\infty} \sum_{j=1}^M (y^*(t_j) - y^*(t_{j-1}))^2 \quad (2.46)$$

dla dowolnego ciągu podziałów $(\Pi_M)_{M=1}^{+\infty}$ odcinka $[0, t]$ takiego, że $\text{diam}(\Pi_M) = 0$. Wzór (2.46) wskazuje, że dobrym oszacowaniem zmienności aktualnej w danym dniu jest suma kwadratów zwrotów śróddziennych, przy czym wraz ze wzrostem częstotliwości powinniśmy otrzymywać co raz dokładniejsze oszacowania dziennej zmienności. Niestety, graniczne własności wariacji kwadratowej (2.46) ograniczają wspomniane wcześniej efekty mikrostruktury rynku.

Definicja 2.5.

Wariancję zrealizowaną procesu logarytmicznych cen w okresie n nazywamy sumę

$$\{y\}_n = \sum_{j=1}^M \left(Y \left((n-1)\Delta + \frac{\Delta j}{M} \right) - Y \left((n-1)\Delta + \frac{\Delta(j-1)}{M} \right) \right)^2, \quad (2.47)$$

gdzie proces logarytmicznych cen instrumentu finansowego $(Y(t))_{t \geq 0}$ jest obserwowany M -krotnie w ciągu jednego okresu czasu n o długości Δ .

Zazwyczaj jako Δ przyjmuje się jeden dzień. Wówczas zwroty wyznaczone w ciągu dnia nazywane są śróddziennymi. Nic jednak nie stoi na przeszkodzie, aby

¹⁵Liu i in. (2015) wykorzystali 400 wariantów mierników zmienności na bazie składającej się z 31 aktywów spośród 5 klas aktywów, dla szeregów czasowych liczących 11 lat.

przyjąć jako Δ np. tydzień a jako zwroty obserwowane w trakcie okresu przyjąć zwroty dzienne.

Przy założeniu braku efektu mikrostruktury rynku wariancja zrealizowana $\{y\}_n$ jest zgodnym estymatorem σ_n^2 (przy $M \rightarrow +\infty$), a przy dodatkowym założeniu, że w równaniu (2.3) dla procesu logarytmu cen $\mu = \beta = 0$ jest również estymatorem nieobciążonym.

Wariancję zrealizowaną można zdekomponować na dwa składniki

$$\{y\}_n = \sigma_n^2 + u_n, \quad (2.48)$$

gdzie u_n nazywane jest błędem wariancji zrealizowanej, dla którego zachodzi warunek $\mathbb{E}(u_n | \sigma_n^2) = 0$. Wówczas

$$\begin{aligned} \mathbb{E}(\{y\}_n) &= \mathbb{E}(\sigma_n^2) = \Delta\xi, \quad \text{Var}(\{y\}_n) = \text{Var}(\sigma_n^2) + \text{Var}(u_n), \\ \text{cov}(\{y\}_n, \{y\}_{n+s}) &= \text{cov}(\sigma_n^2, \sigma_{n+s}^2). \end{aligned}$$

Przyjmując

$$\sigma_{j,n}^2 = \sigma^{2*} \left((n-1)\Delta + \frac{\Delta j}{M} \right) - \sigma^{2*} \left((n-1)\Delta + \frac{\Delta(j-1)}{M} \right), \quad (2.49)$$

otrzymujemy

$$u_n = \sum_{j=1}^M \sigma_{j,n}^2 (\epsilon_{j,n}^2 - 1),$$

gdzie $\epsilon_{j,n}^2 \sim i.i.d N(0, 1)$ są niezależne od $\sigma_{j,n}^2$. Dla ustalonego n oraz $j = 1, \dots, M$ zmienne losowe $\sigma_{j,n}^2$, mają ten sam rozkład, którego parametry można wyznaczyć ze wzorów na parametry rozkładu zmienności aktualnej dla przedziału czasu Δ/M (por. wzór (2.29)). Pozwala to na wyznaczenie wariancji u_n za pomocą wartości oczekiwanej ξ , wariancji ω^2 oraz parametru λ procesu wariancji chwilowej zgodnie ze wzorem

$$\begin{aligned} \text{Var}(u_n) &= 2M \left[\text{Var}(\sigma_{1,n}^2) + (\mathbb{E}\sigma_{1,n}^2)^2 \right] = \\ &= 2M \left[2\omega^2\lambda^{-2} (e^{-\lambda\Delta/M} - 1 + \lambda\Delta M^{-1}) + (\Delta M^{-1})^2 \xi^2 \right]. \end{aligned} \quad (2.50)$$

Można wówczas wyznaczyć funkcję autokorelacji wariancji zrealizowanej

$$\begin{aligned} \text{cor}(\{y\}_n, \{y\}_{n+s}) &= \frac{\text{cov}(\sigma_n^2, \sigma_{n+s}^2)}{\text{Var}(u_n) + \text{Var}(\sigma_n^2)} = \\ &= \frac{\lambda^{-2} (1 - e^{-\lambda\Delta})^2 e^{-\lambda\Delta(s-1)}}{2M \left[2\omega^2 \lambda^{-2} (e^{-\lambda\Delta/M} - 1 + \lambda\Delta M^{-1}) + (\Delta M^{-1})^2 \xi^2 \right] + 2\omega^2 \lambda^{-2} (e^{-\lambda\Delta} - 1 + \lambda\Delta)} \end{aligned} \quad (2.51)$$

Poza wzorem (2.51) na autokorelację nie zostały wykorzystane własności modelu BNS. Przedstawione właściwości wariancji zrealizowanej są prawdziwe dla dowolnego modelu stochastycznej zmienności, dla którego proces wariancji chwilowej jest stacjonarny, o wartości oczekiwanej ξ , wariancji ω^2 i funkcji autokorelacji r , a równanie (2.3) opisuje dynamikę procesu logarytmu cen. We wzorze (2.51) zastosowano postać funkcji autokorelacji procesu wariancji chwilowej dla modelu BNS (taką samą ma model CEV).

Uwzględnienie złożenia procesów zmienności nieznacznie komplikuje wzór na wariancję błędu wariancji zrealizowanej (2.50). Wariancja $\sigma_{1,n}^2$ przyjmuje postać

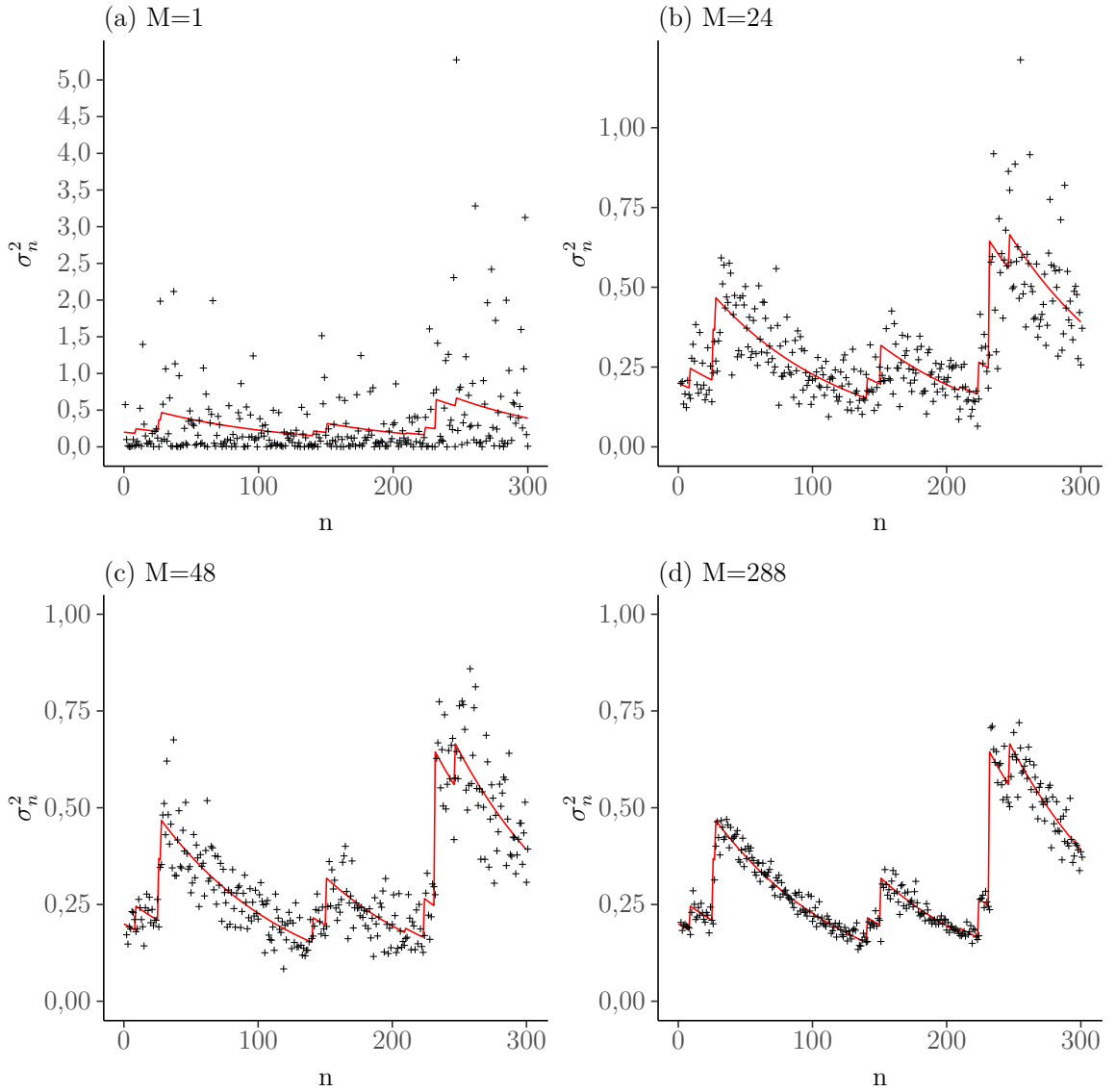
$$\begin{aligned} \text{Var}(\sigma_{1,n}^2) &= \sigma^2(t) = \sum_{j=1}^J \text{Var}(\sigma_{j,1,n}^2) = \\ &= 2\omega^2 \sum_{j=1}^J \frac{w_j}{2\lambda_j^2} (e^{-\lambda_j \Delta M^{-1}} - 1 + \lambda_j \Delta M^{-1}). \end{aligned} \quad (2.52)$$

Rysunek 2.6 ilustruje jak wpływ częstotliwości M na dokładność oszacowania wariancji aktualnej za pomocą wariancji zrealizowanej, przy założeniu braku efektu mikrostruktury rynku. Pojedyncze okresy n można interpretować jako kolejne dni obserwacji. Wówczas wartości $M=1, 24, 48, 288$ oznaczają, że wariancja zrealizowana jest liczona na podstawie obserwacji odpowiednio dziennych (czyli równa jest kwadratowi zwrotu dziennego), godzinnych, półgodzinnych oraz pięciominutowych. Widać wyraźnie, że wraz ze wzrostem M rośnie dokładność oszacowania wariancji aktualnej za pomocą wariancji zrealizowanej. Barndorff-Nielsen i Shephard (2002) wykazali, że gdy $M \rightarrow +\infty$, to

$$\text{Var}(\sqrt{M}u_n) = \text{Var}(\sqrt{M}(\{y\}_n - \sigma_n^2)) \rightarrow 2\Delta^2(\xi^2 + \omega^2).$$

Oznacza to, że wraz ze wzrostem wartości oczekiwanej (parametr ξ) lub wariancji (parametr ω^2) rozkładu stacjonarnego procesu wariancji chwilowej rośnie także

asymptotyczny błąd (przeskalowanej przez $\sqrt{M}$) wariancji zrealizowanej.



Rysunek 2.6: Wariancja aktualna σ_n (linia czerwona) oraz wariancja zrealizowana $\{y\}_n$ (symbol +) dla $M = 1, 12, 48, 288$. Zastosowano proces wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,2$ i wariancji $\omega^2 = 0,02$ oraz parametrze $\lambda = 0,01$.

Źródło: opracowanie własne na podstawie Barndorff-Nielsen i Shephard (2002).

Przedstawione wyniki wartości oczekiwanej i wariancji błędu wariancji zrealizowanej zostały wyznaczone przy założeniu $\mu = \beta = 0$. Jednak wpływ tych param-

trów nie jest istotny, gdy $M \rightarrow +\infty$. Przedstawia to poniższe twierdzenie.

Twierdzenie 2.5. *Barndorff-Nielsen i Shephard (2002) Dla modelu stochastycznej zmienności BNS (ze złożeniem procesów zmienności lub bez), dla którego proces wariancji chwilowej ma lokalnie ograniczone wahania¹⁶ oraz dla każdego Δ i dla $M \rightarrow +\infty$ zachodzi*

$$\frac{\{y\}_n - \sigma_n^2}{\sqrt{2 \sum_{j=1}^M (\sigma_{j,n}^2)^2}} \xrightarrow{d} N(0, 1) \quad (2.53)$$

(zbieżność według rozkładu). Co więcej, przy $M \rightarrow +\infty$ zachodzi zbieżność (prawie na pewno)

$$\Delta^{-1} M \sum_{j=1}^M (\sigma_{j,n}^2)^2 \xrightarrow{p.n.} \sigma_n^{[4]}, \quad (2.54)$$

oraz dla $M \rightarrow +\infty$ zbieżność (według rozkładu)

$$\sqrt{M} \frac{\{y\}_n - \sigma_n^2}{\sigma_n^{[4]}} \xrightarrow{d} N(0, 1), \quad (2.55)$$

gdzie

$$\sigma_n^{[4]} = \int_{(n-1)\Delta}^{n\Delta} \sigma^4(s) ds.$$

Twierdzenie jest prawdziwe dla dowolnego modelu stochastycznej zmienności, dla którego proces wariancji chwilowej jest stacjonarny, o wartości oczekiwanej ξ , wariancji ω^2 i funkcji autokorelacji r , a równanie (2.3) opisuje dynamikę procesu logarytmu cen.

Główny wniosek jaki można wyprowadzić z twierdzenia 2.5 jest taki, że asymptotyczny (przy $M \rightarrow +\infty$) rozkład błędu wariancji zrealizowanej (aż do momentów stopnia drugiego) nie zależy od wartości parametrów μ i β (wzór 2.53). Co oznacza, że dla danych wysokiej częstotliwości wpływ dryfu i premii za ryzyko na wariancję zrealizowaną jest niewielki (co najwyżej od momentów stopnia trzeciego) i można te parametry w wielu przypadkach w analizie pominąć. Twierdzenie implikuje także, że wariancja błędu wariancji zależy od poziomu zmienności wyrażonej przez aktualną kwartyczość $\sigma_n^{[4]}$ (wzór 2.55). Im większa zmienność tym większa wariancja błędu.

¹⁶To znaczy z prawdopodobieństwem 1 trajektorie procesu mają ograniczone wahania na każdym zwartym przedziale $[0, +\infty)$.

Widać to wyraźnie na rysunku (2.6 d): gdy wariancja aktualna maleje do poziomu 0,25, to wartości wariancji zrealizowanej mają niewielką zmienność i są bliskie wartości σ_n^2 . Natomiast, gdy wariancja aktualna rośnie do 0,75, zmienność oszacowań wariancji zrealizowanej rośnie.

Twierdzenie 2.5 opiera się na nieobserwowalnych wartościach kwartyczności aktualnej (wzór 2.55) lub sumie kwadratów wariancji aktualnej wysokiej częstotliwości (wzór 2.53). Można jednak posłużyć się także wartościami obserwowalnymi korzystając z następującej granicy (zbieżność prawie na pewno przy $M \rightarrow +\infty$)

$$\Delta^{-1} \frac{M}{3} \sum_{j=1}^M y_{j,n}^4 \xrightarrow{p.n.} \sigma_n^{[4]},$$

gdzie $y_{j,n}^4$ to czwarte potęgi logarytmicznych zwrotów wysokiej częstotliwości

$$Y_{j,n} = Y \left((n-1)\Delta + \frac{\Delta j}{M} \right) - Y \left((n-1)\Delta + \frac{\Delta(j-1)}{M} \right).$$

Można zatem analogicznie do wzoru (2.53) otrzymać granicę (dla $M \rightarrow +\infty$)

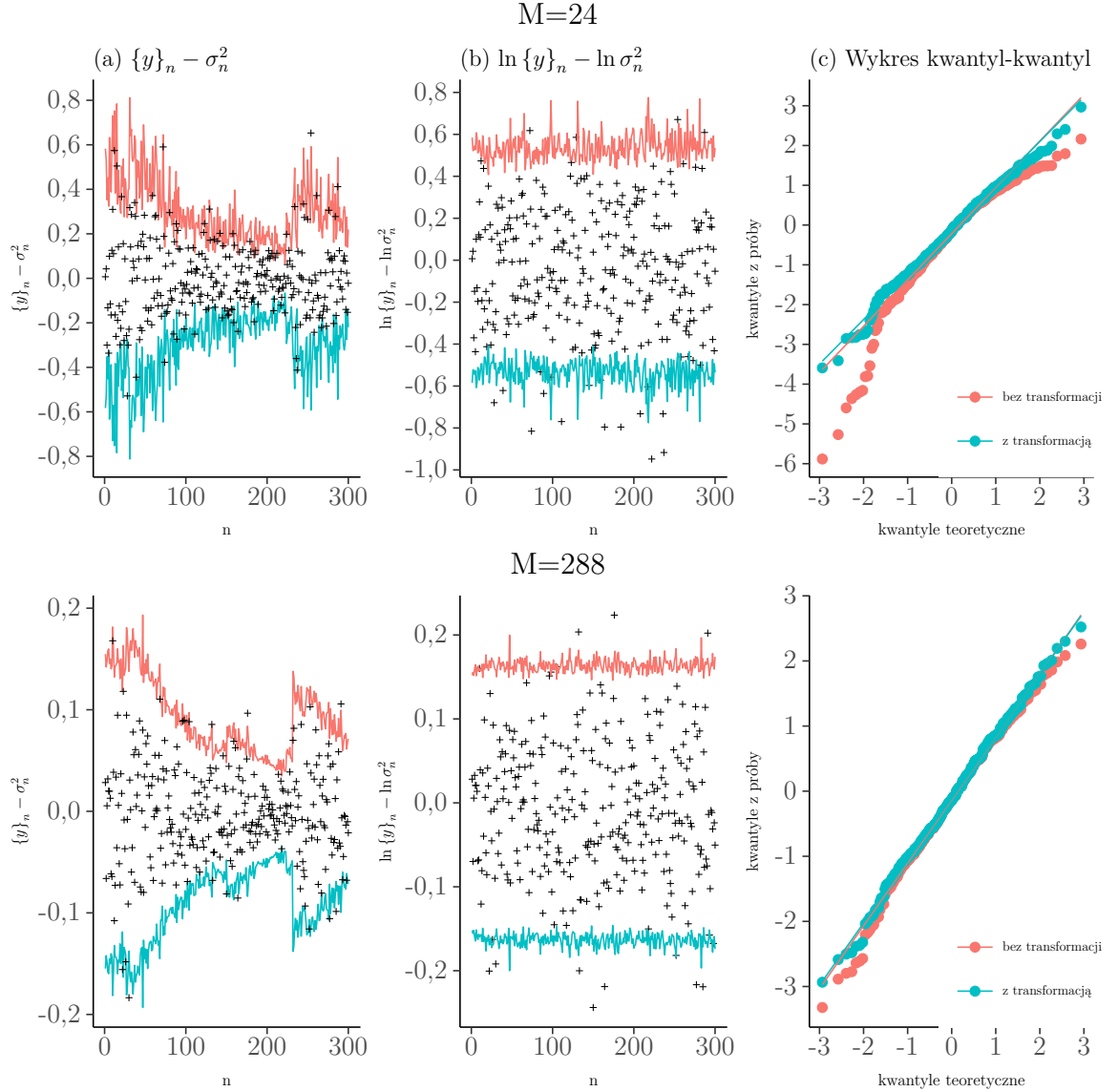
$$\frac{\{y\}_n - \sigma_n^2}{\sqrt{\frac{2}{3} \sum_{j=1}^M y_{j,n}^4}} \xrightarrow{d} N(0, 1). \quad (2.56)$$

Badania Monte Carlo przeprowadzone przez Barndorff-Nielsen i Shephard (2001a) pokazały, że zbieżność we wzorze (2.56) zachodzi dla bardzo dużych M i znacznie lepiej zachodzi zbieżność na skali logarytmicznej postaci

$$\frac{\ln \{y\}_n - \ln \sigma_n^2}{\sqrt{\frac{2}{3} \frac{\sum_{j=1}^M y_{j,n}^4}{\left(\sum_{j=1}^M y_{j,n}^2 \right)^2}}} \xrightarrow{d} N(0, 1), \quad (2.57)$$

która dość dobrze aproksymuje graniczny rozkład dla umiarkowanych wartości parametru M . Prawy segment rysunku 2.7 pokazuje wykresy kwantyl-kwantyl dla wystandaryzowanego błędu wariancji zrealizowanej (wzór (2.56)) oraz po transformacji na skalę logarytmiczną (zgodnie ze wzorem (2.57)). Widać, że dla $M=24$, bez transformacji rozkład błędu wariancji zrealizowanej znacznie odbiega od rozkładu normalnego, natomiast po transformacji punkty leżą niemal na linii wyznaczającej idealną zgodność z rozkładem normalnym. Dla $M=288$ zbieżność błędu do rozkładu normalnego jest lepsza, ale nadal wyniki dla transformacji logarytmicznej wskazują jeszcze lepsze dopasowanie do rozkładu normalnego. Porównując wyniki dla

$M=24$ i $M=288$ widać także znaczny spadek szerokości wstęgi wyznaczonej przez błąd standardowy przemnożony przez 1,96 (co odpowiadałoby liniom kwantylowym dla prawdopodobieństwa 0,95 przy założeniu, że doszło do zbieżności do rozkładu normalnego) wraz ze wzrostem M .



Rysunek 2.7: Błąd wariancji zrealizowanej oryginalny (a) oraz po transformacji logarytmicznej (b) dla symulacji przedstawionej na rysunku 2.6. Linie wyznaczają błąd standardowy przemnożony przez 1,96 (linie kwantylowe dla prawdopodobieństwa 0,95 przy założeniu, że doszło do zbieżności do rozkładu normalnego). Prawy segment (c) przedstawia wykresy kwantyl-kwantyl dla wystandaryzowanego błędu oryginalnego (kolor czerwony) i po transformacji logarytmicznej (kolor turkusowy). Górny panel przedstawia wyniki dla M=24, dolny dla M=288.

Źródło: opracowanie własne na podstawie Barndorff-Nielsen i Shephard (2001a).

Estymacja niegaussowskich modeli zmienności stochastycznej typu Ornsteina-Uhlenbecka

3.1 Uwagi wstępne

Estymacja niegaussowskich modeli stochastycznej zmienności jest trudna i wiele opracowań zostało poświęcone temu zagadnieniu z wykorzystaniem bardzo różnych metod. Podstawowa trudność wspólna wszystkim modelom stochastycznej zmienności jest występowanie w modelu zmiennych ukrytych (por. Kim i in. (1998), Pajor (2003)). Załóżmy, że chcemy wyznaczyć funkcję wiarygodności próby losowej dla podstawowego modelu BNS bez złożenia i efektu dźwigni z rozkładem stacjonarnym zależnym od dwóch parametrów δ i γ . Wówczas pod warunkiem obserwacji logarytmicznych zwrotów $y_{1:N} = (y_1, y_2, \dots, y_N)$ i dla wektora parametrów $\theta = [\mu, \beta, \lambda, \delta, \gamma]^T$ funkcja wiarygodności przyjmuje postać

$$\begin{aligned} p(\theta; y_{1:N}) &= \int p(y_{1:N} | \sigma_{0:N}^2; \mu, \beta) p(\sigma_{1:N}^2; \lambda, \delta, \gamma) d\sigma_{1:N}^2 = \\ &= \int \prod_{n=1}^N p(y_n | \sigma_n^2; \mu, \beta) p(\sigma_{1:N}^2; \lambda, \delta, \gamma) d\sigma_{1:N}^2 \end{aligned} \quad (3.1)$$

Nie można jej wyznaczyć analitycznie, nie jest też znana postać funkcji gęstości $p(\sigma_1^2, \dots, \sigma_N^2; \lambda, \delta, \nu)$ dla żadnego rozkładu stacjonarnego procesu wariancji chwilo-

wej. Uniemożliwia to użycie metody największej wiarygodności w celu wyznaczenia oszacowań parametrów.

W pierwszej swojej pracy Barndorff-Nielsen i Shephard (2001b) w przykładzie empirycznym do wyznaczenia oszacowania parametrów użyli nieliniowej metody najmniejszych kwadratów poprzez porównanie empirycznej funkcji autokowariancji kwadratów z jej teoretycznym odpowiednikiem (por. wzór (2.36)). Metoda ta pozwala jedynie na wyznaczenie parametrów $\lambda_1, \dots, \lambda_P$, wag dla złożenia $w_1, \dots, w_P$ oraz wariancji rozkładu stacjonarnego procesu wariancji chwilowej ω^2 . Nie można za pomocą tej metody wyznaczyć wartości oczekiwanej rozkładu stacjonarnego procesu wariancji chwilowej, ponieważ parametr ten nie wpływa na wartość postać teoretycznej funkcji autokowariancji. Barndorff-Nielsen i Shephard (2001b) zasugerowali użycie innych metod: Monte Carlo opartych na łańcuchach Markowa (*Markov Chain Monte Carlo*, w skrócie MCMC), estymacji pośredniej (Gourieroux i in., 1993), funkcji estymujących (Sørensen, 2000), filtrów Kalmana i cząsteczkowych.

Następnie powstało szereg prac mających na celu implementacji podejścia bayesowskiego, głównie za pomocą metod MCMC: Roberts i in. (2004) oraz Frühwirth-Schnatter i Sögner (2009) zaproponowali metody estymacji MCMCM w przypadku rozkładu stacjonarnego gamma, Gander i Stephens (2007a); Gander i Stephens (2007b) oraz Griffin i Steel (2006) w przypadku temperowanych rozkładów stabilnych i uogólnionych rozkładów odwrotnego Gaussa, Griffin i Steel (2010) dla ciągłego złożenia procesów wariancji chwilowej. Andrieu i in. (2010) zaproponował metodę cząsteczkowych markowskich łańcuchów Monte Carlo (*Particle Markov Chain Monte Carlo*, w skrócie PMCMC), wykorzystuje filtry cząsteczkowe w obrębie markowskich łańcuchów Monte Carlo do wyznaczenia wartości funkcji akceptacji propozycji nowego stanu łańcucha Markowa. Jako jeden z przykładów zastosowania swojej metody podał podstawowy model BNS. Metodę PMCMC zastosowali również James i in. (2018) w modelu z czasem operacyjnym będącym niegaussowskim procesem Ornsteina-Uhlenbecka. Gleim i Pigorsch (2013) użyli metody przybliżonych bayesowskich obliczeń (*Approximate Bayesian Computation*) z użyciem statystyk dostatecznych do estymacji parametrów w podstawowym modelu BNS oraz ze złożeniem dwóch procesów zmienności.

Dużo mniej prac powstało w oparciu o klasyczne (częstościowe) wnioskowanie statystyczne: Lindberg (2008) zaproponował, aby jako aproksymację zmienności użyć wielkości obrotu, co pozwalało na zastosowanie metody największej wiarygodności. Przy tym samym założeniu, że obroty mogą przybliżać zmienność Hubalek i Posedel (2011) zastosowali funkcje estymujące. Taufer i in. (2011) zaproponowali użycie metody empirycznych funkcji charakterystycznych.

W tej pracy zostaną zastosowane metody estymacji modelu BNS oparte o filtry Kalmana i cząsteczkowe. W podrozdziale 3.2 przedstawione zostaną modele przestrzeni stanów oraz zagadnienie filtracji, które stanowią teoretyczne podstawy wspólne dla filtrów Kalmana i filtrów cząsteczkowych. W podrozdziale 3.3 omówiony zostanie filtr Kalmana¹, natomiast w podrozdziale 3.4 specyfikacje filtra Kalmana dla modelu BNS w oparciu o przestrzenie stanów z prac Barndorff-Nielsen i Shephard (2001b) i Barndorff-Nielsen i Shephard (2002) oraz przedstawioną w pracy Szczepocki (2018) przestrzeń stanów dla złożenia procesów zmienności. W podrozdziale 3.5 przedstawiona zostanie zarys teorii filtrów cząsteczkowych² oraz następnie w podrozdziale 3.6 zastosowanie filtra cząsteczkowego do estymacji wariancji aktualnej w modelu BNS. Następnie przedstawiona zostanie metoda iterowanej filtracji do estymacji parametrów modelu BNS. Zastosowanie metody iterowanej filtracji w kontekście modelu BNS jest autorską propozycją. Część wyników, dla modelu bez złożenia procesów zmienności i efektu dźwigni, przy założeniu procesu wariancji chwilowej o rozkładzie gamma została przedstawiona w pracach Szczepocki (2019a) (dla logarytmicznej stopy zwrotu jako zmiennej pomiaru) i Szczepocki (2019b) (dla wariancji zrealizowanej jako zmiennej pomiaru).

3.2 Modele przestrzeni stanów

Modele przestrzeni stanu są modelami statystycznymi stosowanym do opisu dynamiki zjawisk zarówno w naukach technicznych, przyrodniczych jak i ekonomicznych. Modele przestrzeni stanu są również nazywane ukrytymi modelami Markowa³ (*Hid-*

¹Przedstawiona teoria opiera się na pracy Hamilton (1994).

²Na podstawie prac Doucet i in. (2001) oraz Brzozowska-Rup i Dawidowicz (2009).

³Por. Brzozowska-Rup i Dawidowicz (2009) oraz monografia Cappé i in. (2009)

den Markov Model, HMM), częściowo obserwowanymi procesami Markowa⁴ (*Partially Observed Markov Process*, POMP). Stosuje się je wtedy, gdy pewną zmienną nie możemy obserwować bezpośrednio, tylko poprzez inną zmienną będącą funkcją szukanej zmiennej. W modelach SV tą nieobserwowaną zmienną jest zmienność instrumentu finansowego.

Definicja 3.1.

Modelem przestrzeni stanów nazywamy parę procesów stochastycznych $(X_n, Y_n)_{n \in \mathbb{N}}$, gdzie $X_n \in \mathbb{R}^p$ jest procesem Markowa przyjmującym wartości w przestrzeni $\mathbb{R}^p$, z rozkładem początkowym o znanej funkcji gęstości prawdopodobieństwa

$$X_0 \sim p(x_0; \theta) \quad (3.2)$$

oraz znanej funkcji gęstości prawdopodobieństwa

$$X_n | X_{n-1} = x_{n-1} \sim p(x_n | x_{n-1}; \theta) \quad \text{dla } n \in \mathbb{N}_+. \quad (3.3)$$

Proces $(X_n)_{n \in \mathbb{N}}$ nie jest bezpośrednio obserwowalny, tylko poprzez proces $(Y_n)_{n \in \mathbb{N}}$ przyjmujący wartości w przestrzeni $\mathbb{R}^s$, przy czym proces ten jest warunkowo niezależny względem σ -ciała $\Sigma_n^- = \sigma(X_k, Y_k; k \in \{0, 1, \dots, n-1\})$. Ponadto znane są funkcję gęstości prawdopodobieństwa

$$Y_n | X_n = x_n \sim p(y_n | x_n; \theta) \quad \text{dla } n \in \mathbb{N}_+. \quad (3.4)$$

Procesy $(X_n)_{n \in \mathbb{N}}$, $(Y_n)_{n \in \mathbb{N}}$ nazywane są odpowiednio procesem stanu i pomiaru. Natomiast zmienne losowe X_n i Y_n odpowiednio zmiennymi (wektorami) stanu i pomiaru. O wektorze parametrów θ zakładamy, że $\theta \in \Theta \subset \mathbb{R}^r$.

Modele przestrzeni stanu można zapisać również w postaci układu równań (Cappé i in., 2009, str. 3)

$$Y_n = h(X_n, u_n; \theta) \quad (3.5)$$

$$X_n = f(X_{n-1}, v_n; \theta), \quad (3.6)$$

gdzie $(u_n)_{n \in \mathbb{N}}$, $(v_n)_{n \in \mathbb{N}}$ są ciągami niezależnych zmiennych losowych zwanych zaburzeniami lub szumem, natomiast h, f są znanymi funkcjami mierzalnymi. Równanie

⁴Por. Ionides (2005) oraz Ionides i in. (2011)

(3.5) nazywane jest równaniem pomiaru lub opisowym, (3.6) równaniem stanu, dynamiki lub ewolucji. W zależności od postaci funkcji h i f można modele przestrzeni stanów podzielić na liniowe i nieliniowe. W pierwszym przypadku, funkcje f i h są liniowe ze względu na obie zmienne. W przeciwnym przypadku model jest nieliniowy. Modele przestrzeni stanów można również podzielić na gaussowskie, gdy zaburzenia mają (wielowymiarowy) rozkład normalny i niegaussowskie (pozostałe przypadki).

Problem estymacji modelu przestrzeni stanu można podzielić na dwa zagadnienia:

- uzyskanie oszacowania procesu ukrytego $(X_n)_{n \in \mathbb{N}}$ na podstawie procesu pomiaru $(Y_n)_{n \in \mathbb{N}}$,
- estymacja wektora parametrów θ .

W pierwszym zagadnieniu tradycyjnie przyjmuje się, że wektor parametrów θ jest znany. Proces stanu jest ukryty, jego wartości nie są znane. Możemy tylko określić rozkład prawdopodobieństwa dla możliwych wartości tego procesu na podstawie zaobserwowanych wartości procesu pomiaru. Niepewność co do wartości procesu stanu w okresie od 0 do n przy obserwacji procesu pomiaru $y_1, y_2, \dots, y_n$ opisuje rozkład prawdopodobieństwa o gęstości $p(x_{0:n}|y_{1:n}; \theta)$. Zazwyczaj punktem zainteresowania jest rozkład brzegowy tego rozkładu o gęstości $p(x_n|y_{1:n}; \theta)$ zwany także rozkładem filtracji. Tradycyjnie wyznaczenie tego rozkładu lub jego charakterystyk liczbowych nazywane jest zagadnieniem filtracji. Ponadto prognozy na m -kroków do przodu określają gęstości $p(x_{n+m}|y_{1:n}; \theta)$, dla $m \in \mathbb{N}_+$.

Mając do dyspozycji obserwacje procesu pomiaru $y_1, y_2, \dots, y_n, \dots, y_N$ można także dla każdego $n \leq N$ badać rozkład o gęstości $p(x_n|y_{1:N}; \theta)$. Wyznaczenie tego rozkładu lub jego charakterystyk liczbowych nazywane jest wygładzaniem (*smoothing*). Jest to bliższe tradycyjnemu podejściu, gdy na podstawie próby (całej trajektorii) dokonuje się oszacowania procesu stanu.

Kolejną własnością zaczerpniętą z teorii przetwarzania sygnałów jest rekurencyjny charakter wyznaczania kolejnych oszacowań $p(x_{0:n}|y_{1:n}; \theta)$ i $p(x_n|y_{1:n}; \theta)$. Pojedynczy krok rekurencji zostanie przedstawiony na przykładzie gęstości filtracji $p(x_n|y_{1:n}; \theta)$. Załóżmy, że wyznaczono rozkłady filtracji do okresu $n - 1$. Najpierw

na podstawie obserwacji $y_1, y_2, \dots, y_{n-1}$ dokonuje się predykcji procesu stanu na okres n za pomocą gęstości rozkładu predykcyjnego postaci

$$p(x_n|y_{1:n-1}; \theta) = \int p(x_n|x_{n-1}; \theta) p(x_n|y_{1:n}; \theta) dx_{n-1} \quad (3.7)$$

korzystając z faktu, że znana jest gęstość przejścia $p(x_n|x_{n-1}; \theta)$. Wraz z pojawieniem się nowej obserwacji y_n możemy skorygować naszą wiedzę na temat rozkładu prawdopodobieństwa procesu stanu X_n korzystając ze wzoru Bayesa

$$\begin{aligned} p(x_n|y_{1:n}; \theta) &= \frac{p(y_n, x_n|y_{1:n-1}; \theta)}{p(y_n|y_{1:n-1}; \theta)} = \\ &= \frac{p(y_n|x_n, y_{1:n-1}; \theta) p(x_n|y_{1:n-1}; \theta)}{\int p(y_n|x_n; \theta) p(x_n|y_{1:n-1}; \theta) dx_n} = \\ &= \frac{p(y_n|x_n; \theta) p(x_n|y_{1:n-1}; \theta)}{\int p(y_n|x_n; \theta) p(x_n|y_{1:n-1}; \theta) dx_n}, \end{aligned} \quad (3.8)$$

przy czym ostatnia równość $p(x_n|y_{1:n-1}; \theta) = p(y_n|x_n; \theta)$ zachodzi z warunkowej niezależności procesu pomiaru. Wzór (3.8) wyznacza gęstość rozkładu filtracji dla okresu n , co domyka rekurencję. Podobną zależność można wykazać dla gęstości $p(x_{0:n}|y_{1:n}; \theta)$. Należy jednak zwrócić uwagę, że zazwyczaj całkę w mianowniku wzoru (3.8)

$$p(y_n|y_{1:n-1}; \theta) = \int p(y_n|x_n; \theta) p(x_n|y_{1:n-1}; \theta) dx_n \quad (3.9)$$

nie można wyznaczyć w sposób analityczny. Jednym z nielicznych przypadków jest liniowy model z zaburzeniami gaussowskimi, dla którego gęstości predykcji i filtracji można wyznaczyć za pomocą filtru Kalmana. Podejście to zostanie szerzej omówione w podrozdziale 3.3. W pozostałych przypadkach szukane gęstości można przybliżać poprzez aproksymację całek metodą Monte Carlo. Metoda ta jest znana pod nazwą filtru cząsteczkowego lub sekwencyjnej metody Monte Carlo. Podejście to zostanie omówione w podrozdziale 3.5.

Do wygładzania najczęściej stosuje się algorytmy typu „do przodu - do tyłu” (*forward - backward*). Pierwsza część, „do przodu”, to wyznaczenie gęstości predykcji $p(x_n|y_{1:n-1}; \theta)$ i filtracji $p(x_n|y_{1:n}; \theta)$ dla $n = 1, \dots, N$ dla $n = 1, \dots, N$. Następnie następuje druga część, „do tyłu”, która polega na wyznaczeniu iteracyjnie dla

$n = N - 1, \dots, 1$ gęstości wygładzania za pomocą

$$\begin{aligned}
 p(x_n|y_{1:N}; \theta) &= \int p(x_n, x_{n+1}|y_{1:N}; \theta) dx_{n+1} = \\
 &= \int p(x_{n+1}|y_{1:N}; \theta) p(x_n|x_{n+1}, y_{1:N}; \theta) dx_{n+1} = \\
 &= \int p(x_{n+1}|y_{1:N}; \theta) p(x_n|x_{n+1}, y_{1:n}; \theta) dx_{n+1} = \\
 &= p(x_n|y_{1:n}; \theta) \int \frac{p(x_{n+1}|y_{1:N}; \theta) p(x_{n+1}|x_n; \theta)}{p(x_{n+1}|y_{1:n}; \theta)} dx_{n+1} \quad (3.10)
 \end{aligned}$$

Podobnie jak w zagadnieniu filtracji całkę we wzorze (3.10) nie można wyznaczyć analitycznie dla większości przypadków. Istotnym wyjątkiem od tej reguły jest ponownie przypadek liniowego gaussowskiego modelu przestrzeni stanów, dla którego istnieje analityczna postać wygładzania Kalmana, co zostanie pokazane w podrozdziale 3.3. Wygładzanie przy pomocy filtru cząsteczkowego, w którym całkę ze wzoru (3.10) aproksymuje się metodą Monte Carlo zostanie omówione w podrozdziale 3.5.

3.3 Filtr Kalmana

Filtr Kalmana został zaproponowany w 1960 roku przez Rudolfa Emila Kalmana (1930-2016) w pracy Kalman (1961). Następnie został rozszerzony na przypadek z czasem ciągłym w pracy Kalman i Bucy (1961). Algorytmy opierające się na teorii filtru Kalmana znalazły ogromną liczbę zastosowań praktycznych, m.in. w systemach nawigacji GPS, układach elektronicznych i energoelektronicznych, przetwarzaniu sygnałów, inżynierii dźwięku i obrazu, rzeczywistości rozszerzonej, systemów autopilota, układach sterowania napędów elektrycznych (Dróżdż, 2017). Do rozpowszechnienia filtru Kalmana przyczyniło się zastosowanie do estymacji trajektorii lotów podczas misji Apollo na Księżyc (Grewal i Andrews, 2010).

Filtr Kalmana znalazł także liczne zastosowania w ekonomii: w modelach ze zmiennymi w czasie parametrami, do estymacji modeli ARMA oraz stochastycznej zmienności. Stosowanie filtru Kalmana w modelach stochastycznej zmienności upowszechniła praca Harvey i in. (1994). W polskiej literaturze przykłady zastosowania filtru Kalmana można znaleźć w pracach Skrzypek (1986), Dziechciarz (1993), Grzesiak (1995), Żółtowska (1997), a do estymacji parametrów modeli stochastycznej zmienności w Pajor (2003).

Filtr Kalmana jest optymalnym rozwiązaniem zagadnienia filtracji w przypadku, modelu liniowego i gaussowskiego postaci⁵

$$\begin{cases} Y_n = d + HX_n + u_n \\ X_n = c + FX_{n-1} + v_n \end{cases} \quad (3.11)$$

gdzie:

$u_n \in \mathbb{R}^p$ wektor szumu pomiarowego, o którym zakłada się, że ma (wielowymiarowy) rozkład normalny o zerowym wektorze wartości oczekiwanej i znanej macierzy wariancji-kowariancji R ,

$v_n \in \mathbb{R}^s$ wektor szumu procesu, o którym zakłada się, że ma (wielowymiarowy) rozkład normalny o zerowym wektorze wartości oczekiwanej i znanej macierzy wariancji-kowariancji Q ,

$H \in \mathcal{M}(\mathbb{R}^p, \mathbb{R}^s)$ macierz filtru,

$F \in \mathcal{M}(\mathbb{R}^p, \mathbb{R}^p)$ macierz przejścia,

$c \in \mathbb{R}^p$ wektor wyrazów wolnych równania przejścia,

$d \in \mathbb{R}^s$ wektor wyrazów wolnych równania pomiaru.

Algorytm filtru Kalmana ma charakter rekurencyjny. Załóżmy, że $\hat{X}_{n-1|n-1}$ jest optymalnym estymatorem $X_{n-1|n-1}$ względem obserwacji procesu pomiaru do okresu $n-1$ włącznie oraz $P_{n-1|n-1}$ jest odpowiadającą macierzą kowariancji błędu estymacji. Filtr Kalmana wyznacza rekurencyjnie oszacowania $\hat{X}_{n|n-1}$ oraz $\hat{X}_{n|n}$ za pomocą algorytmu składającego się z dwóch faz:

- predykcji (aktualizacji czasowej),
- korekty (aktualizacji pomiarowej).

W pierwszej fazie następuje wyznaczenie prognozy na okres n wektora stanu i macierzy kowariancji błędu oszacowania w oparciu o obserwacje procesu pomiaru do okresu $n-1$ oraz postać równania stanu. Aktualizację czasową wektora stanu wykonuje się zgodnie ze wzorem

$$\hat{X}_{n|n-1} = c + H\hat{X}_{n|n-1}, \quad (3.12)$$

⁵W literaturze rozważa się również bardziej ogólną postać uwzględniającą zmienne egzogeniczne (por. Miszczak (2006, str. 107)).

gdzie: $\hat{X}_{n|n-1}$ to oszacowanie *a priori* (przed pomiarem) wektora stanu w okresie n .

Aktualizacja czasowa wektora kowariancji przyjmuje postać

$$P_{n|n-1} = H P_{n-1|n-1} H^T + Q \quad (3.13)$$

gdzie: $P_{n|n-1}$ to macierz kowariancji *a priori* wektora reszt oszacowania wektora stanu w okresie n , to jest macierz

$$\mathbb{E} \left[(X_n - \hat{X}_{n|n-1}) (X_n - \hat{X}_{n|n-1})^T \right].$$

W drugiej fazie następują korekta uzyskanych w pierwszej fazie oszacowań przy użyciu zaobserwowanej wartości y_n wektora pomiaru Y_n . Siłę tej korekty wyznacza wzmocnienie Kalmana (*Kalman gain*) K_n , które przyjmuje postać

$$K_n = P_{n|n-1} F^T S_k^{-1} = P_{n|n-1} F^T [F P_{n|n-1} F^T + R]^{-1}. \quad (3.14)$$

Korzystając ze wzmocnienia Kalmana uzyskuje się oszacowanie *a posteriori* wektora stanu w okresie n

$$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n (y_n - d - F \hat{X}_{n|n-1}) \quad (3.15)$$

oraz macierz kowariancji *a posteriori* wektora reszt oszacowania wektora stanu w okresie n

$$P_{n|n} = (I_{p^2} - K_n F) P_{n|n-1}, \quad (3.16)$$

gdzie I_{p^2} jest macierzą jednostkową stopnia p^2 . Macierz $P_{n|n}$ to macierz kowariancji *a posteriori* wektora reszt oszacowania wektora stanu w momencie n , czyli macierz postaci

$$\mathbb{E} \left[(X_n - \hat{X}_{n|n}) (X_n - \hat{X}_{n|n})^T \right].$$

Algorytm 1 przedstawia procedurę filtru Kalmana w postaci pseudokodu.

Wzmocnienie Kalmana wyznacza jak mocno należy skorygować wartości teoretyczne uzyskane w aktualizacji czasowej. Dobrze to widać w przypadku, gdy wektory stanu i obserwacji są jednowymiarowe ($p = s = 1$). Wówczas wzmocnienie Kalmana można zapisać w postaci

$$K_n = P_{n|n-1} F^T [F P_{n|n-1} F^T + R]^{-1} = F^{-1} \frac{F P_{n|n-1} F^T}{F P_{n|n-1} F^T + R}$$

Algorytm 1. Filtr Kalmana

Dane są: macierze H, F, Q, R .

1. Jeżeli proces stanu jest stacjonarny przyjmujemy:

$$\hat{X}_{1|0} = 0 \text{ oraz } P_{1|0} = \mathbb{E}(X_n X_n^T),$$

w przeciwnym przypadku $\hat{X}_{1|0}$ i $P_{1|0}$ muszą być zadane przez użytkownika.

2. Dla $n = 1, \dots, N$ wykonujemy:

(i) $\hat{X}_{n|n-1} = c + H\hat{X}_{n-1|n-1},$

(ii) $P_{n|n-1} = HP_{n-1|n-1}H^T + Q,$

(iii) $K_n = P_{n|n-1}F^T [FP_{n|n-1}F^T + R]^{-1},$

(iv) $\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n(y_n - d - F\hat{X}_{n|n-1}),$

(v) $P_{n|n} = (I - K_nF)P_{n|n-1}.$

Rezultat algorytmu:

$$\hat{X}_{n|n}, P_{n|n} \text{ oraz } \hat{X}_{n|n-1}, P_{n|n-1} \text{ dla } n = 1, \dots, N.$$

Licznik ułamka wyznacza niepewność związaną z modelem (macierz kowariancji $P_{n|n-1}$) przeniesioną do równania pomiaru (poprzez pomnożenie obustronnie przez macierz F). Natomiast mianownik to zmienność modelu powiększona o kowariancję szumu pomiaru R . Czynniki F^{-1} stojący przed ułamkiem wynika, z tego że stosunek ma zostać odniesiony do wektora stanu, a nie wektora obserwacji. Tym większa jest wartość szumu pomiaru R , tym mniejsza jest wartość ułamka i wzmocnienia Kalmana. Gdy R dąży do $+\infty$ wzmocnienie Kalmana dąży do 0. W granicznym przypadku, gdy $K_n = 0$, wektor stanu *a posteriori* $\hat{X}_{n|n}$ jest równy wektorowi *a priori* $\hat{X}_{n|n-1}$, czyli cała wiedza o stanie X_n pochodzi z predykcji, a pomiary są ignorowane. Jest to zgodne z intuicją, że przy większej wartości szumu pomiaru powinniśmy mniej wierzyć pomiarom, a bardziej wartościom teoretycznym wyznaczonym z aktualizacji czasowej - i na odwrót: gdy maleje szum pomiaru, czyli kowariancja R dąży do zera wzmocnienie Kalmana dąży do wartości F^{-1} . Oznacza to, że w granicznym przypadku, gdy $R = 0$, oszacowanie wektora *a posteriori* $\hat{X}_{n|n}$ będzie równe y_n , czyli cała wiedza o stanie X_n będzie pochodziła z obserwacji y_n .

Do rozpoczęcia algorytmu konieczne jest podanie wartości stanu $\hat{X}_{1|0}$ oraz macierzy kowariancji $P_{1|0}$. Optymalnym rozwiązaniem w przypadku, gdy proces stanu

jest stacjonarny (gdy wszystkie wartości własne macierzy F leżą wewnątrz koła jednostkowego) jest przyjęcie jako wartości początkowych bezwarunkowych wartości oczekiwanych i kowariancji postaci

$$\hat{X}_{1|0} = 0, \quad (3.17)$$

$$P_{1|0} = \mathbb{E} \left(X_n X_n^T \right) \quad (3.18)$$

Wartość bezwarunkowej kowariancji można wyznaczyć ze wzoru:

$$\text{vec} \left(P_{1|0} \right) = [I_{p^2} - F \otimes F]^{-1} \text{vec}(Q),$$

gdzie $\otimes$ oznacza iloczyn Kroneckera macierzy, vec operator wektoryzacji macierzy w wektor kolumnowy. W przeciwnym przypadku, gdy proces stanu nie jest stacjonarny proponowane są różne rozwiązania. Można jako $\hat{X}_{1|0}$ przyjąć przewidywaną przez badacza wartość na podstawie wiedzy *a priori*, natomiast jako $P_{1|0}$ niepewność badacza co do przyjętej wartości $\hat{X}_{1|0}$. Inna możliwość to przyjęcie, że wartość początkowa jest ustalona (ma rozkład zdegenerowany: $P_{1|0} = 0$). Wówczas wartość początkową wyznacza się na podstawie próby metodą największej wiarygodności (por. Rosenberg (1973) oraz Harvey (1990)). Kolejna możliwość to przyjęcie, że początkowy wektor stanu jest losowy o rozproszonej gęstości i nie są dostępne żadne informacje co do wartości początkowej: $\hat{X}_{1|0} = 0$, $P_{1|0} = I\kappa$, gdzie $\kappa \rightarrow +\infty$ (w praktycznych zastosowaniach przyjmuje się dowolnie duże κ , por. De Jong (1991)).

Optymalność filtru Kalmana polega tym, że oszacowanie $\hat{X}_{n|n}$ minimalizuje błąd średniokwadratowy

$$\mathbb{E} \left[\left(X_n - \hat{X} \right) \left(X_n - \hat{X} \right)^T \right]. \quad (3.19)$$

pośród wszystkich estymatorów $\hat{X}$ powstałych na podstawie procesu obserwacji do okresu n . Każdy inny estymator ma macierz błędu średniokwadratowego, która różni się od macierzy błędu średniokwadratowego $P_{n|n}$ filtru Kalmana o macierz dodatnio zorientowaną.

W przypadku, gdy zaburzenia nie mają rozkładu normalnego, filtr Kalmana przestaje być optymalny, ale minimalizuje błąd średniokwadratowy w klasie prognoz liniowych wyznaczonych na podstawie procesu obserwacji $Y_1, \dots, Y_n$.

Zaletą filtru Kalmana jest rekurencyjny charakter wyznaczania kolejnych wartości $\hat{X}_{n|n-1}$, gdy pojawiają się kolejne obserwacje procesu pomiaru. Często jednak

dysponujemy od razu ciągiem obserwacji aż do momentu N . Możemy wówczas wykorzystać je oszacowania stanu w dowolnym okresie n leżącym wewnątrz próby. Taki estymator wektora stanu nazywamy wygładzaniem i oznaczamy symbolem $\hat{X}_{n|N}$.

W przypadku wygładzania Kalmana w celu otrzymania oszacowań $\hat{X}_{n|N}$ dla $n = 1, \dots, N$ należy najpierw korzystając z filtru Kalmana na podstawie obserwacji $y_1, y_2, \dots, y_N$ wyznaczyć ciągi oszacowań wektora stanu $(\hat{X}_{n|n-1})_{n=1}^N$, $(\hat{X}_{n|n})_{n=1}^N$ oraz kowariancji $(P_{n|n-1})_{n=1}^N, (P_{n|n})_{n=1}^N$. Następnie wyznacza się wartości wygładzone rekurencyjnie w kierunku odwrotnym zgodnie ze wzorem

$$\hat{X}_{n|N} = \hat{X}_{n|n} + J_n (\hat{X}_{n+1|N} - \hat{X}_{n+1|n}), \quad (3.20)$$

dla $n = N - 1, N - 2, \dots, 1$, gdzie $J_n = P_{n|n} F^T P_{n+1|n}$. Podobnie można wyznaczyć macierze kowariancji

$$P_{n|N} = P_{n|n} + J_n (P_{n+1|N} - P_{n+1|n}) J_n^T \quad (3.21)$$

w kierunku odwrotnym $n = N - 1, N - 2, \dots, 1$. Procedurę postępowania wygładzania Kalmana w postaci pseudokodu przedstawia algorytm 2.

Przedstawione rozważania zakładały, że H, F, c, d, Q, R są znane. Zazwyczaj jednak zależą od nieznanymi parametrów, które należy oszacować na podstawie danych. Przyjmijmy, że wektor $\theta \in \Theta \subset \mathbb{R}^d$ zbiera wszystkie parametry, od których zależą macierze H, F, c, d, Q, R . Korzystając z założenia o normalności zakłóceń zachodzi

$$Y_n | X_n \sim N(\mu(\theta), \Sigma(\theta)), \quad (3.22)$$

gdzie:

$$\begin{aligned} \mu(\theta) &= [H(\theta)]^T \hat{X}_{n|n-1}, \\ \Sigma(\theta) &= [H(\theta)]^T P_{n|n-1} [H(\theta)]. \end{aligned}$$

Można zatem wyznaczyć logarytm funkcji wiarygodności

$$\begin{aligned} \ln \mathcal{L}(\theta; y_1, \dots, y_N) &= \sum_{n=1}^N \ln f(y_n | x_n; \theta) = -\frac{sN}{2} \ln(2\pi) - \frac{1}{2} \sum_{n=1}^N \ln |\Sigma(\theta)| \\ &\quad - \frac{1}{2} \sum_{n=1}^N [y_n - \mu(\theta)]^T [\Sigma(\theta)] [y_n - \mu(\theta)]. \end{aligned} \quad (3.23)$$

Algorytm 2. Wygładzanie Kalmana

Dane są: macierze H, F, Q, R .

1. Jeżeli proces stanu jest stacjonarny przyjmujemy:

$$\hat{X}_{1|0} = 0 \text{ oraz } P_{1|0} = \mathbb{E}(X_n X_n^T),$$

w przeciwnym przypadku $\hat{X}_{1|0}$ i $P_{1|0}$ muszą być zadane przez użytkownika.

2. Dla $n = 1, \dots, N$ wykonujemy:

$$(i) \hat{X}_{n|n-1} = c + H \hat{X}_{n|n-1},$$

$$(ii) P_{n|n-1} = H P_{n-1|n-1} H^T + Q,$$

$$(iii) K_n = P_{n|n-1} F^T [F P_{n|n-1} F^T + R]^{-1},$$

$$(iv) \hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n (y_n - d - F \hat{X}_{n|n-1}),$$

$$(v) P_{n|n} = (I - K_n F) P_{n|n-1}.$$

3. Dla $n = N - 1, \dots, 1$ wykonujemy:

$$(i) J_n = P_{n|n} F^T P_{n+1|n},$$

$$(ii) \hat{X}_{n|N} = \bar{X}_{n|n} + J_n (\hat{X}_{n+1|N} - \hat{X}_{n+1|n}),$$

$$(iii) P_{n|N} = P_{n|n} + J_n (P_{n+1|N} - P_{n+1|n}) J_n^T$$

Rezultat algorytmu:

$$\hat{X}_{n|N}, P_{n|N} \text{ dla } n = 1, \dots, N \text{ dla } n = 1, \dots, N.$$

Maksymalizując funkcję logarytmu wiarygodności (3.23) można wyznaczyć estymator θ^{ML} metody największej wiarygodności prawdziwej wartości parametru θ_0 . Przy spełnieniu odpowiednich warunków (por. Caines (1988, roz. 6)) estymator $\hat{\theta}^{ML}$ jest zgodny i asymptotycznie normalny. Warunki te to: identyfikowalność parametru θ , θ_0 nie może leżeć na brzegu zbioru Θ , wartości własne macierzy F muszą leżeć wewnątrz koła jednostkowego. Asymptotyczna normalność estymatora największej wiarygodności dla filtru Kalmana w konkretnych przypadkach modeli stanów wykazali Pagan (1980) oraz Ghosh (1989). Własność ta przyjmuje postać

$$\sqrt{N} \mathcal{I}_{2D,N}^{1/2} (\hat{\theta}^{ML} - \theta_0) \xrightarrow[N \rightarrow +\infty]{d} N(0, I), \quad (3.24)$$

gdzie $\mathcal{I}_{2D,N}$ jest macierzą informacji Fishera daną wzorem

$$\mathcal{I}_{2D,N} = -\frac{1}{N} \mathbb{E} \left[\sum_{n=1}^N \frac{\partial^2 \ln f(y_n | x_n; \theta)}{\partial \theta \partial \theta^T} \Big|_{\theta=\theta_0} \right] \quad (3.25)$$

Zazwyczaj jednak wykorzystuje się własność, że granica przy $N \rightarrow +\infty$ macierzą informacji Fishera $\mathcal{I}_{2D,N}$ jest taka sama jak granica według prawdopodobieństwa wyrażenia

$$\hat{\mathcal{I}} = -\frac{1}{N} \sum_{n=1}^N \frac{\partial^2 \ln f(y_n|x_n; \theta)}{\partial \theta \partial \theta^T} \Big|_{\theta=\hat{\theta}^{ML}}, \quad (3.26)$$

które można wyznaczyć analitycznie lub numerycznie.

Istotnym zagadaniem w kontekście niegaussowskich modeli stochastycznej zmienności jest wpływ złamania założenia o gaussowskiej naturze zakłóceń. Kwestię tę rozważał White (1982). Wykazał, przy pewnych założeniach (m.in. dotyczących regularności funkcji $\ln f(y_n|x_n; \theta)$) tak powstały estymator metody quasi-największej wiarygodności (*quasi-maximum likelihood estimator*) θ^{QML} jest zgodny i asymptotycznie normalny, przy czym

$$\sqrt{N} (\hat{\theta}^{QML} - \theta_0) \xrightarrow[N \rightarrow +\infty]{d} N \left(0, [\mathcal{I}_{2D} \mathcal{I}_{0P}^{-1} \mathcal{I}_{2D}]^{-1} \right), \quad (3.27)$$

gdzie $\mathcal{I}_{2D}$, to granica według prawdopodobieństwa (3.26) dla prawdziwej wartości parametru θ_0 , natomiast $\mathcal{I}_{0P}$, to granica (przy $N \rightarrow +\infty$) $1/N \sum_{n=1}^N [s_n(\theta)] [s_n(\theta)]^T$, gdzie

$$s_n(\theta) = \frac{\partial \ln f(y_n|x_n; \theta)}{\partial \theta} \Big|_{\theta=\theta_0}.$$

Warunki podane przez White (1982) są jednak w praktyce trudne do zweryfikowania. W takim przypadku należy badać własności estymatora metodami Monte Carlo.

3.4 Liniowe modele przestrzeni dla nieguassowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka

Barndorff-Nielsen i Shephard (2001b) przedstawili model stochastyczny BNS dany równaniami (2.3) oraz (2.4) w postaci przestrzeni stanów w przypadku, gdy $\beta = 0$.

Model przestrzeni 3.1.

Dla podstawowego modelu stochastycznej zmienności BNS (definicja (2.1)), dla któ-

rego $\beta = 0$ model przestrzeni stanów można zapisać jako

$$\begin{cases} Y_n = d + HX_n + u_n \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.28)$$

Wektory stanu i pomiaru przyjmują odpowiednio postać

$$X_n = \begin{bmatrix} \lambda\sigma_n^2 \\ \sigma^2(\Delta n) \end{bmatrix}, \quad Y_n = \begin{bmatrix} y_n \\ y_n^2 \end{bmatrix}. \quad (3.29)$$

Wektor c oraz macierze H i F są równe odpowiednio

$$c = \begin{bmatrix} \mu\Delta \\ \mu^2\Delta^2 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 0 \\ \lambda^{-1} & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 1 - e^{-\lambda\Delta} \\ 0 & e^{-\lambda\Delta} \end{bmatrix}. \quad (3.30)$$

Wektor szumu pomiarowego u_n ma wartość oczekiwaną i macierz wariancji-kowariancji równe odpowiednio

$$\mathbb{E}(u_n) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad R = \text{Var}(u_n) = \begin{bmatrix} \xi\Delta & 2\mu\Delta^2\xi \\ 2\mu\Delta^2\xi & 4\mu^2\Delta^3\xi + 2(2\omega^2r^{**}(\Delta) + \xi^2\Delta^2) \end{bmatrix}. \quad (3.31)$$

Wektor szumu procesu v_n przyjmuje postać

$$v_n = \begin{bmatrix} \eta_{2n} - \eta_{1n} \\ \eta_{1n} \end{bmatrix}. \quad (3.32)$$

Wektor losowy $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ ma wartość oczekiwaną i macierz wariancji kowariancji postaci

$$\mathbb{E}(\eta_n) = \xi \begin{bmatrix} 1 - e^{-\lambda\Delta} \\ \lambda\Delta \end{bmatrix}, \quad \text{Var}(\eta_n) = 2\omega^2 \begin{bmatrix} \frac{1}{2}(1 - e^{2\lambda\Delta}) & 1 - e^{-\lambda\Delta} \\ 1 - e^{-\lambda\Delta} & \lambda\Delta \end{bmatrix}. \quad (3.33)$$

Parametry ξ, ω^2 , to odpowiednio wartość oczekiwana i wariancja rozkładu stacjonarnego procesu wariancji chwilowej, a funkcja r^{**} jest dana wzorem (2.28) i zależy od parametru λ i długości okresu pomiędzy obserwacjami Δ . Wartości własne macierzy F są równe 0 i $e^{-\lambda\Delta}$. Dodatnia wartość parametru λ implikuje, że obie wartości własne leżą wewnątrz kręgu jednostkowego i można jako początkowe wartości przyjąć bezwarunkową wartość oczekiwaną i kowariancję zgodnie ze wzorami (3.17) i (3.18).

W prezentowanej w podrzdziale 3.3 teorii filtrów Kalmana wektory zaburzeń losowych miały zerową wartość oczekiwaną. Natomiast, w przedstawionej przez Barndorff-Nielsen i Shephard (2001b) przestrzeni stanów wartość oczekiwana wektora szumu pomiaru nie jest równa wektorowi zerowemu. Można jednak przekształcić równanie stanu do równoważnej postaci spełniającej warunek zerowej wartości oczekiwanej wektorów losowych. Równanie stanu jest równoważne równaniu

$$\begin{aligned} X_n = & \begin{bmatrix} \xi (\lambda \Delta - 1 + e^{-\lambda \Delta}) \\ \xi (1 - e^{-\lambda \Delta}) \end{bmatrix} + \begin{bmatrix} 0 & 1 - e^{-\lambda \Delta} \\ 0 & e^{-\lambda \Delta} \end{bmatrix} X_{n-1} \\ & + \begin{bmatrix} \eta_{2n} - \eta_{1n} - \xi (\lambda \Delta - 1 + e^{-\lambda \Delta}) \\ \eta_{1n} - \xi (1 - e^{-\lambda \Delta}) \end{bmatrix}. \end{aligned} \quad (3.34)$$

Równanie (3.34) zawiera wyraz wolny. Wstawiając w miejsce X_n wektor

$$\tilde{X}_n = \begin{bmatrix} \lambda \sigma_n^2 - \xi \Delta \lambda \\ \sigma^2(\Delta n) - \xi \end{bmatrix} = X_n - \begin{bmatrix} \xi \Delta \lambda \\ \xi \end{bmatrix}. \quad (3.35)$$

otrzymujemy równanie

$$\begin{aligned} \tilde{X}_n = & \begin{bmatrix} 0 & 1 - e^{-\lambda \Delta} \\ 0 & e^{-\lambda \Delta} \end{bmatrix} \tilde{X}_{n-1} + \begin{bmatrix} \eta_{2n} - \eta_{1n} - \xi (\lambda \Delta - 1 + e^{-\lambda \Delta}) \\ \eta_{1n} - \xi (1 - e^{-\lambda \Delta}) \end{bmatrix} \\ = & F \tilde{X}_{n-1} + \tilde{v}_n, \end{aligned} \quad (3.36)$$

bez wyrazu wolnego. Wartość oczekiwana $\tilde{v}_n$ jest równa wektorowi zerowemu, natomiast wariancja przyjmuje postać

$$\begin{aligned} \text{Var}(\tilde{v}_n) = & \text{Var} \left(\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \eta_n \right) = \\ = & 2\omega^2 \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2}(1 - e^{2\lambda \Delta}) & 1 - e^{-\lambda \Delta} \\ 1 - e^{-\lambda \Delta} & \lambda \Delta \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}^T = \\ = & 2\omega^2 \begin{bmatrix} -\frac{3}{2} - \frac{1}{2}e^{-2\lambda \Delta}2e^{-\lambda \Delta} + \lambda \Delta & 1 - e^{-\lambda \Delta} - \frac{1}{2}(1 - e^{2\lambda \Delta}) \\ 1 - e^{-\lambda \Delta} - \frac{1}{2}(1 - e^{2\lambda \Delta}) & \frac{1}{2}(1 - e^{2\lambda \Delta}) \end{bmatrix} = Q. \end{aligned} \quad (3.37)$$

Należy wtedy także skorygować równanie pomiaru

$$\begin{aligned} Y_n &= \begin{bmatrix} \mu\Delta \\ \mu^2\Delta^2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \lambda^{-1} & 0 \end{bmatrix} X_n + u_n = \\ &= \begin{bmatrix} \mu\Delta \\ \mu^2\Delta^2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \lambda^{-1} & 0 \end{bmatrix} X_n - \begin{bmatrix} 0 \\ \xi\Delta \end{bmatrix} + \begin{bmatrix} 0 \\ \xi\Delta \end{bmatrix} u_n = \\ &= \begin{bmatrix} \mu\Delta \\ \mu^2\Delta^2 + \xi\Delta \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \lambda^{-1} & 0 \end{bmatrix} \tilde{X}_n + u_n = d + H\tilde{X}_n + u_n. \end{aligned}$$

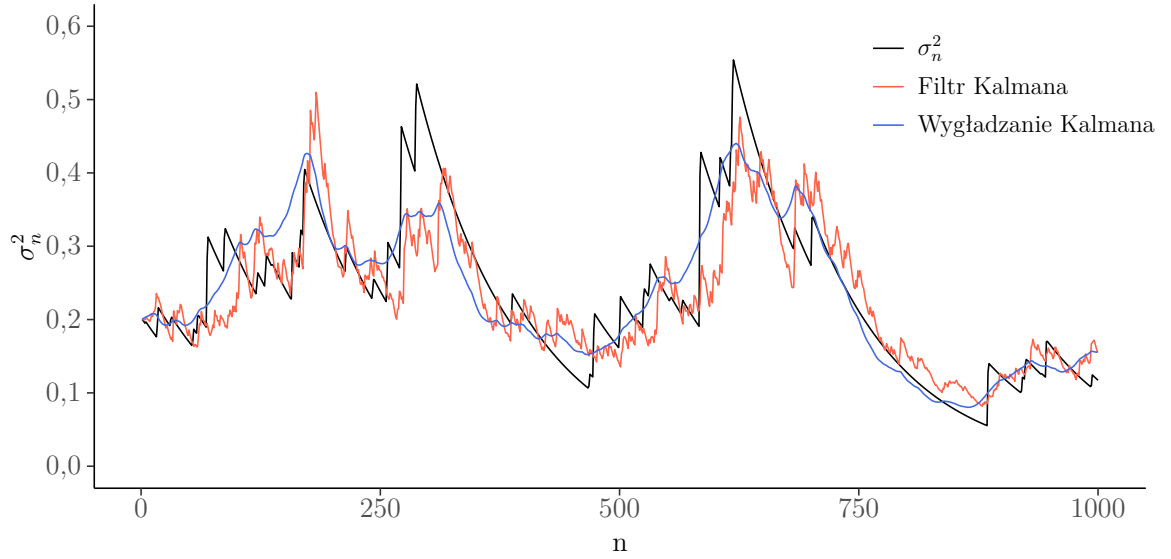
Otrzymujemy zatem następujący model przestrzeni stanów

$$\begin{cases} Y_n = \begin{bmatrix} \mu\Delta \\ \mu^2\Delta^2 + \xi\Delta \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \lambda^{-1} & 0 \end{bmatrix} \tilde{X}_n + u_n, \\ \tilde{X}_n = \begin{bmatrix} 0 & 1 - e^{-\lambda\Delta} \\ 0 & e^{-\lambda\Delta} \end{bmatrix} \tilde{X}_{n-1} + \tilde{\eta}_n, \end{cases} \quad (3.38)$$

z zaburzeniami u_n i $\tilde{v}_n$ o wartościach oczekiwanych równych wektorowi zerowemu i macierzami kowariancji R i Q danymi wzorami odpowiednio (3.31) i (3.37). Przedstawiona postać przestrzeni stanów jest liniowa, ale nie jest guassowska. Zaburzenia nie mają rozkładu normalnego, zatem za pomocą filtru Kalmana otrzymujemy najlepsze liniowe oszacowania wariancji aktualnej σ_n^2 w okresie n na podstawie obserwacji logarytmicznych zwrotów y_k i ich kwadratów y_k^2 dla $k = 1, \dots, n$. Nie jest to jednak estymator minimalizujący błąd średniokwadratowy. Stąd w dalszej części rozdziału zostanie zaprezentowane podejście z wykorzystaniem filtru cząsteczkowego. Mając do dyspozycji obserwacje z całej próby można także wykorzystać wygładzanie Kalmana, co poprawia dokładność estymacji. Na rysunku (3.1) przedstawiono przykład estymacji wariancji aktualnej za pomocą filtru oraz wygładzania Kalmana. Wariancja aktualna została wyznaczona na podstawie procesu wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,2$ oraz odchyleniu standardowym $\omega = 0,1$. Pozostałe parametry: $\lambda = 0,01, \mu = \beta = 0$.

Uwzględnienie parametru β różnego od zera jest w przypadku liniowej przestrzeni stanów niemożliwe. Zgodnie ze wzorem 2.24 ze strony 61 rozkład warunkowy logarytmicznych stóp zwrotu przyjmuje postać

$$y_n | \sigma_n^2 \sim N(\mu\Delta + \beta\sigma_n^2, \sigma_n^2),$$



Rysunek 3.1: Trajektoria wariacji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru Kalmana (czerwona linia) i wygładzania Kalmana (linia niebieska).

Źródło: opracowanie własne.

co można zapisać

$$y_n = \mu\Delta + \beta\sigma_n^2 + \sqrt{\sigma_n^2}\epsilon_n,$$

gdzie $\epsilon_n \sim N(0,1)$. Model jest nieliniowy ze względu na wariację aktualną σ_n^2 . Można by było wykorzystać jedno z uogólnień filtru Kalmana dla przypadku nieliniowego: rozszerzonego filtru Kalmana (*Extended Kalman Filter*, EKF) lub bezśladowego filtru Kalmana (*Unscented Kalman Filter*, UKF), ale w przypadku jednocześnie nieliniowym i niegaussowskim jest to rozwiązanie mniej optymalne niż zastosowanie filtrów cząsteczkowych (Brzozowska-Rup i Dawidowicz, 2009). Uwzględnienie parametru β różnego od zera i włączenie efektu dźwigni do modelu przestrzeni stanów zostanie przedstawione w specyfikacjach filtru cząsteczkowego w podrozdziale 3.6.

Barndorff-Nielsen i Shephard (2001b) zasugerowali, że można również zastosować filtr Kalmana w przypadku złożenia złożenia kilku procesów wariacji chwilowej. Nie przedstawili jednak odpowiedniego modelu przestrzeni stanów. Poniższa postać przestrzeni stanów została przedstawiona w pracy Szczepocki (2018)⁶.

⁶W pracy Szczepocki (2018) zastosowano nieco inną metodę wyznaczania wag w złożeniu procesów zmienności w oparciu na propozycji Griffin i Steel (2006)

Model przestrzeni 3.2.

Dla modelu stochastycznej zmienności BNS ze złożeniem procesów zmienności (definicja (2.2) ze strony 67), dla którego $\beta = 0$ model przestrzeni stanów można zapisać jako

$$\begin{cases} Y_n = d + HX_n + u_n \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.39)$$

Wektor pomiaru przyjmuje postać

$$Y_n = \begin{bmatrix} y_n \\ y_n^2 \end{bmatrix}_{2 \times 1}, \quad (3.40)$$

natomiast wektor stanu

$$X_n = \begin{bmatrix} \lambda_1 \sigma_{1,n}^2 - w_1 \xi \Delta \lambda_1 \\ \sigma_1^2(\Delta n) - w_1 \xi \\ \lambda_2 \sigma_{2,n}^2 - w_2 \xi \Delta \lambda_2 \\ \sigma_2^2(\Delta n) - w_2 \xi \\ \vdots \\ \lambda_P \sigma_{P,n}^2 - w_P \xi \Delta \lambda_P \\ \sigma_P^2(\Delta n) - w_P \xi \end{bmatrix}_{2P \times 1} \quad (3.41)$$

Wektor c oraz macierz $H_{2 \times 2P}$ są równe odpowiednio

$$c = \begin{bmatrix} \mu \Delta \\ \mu^2 \Delta^2 + \xi \Delta \end{bmatrix}_{2 \times 1}, \quad H = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \lambda_1^{-1} & 0 & \lambda_2^{-1} & 0 & \cdots & \lambda_P^{-1} & 0 \end{bmatrix}_{2 \times 2P}. \quad (3.42)$$

Macierz $F_{2P \times 2P}$ jest blokowo-diagonalna o blokach

$$F^{(p)} = \begin{bmatrix} 0 & 1 - e^{-\lambda_p \Delta} \\ 0 & e^{-\lambda_p \Delta} \end{bmatrix}.$$

dla $p = 1, \dots, P$. Wektor losowy u_n ma wartość oczekiwaną równą wektorowi zerowemu (w $\mathbb{R}^2$) i macierz wariancji-kowariancji postaci

$$R = \text{Var}(u_n) = \begin{bmatrix} \xi \Delta & 2\mu \Delta^2 \xi \\ 2\mu \Delta^2 \xi & 4\mu^2 \Delta^3 \xi + 2 \left(2\omega^2 \sum_{p=1}^P w_p r_p^{**}(\Delta) + \xi^2 \Delta^2 \right) \end{bmatrix}_{2 \times 2}. \quad (3.43)$$

Wektor losowy v_n przyjmuje postać

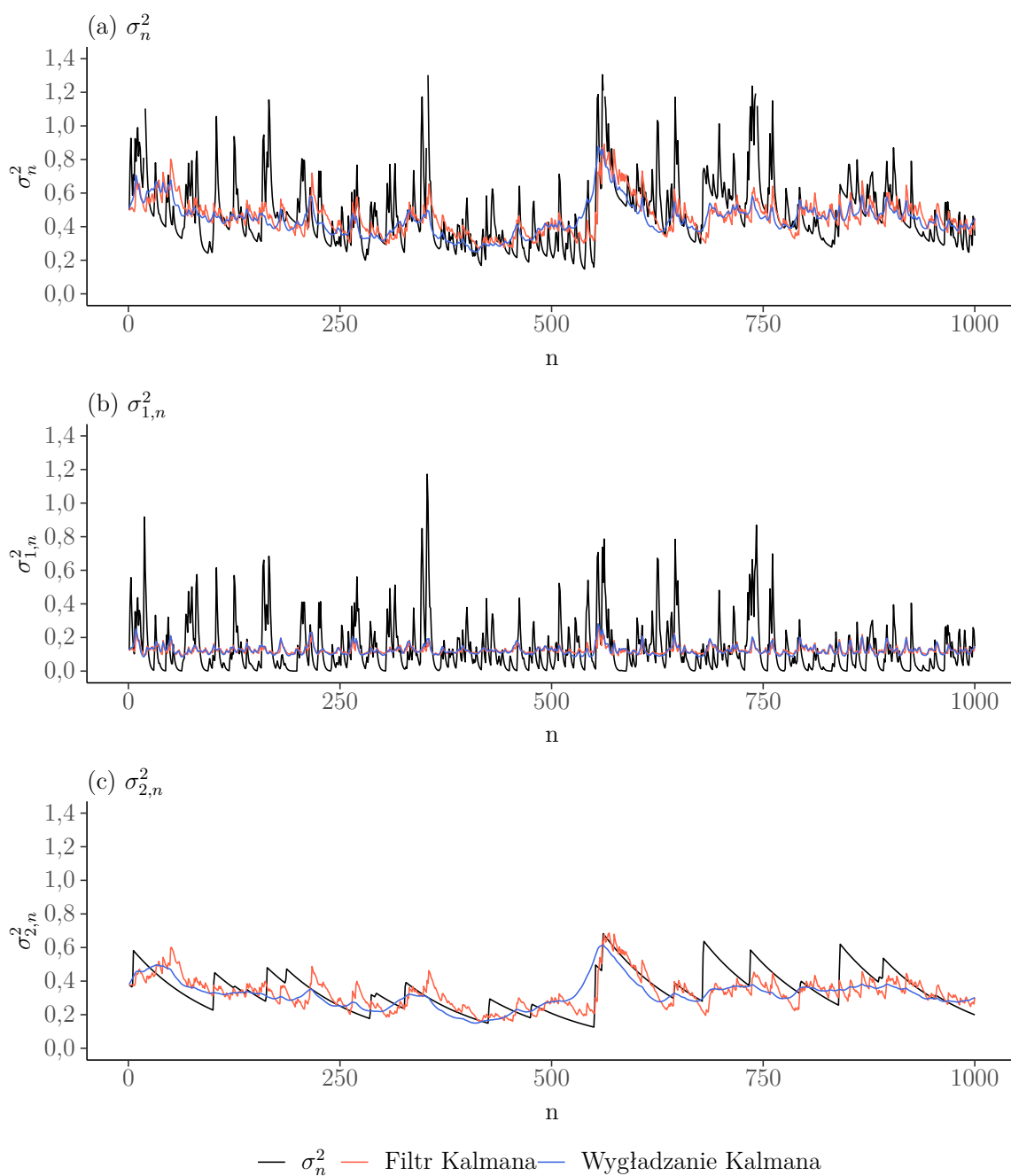
$$\begin{bmatrix} \eta_{2n}^{(1)} - \eta_{1n}^{(1)} - w_1 \xi (\lambda_1 \Delta - 1 + e^{-\lambda_1 \Delta}) \\ \eta_{1n}^{(1)} - w_1 \xi (1 - e^{-\lambda_1 \Delta}) \\ \eta_{2n}^{(2)} - \eta_{1n}^{(2)} - w_2 \xi (\lambda_2 \Delta - 1 + e^{-\lambda_2 \Delta}) \\ \eta_{1n}^{(2)} - w_2 \xi (1 - e^{-\lambda_2 \Delta}) \\ \vdots \\ \eta_{2n}^{(P)} - \eta_{1n}^{(P)} - w_P \xi (\lambda_P \Delta - 1 + e^{-\lambda_P \Delta}) \\ \eta_{1n}^{(P)} - w_P \xi (1 - e^{-\lambda_P \Delta}) \end{bmatrix}_{2P \times 1},$$

Wartość oczekiwana wektora v_n jest równa wektorowi zerowemu (w $\mathbb{R}^{2P}$), natomiast macierz wariancji-kowariancji $Q_{2P \times 2P}$ jest macierzą blokowo-diagonalną o blokach

$$2w_p \omega^2 \begin{bmatrix} -\frac{3}{2} - \frac{1}{2} e^{-2\lambda_p \Delta} 2e^{-\lambda_p \Delta} + \lambda_p \Delta & 1 - e^{-\lambda_p \Delta} - \frac{1}{2} (1 - e^{2\lambda_p \Delta}) \\ 1 - e^{-\lambda_p \Delta} - \frac{1}{2} (1 - e^{2\lambda_p \Delta}) & \frac{1}{2} (1 - e^{2\lambda_p \Delta}) \end{bmatrix}.$$

W efekcie otrzymujemy liniową postać przestrzeni stanu, która umożliwia użycie filtru Kalmana, ale podobnie jak w przypadku podstawowego modelu BNS zaburzenia nie mają rozkładu gaussowskiego. Zatem oszacowania wariancji aktualnej nie są optymalne (nie minimalizują błędu średniokwadratowego). Na rysunku (3.2) przedstawiono przykład zastosowania filtru Kalmana, w przypadku złożenia dwóch procesów wariancji chwilowej rozkładzie brzegowym gamma (proces Ga-supOU₂) o wartości oczekiwanej $\xi = 0,5$ oraz odchyleniu standardowym $\omega = 0,3$ oraz wagach $w_1 = 0,25$ i $w_2 = 0,75$ i parametrach persystencji równych odpowiednio $\lambda_1 = 0,5$, $\lambda_2 = 0,01$. Ponadto przyjęto $\mu = \beta = 0$. Widać wyraźnie, że zarówno filtr jak i wygładzanie Kalmana dają dużo lepsze oszacowanie drugiej składowej o większej wadze w złożeniu.

Filtr Kalmana można wykorzystać również wykorzystany w przypadku, gdy zamiast logarytmicznych zwrotów obserwowane są wartości wariancji zrealizowanej. Barndorff-Nielsen i Shephard (2002) przedstawili model przestrzeni stanów dla modelu stochastycznej zmienności w oparciu o dane wysokiej częstotliwości i estymator wariancji zrealizowanej (por. podrozdział 2.6). Wykorzystali przy tym własność wariancja aktualnej σ_n^2 pozwalającą zapisać ją w reprezentacji stanu modelu ARMA(1,1).



Rysunek 3.2: Trajektoria wariancji aktualnej σ_n^2 (a) powstałej ze złożenie dwóch wariancji aktualnych $\sigma_{1,n}^2$ (b) i $\sigma_{2,n}^2$ (c) oraz ich oszacowania za pomocą filtru Kalmana (czerwona linia) i wygładzania Kalmana (linia niebieska).

Źródło: opracowanie własne.

Model przestrzeni 3.3.

Dla podstawowego modelu stochastycznej zmienności BNS (definicja (2.1) ze strony 51), dla którego $\mu = \beta = 0$ model przestrzeni stanów można zapisać jako

$$\begin{cases} Y_n = d + HX_n + u_n \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.44)$$

Proces stanu i pomiaru przyjmują odpowiednio postać

$$X_n = \begin{bmatrix} \sigma_n^2 - \xi\Delta \\ \epsilon_n \end{bmatrix}, \quad Y_n = \{y_n\}. \quad (3.45)$$

Wektor c jest stałą liczbową $c = \mu\Delta$, macierze H i F są równe

$$H = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} \phi & 1 \\ 0 & 0 \end{bmatrix}. \quad (3.46)$$

Wektor szumu pomiarowego u_n jest jednowymiarową zmienną losową o wartości oczekiwanej 0 i wariancji (por. wzór (2.50))

$$Q = 2M \left(2\omega^2 \lambda^{-2} \left(e^{-\lambda\Delta/M} - 1 + \lambda\Delta M^{-1} \right) + \left(\Delta M^{-1} \right)^2 \xi^2 \right). \quad (3.47)$$

Natomiast wektor szumu procesu v_n ma wartość oczekiwaną równą wektorowi zerowemu (w $\mathbb{R}^2$) i macierz wariancji-kowariancji

$$R = \begin{bmatrix} \sigma_\sigma^2 & \sigma_\sigma^2 \vartheta \\ \sigma_\sigma^2 \vartheta & \sigma_\sigma^2 \vartheta^2 \end{bmatrix} \quad (3.48)$$

Parametry $\phi, \vartheta, \sigma_\sigma^2$ odpowiadają składnikowi autoregresyjnemu, średniej ruchomej i wariancji procesu ARMA(1,1) dla wariancji aktualnej σ_n^2 , natomiast ϵ_n jest składnikiem losowym tego procesu.

Model przestrzeni stanu przyjmuje zatem postać

$$\begin{cases} Y_n = \Delta\xi + \begin{bmatrix} 1 & 0 \end{bmatrix} X_n + u_n \\ X_n = \begin{bmatrix} \phi & 1 \\ 0 & 0 \end{bmatrix} X_{n-1} + \begin{bmatrix} \sigma_\sigma \\ \sigma_\sigma \vartheta \end{bmatrix} \eta_n, \end{cases} \quad (3.49)$$

Zaburzenia nie mają rozkładu normalnego, zatem za pomocą filtru Kalmana otrzymujemy najlepsze liniowe oszacowania wariancji aktualnej σ_n^2 (przesunięte o $\xi\Delta$) w okresie n na podstawie estymatora wariancji zrealizowanej obserwowanego do okresu n . Na rysunku 3.3 przedstawiono przykład wykorzystania modelu przestrzeni stanów postaci (3.49) do estymacji wariancji aktualnej przy pomocy filtru i wygładzania Kalmana. Widać wyraźnie, że wraz ze wzrostem M rośnie dokładność estymacji. Wariancja aktualna została wyznaczona na podstawie procesu wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,2$ oraz odchyleniu standardowym $\omega = 0,1$. Pozostałe parametry: $\lambda = 0,01$, $\mu = \beta = 0$.

Modele przestrzeni stanów opisane przez twierdzenia 3.1 oraz 3.7 choć różnią się zarówno równaniami przejścia jak i procesu, to mają pewne elementy wspólne. Po pierwsze model określony przez twierdzenie 3.7 można także wykorzystać w przypadku, gdy zamiast wariancji zrealizowanej obserwowane są zwroty dzienne, pod warunkiem, że $\mu = \beta = 0$. Można wówczas przyjąć $M = 1$, co oznacza, że wariancja zrealizowana jest równa kwadratowi dziennym logarytmicznych stóp zwrotu. Po drugie równania przejścia w modelach przestrzeni stanów opisanych przez twierdzenia 3.1 oraz 3.7 dostarczają identycznego oszacowania *a priori* $\hat{\sigma}_{n|n-1}^2$ wariancji aktualnej (przy danym oszacowaniu $\hat{\sigma}_{n-1|n-1}^2$). Choć drugie składowe wektorów stanów są inne to i tak nie one biorą udziału w równaniach pomiaru. Zatem równania stanu mogą być w tych dwóch modelach przestrzeni stanów stosowane zamiennie⁷. Wynika to z faktu, że obie reprezentacje korzystają z tej samej informacji o procesie wariancji chwilowej: wartości oczekiwanej ξ , wariancji ω^2 oraz funkcji autokorelacji r .

Również model przestrzeni stanów (3.49) można uogólnić na przypadek złożenia P procesów zmienności. Barndorff-Nielsen i Shephard (2002) w zarysie przedstawili jak uwzględnić złożenie procesów zmienności i przedstawili wyniki badania empirycznego dla złożenia dwóch i trzech procesów zmienności.

Model przestrzeni 3.4.

Dla modelu stochastycznej zmienności BNS ze złożeniem procesów zmienności (definicja (2.2) ze strony 67), w którym parametr $\beta = 0$ model przestrzeni stanów a

⁷Zamiana równań stanów pociąga za sobą korektę równań pomiaru.

procesem pomiaru jest wariancja zrealizowana przyjmuje postać

$$\begin{cases} Y_n = d + HX_n + u_n \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.50)$$

Proces stanu i pomiaru przyjmują odpowiednio postać

$$X_n = \begin{bmatrix} \sigma_{1,n}^2 - w_1 \xi \Delta \\ \epsilon_{1,n} \\ \sigma_{2,n}^2 - w_2 \xi \Delta \epsilon_{2,n} \\ \vdots \\ \sigma_{P,n}^2 - w_P \xi \Delta \epsilon_{P,n} \end{bmatrix}_{2P \times 1}, \quad Y_n = \{y_n\} \quad (3.51)$$

Wektor c jest stałą liczbową $c = \mu \Delta$, macierze H jest równa

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad (3.52)$$

natomiast macierz $F_{2P \times 2P}$ jest blokowo-diagonalna o blokach

$$F = \begin{bmatrix} \phi_p & 1 \\ 0 & 0 \end{bmatrix}_{2 \times 2}. \quad (3.53)$$

Wektor szumu pomiarowego u_n jest jednowymiarową zmienną losową o wartości oczekiwanej 0 i wariancji (por. wzór (2.52))

$$Q = 2M \left(= 2\omega^2 \sum_{p=1}^P \frac{w_p}{2\lambda_p^2} \left(e^{-\lambda_p \Delta M^{-1}} - 1 + \lambda_p \Delta M^{-1} \right) + \left(\Delta M^{-1} \right)^2 \xi^2 \right). \quad (3.54)$$

Natomiast wektor szumu procesu v_n ma wartość oczekiwaną równą wektorowi zerowemu (w $\mathbb{R}^{2P}$) i macierzy wariancji-kowariancji $R_{2P \times 2P}$ blokowo-diagonalnej o blokach

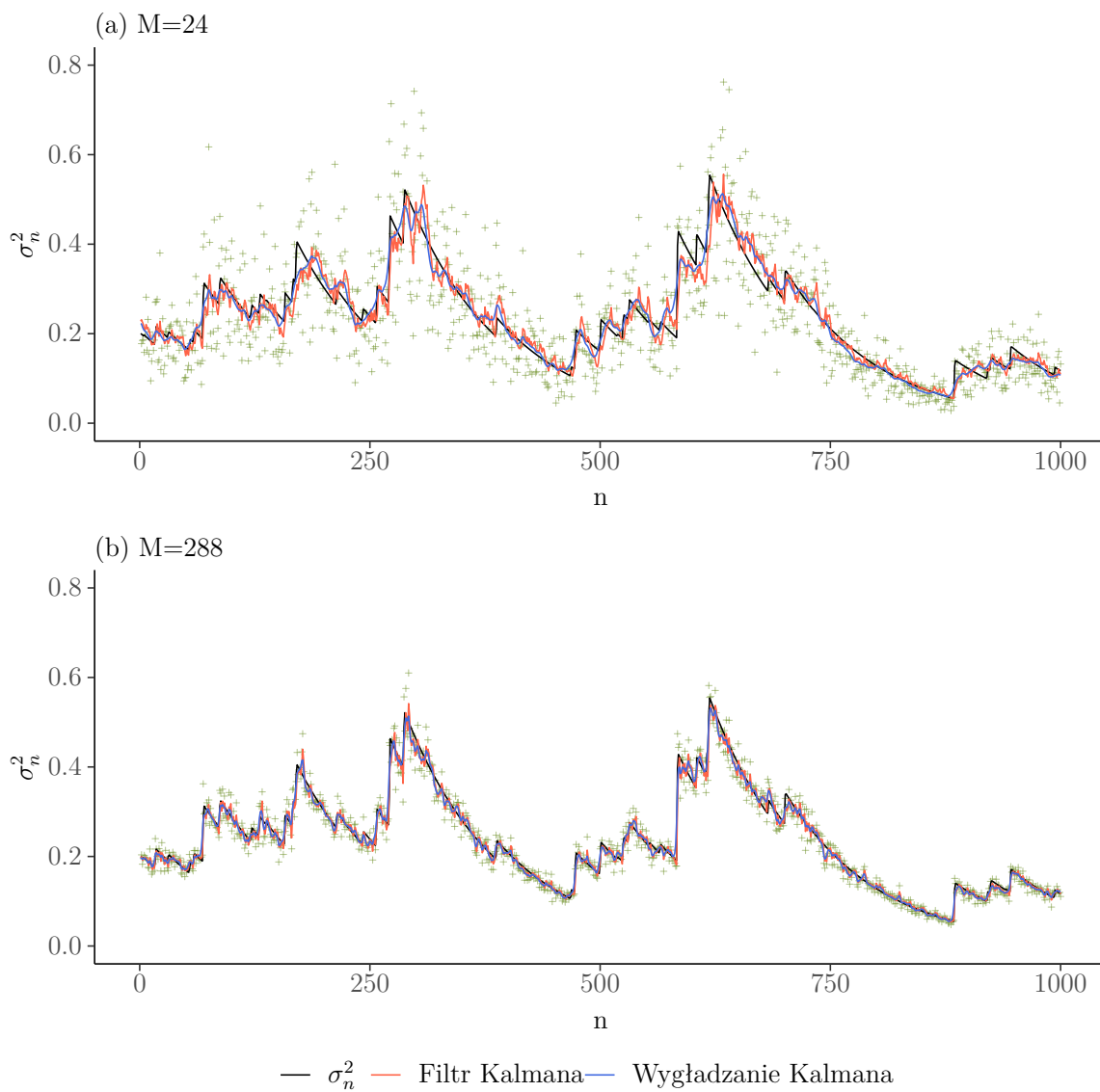
$$R_p = \begin{bmatrix} \sigma_{\sigma,p}^2 & \sigma_{\sigma,p}^2 \vartheta_p \\ \sigma_{\sigma,p}^2 \vartheta_p & \sigma_{\sigma,p}^2 \vartheta_p^2 \end{bmatrix}. \quad (3.55)$$

Dla $p = 1, \dots, P$ parametry $\phi_p, \vartheta_p, \sigma_{\sigma,p}^2$ odpowiadają składnikowi autoregresyjnemu, średniej ruchomej i wariancji procesu ARMA(1,1) dla wariancji aktualnej $\sigma_{p,n}^2$, a $\epsilon_{p,n}$ jest składnikiem losowym tego procesu.

Podobnie jak w przypadku twierdzeń 3.1 i 3.7 można zamiennie stosować równania przestrzeni stanów z twierdzeń 3.4 oraz 3.2.

Przedstawione modele przestrzeni stanów nie zależą od wyboru rozkładu stacjonarnego procesu wariacji chwilowej, a jedynie od parametrów tego rozkładu: wartości oczekiwanej i wariancji oraz parametru λ . Jest to zatem podejście semi-parametryczne (Barndorff-Nielsen i Shephard, 2001b).

Linijowe przestrzenie stanów umożliwiają zastosowanie metody quasi-największej wiarygodności do wyznaczenia estymatorów parametrów. Wartość estymatora można wyznaczyć maksymalizując numerycznie funkcję quasi-wiarygodności daną wzorem (3.23). Podejście to zostało zaprezentowane w Barndorff-Nielsen i Shephard (2002) do modeli przestrzeni stanów 3.7 oraz 3.4 dla szeregu obserwacji wariancji zrealizowanej wyznaczonych na podstawie kursu USD/DM (dolara wyrażonego w markach niemieckich). W pracy Szczepocki (2018) zastosowano metodę quasi-największej wiarygodności na podstawie modelu przestrzeni stanów 3.2 do modelowania zmienności kursu EUR/PLN (kursu Euro wyrażonego w złotych). W pracy tej do numerycznego maksymalizowania wartości funkcji quasi-wiarygodności zastosowano algorytm *L-BFGS-B* (*Limited Memory BFGS*) zaimplementowany w programie R. Jest to wariant metody quasi-Newtona z pracy Byrd i in. (1995).



Rysunek 3.3: Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru Kalmana (czerwona linia) i wygładzania Kalmana (linia niebieska) dla $M = 24, 288$. Punktami + oznaczono oszacowania wariancji zrealizowanej.

Źródło: opracowanie własne.

3.5 Filtry cząsteczkowe

Filtry cząsteczkowe, zwane także sekwencyjnymi metodami Monte Carlo, zostały opracowane niezależnie przez różnych autorów (Pitt i Shephard, 1999). Jako pierwszą wkszuje się pracę Gordon i in. (1993), w której zaproponowano algorytm, który obecnie nazywa się bootstrapowym filtrem cząsteczkowym, do estymacji niegaussowskich modeli przestrzeni stanów. Kitagawa (1996) przedstawił podobny, ale bardziej ogólny algorytm do modelowania szeregów czasowych. Po raz pierwszy termin filtr cząsteczkowy został użyty w pracy Del Moral (1996).

Filtry cząsteczkowe aproksymują rekurencyjnie gęstości $p(x_{0:n}|y_{1:n}; \theta)$ dla kolejnych wartości $n = 1, \dots, N$. Ponadto dostarczają także aproksymacje gęstości filtracji $p(x_n|y_{1:n}; \theta)$. Metody te przybliżają rozkłady ciągłe $p(x_{0:n}|y_{1:n}; \theta)$ rozkładami dyskretnymi wyznaczonymi za pomocą ważonych losowych prób $(x_{0:n}^{(i)}, w_n^{(i)})_{i=1}^K$, zwanych cząsteczkami. Nie można jednak stosować klasycznej metody Monte Carlo i losować bezpośrednio z rozkładu $p(x_{0:n}|y_{1:n}; \theta)$, ponieważ rozkład ten nie jest znany. Zamiast z tego stosuje się losowanie z rozkładu ważności (*importance sampling*) o funkcji gęstości (zwanej funkcją ważności) $q_{0:n}(x_{0:n}|y_{1:n}; \theta)$, który przybliży nieznany rozkład $p(x_{0:n}|y_{1:n}; \theta)$. Rozkład ważności dobiera się tak, aby był jak najbardziej podobny aproksymowanemu rozkładowi i jednocześnie była możliwość generowania zmiennych losowych z tego rozkładu. Wylosowana w ten sposób próba musi zatem zostać „przeważona” ze względu na fakt losowanie z niepoprawnego rozkładu. Wagi, zwane wagami ważności (*importance weights*) definiuje się jako iloraz

$$w_n^{(i)} = \frac{p(x_{0:n}^{(i)}|y_{1:n}; \theta)}{q_{0:n}(x_{0:n}^{(i)}|y_{1:n}; \theta)}, \quad (3.56)$$

dla $i = 1, \dots, K$. Rozkład ważności w ogólności może zmieniać się wraz czasem (stąd $0 : n$ w indeksie dolnym). Zazwyczaj w celu przyspieszenia obliczeń funkcję ważności dekomponuje się na dwie części

$$q_{0:n}(x_{0:n}|y_{1:n}; \theta) = q_n(x_n|x_{0:n-1}, y_{1:n}; \theta) q_{0:n-1}(x_{0:n-1}|y_{1:n-1}; \theta). \quad (3.57)$$

Dla każdej cząsteczki $(x_{0:n}^{(i)})_{i=1}^K$ drugi składnik $q_{0:n-1}(x_{0:n-1}^{(i)}|y_{1:n-1}; \theta)$ jest deltą Diraca $\delta_{x_{0:n-1}^{(i)}}$, czyli rozkładem przyjmującym wartość 1 dla wylosowanej w okresie

$n - 1$ cząsteczek $x_{0:n-1}^{(i)}$ i przyjmujący wartość 0 dla wszystkich innych możliwych prób $x_{0:n-1}$. W celu otrzymania dla okresu n nowego zestawu cząsteczek $(x_{0:n}^{(i)})_{i=1}^K$ losuje się tylko $(x_n^{(i)})_{i=1}^K$ z rozkładu o funkcji ważności $q_n(x_n|x_{0:n-1}, y_{1:n}; \theta)$ i dołącza do otrzymanych do okresu $n - 1$ cząsteczek $(x_{0:n-1}^{(i)})_{i=1}^K$ nowe składowe

$$(x_{0:n}^{(i)})_{i=1}^K = (x_{0:n-1}^{(i)}, x_n^{(i)})_{i=1}^K.$$

Dodatkowo można zdekomponować gęstość

$$\begin{aligned} p(x_{0:n}|y_{1:n}; \theta) &= \frac{p(y_n|x_{0:n}, y_{1:n-1}; \theta) p(x_{0:n}|x_{0:n}, y_{1:n-1}; \theta)}{p(x_{0:n}|y_{1:n-1}; \theta)} = \\ &= \frac{p(y_n|x_{0:n}, y_{1:n-1}; \theta) p(x_n|x_{0:n-1}, y_{1:n-1}; \theta) p(x_{0:n-1}|y_{1:n-1}; \theta)}{p(y_n|y_{1:n-1}; \theta)} = \\ &= \frac{p(y_n|x_n; \theta) p(x_n|x_{n-1}; \theta) p(x_{0:n-1}|y_{1:n-1}; \theta)}{p(y_n|y_{1:n-1}; \theta)} \end{aligned} \quad (3.58)$$

Wstawiając (3.57) i (3.58) do (3.56) otrzymujemy

$$\begin{aligned} w_n^{(i)} &= \frac{p(y_n|x_n^{(i)}; \theta) p(x_n^{(i)}|x_{n-1}^{(i)}; \theta) p(x_{0:n-1}^{(i)}|y_{1:n-1}; \theta)}{p(y_n|y_{1:n-1}; \theta) q_n(x_n^{(i)}|x_{0:n-1}^{(i)}, y_{1:n}; \theta) q_{0:n-1}(x_{0:n-1}^{(i)}|y_{1:n-1}; \theta)} \\ &\propto w_{n-1}^{(i)} \frac{p(y_n|x_n^{(i)}; \theta) p(x_n^{(i)}|x_{n-1}^{(i)}; \theta)}{q_n(x_n^{(i)}|x_{0:n-1}^{(i)}, y_{1:n}; \theta)}, \end{aligned} \quad (3.59)$$

dla $i = 1, \dots, K$ (symbol $\propto$ oznacza "wprost proporcjonalne do"). Pozwala to wyznaczyć wagi w okresie n tylko na podstawie wag z poprzedniego okresu $n - 1$ oraz iloczynu gęstości $p(x_n|x_{n-1}; \theta)$ i $p(y_n|x_n; \theta)$ określonych przez dany model przestrzeni stanów podzielonych przez funkcję ważności $q_n(x_n|x_{0:n-1}, y_{1:n}; \theta)$. Iloraz ten nazywany jest współczynnikiem przyrostu wagi i jest niezależny od przeszłych obserwacji zmiennej pomiaru i od trajektorii procesu stanu do okresu $n - 2$ włącznie. Do rozpoczęcia rekurencji potrzebna jest losowanie z rozkładu ważności dla $n = 0$. Zazwyczaj jest to losowanie ze znanego rozkładu $p(x_0; \theta)$.

Mając do dyspozycji zbiór cząsteczek $(x_{0:n}^{(i)}, w_n^{(i)})_{i=1}^K$ można dokonać aproksymacji rozkładu $p(x_{0:n}|y_{1:n}; \theta)$. Zazwyczaj dokonuje się to przy pomocy znormalizowanych wag

$$\hat{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{i=1}^K w_n^{(i)}}$$

w celu zwiększenia stabilności estymatora. Wówczas

$$\hat{p}(x_{0:n}|y_{1:n};\theta) = \sum_{i=1}^K \hat{w}_n^{(i)} \delta_{x_{0:n}^{(i)}}(x_{0:n}) \approx p(x_{0:n}|y_{1:n};\theta). \quad (3.60)$$

Ponadto można także uzyskać oszacowanie wartości oczekiwanej dla pewnej funkcji $f(x_{0:n})$ zmiennej przestrzeni stanu postaci

$$\sum_{i=1}^K \hat{w}_n^{(i)} f(x_{0:n}^{(i)}) \approx \mathbb{E}[f(X_{0:n})|y_{1:n}]. \quad (3.61)$$

Przedstawiona procedura wyznaczania oszacowań gęstości $p(x_{0:n}|y_{1:n};\theta)$ dla kolejnych wartości $n \in \mathbb{N}$ nazywana jest sekwencyjną metodą ważności (*sequential importance sampling*, SIS) i jest podstawowym wariantem filtru cząsteczkowego. Zapis w postaci pseudokodu przedstawia algorytm 3.

Algorytm 3. Filtr cząsteczkowy - sekwencyjna metoda ważności (SIS)

1. W okresie $n = 0$.

Dla $i = 1, \dots, K$:

- (a) Losujemy $x_0^{(i)}$ z rozkładu $q_0(x_0)$
- (b) Wyznaczamy wagi $w_n^{(i)} = \frac{p(x_0)}{q_0(x_0)}$

2. Dla okresów $n = 1, \dots, N$ wykonujemy:

(i) Dla $i = 1, \dots, K$:

- (a) Losujemy $x_n^{(i)}$ z rozkładu $q_n(x_n|x_{0:n-1}, y_{1:n};\theta)$
- (b) Wyznaczamy wagi $w_n^{(i)} = w_{n-1}^{(i)} \frac{p(y_n|x_n^{(i)};\theta) p(x_n^{(i)}|x_{n-1}^{(i)};\theta)}{q_n(x_n^{(i)}|x_{0:n-1}, y_{1:n};\theta)}$

(ii) Dla $i = 1, \dots, K$:

Normalizujemy wagi: $\hat{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{i=1}^K w_n^{(i)}}$

(iii) Wyznaczamy oszacowania wartości $\mathbb{E}[f(x_n)] \approx \sum_{i=1}^K \hat{w}_n^{(i)} f(x_n^{(i)})$

Rezultat algorytmu:

Oszacowania wartości $\mathbb{E}[f(x_n)]$ dla okresów $n = 1, \dots, N$.

Istotną wadą ograniczającą lub wręcz uniemożliwiającą stosowanie filtru cząsteczkowego SIS w praktyce jest problem degeneracji wag. W praktyce, po kilku

iteracjach algorytmu SIS niemal wszystkie wagi będą równe 0 poza jedną wagą o wartości bliskiej 1. Gordon i in. (1993) zaproponowali, aby w algorytmie SIS dodać jeszcze jeden krok: po wyznaczeniu oszacowań należy dokonać repróbkiowania (*resampling*), czyli losowania ze zwracaniem N cząsteczek według prawdopodobieństw wyznaczonych przez znormalizowane wagi. Ma to na celu powieleniu cząsteczek o dużych wagach i wyeliminowanie cząsteczek o małych wagach. Po wylosowaniu ze zwracaniem wagi są zrównane do odwrotności liczby cząsteczek: $w_n^{(i)} = 1/K$, dla $i = 1, \dots, K$. Tak powstały algorytm filtru cząsteczkowego jest nazywany *sequential importance resampling* (SIR), a w polskiej literaturze klasycznym filtrem cząsteczkowym (Brzozowska-Rup i Dawidowicz, 2009). Algorytm 4 przedstawia zapis klasycznego filtru cząsteczkowego w pseudokodzie.

Algorytm 4. Filtr cząsteczkowy - klasyczny algorytm (SIR)

1. W okresie $n = 0$.

Dla $i = 1, \dots, K$:

(a) Losujemy $x_0^{(i)}$ z rozkładu $q_0(x_0)$

(b) Wyznaczamy wagi $w_n^{(i)} = \frac{p(x_0)}{q_0(x_0)}$

2. Dla okresów $n = 1, \dots, N$ wykonujemy:

(i) Dla $i = 1, \dots, K$:

(a) Losujemy $x_n^{(i)}$ z rozkładu $q_n(x_n | x_{0:n-1}, y_{1:n}; \theta)$

(b) Wyznaczamy wagi $w_n^{(i)} = w_{n-1}^{(i)} \frac{p(y_n | x_n^{(i)}; \theta) p(x_n^{(i)} | x_{n-1}^{(i)}; \theta)}{q_n(x_n^{(i)} | x_{0:n-1}, y_{1:n}; \theta)}$

(ii) Dla $i = 1, \dots, K$:

Normalizujemy wagi: $\hat{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{i=1}^K w_n^{(i)}}$

(iii) Wyznaczamy oszacowania wartości $\mathbb{E}[f(x_n)] \approx \sum_{i=1}^K \hat{w}_n^{(i)} f(x_n^{(i)})$

(iv) Losujemy ze zwracaniem K cząsteczek z prawdopodobieństwami $(\hat{w}_n^{(i)})_{i=1}^K$.

(v) Przyjmujemy $w_n^{(i)} = 1/K$ dla $i = 1, \dots, K$

Rezultat algorytmu:

Oszacowania wartości $\mathbb{E}[f(x_n)]$ dla okresów $n = 1, \dots, N$.

Stosowanie repróbkiowania ma również wady. Przede wszystkim eliminacja cza-

steczek o najmniejszych wagach i zwielokrotnienie cząsteczek powoduje degenerację różnorodności próby oraz utratę niezależności cząsteczek. Ponadto zwiększa także wariancję estymatorów. Dlatego część autorów sugeruje stosowanie repróbkowania tylko w przypadku, gdy wagi ważności są niezbilansowane. Trzy popularne miary zbilansowania wag to współczynnik wariacji (*coefficient of variation*, CV (Kong i in., 1994)), efektywna liczebność próby (*effective sample size*, ESS (Liu, 1996)) i entropia Shannona (*Shannon Entropy*, SE). W przypadku, gdy dla pewnego n miara zbilansowania wag spadnie poniżej pewnej z góry ustalonej wartości progowej, to następuje w danej iteracji repróbkowanie. W konsekwencji losowanie ze zwracaniem odbywa się w losowych iteracjach algorytmu.

Najczęściej jako miarę zbilansowania stosuje się efektywną liczebność próby definiowaną za pomocą wariancji wag ważności

$$K_{ESS} = \frac{K}{1 + \text{Vec}_q(\hat{w}_n^{(i)})} = \frac{K}{\mathbb{E} \left[\left(\hat{w}_n^{(i)} \right)^2 \right]}$$

Współczynnik ten przyjmuje wartości $1 \leq K_{ESS} \leq K$. W przypadku, gdy $K_{ESS} = K$ próba jest doskonale zbilansowana i wszystkie cząsteczki dokładają się do estymatora (3.60). Nie można go jednak zazwyczaj policzyć analitycznie. Zamiast tego stosuje się oszacowanie

$$\hat{K}_{ESS} = \frac{1}{\sum_{i=1}^K \hat{w}_n^{(i)}}, \quad (3.62)$$

które również przyjmuje wartości $1 \leq \hat{K}_{ESS} \leq K$. Najczęściej wskazuje się jako progową wartość połowę liczby cząstek $\tilde{K} = K/2$.

Współczynnik wariacji CV definiuje się jako

$$CV = \sqrt{\frac{1}{K} \sum_{i=1}^K \left(K \hat{w}_n^{(i)} - 1 \right)^2}. \quad (3.63)$$

Miara ta przyjmuje wartości $0 \leq CV \leq \sqrt{K-1}$. Jeśli wszystkie wagi mają identyczną wartość to miara przyjmuje wartość 0. Jeżeli jedna cząsteczka ma wagę równą 1, natomiast wagi pozostałych cząsteczek są zerowe, to miara przyjmuje wartość $\sqrt{K-1}$. W porównaniu do pozostałych dwóch miar jako jedyna przyjmuje tym większą wartość im gorsze jest zbilansowanie wag, czyli mierzy niezbilansowanie wag.

Entropie Shannona wyznacza się ze wzoru

$$SE = - \sum_{i=1}^K \hat{w}_n^{(i)} \log_2 \hat{w}_n^{(i)}. \quad (3.64)$$

Minimalna wartość tej miary to 0, w przypadku, gdy cała masa prawdopodobieństwa jest skoncentrowana w jednej cząsteczce. Maksymalną wartość jaką przyjmuje entropia Shannona, to $\log_2 K$, kiedy wszystkie wagi mają identyczną wartość.

Kolejnym istotnym zagadnieniem jest wybór funkcji ważności $q_n(x_n|x_{0:n-1}, y_{1:n}; \theta)$. Najprostszym wyborem jest wybranie jako funkcji ważności gęstości przejścia

$$q_n(x_n|x_{0:n-1}, y_{1:n}; \theta) = p(x_n|x_{0:n-1}, y_{1:n}; \theta). \quad (3.65)$$

Tak dobrana funkcja ważności ze względu na bayesowską interpretację równania (3.7) jest przez wielu autorów (np. Brzozowska-Rup i Dawidowicz (2009)) nazywana funkcją ważności rozładu a priori. Podejście to zostało wykorzystane w pracy wprowadzającej filtr cząsteczkowy (Gordon i in., 1993) i jest nazywane filtrem bootstrapowym⁸ (*bootstrap filter*). W tym wariancie funkcji ważności wzór na wagi ważności (3.59) redukuje się do wzoru

$$w_n^{(i)} = w_{n-1}^{(i)} p(y_n|x_n^{(i)}; \theta). \quad (3.66)$$

Ponadto przy losowaniu ze zwracaniem w każdej iteracji algorytmu wzór (3.66) upraszcza się do postaci

$$w_n^{(i)} = p(y_n|x_n^{(i)}; \theta),$$

ponieważ wagi po losowaniu ze zwracaniem dla wszystkich cząsteczek są równe stałej $1/K$. Uproszczenie wag ma istotne znaczenie nie tylko ze względu na przyspieszenie obliczeń. Pozwala także na zastosowanie filtru cząsteczkowego w przypadkach, gdy nie jest znana postać funkcji przejścia, ale można losować zmienne losowe z rozkładu przejścia. Jak okaże się w dalszej części rozdziału jest to akurat przypadek modeli stochastycznej zmienności skonstruowanych w oparciu o niegaussowskie procesy Ornsteina-Uhlenbecka. Wadą filtru bootstrapowego jest to, że tak wybrana

⁸Nazwa pochodzi od wprowadzonej przez Bradleya Efrona metody *bootstrap* — szacowania rozkładu błędów estymacji, za pomocą wielokrotnego losowania ze zwracaniem z próby (por. Domański i Pruska (2000, roz. 8)).

funkcja ważności jest niezależna od obserwacji zmiennej pomiaru w danej iteracji algorytmu y_n , co powoduje, że filtr bootstrapowy jest bardzo wrażliwy na obserwacje nietypowe.

W literaturze proponowane są także inne funkcje ważności. W większości przypadków są to metody trudne do zaimplementowania w praktycznych zastosowaniach. Przykładowo Liu i Chen (1995) zaproponowali metodę wyboru funkcji ważności minimalizującą warunkową wariancję wag ważności. Metoda ta jednak wymaga losowania z rozkładu warunkowego $p(x_n|x_{n-1}, y_n; \theta)$ i obliczenia wartości funkcji $p(y_n|x_{n-1}; \theta)$ do wyznaczania wartości wag ważności. Jest możliwe tylko w nielicznych przypadkach np. modelu liniowego gaussowskiego (Creal, 2012). Kolejną alternatywą jest pomocniczy filtr cząsteczkowy (*auxiliary particle filter*, APF) zaproponowany przez Pitt i Shephard (1999). W metodzie tej następuje odwrócenie kolejności: najpierw dokonuje się repróbkiwanie według wag wyznaczonych według pewnej zmiennej pomocniczej, a następnie losowanie cząsteczek z reprobkiwanych wartości.

Głównym celem filtrów cząsteczkowych jest wyznaczenie estymatorów stanu zmiennej ukrytej oraz rozkładu warunkowego $p(x_n|y_{1:n}; \theta)$. Wybór estymatora punktowego zależy od przyjętego kryterium optymalności (funkcji straty, por. Lehmann (1991, roz. 1)). W przypadku filtru cząsteczkowego jako estymator zmiennej ukrytej przyjmuje się wartość oczekiwaną rozkładu warunkowego $p(x_n|y_{1:n}; \theta)$

$$\hat{x}_n = \mathbb{E}(X_n|y_{1:n}), \quad (3.67)$$

ponieważ taki wybór jest optymalny ze względu na minimalizację błędu średniokwadratowego (Brzozowska-Rup i Dawidowicz, 2009). Wartość oczekiwaną stanu ukrytego (wzór (3.67)) można przybliżyć za pomocą estymatora

$$\hat{x}_n^K = \sum_{i=1}^K \hat{w}_n^{(i)} x_n^{(i)}. \quad (3.68)$$

Wzór (3.67) jest specjalnym przypadkiem bardziej ogólnego zagadnienia wyznaczania oszacowania

$$\hat{f}(x_n) = \mathbb{E}(f(X_n)|y_{1:n}), \quad (3.69)$$

dla pewnej funkcji f . Pojawia się pytanie o zachowanie się wyznaczonych estymatorów, przy liczbie cząsteczek $K \rightarrow +\infty$. W literaturze można znaleźć bardzo wiele

dowodów zbieżności przy różnych założeniach dotyczących regularności funkcji f , rozkładów ważności, postaci modelu przestrzeni stanów. Historycznie pierwszy dowód zbieżności można znaleźć w pracy Del Moral (1996). W pracy Crisan i Doucet (2002) przedstawiono przegląd twierdzeń dotyczących zbieżności filtrów cząsteczkowych. Przedstawiono m.in. dowód zbieżności średniokwadratowej dla dowolnej ograniczonej borelowskiej funkcji f i ograniczonej funkcji gęstości $p(y_n|x_n; \theta)$.

Filtry cząsteczkowe dostarczają także oszacowań funkcji wiarygodności

$$L(\theta|y_{1:N}) = p(y_1, \dots, y_N; \theta) = p(y_1; \theta) \prod_{n=2}^N p(y_n|y_{n-1}; \theta). \quad (3.70)$$

Zazwyczaj dla modeli przestrzeni stanu funkcję wiarygodności (3.70) nie można wyznaczyć analitycznie, ze względu na to, że wyznaczenie gęstości $p(y_n|y_{1:n-1}; \theta)$ wymaga policzenia wielowymiarowej całki (por. wzór (3.9)). Filtry cząsteczkowe umożliwiając oszacowanie $p(y_n|y_{n-1}; \theta)$ za pomocą

$$\hat{p}(y_1, \dots, y_N; \theta) = \hat{p}(y_1; \theta) \prod_{n=2}^N \hat{p}(y_n|y_{n-1}; \theta), \quad (3.71)$$

gdzie poszczególne oszacowania $\hat{p}(y_n|y_{1:n-1}; \theta)$ wyznaczają się jako sumę nieunormowanych wag dla danej iteracji algorytmu

$$\hat{p}(y_n|y_{1:n-1}; \theta) = \frac{1}{K} \sum_{i=1}^K \frac{p(x_n^{(i)}|y_{1:n}; \theta)}{q_n(x_n^{(i)}|y_{1:n}; \theta)} = \frac{1}{K} \sum_{i=1}^K w_n^{(i)}. \quad (3.72)$$

W konsekwencji otrzymujemy następujące oszacowanie funkcji wiarygodności

$$\hat{L}(\theta|y_{1:N}) = \prod_{n=2}^N \left(\frac{1}{K} \sum_{i=1}^K w_n^{(i)} \right). \quad (3.73)$$

Jest to estymator nieobciążony dla dowolnej wartości K i zgodny przy $K \rightarrow +\infty$ (Kantas i in., 2015). Natomiast estymator logarytmu funkcji wiarygodności postaci

$$\widehat{\ln L}(\theta|y_{1:N}) = \sum_{n=2}^N \ln \hat{p}(y_n|y_{1:n-1}; \theta) = \sum_{n=2}^N \ln \left(\frac{1}{K} \sum_{i=1}^K w_n^{(i)} \right) \quad (3.74)$$

jest obciążony, co wynika z nierówności Jensena. Można jednak użyć skorygowanej (w oparciu o rozwinięcie logarytmu funkcji wiarygodności w szereg Taylora) wersji estymatora postaci

$$\widehat{\ln L}(\theta|y_{1:N}) = \sum_{n=2}^N \ln \left(\sum_{i=1}^K w_n^{(i)} \right) + \frac{1}{2} \frac{\hat{\sigma}^2}{K \exp \left[2 \sum_{n=2}^N \ln \left(\frac{1}{K} \sum_{i=1}^K w_n^{(i)} \right) \right]}, \quad (3.75)$$

gdzie

$$\hat{\sigma}^2 = \frac{1}{K} \sum_{i=1}^K \left(w_n^{(i)} \right)^2 - \left(\frac{1}{K} \sum_{i=1}^K w_n^{(i)} \right)^2.$$

Taką korektę do oszacowania logarytmu funkcji wiarygodności użyto w m.in. Kim i in. (1998), Chib i in. (2002).

Do wygładzania przy pomocy filtru cząsteczkowego najczęściej stosuje się algorytm typu „do przodu - do tyłu” (*forward - backward*). Pierwsza część („do przodu”) to wyznaczenie aproksymacji gęstości filtracji $p(x_n|y_{1:n}; \theta)$ dla $n = 1, \dots, N$ przy pomocy filtru cząsteczkowego. Następnie następuje druga część („do tyłu”), która polega na wyznaczeniu iteracyjnie dla $n = N - 1, \dots, 1$ aproksymacji gęstości wygładzania $p(x_n|y_{1:N}; \theta)$ (wzór 3.10) za pomocą

$$\hat{p}(x_{0:n}|y_{1:N}; \theta) = \sum_{i=1}^K w_{n|N}^{(i)} \delta_{x_n^{(i)}}(x_{0:n}). \quad (3.76)$$

Wagi ważności $w_{n|N}^{(i)}$ wyznacza się ze wzoru

$$w_{n|N}^{(i)} = \hat{w}_n^{(i)} \left[\sum_{j=1}^K w_{n+1|N}^{(j)} \frac{p(x_n^{(j)}|x_{n-1}^{(i)}; \theta)}{\sum_{k=1}^K \hat{w}_n^{(k)} p(x_n^{(j)}|x_{n-1}^{(k)}; \theta)} \right], \quad (3.77)$$

gdzie $(\hat{w}_n^{(i)})_{i=1}^K$ są znormalizowanymi wagami ważności wyznaczonym dla okresu n wyznaczone w części „do przodu”. Ponadto iteracje „do tyłu” inicjuje się przyjmując $w_{N|N}^{(i)} = \hat{w}_N^{(i)}$, dla $i = 1, \dots, K$.

Zaprezentowana metoda wygładzania filtrem cząsteczkowym wymaga znajomości gęstości przejścia $p(x_n|x_{n-1}; \theta)$. Kitagawa (1996) zaproponował alternatywną metodę wygładzania, dla której gęstość ta nie jest wykorzystywana. Propozycja ta wykorzystuje własności klasycznej filtracji z losowaniem ważności z rozkładu a priori (bootstrap). Mając do dyspozycji obserwacje do okresu $n + L$ otrzymujemy oszacowanie gęstości $p(x_{0:n+L}|y_{1:n+L}; \theta)$ za pomocą

$$\hat{p}(x_{0:n}|y_{1:n+L}; \theta) = \sum_{i=1}^K \hat{w}^{(i)} \delta_{x_{0:n}^{(i)}}(x_{0:n}).$$

Zatem oszacowaniem rozkładu brzegowego $p(x_n|y_{1:n+L}; \theta)$ jest

$$\hat{p}(x_n|y_{1:n+L}; \theta) = \sum_{i=1}^K \hat{w}^{(i)} \delta_{x_{0:n}^{(i)}}(x_n). \quad (3.78)$$

W efekcie zamiast wygładzania typu „nieruchomy przedział” (*fixed interval*) jak w przypadku metody „do przodu - do tyłu” otrzymuje wygładzanie typu „nieruchome opóźnienie” (*fixed lag*).

Podójście te ma jednak pewne ograniczenia. Wymaga zapisania, a następnie wykorzystania cząsteczek opóźnionych o L okresów. Pomiedzy okresem n , a $n + L$ następuje L -razy próbkiowanie. Powoduje to istotne zmniejszenie różnorodności trajektorii, co w konsekwencji może powodować słabą aproksymację gęstości $p(x_n|y_{1:n+L}; \theta)$. Problem spadku różnorodności pogłębia się wraz ze wzrostem L , dlatego należy ustalać długość opóźnienia z rozważą. Kitagawa (2014) sugeruje użycie L równego 20 i odradza użycie L większego niż 50.

3.6 Nieliniowe modele przestrzeni dla niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka

Barndorff-Nielsen i Shephard (2001b) w pierwszej swojej pracy poświęconej niegaussowskim procesom Ornsteina-Uhlenbecka zaproponowali użycie filtru cząsteczkowego do oszacowania zmienności aktualnej, gdy parametry procesu są znane. Przedstawili algorytm generowania cząsteczek dla filtru bootstrapowego, w którym zastosowali model przestrzeni stanów 3.1 upraszczając równanie pomiaru do postaci

$$y_n = \sqrt{\sigma_n^2} u_n,$$

gdzie $u_n \sim N(0, 1)$.

Filtry cząsteczkowe nie wymagają, aby model był liniowy lub gaussowski, zatem model przestrzeni 3.1 można łatwo uogólnić na przypadek $\beta \neq 0$ i uwzględnić efekt dźwigni.

Model przestrzeni 3.5.

Dla modelu stochastycznej zmienności BNS z efektem dźwigni (definicja 2.3) model

przestrzeni stanów można zapisać jako

$$\begin{cases} Y_n = h(X_n, u_n) \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.79)$$

Wektor stanu przyjmuje postać

$$X_n = \begin{bmatrix} \lambda\sigma_n^2 \\ \sigma^2(\Delta n) \end{bmatrix}, \quad (3.80)$$

natomiast zmienną pomiaru Y_n jest logarytmiczna stopy zwrotu y_n . Funkcja $h(X_n, u_n)$ przyjmuje postać

$$\begin{aligned} h(X_n, u_n) &= \mu\Delta + \frac{\beta}{\lambda} (\lambda\sigma_n^2) + \sqrt{\frac{1}{\lambda}} (\lambda\sigma_n^2)u_n + \rho\bar{z}_n = \\ &= \mu\Delta + \beta\sigma_n^2 + \sqrt{\sigma_n^2}u_n + \rho\bar{z}_n, \end{aligned} \quad (3.81)$$

gdzie $u_n \sim N(0, 1)$ oraz

$$\bar{z}_n = Z(\lambda n\Delta) - Z(\lambda(n-1)\Delta) - \lambda\Delta\xi = \eta_{1n} - \lambda\Delta\xi.$$

Macierz F jest równa

$$F = \begin{bmatrix} 0 & 1 - e^{-\lambda\Delta} \\ 0 & e^{-\lambda\Delta} \end{bmatrix}. \quad (3.82)$$

Wektor szumu procesu przyjmuje postać

$$v_n = \begin{bmatrix} \eta_{2n} - \eta_{1n} \\ \eta_{1n} \end{bmatrix}. \quad (3.83)$$

Składowe wektora losowego $\eta_n = [\eta_{1n} \ \eta_{2n}]^T$ wyznacza się ze wzorów (2.22) oraz (2.23).

Model przestrzeni stanów 3.5 różni się od modelu 3.1 postacią równania pomiaru zostawiając równanie stanu bez zmian. W odróżnieniu od filtru Kalmana nie trzeba stosować założeń o normalności zaburzeń procesu. Nie jest natomiast znana postać funkcji przejścia $p(x_n|x_{n-1}; \theta)$. Jednak w przypadku stosowania filtru bootstrapowego wystarczy, że można generować próbę losową z gęstości $p(x_n|x_{n-1}; \theta)$, co w

przypadku modelu 3.5 jest możliwe. Zastosowanie filtru cząsteczkowego w odróżnieniu od filtru Kalmana wymaga określenia rozkładu stacjonarnego procesu wariancji chwilowej. W zależności od wyboru tego rozkładu różnią się metody generowania próby losowej z rozkładu $p(x_n|x_{n-1};\theta)$. Zagadnienie to zostanie omówione w kolejnym podrozdziale.

Gęstość rozkładu warunkowego $p(y_n|x_n;\theta)$ przyjmuje w modelu przestrzeni stanów 3.5 zgodnie ze wzorem (2.42) postać

$$Y_n|X_n = x_n \sim N(\mu\Delta + \beta\sigma_n^2 + \rho z_n, \sigma_n^2). \quad (3.84)$$

Na rysunku 3.4 przedstawiono przykład wykorzystania klasycznego filtra cząsteczkowego i wygładzania cząsteczkowego do estymacji wariancji aktualnej. Wariancja aktualna została wyznaczona na podstawie procesu wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,5$ oraz odchyleniu standardowym $\omega = 0,2$ (pozostałe parametry: $\lambda = 0,01, \mu = \beta = 0$).

Również model przestrzeni stanów ze złożeniem 3.2 można uogólnić na przypadek $\beta \neq 0$ oraz z uwzględnieniem efektu dźwigni z oddzielnymi parametrami $\rho_1, \dots, \rho_P$ dla każdego składnika złożenia.

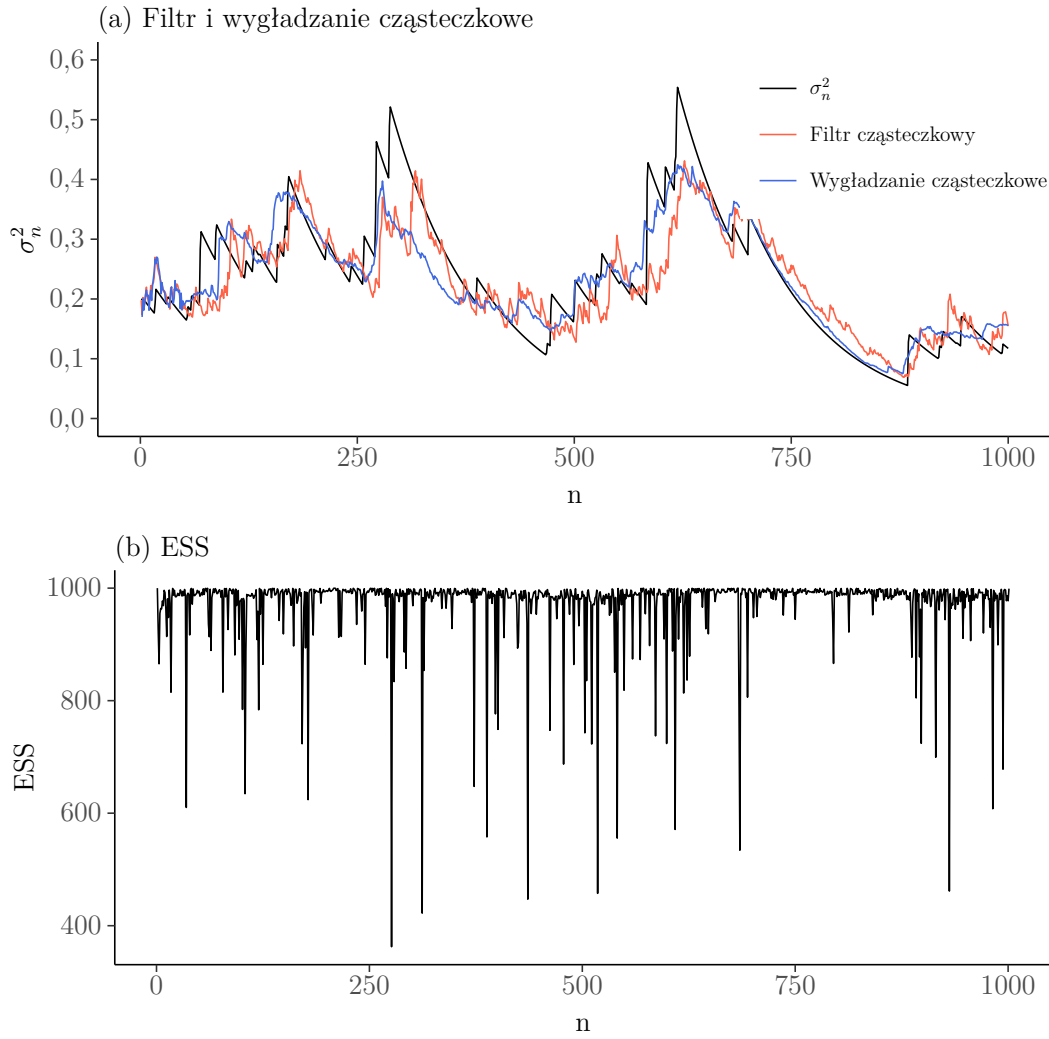
Model przestrzeni 3.6.

Dla modelu stochastycznej zmienności BNS ze złożeniem i efektem dźwigni (definicja (2.4)) model przestrzeni stanów można zapisać jako

$$\begin{cases} Y_n = h(X_n, u_n) \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.85)$$

Wektor stanu przyjmuje postać

$$X_n = \begin{bmatrix} \lambda_1\sigma_{1,n}^2 \\ \sigma_1^2(\Delta n) \\ \lambda_2\sigma_{2,n}^2 \\ \sigma_2^2(\Delta n) \\ \vdots \\ \lambda_P\sigma_{P,n}^2 \\ \sigma_P^2(\Delta n) \end{bmatrix}_{2P \times 1} \quad (3.86)$$



Rysunek 3.4: Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru cząsteczkowego (czerwona linia) i wygładzania cząsteczkowego (linia niebieska) dla $K = 1000$ i $L = 20$ (a) oraz efektywna liczebność próby ESS (b).

Źródło: opracowanie własne.

natomiast procesem pomiaru Y_n są logarytmiczne stopy zwrotu y_n . Funkcja $h(X_n, u_n)$ przyjmuje postać

$$\begin{aligned} h(X_n, u_n) &= \mu\Delta + \beta \sum_{p=1}^P \frac{1}{\lambda_p} (\lambda_p \sigma_{p,n}^2) + \sqrt{\sum_{p=1}^P \frac{1}{\lambda_p} (\lambda_p \sigma_{p,n}^2)} u_n + \sum_{p=1}^P \rho_p \bar{z}_{p,n} = \\ &= \mu\Delta + \beta \sum_{p=1}^P \sigma_{p,n}^2 + \sqrt{\sum_{p=1}^P \sigma_{p,n}^2} u_n + \sum_{p=1}^P \rho_p \bar{z}_{p,n}, \end{aligned} \quad (3.87)$$

gdzie $u_n \sim N(0, 1)$ oraz dla $p = 1, \dots, P$

$$\bar{z}_{p,n} = Z_p(\lambda_p n \Delta) - Z_p(\lambda_p (n-1) \Delta) - w_p \lambda_p \Delta \xi = \eta_{1n}^{(p)} - w_p \lambda_p \Delta \xi.$$

Natomiast macierz $F_{2P \times 2P}$ jest blokowo-diagonalna o blokach

$$F^{(j)} = \begin{bmatrix} 0 & 1 - e^{-\lambda_p \Delta} \\ 0 & e^{-\lambda_p \Delta} \end{bmatrix},$$

dla $p = 1, \dots, P$. Wektor szumu procesu przyjmuje postać

$$\begin{bmatrix} \eta_{2n}^{(1)} - \eta_{1n}^{(1)} \\ \eta_{1n}^{(1)} \\ \eta_{2n}^{(2)} - \eta_{1n}^{(2)} \\ \eta_{1n}^{(2)} \\ \vdots \\ \eta_{2n}^{(P)} - \eta_{1n}^{(P)} \\ \eta_{1n}^{(P)} \end{bmatrix}_{2P \times 1},$$

Składowe wektora losowego $\eta_n^{(p)} = [\eta_{1n}^{(p)} \quad \eta_{2n}^{(p)}]^T$ dla $p = 1, \dots, P$ wyznacza się ze wzorów (2.22) oraz (2.23).

Ponownie choć nie jest znana postać przejścia $p(x_n | x_{n-1}; \theta)$, to można generować próbę losową z tej gęstości, co umożliwia zastosowanie filtru bootstrapowego. Gęstość rozkładu warunkowego $p(y_n | x_n; \theta)$ w modelu przestrzeni stanów 3.6 przyjmuje natomiast postać

$$Y_n | X_n = x_n \sim N \left(\mu \Delta + \beta \sum_{j=1}^P \sigma_{j,n}^2 + \sqrt{\sum_{j=1}^P \sigma_{j,n}^2}, \sum_{j=1}^P \sigma_{j,n}^2 \right). \quad (3.88)$$

Filtr cząsteczkowy można także zastosować w przypadku, gdy zmienną obserwowaną jest wariancja zrealizowana.

Model przestrzeni 3.7.

Dla podstawowego modelu stochastycznej zmienności BNS (definicja (2.1)), model przestrzeni stanów, w którym procesem pomiaru jest wariancja zrealizowana można

zapisać jako

$$\begin{cases} Y_n = h(X_n, u_n) \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.89)$$

Wektor stanu przyjmuje postać

$$X_n = \begin{bmatrix} \lambda\sigma_n^2 \\ \sigma^2(\Delta n) \end{bmatrix}, \quad (3.90)$$

natomiast procesem pomiaru Y_n jest estymator wariancji zrealizowanej $\{y\}_n$. Funkcja $h(X_n, u_n)$ przyjmuje postać

$$h(X_n, u_n) = \frac{1}{\lambda} (\lambda\sigma_n^2) + u_n = \sigma_n^2 + u_n, \quad (3.91)$$

Wektor szumu pomiarowego u_n jest jednowymiarową zmienną losową o rozkładzie normalnym z wartością oczekiwaną 0 i wariancją (por. wzór (2.50))

$$Q = 2M \left(2\omega^2 \lambda^{-2} (e^{-\lambda\Delta/M} - 1 + \lambda\Delta M^{-1}) + (\Delta M^{-1})^2 \xi^2 \right). \quad (3.92)$$

Macierz F jest równa

$$F = \begin{bmatrix} 0 & 1 - e^{-\lambda\Delta} \\ 0 & e^{-\lambda\Delta} \end{bmatrix}. \quad (3.93)$$

Wektor szumu procesu przyjmuje postać

$$v_n = \begin{bmatrix} \eta_{2n} - \eta_{1n} \\ \eta_{1n} \end{bmatrix}. \quad (3.94)$$

Składowe wektora losowego $\eta_n = [\eta_{1n} \ \eta_{2n}]^T$ wyznacza się ze wzorów (2.22) oraz (2.23).

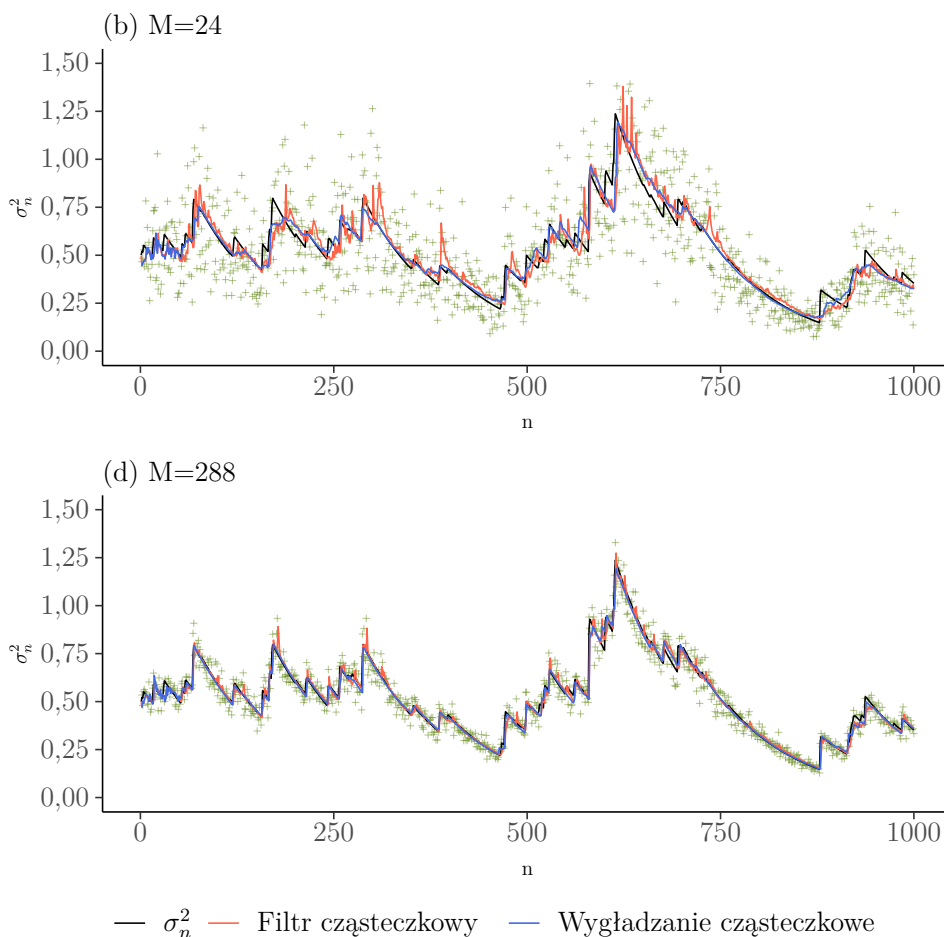
Gęstość rozkładu warunkowego $p(y_n|x_n; \theta)$ przyjmuje natomiast postać

$$Y_n|X_n = x_n \sim N(\sigma_n^2, u_n). \quad (3.95)$$

Należy przy tym pamiętać, że założenie o normalności wynika wyłącznie z rozkładu asymptotycznego, dla $M \rightarrow +\infty$, ale jak pokazują wyniki symulacyjne (Barndorff-Nielsen i Shephard, 2001a)⁹ już dla umiarkowanych wartości M założenie o normalności błędu wariancji zrealizowanej dość dobrze oddaje prawdziwy rozkład. Na

⁹Pokazuje to również przykład przedstawiony na rysunku 2.7.

rysunku 3.5 przedstawiono przykład estymacji wariancji aktualnej za pomocą filtru cząsteczkowego. Wariancja aktualna została wyznaczona na podstawie procesu wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,5$ oraz odchyleniu standardowym $\omega = 0.2$ (pozostałe parametry: $\lambda = 0,01, \mu = \beta = 0$).



Rysunek 3.5: Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru cząsteczkowego (czerwona linia) i wygładzania cząsteczkowego (linia niebieska) dla $M = 12, 288$. Punktami + oznaczono oszacowania wariancji zrealizowanej.

Źródło: opracowanie własne.

Reprezentację przestrzeni stanów 3.7 można uogólnić na przypadek złożenia procesów zmienności.

Model przestrzeni 3.8.

Dla modelu stochastycznej zmienności ze złożeniem procesów zmienności BNS (de-

finicja 2.2), model przestrzeni stanów, w którym procesem pomiaru jest wariancja zrealizowana można zapisać jako

$$\begin{cases} Y_n = h(X_n, u_n) \\ X_n = FX_{n-1} + v_n. \end{cases} \quad (3.96)$$

Wektor stanu przyjmuje postać

$$X_n = \begin{bmatrix} \lambda_1 \sigma_{1,n}^2 \\ \sigma_1^2(\Delta n) \\ \lambda_2 \sigma_{2,n}^2 \\ \sigma_2^2(\Delta n) \\ \vdots \\ \lambda_P \sigma_{P,n}^2 \\ \sigma_P^2(\Delta n) \end{bmatrix}_{2P \times 1}. \quad (3.97)$$

Natomiast zmienną pomiaru Y_n jest estymator wariancji zrealizowanej $\{y\}_n$. Funkcja $h(X_n, u_n)$ przyjmuje postać

$$\begin{aligned} h(X_n, u_n) &= \sum_{p=1}^P \frac{1}{\lambda_p} (\lambda_p \sigma_{p,n}^2) u_n + = \\ &= \sum_{p=1}^P \sigma_{p,n}^2 + u_n. \end{aligned} \quad (3.98)$$

Wektor szumu pomiarowego u_n jest jednowymiarową zmienną losową o rozkładzie normalnym z wartością oczekiwaną 0 i wariancją (por. wzór (2.50))

$$Q = 2M \left(= 2\omega^2 \sum_{p=1}^P \frac{w_p}{2\lambda_p^2} (e^{-\lambda_p \Delta M^{-1}} - 1 + \lambda_p \Delta M^{-1}) + (\Delta M^{-1})^2 \xi^2 \right). \quad (3.99)$$

Macierz $F_{2P \times 2P}$ jest blokowo-diagonalna o blokach

$$F^{(j)} = \begin{bmatrix} 0 & 1 - e^{-\lambda_p \Delta} \\ 0 & e^{-\lambda_p \Delta} \end{bmatrix},$$

dla $p = 1, \dots, P$. Wektor szumu procesu przyjmuje postać

$$\begin{bmatrix} \eta_{2n}^{(1)} - \eta_{1n}^{(1)} \\ \eta_{1n}^{(1)} \\ \eta_{2n}^{(2)} - \eta_{1n}^{(2)} \\ \eta_{1n}^{(2)} \\ \vdots \\ \eta_{2n}^{(P)} - \eta_{1n}^{(P)} \\ \eta_{1n}^{(P)} \end{bmatrix}_{2P \times 1},$$

Składowe wektora losowego $\eta_n^{(p)} = [\eta_{1n}^{(p)} \ \eta_{2n}^{(p)}]^T$ dla $p = 1, \dots, P$ wyznacza się ze wzorów (2.22) oraz (2.23).

3.7 Implementacja filtrów cząsteczkowych

W algorytmach filtrów cząsteczkowych konieczne jest wyznaczanie kolejnych generacji cząsteczek. W modelach przestrzeni stanów dla niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka wiąże się to generowaniem zmiennych losowych o rozkładach wyznaczonych przez dwie całki stochastyczne.

Pierwsza całka stochastyczna związana jest z rekurencją dla procesu wariancji chwilowej

$$\sigma^2(\Delta n) = \sigma^2(\Delta(n-1))e^{-\lambda\Delta} + e^{-\lambda\Delta} \int_0^{\lambda\Delta} e^s dZ(s) = \sigma^2(\Delta(n-1))e^{-\lambda\Delta} + \eta_{1n}. \quad (3.100)$$

Druga całka stochastyczna wynika z rekurencji dla prowadzącego procesu Lévy'ego ukrytego w tle

$$Z(\lambda\Delta n) = Z(\lambda\Delta(n-1)) + \int_0^{\lambda\Delta} dZ(s) = Z(\lambda\Delta(n-1)) + \eta_{2n}. \quad (3.101)$$

W poprzednim podrozdziale wektor losowy $\eta_n = [\eta_{1n} \ \eta_{2n}]^T$ stanowił wektor szumu procesu w przedstawionych modelach przestrzeni stanów. W filtrach cząsteczkowych konieczna jest zatem symulacja zmiennych losowych wyznaczonych przez całki stochastyczne typu

$$\int_0^{\lambda\Delta} f(s) dZ(s) \quad (3.102)$$

dla pewnej ustalonej funkcji f . Pierwsza możliwość to symulacja procesu BDLP a następnie aproksymacja całki schematem Eulera. Podobne podejście zaprezentowali Janicki i Izydorczyk (2001) w kontekście całek względem procesów α -stabilnych. Jest ono trudne ze względu na skokowy charakter procesów BDLP. Druga możliwość to wykorzystanie rozwijania całki (3.102) w nieskończony szereg zmiennych losowych. Metoda ta bazuje na pracach Marcus (1987) oraz Rosinski (1990), a za sprawą pracy Barndorff-Nielsen i Shephard (2001b) zdominowała podejście do symulacji trajektorii procesów stochastycznej zmienności typu Ornsteina-Uhlenbecka.

Rozwijania całki w nieskończony szereg zmiennych losowych

Wprowadźmy oznaczenie na funkcję gęstości ogona rozkładu (*tail mass function*) stacjonarnego procesu wariancji chwilowej

$$W^+(x) = \int_x^\infty w(y)dy. \quad (3.103)$$

Funkcję $W^+(x)$ można wyznaczyć korzystając z gęstości Lévy'ego procesu wariancji chwilowej zgodnie ze wzorem

$$W^+(x) = x u(x) \quad \text{dla } x > 0. \quad (3.104)$$

Oznaczmy przez W^{-1} odwrotność funkcji W^+ , czyli funkcję

$$W^{-1}(x) = \inf \{y > 0 : W^+(y) \leq x\}$$

Wówczas całkę stochastyczną (3.102) można zapisać w postaci

$$\int_0^{\lambda\Delta} f(s)dZ(s) = \sum_{i=1}^{\infty} W^{-1}(a_i/(\lambda\Delta)) f(\lambda\Delta r_i), \quad (3.105)$$

dla niezależnych ciągów zmiennych losowych $(a_i)_{i \in \mathbb{N}}$ oraz $(r_i)_{i \in \mathbb{N}}$, gdzie $a_1 < a_2 < \dots$ jest ciągiem kolejnych zgłoszeń procesu Poissona o intensywności 1, a $(r_i)_{i \in \mathbb{N}}$ jest ciągiem niezależnych zmiennych losowych o rozkładzie równomiernym na przedziale $[0, 1]$.

Jedynym znanym przykładem rozkładem stacjonarnym procesu wariancji chwilowej, dla którego szereg nieskończony (3.105) redukuje się do skończonej ilości wyrazów jest proces Ornsteina-Uhlenbecka o rozkładzie stacjonarnym gamma (proces

Ga-OU). Na podstawie wzoru (1.13) gęstość Lévy'ego przyjmuje postać

$$u(x) = \frac{\alpha}{x} e^{-\beta x} \mathbb{1}_{(0,+\infty)}(x)$$

Zgodnie ze wzorem (3.104) otrzymujemy

$$W_{Ga}^+(x) = \alpha e^{-\beta x} \quad (3.106)$$

oraz

$$W_{Ga}^{-1}(x) = \max \left\{ 0, -\frac{1}{\alpha} \ln \left(\frac{x}{\nu} \right) \right\}. \quad (3.107)$$

Wówczas korzystając ze wzoru (3.105) otrzymujemy

$$\begin{aligned} \int_0^{\lambda\Delta} f(s) dZ(s) &= \sum_{i=1}^{\infty} W_{Ga}^{-1}(a_i/(\lambda\Delta)) f(\lambda\Delta r_i) = \\ &= -\frac{1}{\alpha} \sum_{i=1}^{\infty} \mathbb{1}_{[0,\nu]}(a_i/(\lambda\Delta)) \ln(a_i/(\lambda\Delta\nu)) f(\lambda\Delta r_i) = \\ &= \frac{1}{\alpha} \sum_{i=1}^{\infty} \mathbb{1}_{[0,1]}(c_i) \ln(c_i^{-1}) f(\lambda\Delta r_i) = \\ &= \frac{1}{\alpha} \sum_{i=1}^{N(1)} c_i \ln(c_i^{-1}) f(\lambda\Delta r_i) \end{aligned}$$

gdzie $c_1 < c_2 < \dots < c_{N(1)}$ jest ciągiem kolejnych zgłoszeń procesu Poissona o intensywności $\lambda\Delta\nu$, a $N(1)$ jest liczbą zgłoszeń do czasu 1 w tym ciągu. Zauważmy, że $N(1)$ jest zmienną losową o rozkładzie Poissona również o intensywności $\lambda\Delta\nu$. Zatem w tym przypadku nieskończona suma (3.105) ma skończoną (ale losową) ilość składników. W szczególności ciąg może być pusty, to znaczy pierwsze zgłoszenie c_1 może być większe od 1. Odpowiada to sytuacji, gdy w przedziale czasu $[0, \lambda\Delta]$, proces BDLP jest stały, różniczka $dZ(s)$ jest równa 0, a proces zmienności maleje zgodnie z częścią deterministyczną (por. rysunek 2.1). W dodatku B przedstawiono kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ dla procesu Ga-OU.

Przykładami procesów wariancji chwilowej, dla którego szereg (3.105) nie redukuje się do skończonej ilości wyrazów są rodziny uogólnionych rozkładów odwrotnych Gaussa i temperowanych stabilnych. Wówczas należy wybrać wartość punktu odcięcia $i_{\max}$ w sumie (3.105), od którego kolejne przyrosty w sumie są dostatecznie małe.

W przypadku uogólnionej rodziny rozkładu odwrotnych Gaussa ze wzoru (1.14) oraz wzoru (3.104) wynika, że

$$W_{GIG}^+(x) = \exp(-\gamma^2 x/2) \left(\frac{1}{2} \int_0^{+\infty} \exp(-\frac{1}{2} \delta^{-2} x z) g(z) dz + \max\{0, \nu\} \right), \quad (3.108)$$

gdzie

$$g(x) = \frac{2}{\pi^2 x} \left(J_{|\gamma|}^2(\sqrt{x}) + N_{|\gamma|}^2(\sqrt{2x}) \right)^{-1},$$

gdzie J_ν i N_ν są funkcjami Besella odpowiednio pierwszego i drugiego rodzaju. Funkcję odwrotną W_{GIG}^{-1} nie można podać w postaci analitycznej poza specjalnym przypadkiem rozkładu gamma (wzór (3.107)). Można jednak wyznaczyć ją numerycznie. W dodatku B przedstawiono kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \ \eta_{2n}]^T$ dla procesu GIG-OU. Wybór punktu odcięcia $i_{\max}$ zależy od wyboru podrodziny rozkładów GIG. Wykonane przez autora pracy obliczenia wskazują, że wartość $i_{\max}$ powinna być nie mniejsza niż 100.

W szczególności dla procesu Ornsteina-Uhlenbecka o rozkładzie stacjonarnym odwróconym Gaussa (IG-OU) z parametrami $IG(\delta, \gamma)$ funkcja W^+ przyjmuje postać

$$W_{IG}^+(x) = \frac{\delta x^{-1/2}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \gamma^2 x\right). \quad (3.109)$$

Ze wzoru (3.108) można również wyznaczyć postać funkcji W^+ dla rozkładów stacjonarnych procesu Ornsteina-Uhlenbeck dodatniego hiperbolicznego i odwrotnego gamma.

W przypadku procesu Ornsteina-Uhlenbecka o rozkładzie stacjonarnym pochodzącym z rodziny rozkładów temperowanych stabilnych $TS(\kappa, \nu, \alpha)$ funkcja W^+ przyjmuje postać (Barndorff-Neilsen i Shephard, 2001)

$$W_{TS}^+(x) = \frac{\nu \kappa 2^\kappa}{\Gamma(1 - \kappa)} x^{-\kappa} \exp\left(-\alpha^{1/\kappa} x/2\right). \quad (3.110)$$

Następującą postać W_{TS}^{-1} wyznaczali Gander i Stephens (2007b)

$$W_{TS}^{-1}(x) = \left(\frac{A}{x}\right)^{1/B} \exp\left[-L_W\left(C/B (A/x)^{1/B}\right)\right] \quad (3.111)$$

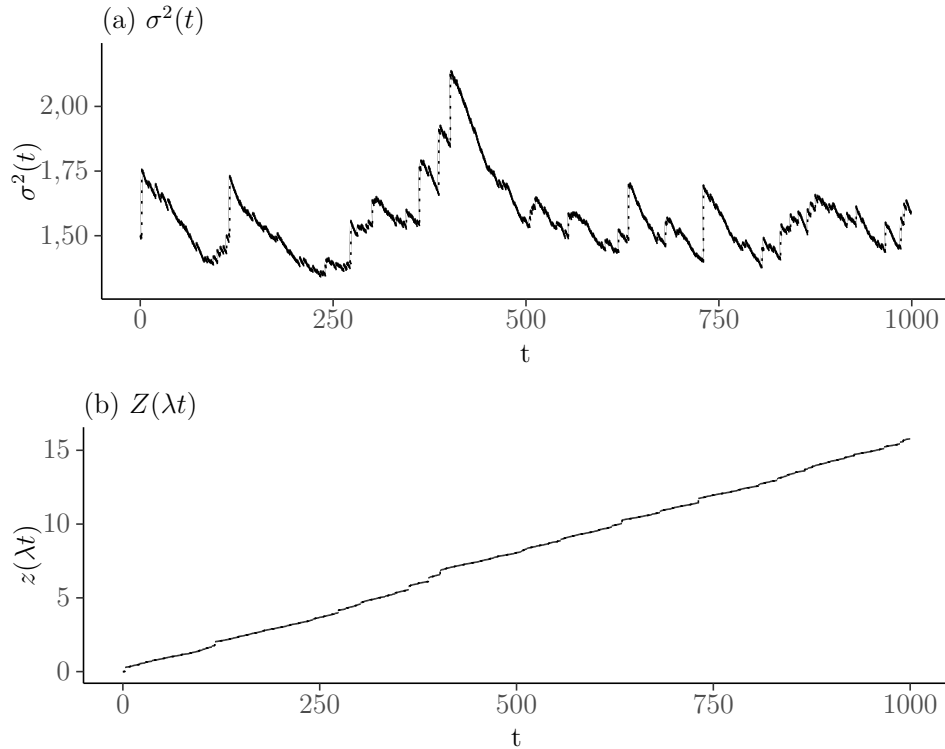
gdzie $A = \nu \kappa 2^\kappa / \Gamma(1 - \kappa)$, $B = \kappa$, $C = \alpha^{1/\kappa} / 2$. Natomiast L_W jest funkcją W Lamberta, to jest funkcją spełniającą warunek $L_W(x) \exp[L_W(x)] = x$.

W szczególności dla rozkładu odwrotnego Gaussa, który jest specjalnym przypadkiem temperowanego rozkładu stabilnego (dla $\kappa = 1/2$) otrzymuje się

$$W_{IG}^{-1}(x) = \frac{1}{\gamma^2} L_W \left(\frac{\delta^2 \gamma^2}{2\pi x^2} \right). \quad (3.112)$$

W przypadku rozkładu stacjonarnego odwrotnego Gaussa bardziej efektywny jest bezpośrednie stosowanie wzoru (3.112) niż numerycznie odwracanie funkcji (3.109).

W dodatku B przedstawiono kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \ \eta_{2n}]^T$. Na rysunku 3.6 przedstawiono symulację trajektorii procesu IG-OU otrzymaną za pomocą wzoru (3.112). Obliczenia autora wskazują, że dla rozkładu odwrotnego Gaussa jako punkt odcięcia można przyjąć $i_{\max} = 50$.



Rysunek 3.6: Trajektoria procesu wariancji chwilowej o rozkładzie stacjonarnym odwrotnym Gaussa z parametrami $\delta = 1$ i $\gamma = 2$ (a) oraz odpowiadający mu proces Lévy’ego ukrytym w tle (BDLP) z parametrem $\lambda = 0,01$ (b).

Źródło: opracowanie własne.

Schematu Eulera

Korzystając z funkcji prostych Valdivieso i in. (2009) uwodnili dla dowolnego procesu podporządkowanego $Z = (Z(\lambda t))_{t \geq 0}$ zbieżność według rozkładu

$$\sum_{j=1}^{\lfloor \frac{\lambda \Delta}{h} \rfloor} e^{jh} Z(jh) \xrightarrow[h \rightarrow 0]{d} \int_0^{\lambda \Delta} e^s dZ(s), \quad (3.113)$$

co uzasadnia użycie schematu Eulera do aproksymowania całki stochastycznej η_{1n} dla dostatecznie małego h . Pozostaje jeszcze kwestia wygenerowania trajektorii procesu BDLP do sumy we wzorze (3.113). Nie jest to zadanie proste, ponieważ procesy te mogą mieć nieskończoną aktywność. W przypadku TS-OU istnieje możliwość wygenerowania procesu BDLP za pomocą sumy $Z = Z^{(1)} + Z^{(2)}$, gdzie $Z^{(1)} = (Z^{(1)}(t))_{t \geq 0}$ jest temperowanym stabilnym procesem $Z^{(1)} = \text{TS}(\kappa, \kappa\nu, \alpha)$, natomiast $Z^{(2)} = (Z^{(2)}(t))_{t \geq 0}$ jest złożonym procesem Poissona postaci

$$Z^{(2)}(t) = \sum_{j=1}^{N(t)} Y_j,$$

gdzie $(N(t))_{t \geq 0}$ jest procesem Poissona o intensywności $\kappa\nu\alpha$, $Y_1, Y_2, \dots$ jest ciągiem niezależnych zmiennych losowych o rozkładzie gamma $\text{Ga}(1 - \kappa, \alpha^{1/\kappa}/2)$. Wówczas

$$\sum_{j=1}^{\lfloor \frac{\lambda \Delta}{h} \rfloor} e^{jh} Z(jh) = \sum_{j=1}^{\lfloor \frac{\lambda \Delta}{h} \rfloor} e^{jh} Z^{(1)}(h) + \sum_{n=1}^{N(\lambda \Delta)} e^{T_j} Y_j, \quad (3.114)$$

gdzie $T_1, T_2, \dots$ kolejnymi zgłoszeniami złożonego procesu Poissona $Z^{(2)}$. Jednak poza przypadkiem procesu IG-OU, dla którego można generować zmienne losowe o rozkładzie odwrotnym Gaussa (odpowiedni algorytm przedstawił Michael i in. (1976)) trudno jest generować zmienne losowe o rozkładzie temperowanym stabilnym, potrzebne do wyznaczenia $Z^{(1)}(h)$.

W przypadku procesów temperowanych stabilnych odpowiednie rozwiązanie zaproponował Rosinski (2001): proces $(J_M(t))_{[0, T]}$ postaci

$$J_M(t) = 2 \sum_{j=1}^M \min \left\{ \left(\frac{\alpha \lambda T \kappa}{b_j \Gamma(1 - \kappa)} \right)^{1/\kappa}, \frac{e_j v_j^{1/\kappa}}{\beta^{1/\kappa}} \right\} e^{\lambda T u_j} \mathbb{1}_{(T u_j \leq t)}, \quad (3.115)$$

zbiega prawie na pewno i jednostajnie do procesu

$$\left(J(t) = \int_0^{\lambda t} e^s dZ^{(1)}(s) \right)_{[0, T]}, \quad (3.116)$$

gdzie $(u_j)_{j \in \mathbb{N}}$, $(v_j)_{j \in \mathbb{N}}$ są ciągami niezależnych zmiennych losowych o rozkładzie równomiernym na przedziale $[0, 1]$, $(e_j)_{j \in \mathbb{N}}$ ciągiem zmiennych losowych o rozkładzie wykładniczym z parametrem 1, $b_1 < b_2 < \dots$ ciągiem zgłoszeń procesu Poissona o intensywności 1. W konsekwencji

$$\int_0^{\lambda \Delta} e^s dZ(s) \approx J_M(\Delta) + \sum_{n=1}^{N(\lambda \Delta)} e^{T_j} Y_j, \quad (3.117)$$

gdzie M jest dostatecznie duże. W praktyce dla rozkładu odwrotnego Gaussa należy przyjąć M równe co najmniej 100.

Symulację trajektorii za pomocą schematu Eulera można również wykorzystać w przypadku procesów OU- $\mathcal{D}$, dla których znany jest proces Lévy'ego ukrytym w tle (BDLP) np. OU-Ga lub OU-IG.

W dodatku B przedstawiono kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ dla IG-OU i OU-Ga przy pomocy schematu Eulera.

3.8 Iterowana filtracja

Podstawowym ograniczeniem filtrów cząsteczkowych jest to, że są w stanie jedynie oszacować ukryty proces stanu przy znanych parametrach. Nie są jednak w stanie oszacować parametrów, od których te procesy zależą. W przypadkach filtrów cząsteczkowych nie można zastosować metod opartych o maksymalizację aproksymacji funkcji wiarygodności mimo że filtry cząsteczkowe dostarczają zgodne i nieobciążone oszacowanie funkcji wiarygodności. Problemem jest to, że oszacowanie dostarczane przez filtry cząsteczkowe nie jest ciągle względem parametrów. Nieciągłość wynika z repróbkiowania. Niewielka zmiana parametrów może spowodować, że skład cząsteczek po repróbkiowaniu stanie się zupełnie inny niż przed zmianą parametrów. Pewnym rozwiązaniem może być zaproponowana przez Malik i Pitt (2011) „ciągła aproksymacja”, ale może być stosowana tylko w przypadku, gdy proces stanu jest jednowymiarowy.

Alternatywne podejście zaproponowali Liu i West (2001). W klasycznym modelu przestrzeni stanu z ustalonym wektorem parametrów θ , rozszerzyli przestrzeń stanów tak, że cząsteczki składają się z par $(x_{0:n}^{(i)}, \theta_n^{(i)})$ i wag $w_n^{(i)}$, dla $i = 1, \dots, K$.

Cząsteczki tworzą dyskretny rozkład, który przybliża rozkład wyznaczony przez gęstość $p(x_{1:n}, \theta | y_{1:n})$. Indeks n przy wektorze θ nie oznacza jednak, że θ jest zmienna w czasie, ale że jest to oszacowanie z okresu n . Korzystając z twierdzenia Bayesa otrzymujemy

$$\begin{aligned} p(x_{1:n}, \theta | y_{1:n}) &\propto p(y_n | x_{1:n}, \theta) p(x_n, \theta | y_{1:n-1}) \\ &\propto p(y_n | x_{1:n}, \theta) p(x_n | \theta, y_{1:n-1}) p(\theta | y_{1:n-1}) \end{aligned} \quad (3.118)$$

Liu i West wykorzystali następujące oszacowanie $p(\theta | y_{1:n-1})$ poprzez mieszaninę rozkładów normalnych

$$\hat{p}(\theta | y_{1:n-1}) = \frac{1}{K} \sum_{i=1}^K f_N(\theta_{n|n-1}^{(i)}, m_{n-1}^{(i)}, h^2 V_{n-1}), \quad (3.119)$$

gdzie $f_N(m, S)$ jest gęstością wielowymiarowego rozkładu normalnego o wektorze wartości oczekiwanej m i macierzy wariancji-kowariancji S . Wektory wartości średniej $\theta_{n|n-1}^{(i)}$ wyznacza się ze wzoru

$$\theta_{n|n-1}^{(i)} = a\theta_{n-1}^{(i)} + (1-a)\bar{\theta}_{n-1}, \quad (3.120)$$

gdzie $\bar{\theta}_{n-1} = 1/K \sum_{i=1}^K \theta_{n|n-1}^{(i)}$, to średnie oszacowanie parametru θ w okresie $n-1$, a współczynnik zawężania (*shrinkage coefficient*), h czynnik skalujący wariancję. Natomiast macierze wariancji kowariancji V_{n-1} wyznacza się ze wzoru

$$V_{n-1} = \frac{1}{K} \sum_{i=1}^K (\theta_{n|n-1}^{(i)} - \bar{\theta}_{n-1}) (\theta_{n|n-1}^{(i)} - \bar{\theta}_{n-1})^T. \quad (3.121)$$

Parametry zawężający a i skalujący h są w filtrze Liu i Westa ściśle powiązane poprzez czynnik dyskontujący $d \in (0, 1]$: $a = (3d - 1)/(2d)$ oraz $h^2 = 1 - a^2$. Liu i West (2001) sugerowali wybierać d z przedziału od 0,95 do 0,99.

Filtr Liu i Westa został zastosowany do estymacji modelu zmienności stochastycznej Taylora (Zanini i in., 2012) oraz z przełączeniami typu Markowa (Carvalho i Lopes, 2007).

Rozwinięciem intuicji zaprezentowanych w filtrze Liu i Westa jest metoda estymacji zwana iterowaną filtracją zaproponowana przez Ionides i in. (2006). W iterowanej filtracji estymator wektora parametrów θ uzyskuje się wykonując sekwencję filtracji przestrzeni stanów poszerzonej o wektor zmiennych w czasie parametrów,

zmniejszając z kroku na krok zmienność parametrów, tak aby w rezultacie otrzymać oszacowanie parametru maksymalizujące logarytm funkcji wiarygodności modelu ze stałymi parametrami.

Metoda iterowanej filtracji ma dwie generacje oznaczane skrótami IF1 i IF2. Pierwsza generacja została zainicjowana w Ionides i in. (2006), a jej teoretyczne uzasadnienie zostało przedstawione w Ionides i in. (2011). Algorytm IF1 w kolejnych iteracjach aproksymuje gradient logarytmu funkcji wiarygodności próby, tak aby kolejne oszacowania zmierzały w kierunku rosnącej wartości logarytmu funkcji wiarygodności. Lindström i in. (2012) zaproponował modyfikację IF1 zwiększającą szybkość zbieżności algorytmu. Natomiast Doucet i in. (2013) przedstawili własną wersję algorytmu o podobnej zasadzie jak IF1, którą nazwali iterowanym wygładzaniem. W odróżnieniu od algorytmu IF1 iterowane wygładzanie aproksymuje nie tylko gradient, ale i hesjan logarytmu funkcji wiarygodności. Metoda ta jest jednak trudna do implementacji i ma głównie znaczenie teoretyczne, a w mniejszym stopniu praktyczne (Nguyen i Ionides, 2017). Drugą generację iterowanej filtracji została wprowadzona przez prace Ionides i in. (2015) i Nguyen (2016). Choć metoda działania jest podobna – algorytm IF2 również dokonuje sekwencji filtracji na rozszerzonej o wektor zmiennych w czasie parametrów przestrzeni stanów, to uzasadnienie teoretyczne jest inne: sekwencje filtracji przybliżają odwzorowanie Bayesa, które jest zbieżne do rozkładu jednopunktowego skupionego w wartości parametru maksymalizującego logarytm funkcji wiarygodności. Przedstawili także na przykładzie silnie nieliniowego modelu przestrzeni stanów, że IF2 daje znacznie dokładniejsze oszacowanie prawdziwej wartości parametru niż IF1. Ponadto Nguyen i Ionides (2017) zaproponowali modyfikację iterowanego wygładzania, którą oznaczyli skrótem IS2. Pokazali na dwóch przykładach, że IS2 osiąga porównywalną dokładność oszacowania parametrów jak IF2.

Metoda iterowanej filtracji została zastosowana w szeregu prac głównie w kontekście modeli ekologicznych i epidemiologicznych: Bhadra i in. (2011), Roy i in. (2013), Blackwood i in. (2013), ale także do estymacji parametrów w modelu stochastycznej zmienności ze stochastycznym parametrem mierzącym efekt dźwigni Bretó (2014). W komentarzu do pracy Andrieu i in. (2010), w którym jako przykład użycia me-

tody PMCMC podano model BNS, Anindya Bhadra (jeden ze współautorów pracy Ionides i in. (2011)) zwrócił uwagę, że alternatywnym rozwiązaniem dla estymacji parametrów modelu BNS na gruncie klasycznego wnioskowania byłoby użycie metody iterowanej filtracji. W pracy Szczepocki (2019a) przedstawiono zastosowanie algorytmu IF2 do podstawowego modelu BNS, w którym zmienną obserwowaną są logarytmiczne stopy zwrotu, a Szczepocki (2019b) w oparciu o wariancję zrealizowaną. W obu pracach wykorzystano rozkład gamma jako rozkład stacjonarny procesu wariancji chwilowej.

W pierwszej generacji iterowanej filtracji oryginalny model przestrzeni stanu jest zastępowany przez niemal identyczny model, w którym stały wektor parametrów $\theta \in \mathbb{R}^r$ jest zastąpiony zmiennym w czasie wektorem θ_n . Gęstości $p(x|x_{n-1}; \theta)$, $p(y|x_n; \theta)$ i $p(x_0; \theta)$ zostają zastąpione przez $p(x|x_{n-1}; \theta_n)$, $p(y|x_n; \theta_n)$ i $p(x_0; \theta_n)$.

Jako proces generującym ewolucję Ionides i in. (2006) przyjęli proces błędzenia losowego w przestrzeni $\mathbb{R}^r$ spełniający następujące warunki:

$$\begin{aligned} \mathbb{E}[\theta_n|\theta_{n-1}] &= \theta_{n-1}, & \mathbb{E}[\theta_0] &= \theta, \\ \text{Var}[\theta_n|\theta_{n-1}] &= \tau^2 \Sigma, & \text{Var}[\theta_0] &= \tau^2 c \Sigma, \end{aligned} \quad (3.122)$$

gdzie c i τ są pewnymi skalarami, Σ dodatnio zorientowaną macierzą (najczęściej diagonalną, gdzie poszczególne elementy odpowiadają za zmienność odpowiednich składowych wektora θ_n). W przypadku, gdy stała τ przyjmuje wartość 0, model o zmiennych parametrach redukuje się do modelu o stałych parametrach.

Algorytm IF1 ma na celu znalezienie oszacowania wektora parametru θ przy $\tau \rightarrow 0$. Procedurę można w skrócie opisać następująco. Ustala się liczbę iteracji J , czynnik dyskontujący d i początkowy współczynnik skalujący wariancję c oraz przyjmuje pewne początkowe oszacowanie $\hat{\theta}_0$ wektora parametrów θ . Następnie dla kolejnych iteracji ($j = 1, \dots, J$) wykonuje się dwa kroki. W pierwszym za pomocą klasycznego filtru cząsteczkowego typu bootstrap wyznacza się oszacowania $\hat{\theta}_n^{(j)} = \hat{\theta}_n(\theta^{(j)}, \tau^{(j)})$ oraz $V_n^{(j)} = V_j(\theta^{(j)}, \tau^{(j)})$ dla $n = 1, \dots, N$, przyjmując $\tau^{(j)} = d^{j-1}$, gdzie

$$\begin{aligned} \hat{\theta}_n &= \hat{\theta}_n(\theta, \tau) = \mathbb{E}[\theta|y_{1:n}], \\ V_n &= V_j(\theta, \tau) = \text{V}[\theta|y_{1:n-1}]. \end{aligned} \quad (3.123)$$

W drugim kroku przyjmujemy:

$$\hat{\theta}^{(j+1)} = \hat{\theta}^{(j)} + V_1^{(j)} \sum_{n=1}^N \left(V_n^{(j)} \right)^{-1} \left(\hat{\theta}_n^{(j)} - \hat{\theta}_{n-1}^{(j)} \right), \quad \hat{\theta}_0^{(j)} = \hat{\theta}^{(j)}.$$

Wówczas wartość $\hat{\theta}^{(J+1)}$ jest szukanym estymatorem największej wiarygodności wyznaczonym metodą iterowanej filtracji (*maximum likelihood via Iterated Filtering*). Zbieżność tak skonstruowanego estymatora przy $J \rightarrow +\infty$ do estymatora największej wiarygodności przedstawia poniższe twierdzenie.

Twierdzenie 3.1. (*Bretó, 2007, str. 22*) *Przy spełnieniu odpowiednich warunków regularności zachodzi warunek*

$$\lim_{\tau \rightarrow 0} \sum_{n=1}^N V_n^{-1} \left(\hat{\theta}_n - \hat{\theta}_{n-1} \right) = \nabla \ln p(y_{1:N} | \theta), \quad (3.124)$$

gdzie ∇g jest gradientem funkcji g oraz $\hat{\theta}_0 = \theta$. Ponadto dla ciągu $(\tau^{(j)})_{j=1}^{\infty}$, takiego że $\tau^{(j)} \rightarrow 0$ określmy rekurencyjnie ciąg $(\hat{\theta}^{(j)})_{j=1}^{\infty}$ za pomocą wzoru

$$\hat{\theta}^{(j+1)} = \hat{\theta}^{(j)} + V_1^{(j)} \sum_{n=1}^N \left(V_n^{(j)} \right)^{-1} \left(\hat{\theta}_n^{(j)} - \hat{\theta}_{n-1}^{(j)} \right), \quad (3.125)$$

gdzie $\hat{\theta}_n^{(j)} = \hat{\theta}_n(\theta^{(j)}, \tau^{(j)})$, $V_n^{(j)} = V_j(\theta^{(j)}, \tau^{(j)})$. Jeżeli istnieje $\hat{\theta}$, takie że $|\hat{\theta} - \hat{\theta}_n^{(j)}|/\tau^{(j)} \rightarrow 0$, to $\nabla \ln p(y_{1:N} | \theta = \hat{\theta}) = 0$.

Twierdzenie 3.1 pokazuje, że dla dostatecznie małego $\tau^{(j)}$ procedura iterowanej filtracji dokonuje rekurencyjnie kolejnych oszacowań θ w kierunku rosnącej wartości logarytmu funkcji wiarygodności z punktem stałym będącym lokalnym maksimum logarytmu funkcji wiarygodności. Algorytm 5 przedstawia kolejne korki postępowanie w algorytmie iterowanej filtracji pierwszej generacji (IF1).

Druga generacja iterowanej filtracji opiera się na pomysłe klonowania danych (*data cloning*) zaproponowanym przez Lele i in. (2007). Jest to metoda, która wykorzystuje algorytm Monte Carlo opartych na łańcuchach Markowa do wnioskowania klasycznego (częstościowego). Autorzy tej pracy zauważyli, że często korzysta się z wnioskowania bayesowskiego wyłącznie z powodów pragmatycznych – klasyczne wnioskowanie jest trudne, wręcz niemożliwe. Problem ten dotyczy szczególnie nieliniowe i niegaussowskie modele przestrzeni stanów, gdzie użycie metod MCMC stało

Algorytm 5. Iterowana filtracja (IF1)

Dane są: liczba iteracji J , czynnik dyskontujący $0 < d < 1$, $b > 0$, początkowa wartość $\hat{\theta}^{(1)}$, macierz kowariancji Σ .

Dla $j = 1, \dots, J$ wykonujemy:

1. W okresie $n = 0$.

(i) Dla $i = 1, \dots, K$:

(a.) Losujemy $\theta_0^{(j,i)}$ z rozkładu $N(\hat{\theta}^{(j)}, cd^{j-1}\Sigma)$

(b.) Losujemy $x_0^{(i)}$ z rozkładu $q_0(x_0; \theta_0^{(j,i)})$

(c.) Przyjmujemy wagi $w_n^{(i)} = 1/K$

2. Dla okresów $n = 1, \dots, N$ wykonujemy:

(i) Dla $i = 1, \dots, K$:

(a.) Losujemy $\theta_n^{(j,i)}$ z rozkładu $N(\hat{\theta}_n^{(j)}, d^{j-1}\Sigma)$

(b.) Losujemy $x_n^{(i)}$ z rozkładu o gęstości $p(x_n | x_{n-1}^{(i)}; \theta_n^{(j,i)})$

(c.) Wyznaczamy wagi $w_n^{(i)} = p(y_n | x_n^{(i)}; \theta_n^{(j,i)})$

(ii) Normalizujemy wagi: $\hat{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{k=1}^K w_n^{(k)}}$

(iii) Losujemy ze zwracaniem K cząsteczek z prawdopodobieństwami $(\hat{w}_n^{(i)})_{i=1}^K$.

(iv) Przyjmujemy $w_n^{(i)} = 1/K$ dla $i = 1, \dots, K$

(v) Wyznaczamy średnią wartość $\bar{\theta}_n^{(j)} = 1/K \sum_{i=1}^K \theta_n^{(j,i)}$

(vi) Wyznaczamy macierz kowariancji $V_n^{(j)} = 1/K \sum_{i=1}^K (\theta_n^{(j,i)} - \bar{\theta}_n^{(j)}) (\theta_n^{(j,i)} - \bar{\theta}_n^{(j)})^T$

3. Przyjmujemy $\hat{\theta}^{(j+1)} = \hat{\theta}^{(j)} + V_1^{(j)} \sum_{n=1}^N (V_n^{(j)})^{-1} (\bar{\theta}_n^{(j)} - \bar{\theta}_{n-1}^{(j)})$

Estymator największej wiarygodności wyznaczonym metodą iterowanej

filtracji $\hat{\theta}^{MIF} = \theta^{J+1}$

się dominujące. Było także widać w przeglądzie literatury na temat estymacji parametrów modelu BNS podanym na początku rozdziału, w której większość pozycji była oparta na algorytmach MCMC.

Wnioskowanie bayesowskie choć atrakcyjne ze względu na możliwość do zastosowania w wielu trudnych problemach wiąże się także z pewnymi problematycznymi kwestiami. Lele i in. (2007) wskazują na cztery trudności. Po pierwsze, wyniki wnioskowania zależą od w dużej mierze subiektywnego wyboru rozkładu a priori. Po dru-

gie, interpretacja i użycie rozkładów nieinformacyjnych ciągle wzbudza kontrowersje¹⁰ (Irony i Singpurwalla, 1997). Po trzecie, bayesowskie przedziały wiarygodności nie mają korzystnej z punktu widzenia nauki interpretacji częstościowej (replikacji eksperymentu), ale raczej reprezentują przekonania badacza co do wartości parametru¹¹ (por. Grzenda (2016, str. 110)). Po czwarte, interpretacja bayesowskich przedziałów wiarygodności w przypadku nieinformacyjnych rozkładów a priori, ze względu na kontrowersje związane z interpretacją tych rozkładów i interpretacją przedziałów wiarygodności jest tym bardziej problematyczna.

Idea metody klonowania danych wynika ze znanego wyniku teoretycznego: wraz ze wzrostem liczebności próby rozkład a posteriori $\pi(\theta|y_{1:N})$ zbiega (przy pewnych założeniach co do regularności) do rozkładu normalnego o wartości oczekiwanej równej estymatorowi największej wiarygodności i macierzy wariancji-kowariancji równej odwrotności drugiej pochodnej logarytmu funkcji wiarygodności wyznaczonej w punkcie wyznaczonym przez estymator największej wiarygodności, bez względu na wybór rozkładu a priori $\pi(\theta)$ (por. Walker (1969)). W praktycznych rozważaniach mamy jednak tylko skończoną ilość obserwacji N . Lele i in. (2007) zaproponowali, aby w takim przypadku replikować (dosłownie klonować) dane tak jakby pochodziły z k niezależnych eksperymentów, dla których wszystkie skończyły się identycznymi wynikami $y_{1:N}$. Funkcja wiarygodności tak powstałej próby powstaje poprzez podniesienie do j -tej potęgi: $p(\theta, y_{1:N})^j$. Wówczas, gdy j rośnie występuje podobny rezultat jak przy wzroście liczebności próby N : rozkład a posteriori $\pi^{(j)}(\theta)$ zbiega do oszacowania największej wiarygodności.

W algorytmie iterowanej filtracji IF2 zaczerpnięte jest także z pracy Lele i in. (2007) pojęcie iterowanego odwzorowania bayesowskiego (*iterated Bayes map*). Oznaczmy przez $\pi^{(1)}(\theta)$ rozkład a posteriori parametru θ dowolnego modelu prze-

¹⁰Rozkłady nieinformacyjne a priori to rozkłady mające minimalny wpływ na rozkład a posteriori (Grzenda, 2016, str. 42).

¹¹Argument ten ma znaczenie w naukach eksperymentalnych. W ekonomii, w której zazwyczaj nie można powtarzać eksperymentów część autorów uważa to za zaletę podejścia bayesowskiego, na przykład Grzenda (2015).

strzeni stanów (definicja 3.1) przed klonowaniem danych. Przyjmuje on postać

$$\pi^{(1)}(\theta|y_{1:N}) = \frac{\left\{ \int p(y_{1:N}|x_{1:N}; \theta) dx_{1:N} \right\} \pi(\theta)}{\int \left\{ \int p(y_{1:N}|x_{1:N}) \pi(\theta) dx_{1:N} d\theta \right\}}. \quad (3.126)$$

Wstawiając do wzoru (3.126) w miejsce rozkładu a priori $\pi(\theta)$ rozkład $\pi^{(1)}(\theta)$ dany wzorem (3.126) otrzymujemy

$$\pi^{(2)}(\theta|y_{1:N}) = \frac{\left\{ \int p(y_{1:N}|x_{1:N}; \theta) dx_{1:N} \right\}^2 \pi(\theta)}{\int \left\{ \int p(y_{1:N}|x_{1:N}) dx_{1:N} \right\}^2 \pi(\theta) d\theta}. \quad (3.127)$$

Jest to rozkład a posteriori dla podwójnego zestawu danych. Kontynuując indukcyjnie postępowanie otrzymujemy rozkład a posteriori $\pi^{(j)}(\theta|y_{1:N})$ dla zestawu j -kopii danych. W ten sposób powstaje iterowane odwzorowanie $T(\pi^{(j)}(\theta|y_{1:N})) = \pi^{(j+1)}(\theta|y_{1:N})$, dla $j = 1, 2, \dots$. Powstaje pytanie czy tak powstałe odwzorowanie ma punkt stały. Lele i in. (2007) wykazali, że jest tak w istocie i jest to rozkład zdegenerowany jednopunktowy w punkcie wyznaczonym przez estymator funkcji największej wiarygodności. Praktyczny wniosek jest taki, że wyznaczając wielokrotnie rozkład a posteriori (na przykład poprzez MCMC) to średnia ze średnich wartości tych rozkładów zbiega do oszacowania wartości największej wiarygodności (Lele i in., 2007).

Kolejne kroki postępowania w metodzie IF2 dla dowolnego modelu przestrzeni stanów (definicja 3.1) przedstawia algorytm 6. Tak jak w przypadku pierwszej generacji iterowanej filtracji wykonywany jest sekwencja J filtracji na rozszerzonej o wektor parametrów przestrzeni stanów. W przypadku, gdy $J = 1$ oraz dla dowolnego $n = 0, \dots, N$ rozkład zaburzeń $h_n(\theta; \varphi, \tau)$ jest zdegenerowany do rozkładu jednopunktowego skoncentrowanego w punkcie φ , to algorytm redukuje się do klasycznego filtru bootstrapowego. Wprowadzone w tym algorytmie zapis $X_{n,k}^{F,j}$ oraz $\Theta^{F,j}$ oznacza składowe cząsteczki wyznaczone na podstawie filtracji, natomiast $X_{n,k}^{P,j}$ oraz $\Theta^{P,j}$ predykcji. Losowanie z rozkładu wielomianowego indeksów $l_{1:K}$ cząsteczek jest równoważne losowaniu ze zwracaniem. Można także wykorzystać inne metody reprób-kowania na przykład losowanie systematyczne. Jako rozkład zaburzeń $h_n(\theta; \varphi, \tau)$ wybiera się zazwyczaj wielowymiarowy rozkład normalny o wartości oczekiwanej φ i macierzy wariancji-kowariancji $\tau\Sigma$, ale może to być dowolny rozkład, dla którego

wariancja będzie proporcjonalna do pewnej stałej τ . Łączny rozkład zaburzeń można zapisać w postaci

$$h(\theta_{0:N}|\varphi; \tau) = h_0(\theta_0|\varphi; \tau) \prod_{n=1}^N h_n(\theta_n|\theta_{n-1}; \tau). \quad (3.128)$$

Każda iteracja IF2 jest aproksymacją Monte Carlo bayesowskiego odwzorowania

$$T_\tau f(\theta_N) = \frac{\int p(\theta_{0:N}; y_{1:N}) h(\theta_{0:N}|\varphi; \tau) f(\varphi) d\varphi d\theta_{0:N-1}}{\int p(\theta_{0:N}; y_{1:N}) h(\theta_{0:N}|\varphi; \tau) f(\varphi) d\varphi d\theta_{0:N}}. \quad (3.129)$$

Funkcja f jest rozkładem początkowego zbioru cząsteczek, natomiast $T_\tau f(\theta_N)$ końcowego zbioru cząsteczek dla ostatniego okresu czasu N . Funkcja f jest odpowiednikiem rozkładu a priori z metody klonowania danych. Kolejne iteracje algorytmu IF2 odpowiadają złożeniom odwzorowania $T_\tau f(\theta_N)$, co natomiast odpowiada zwielokrotnieniu funkcji wiarygodności w metodzie klonowania danych. Jest jednak pewna istotna różnica pomiędzy metodą klonowania danych a IF2: odwzorowanie $T_\tau f(\theta_N)$ zależy od rozkładu zaburzenia $h(\theta_{0:N}|\varphi; \tau)$. Ionides i in. (2015) nazwali algorytm IF2 wnioskowaniem poprzez iterowane zaburzone odwzorowane Bayesa (*inference via iterated, pertubated Bayes map*). Ponadto rozkład zaburzeń w algorytmie IF2 podlega malejącemu ciągowi $(\tau_j)_{j=1}^J$ skalującego wariancję zaburzeń. Nie pozwala to na skorzystanie zaproponowanego przez Lele i in. (2007) granicznego twierdzenia dla metody klonowania danych. Zamiast tego Ionides i in. (2015) przedstawili warunki, dla których zachodzą:

- (A1) dla każdego ustalonego $\tau > 0$, granica $\lim_{m \rightarrow +\infty} T_\tau^m f = f_\tau$ istnieje,
- (A2) wraz ze wzrostem K i J algorytm IF2 numerycznie aproksymuje f_τ ,
- (A3) wraz z malejącą wariancją zaburzeń granica $\lim_{\tau \rightarrow 0} f_\tau$ zbiega do rozkładu jednopunktowego skoncentrowanego w punkcie będącym estymatorem największej wiarygodności.

W celu przedstawienia odpowiednich twierdzeń z pracy Ionides i in. (2015) należy przedstawić konieczne założenia. Niech $(\check{\Theta}_{0:N}^j)_{j=1}^{+\infty}$ będzie łańcuchem Markowa przyjmujące wartości w Θ^{N+1} , takim że $\check{\Theta}_{0:N}^1$ ma gęstość $\int h(\theta_{0:N}|\varphi; \tau) f(\varphi) d\varphi$ oraz $\check{\Theta}_{0:N}^j$ ma rozkład warunkowy $h(\theta_{0:N}|\varphi_N; \tau)$ pod warunkiem $\check{\Theta}_{0:N}^{j-1} = \varphi_{0:N}$ dla

$j \geq 2$. Proces $\left(\check{\Theta}_{0:N}^j\right)_{j=1}^{+\infty}$ jest skonstruowany na przestrzeni probabilistycznej $\Omega_\theta = \{(\theta_{0:N}^1, \theta_{0:N}^2, \dots)\}$, gdzie $\theta_{0:N}^j = \check{\Theta}_{0:N}^j(v)$ dla $v = (\theta_{0:N}^1, \theta_{0:N}^2, \dots) \in \Omega_\theta$. W celu rozważania przeskalowanych czasowo granic procesu $\left(\check{\Theta}_{0:N}^j\right)_{j=1}^{+\infty}$ przy $\tau \rightarrow 0$ oznaczmy przez $(W_\tau(t))_{0 \leq t \leq 1}$ proces prawostronnie ciągły z lewostronnymi granicami, przedziałami stały, który dla punktów nieciągłości przyjmuje wartość $W_\tau(j\tau^2) = \check{\Theta}_{0:N}^{j+1}$, gdzie $j = 0, 1, 2, \dots$.

Ionides i in. (2015) przyjęli następujące założenia.

Algorytm 6. Iterowana filtracja (IF2)

Dane są:

Liczba obserwacji: N , liczba cząsteczek: K , liczba iteracji: J

Zbiór K początkowych oszacowań wektora parametrów θ : $\{\Theta_k^0$ dla $k = 1, \dots, K\}$

Rozkład zaburzeń $h_n(\theta; \varphi, \tau)$, dla $n = 1, \dots, N$

Ciąg schładzający: τ_j , dla $j = 1, \dots, J$

Dla $j = 1, \dots, J$ wykonujemy:

1. W okresie $n = 0$.

(i) Losujemy $\Theta_{0,k}^{F,j}$ z rozkładu $h_0(\theta | \Theta_k^{j-1}; \tau_j)$ dla $i = 1, \dots, K$

(ii) Losujemy $X_{0,k}^{F,j}$ z rozkładu $p(x_0; \Theta_{0,k}^{F,j})$ dla $i = 1, \dots, K$

2. Dla okresów $n = 1, \dots, N$ wykonujemy:

(i) Losujemy $\Theta_{n,k}^{P,j}$ z rozkładu $h_n(\theta | \Theta_{n-1,k}^{F,j}; \tau_j)$ dla $i = 1, \dots, K$

(ii) Losujemy $X_{n,k}^{P,j}$ z rozkładu o gęstości $p(x_n | X_{n-1,k}^{F,j}; \Theta_{n,k}^{P,j})$ dla $i = 1, \dots, K$

(iii) Wyznaczamy wagi $w_{n,k}^j = p(y_n | X_{n,k}^{P,j}; \Theta_{n,k}^{P,j})$ dla $i = 1, \dots, K$

(iv) Normalizujemy wagi: $\hat{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{k=1}^K w_n^{(k)}}$ dla $i = 1, \dots, K$

(v) Losujemy ciąg $l_{1:K}$ z rozkładu wielomianowego z K zdarzeniami

oraz prawdopodobieństwami wyznaczonymi przez znormalizowane wagi.

(vi) Przyjmujemy: $\Theta_{n,k}^{F,j} = \Theta_{n,l_k}^{P,j}$ oraz $X_{n,k}^{F,j} = X_{n,l_k}^{P,j}$ dla $i = 1, \dots, K$

3. Przyjmujemy $\Theta_{N,k}^j$ dla $i = 1, \dots, K$

Wynik algorytmu:

zbiór K końcowych oszacowań wektora parametrów θ : $\{\Theta_k^J$ dla $k = 1, \dots, K\}$.

- (B1) $(W_\tau(t))_{0 \leq t \leq 1}$ zbiega słabo $\tau \rightarrow 0$ do procesu dyfuzyjnego $(W(t))_{0 \leq t \leq 1}$ w przestrzeni procesów prawostronnie ciągłych z lewostronnymi granicami $\mathbb{D}$ z topologią jednostajnej zbieżności. Dla każdego otwartego $A \subset \Theta$ z dodatnią miarą Lebesguea i $\epsilon > 0$, istnieje $\delta(A, \epsilon) > 0$ takie że $P(W(t) \in A \text{ dla każdego } \epsilon \leq t \leq 1 | W(0)) > \delta$.
- (B2) Dla każdego t_0 i $\tau_0 > 0$, $W_\tau(t)$ ma dodatni rozkład na Θ jednostajnie powyżej $W(0)$ dla $t > t_0$ i $\tau < \tau_0$.
- (B3) Funkcja $p(\theta | y_{1:N})$ jest ciągła w otoczeniu $\{\theta : p(\theta | y_{1:N}) > \lambda_1\}$ dla pewnego $\lambda_1 < \sup_{\varphi} p(\varphi | y_{1:N})$.
- (B4) Istnieje $c > 0$ z $c^{-1} > p(y_n | x_n \theta; \cdot) > c$ dla każdego $1 \leq n \leq N$ i $\theta \in \Theta$.
- (B5) Istnieje stała C_1 taka że $h_n(\theta | \varphi; \tau) = 0$, gdy $|\theta - \varphi| > C_1 \tau$, dla dowolnego τ .
- (B6) Istnieje stała C_2 taka że $\sup_{1 \leq n \leq N} |\theta_n - \theta_{n-1}| < C_1 \tau$ implikuje $|p(\theta_{0:N} | y_{1:N}) - p(\theta_N | y_{1:N})| < C_2 \tau$ dla każdego τ i dla każdego n .

Ionides i in. (2015) podali również warunki, dla których sześć powyższych założeń może być spełnionych. Warunki (B1) i (B2) są spełnione, gdy proces zaburzeń $h_n(\theta | \varphi; \tau)$ dla $n \in 0, \dots, N$ jest gaussowskim błędzeniem losowym we wnętrzu pewnego prostokątnego obszaru w $\mathbb{R}^{\dim(\Theta)}$ odbijającym się od brzegu tego zbioru (*reflected Gaussian random walk*). Wówczas $(W(t))_{0 \leq t \leq 1}$ jest również gaussowskim błędzeniem losowym. Ponadto, gdy $h_n(\theta | \varphi; \tau)$ dla $n \in 0, \dots, N$ należy do rodziny rozkładów z parametrem skali i położenia z wartością oczekiwaną φ z dala od brzegu, to $(W(t))_{0 \leq t \leq 1}$ będzie zachowywał się jak gaussowskie błędzenie losowe we wnętrzu Θ . Warunek (B5) może być spełniony przez konstrukcje rozkładu zaburzeń. Warunki (B3) i (B6) są mało wymagającymi warunkami dla odpowiednio funkcji wiarygodności próby w przypadku zwykłej i rozszerzonej przestrzeni stanów.

Poniższe twierdzenie szczegółowo opisuje spełnienie warunków (A1) i (A2) dla algorytmu iterowanej filtracji drugiej generacji (IF2).

Twierdzenie 3.1.

(Ionides i in., 2015) Niech T_τ będzie odwzorowaniem danym wzorem 3.129. Załóżmy,

że spełnione są założenia (B2) i (B4). Wówczas istnieje unikalny rozkład gęstości f_τ na przestrzeni parametrów Θ , taki że dla każdej gęstości rozkładu f na Θ zachodzi

$$\lim_{j \rightarrow +\infty} \|T_\tau^j f - f_\tau\|_1 = 0, \quad (3.130)$$

gdzie $\|f\|_1$ jest norma w przestrzeni funkcji całkowalnych L^1 .

Niech $\{\Theta_k^J$ dla $k = 1, \dots, K\}$ będzie wynikiem algorytmu IF2, dla którego $\tau_j = \tau > 0$. Istnieje stała $C > 0$ taka że dla każdej funkcji $\phi : \Theta \rightarrow \mathbb{R}$ i dla każdego J zachodzi

$$\mathbb{E} \left\{ \left| \frac{1}{K} \sum_{k=1}^K \phi(\Theta_k^J) - \int \phi(\theta) f_\tau(\theta) d\theta \right| \right\} \leq \frac{C \sup_{\theta} |\phi(\theta)|}{\sqrt{K}}. \quad (3.131)$$

Warunki konieczne do spełnienia (A3) formalizuje następujące twierdzenie.

Twierdzenie 3.2.

(Ionides i in., 2015) Załóżmy, że zachodzą warunki (B1)-(B6). Dla $\lambda_2 < \sup_{\theta} \ell(\theta)$ zachodzi

$$\lim_{\tau \rightarrow 0} \int f_\tau(\theta) \mathbb{1}_{\{\ell(\theta) < \lambda_2\}}(\theta) d\theta = 0. \quad (3.132)$$

Metoda iterowanej filtracji, podobnie jak większość metod numerycznych nie gwarantuje znalezienia globalnego maksimum w przypadku, gdy logarytm funkcji wiarygodności nie jest funkcją wypukłą. Algorytmy iterowanej filtracji umożliwiają zlokalizować jedynie lokalne minimum. Twierdzenie 3.1 dla IF1 oraz twierdzenia 3.1 i 3.2 dla IF2 także nie mówią nic o szybkości zbieżności do maksimum funkcji wiarygodności. Ponadto założenia (B1)-(B6) dla IF2 są trudne do zweryfikowania w praktycznych zastosowaniach. Powoduje to konieczność weryfikowania otrzymanych wyników w celu upewnienia się, że zbieżność została osiągnięta, a otrzymane maksimum jest globalne.

Korzystając z algorytmu iterowanej filtracji IF2 należy wybrać szereg ustawień, które mają wpływ na wyniki algorytmu (porównaj początek algorytmu 6). Po pierwsze należy wybrać liczbę cząsteczek M oraz liczbę iteracji J . Zazwyczaj jako liczbę cząsteczek można wybrać wartość podobną do wartości, która pozwala efektywnie w danej przestrzeni stanów dokonywać oszacowań zmiennej stanu za pomocą klasycznego filtru bootstrapowego, gdy znane są parametry Ionides (2019). Należy zwrócić uwagę, że algorytmy iterowanej filtracji wykonują filtrację na rozszerzonej o wektor

parametrów przestrzeni stanów. Efektywność filtrów cząsteczkowych spada wraz ze wzrostem wymiaru przestrzeni stanów (Vaswani, 2008). Z drugiej strony dodatkowa zmienność zmniejsza problem degeneracji różnorodności próby charakterystyczny dla filtrów cząsteczkowych, ponieważ każda cząsteczka ma wówczas większą szansę na trafienie na prawdziwą trajektorię danych (Ionides i in., 2015). Nie jest więc jasne co może w danej sytuacji przeważać czy spadek efektywności związany ze wzrostem wymiaru, czy korzyść związana z dodatkowej zmienności. Pewną wskazówką mogą być wyniki przedstawione w pracy Ionides i in. (2015), w której autorzy porównali efektywną wielkość próby ESS w filtrze cząsteczkowym z i bez dodatkowych zaburzeń parametrów dla modelu epidemiologicznego rozważanego w tamtej pracy. Okazało się, że w przypadku, gdy wektor parametrów odpowiadał wartości estymatora największej wiarygodności to wartość ESS spadała średnio o 5% w przypadku dodatkowych zaburzeń na przestrzeni parametrów. Natomiast, gdy wektor parametrów różnił się znacząco od estymatora największej wiarygodności to wartość ESS rosła średnio o 13% w przypadku dodatkowych zaburzeń na przestrzeni parametrów. Widać zatem, że dodatkowa zmienność rzeczywiście pozwala znacząco zmniejszyć degenerację różnorodności próby wtedy, gdy w przeszukiwaniu przestrzeni parametrów algorytm iterowanej filtracji jest jeszcze daleko od wartości maksymalizującej logarytm funkcji wiarygodności, a cząsteczki nie trafiają w regiony prawdziwej trajektorii wyznaczonej przez dane. W przypadku, gdy algorytm zbliża się do wartości estymatora największej wiarygodności dodatkowa zmienność staje się co raz mniej potrzebna, a w przypadku osiągnięcia estymatora największej wiarygodności powoduje nieznaczny spadek efektywnej liczebności próby. Wraz z kolejnymi iteracjami maleje jednak wariancja perturbacji kontrolowana przez ciąg schładzający $(\tau_j)_{j=1}^J$. Co pozwala wykorzystać zalety zaburzeń w początkowych iteracjach i zminimalizować negatywny efekt w pobliżu maksimum.

Jako ciąg schładzający (*cooling schedule*) zmniejszający wariancję zaburzeń Ionides i in. (2015) proponują stosować albo ciąg geometryczny ($\tau_j = \tau_1 q^j$, gdzie $\tau_1 > 0$, $q \in (0, 1)$) albo ciąg o postaci funkcji hiperbolicznej ($\tau_j = 1/(1 + \alpha j)$, gdzie $\alpha > 1$). W pracy Ionides i in. (2015) w przykładach zastosowano ciągi geometryczne o parametrach $\tau_1 = 0, 1$ i $\tau_1 = 0, 2$ oraz $q = 10^{-1/99}$. Wartość q została tak dobrana, aby dla

ostatniej, setnej iteracji otrzymać odpowiednio wartość $\tau_{100} = 0, 1$ oraz $\tau_{100} = 0, 2$.

Liczbę iteracji J w algorytmie iterowanej filtracji należy dobierać metodą prób i błędów obserwując wykresy diagnostyczne (zostaną przedstawione w dalszej części rozdziału). Na wybór J wpływ ma także dobór ciągu schładzającego. Gdy wartość ciągu schładzającego $(\tau_j)_{j=1}^J$ jest bliska zera dalsze zwiększanie iteracji nie ma sensu, ponieważ nie spowodują istotnych przesunięć w przestrzeni parametrów. Ionides i in. (2015) w prezentowanych przykładach przyjmował $J = 100$ z opisanymi powyżej ciągami schładzającymi, natomiast Bretó (2014) stosując algorytm IF1 przyjął $J = 150$.

Jako rozkład zaburzeń stosuje się najczęściej wielowymiarowy rozkład normalny (Ionides i in., 2015)

$$h_n(\theta; \varphi, \tau_j) \sim N(\varphi, \tau_j \mathbb{V}_n). \quad (3.133)$$

Zazwyczaj macierz $\mathbb{V}_n$ jest diagonalna o początkowych wartościach zaburzeń dla poszczególnych parametrów na głównej przekątnej. Wówczas jako τ_1 można przyjąć 1.

Ostatnim ustawieniem algorytmu iterowanie filtracji IF2 jest wybór zbioru K początkowych oszacowań wektora parametrów θ : $\{\Theta_k^0$ dla $k = 1, \dots, K\}$. Zazwyczaj przyjmuje się K replikacji tej samej ustalonej wartości startowej Ionides (2019).

Podstawowym narzędziem diagnostycznym są wykresy, które dla kolejnych kroków iteracji $j = 1, \dots, J$ przedstawiają składowe wektora

$$\bar{\Theta}^j = \frac{1}{K} \sum_{k=1}^K \Theta_{N,k}^j. \quad (3.134)$$

Zbieżność do maksimum globalnego wskazuje sytuacja, gdy trajektorie składowych wektora $\bar{\Theta}^j$ dla różnych początkowych wartości zbiegają do jednoznacznie wyznaczonych granic. Na rysunku 3.7 (a) przedstawiono wykresy 10 trajektorii algorytmu IF2 dla podstawowego modelu BNS o losowo wybranych wartościach początkowych. Widać, że poszczególne składowe wraz ze wzrostem iteracji choć nie zbiegają do wspólnej granicy, to jednak zbiegają do wyraźnie zaznaczonych obszarów.

W opublikowanej pracy doktorskiej Bretó (2007) jednego ze współautorów pracy Ionides i in. (2006) została przedstawiona następująca procedura przeszukiwania przestrzeni parametrów w poszukiwaniu globalnego maksimum funkcji wiarygodności:

- Krok 1 Wybieramy S początkowych wartości (na przykład poprzez losowanie każdej składowej wektora θ z pewnego akceptowalnego przedziału wartości) i wyznaczamy metodą iterowanej filtracji S par $(\hat{\theta}_s, \hat{\ell}_s)$ oszacowań wektora parametru θ i wartości logarytmu funkcji wiarygodności.
- Krok 2 Jeśli jest wyraźne maksimum globalne, to znaczy jest wiele par $(\hat{\theta}_s, \hat{\ell}_s)$ takich, że $\left(\max_{1 \leq r \leq S} \hat{\ell}_r - \hat{\ell}_s\right)$ jest małe oraz $\|\hat{\theta}_{\underset{1 \leq r \leq S}{\operatorname{argmax}} \hat{\ell}_r} - \hat{\theta}_s\|$ jest również małe, należy wybrać jako oszacowanie wektora parametrów θ średnią z oszacowań bliskich globalnego maksimum¹².
- Krok 3 Jeśli maksimum globalne nie jest wyraźne, to znaczy jest wiele par $(\hat{\theta}_s, \hat{\ell}_s)$ takich, że $\left(\max_{1 \leq r \leq S} \hat{\ell}_r - \hat{\ell}_s\right)$ jest małe, natomiast $\|\hat{\theta}_{\underset{1 \leq r \leq S}{\operatorname{argmax}} \hat{\ell}_r} - \hat{\theta}_s\|$ jest duże, to niektóre kombinacje parametrów są słabo identyfikowalne. Prawdopodobnie należy uwzględnić dodatkowe założenia w celu poprawienia identyfikowalności modelu i powrócić do kroku 1.

Należy określić jednak co oznacza „małe” w kroku 2 i 3. W przypadku logarytmu funkcji wiarygodności według Bretó (2007) za „małą” można uznać różnicę rzędu jednej jednostki logarytmu funkcji wiarygodności. Trudniej jednak wskazać co można uznać za „małą” różnicę w przypadku wektora parametrów. Na rysunku 3.7 wszystkie dziesięć oszacowań logarytmu funkcji wiarygodności wyznaczonych na podstawie wartości parametrów z ostatniej iteracji IF2 różnią się co do wartości bezwzględnej o mniej niż jedną jednostkę logarytmu funkcji wiarygodności.

W kroku 2 Bretó (2007) preferuje wyznaczyć średnią z oszacowań wektora parametru θ bliskich globalnego maksimum zamiast jedną wartość o największej wartości oszacowania logarytmu największej wiarygodności, ponieważ błąd oszacowania logarytmu największej wiarygodności jest zazwyczaj większy niż różnice pomiędzy rzeczywistymi wartościami funkcji wiarygodności w pobliżu globalnego ekstremum. Inne podejście do wyboru oszacowania parametru θ wartości preferuje Ionides (2019) inny ze współautorów pracy Ionides i in. (2006). Zaleca by dla każdego oszacowania $\hat{\theta}_s$ powtórzyć wielokrotnie oszacowanie funkcji wiarygodności za pomocą filtru cząsteczkowego i wyznaczyć średnie oszacowanie, które następnie zostaje poddane

¹²Symbol $\|\cdot\|$ oznacza w tym przypadku dowolną normę w przestrzeni $\mathbb{R}^{\dim(\theta)}$.

transformacji logarytmicznej i staje się oszacowaniem logarytmu funkcji wiarygodności. Jako oszacowanie parametru θ w drugim kroku wybiera się zatem to oszacowanie $\hat{\theta}_s$, dla którego $\hat{\ell}_s$ wyznaczone jako logarytm średnich oszacowań funkcji wiarygodności jest największe¹³.

W kroku trzecim można również posłużyć się wykresami korelacji pomiędzy poszczególnymi składowymi $\hat{\theta}_s$ w próbie złożonej z S oszacowań parametru θ . Silne korelacje pomiędzy składowymi oszacowań wektora parametru modelu są wskazówką, że parametry modelu są słabo identyfikowalne (Bretó, 2007).

Kolejną możliwością diagnostyki jest tworzenie wykresów jednowymiarowych cięć funkcji wiarygodności względem poszczególnych składowych wektora θ postaci $\ell(\hat{\theta} + h\mathbf{e}_i)$ w pewnym otoczeniu wektora $\hat{\theta}$, gdzie $\mathbf{e}_i$ jest wektorem mającym same zera oprócz i -tej składowej dla której przyjmuje wartość 1. Jeśli poszczególne składowe $\hat{\theta}$ są maksymami lokalnym wszystkich jednowymiarowych cięć funkcji wiarygodności, to $\hat{\theta}$ jest również maksimum lokalnym $\ell(\hat{\theta})$ jeśli tylko funkcja ℓ jest ciągle różniczkowalna w pewnym otoczeniu $\hat{\theta}$. Jednakże w modelach przestrzeni stanów zazwyczaj nie dysponujemy postacią analityczną funkcji wiarygodności ℓ , więc cięcia funkcji wiarygodności należy wyznaczać korzystając z oszacowań funkcji wiarygodności wyznaczonych za pomocą filtru cząsteczkowego, co wiąże się z koniecznością uwzględnienia błędu Monte Carlo oszacowania logarytmu funkcji wiarygodności. Z powodu błędu Monte Carlo oszacowania $\hat{\ell}(\hat{\theta} + h\mathbf{e}_i)$ nie są ciągle względem zmiennej h i należy poprowadzić dopasowanie lokalnie sklejanymi wielomianami drugiego stopnia (Ionides, 2005). Na rysunku 3.7 (b) przedstawiono jednowymiarowe cięcia logarytmu funkcji wiarygodności dla podstawowego modelu BNS. Przyjęto wygładzanie lokalnie sklejanymi wielomianami drugiego stopnia z parametrem wygładzania 0,5.

Jednowymiarowe cięcia funkcji wiarygodności można także wykorzystać do wyznaczenia błędu standardowego oszacowania $\hat{\theta}$. Błąd standardowy i -tej składowej wyznacza się ze wzoru: $\text{SE}(\hat{\theta}_i) = [\mathcal{I}^{-1/2}]_{ii}$, gdzie $\mathcal{I}$ jest macierzą informacyjną Fi-

¹³Filtry cząsteczkowe dostarczają nieobciążonego oszacowania funkcji wiarygodności, ale oszacowanie logarytmu jest obciążone. Dlatego należy uśredniać oszacowania funkcji wiarygodności, a nie logarytmu funkcji wiarygodności.

shera, którą można oszacować za pomocą

$$\hat{\mathcal{I}}_{i,j} = \sum_{n=2}^N \left[\frac{\partial}{\partial \theta_i} \ln p(y_n | y_{1:n-1}, \hat{\theta}) \right] \left[\frac{\partial}{\partial \theta_j} \ln p(y_n | y_{1:n-1}, \hat{\theta}) \right]^T. \quad (3.135)$$

Bretó (2007) podał procedurę jak można wyznaczyć numerycznie oszacowanie macierzy informacyjnej Fishera:

Krok 1 Obliczamy $\ell_{t,i,j} = \ln p(y_n | y_{1:n-1}, \hat{\theta} + h_{ij} \mathbf{e}_i)$ dla $1 \leq i \leq \dim(\theta)$ oraz $1 \leq j \leq q$, gdzie $\mathbf{e}_i$ jest wektorem mającym same zera oprócz i -tej składowej dla której przyjmuje wartość 1, q dolną liczbą całkowitą dodatnią.

Krok 2 Dla każdego n wyznaczamy regresję liniową $\ell_{t,i,j}$ względem h_{ij} dla każdego i otrzymując współczynnik regresji $\dot{\ell}_{t,i}$ z oszacowaniem wariancji współczynnika $\hat{\text{Var}}(\dot{\ell}_{t,i})$.

Krok 3 Wyznaczamy oszacowanie macierzy informacyjnej Fishera $\hat{\mathcal{I}}_{ij} = \sum_{n=2}^N \dot{\ell}_{t,i} \dot{\ell}_{t,j}^T$.

W kroku 2 powyższej procedury wyznaczone są numeryczne aproksymacje pochodnych cząstkowych względem pewnego otoczenia h_{ij} . Jeśli otoczenie h_{ij} będzie za małe błąd Monte Carlo może zdominować wyznaczoną wartość oszacowania pochodnej. Błąd oszacowania $\hat{\mathcal{I}}_{ii}$ można przybliżyć poprzez $\sum_{n=2}^N \hat{\text{Var}}(\dot{\ell}_{t,i})$ (Bretó, 2007). Dlatego jeśli wartość $\hat{\mathcal{I}}_{ii}$ nie jest znacznie większa niż $\sum_{n=2}^N \hat{\text{Var}}(\dot{\ell}_{t,i})$, to należy zwiększyć otoczenie h_{ij} poprzez zwiększenie w kroku 1 wartości stałej q .

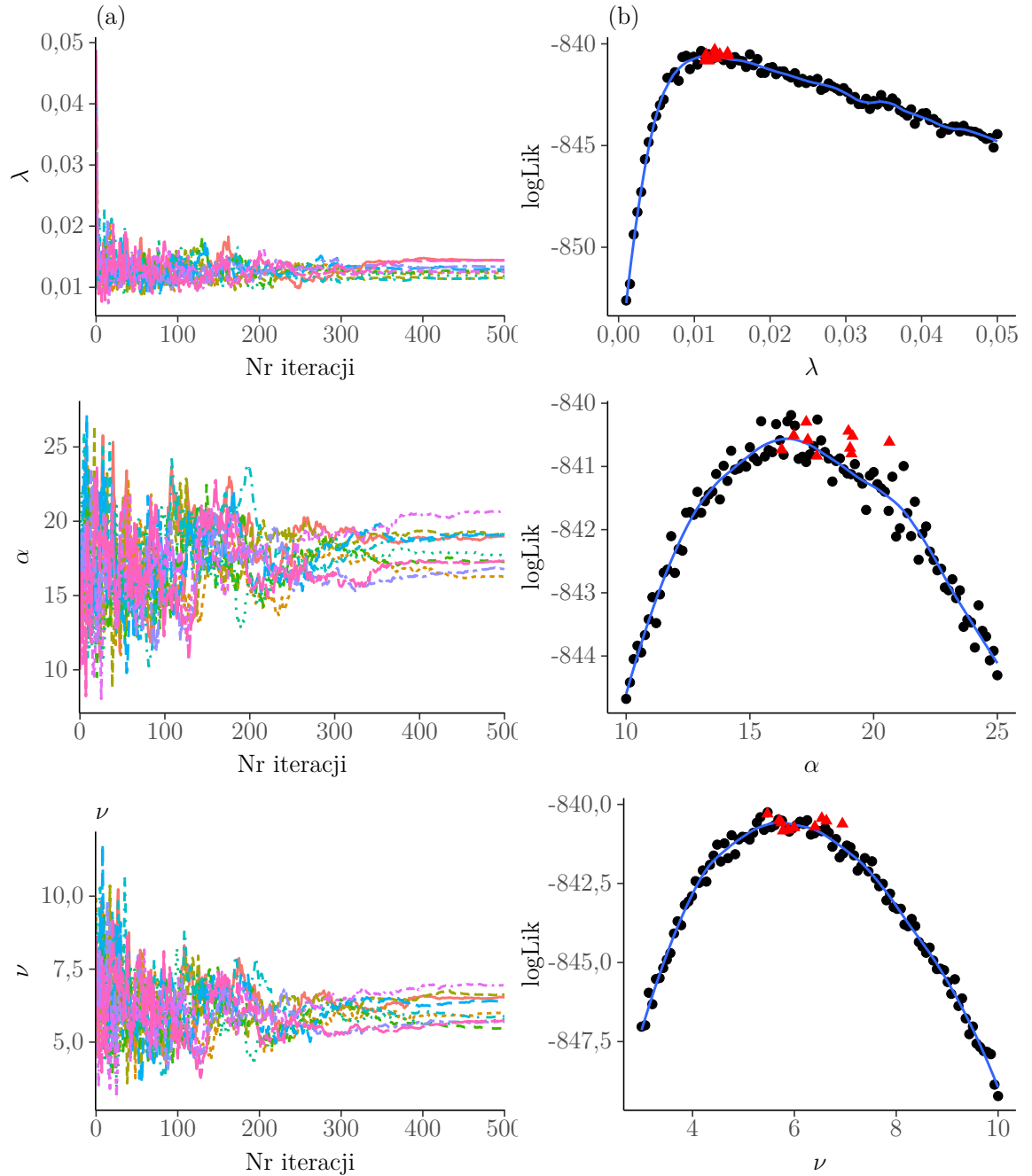
Błędy standardowe można wykorzystać do wyznaczenia przedziałów ufności w oparciu o asymptotyczną zbieżność estymatorów metody największej wiarygodności do rozkładu normalnego. Przykładowo przedział $\hat{\theta}_i \pm 1,96 \text{SE}(\hat{\theta}_i)$ stanowi przybliżenie 95% przedziału ufności dla oszacowania i -tej składowej wektora θ .

Przedziały ufności można także wyznaczać metodą profili funkcji wiarygodności zaproponowanej przez Barndorff-Nielsen i Cox (1994). Jeśli wektor θ podzieli się na dwie składowe d_ζ i d_η to d_η -wymiarowy profil logarytmu funkcji wiarygodności wektora d_η wyznacza się za pomocą $\ell_{(p)}(\eta) = \sup_{\zeta} \ell(\zeta, \eta)$. Przybliżeniem 95% przedziału ufności dla η jest wówczas przedział

$$\left\{ \eta : 2 \left[\ell_{(p)}(\hat{\eta}) - \ell_{(p)}(\eta) \right] < \chi_{0,95}^2(d_\eta) \right\}, \quad (3.136)$$

gdzie $\chi^2_{0,95}(d_\eta)$ jest kwantylem rzędu 0,95 rozkładu chi-kwadrat o d_η stopniach swobody, $\hat{\eta} = \underset{\eta}{\operatorname{argmax}} \ell_{(p)}(\eta)$. Wykonanie takich przedziałów ufności wymaga jednak znacznie większego nakładu czasu do wykonania obliczeń niż metoda w oparciu o cięcia logarytmu funkcji wiarygodności.

Zastosowanie metody iterowanej filtracji do estymacji parametrów niegaussowskich modeli stochastycznej zmienności wymaga przedstawienia modelu w postaci przestrzeni stanów oraz opracowania symulacji kolejnych generacji cząsteczek (krok 2 (ii) w algorytmie 6). Zostało to przedstawione odpowiednio w podrozdziałach 3.6 oraz 3.7. W odróżnieniu od wyników zaprezentowanych w pracach Szczepocki (2019a) oraz Szczepocki (2019b) przestrzenie stanów z podrozdziału 3.6 pozwalają na estymację parametrów dla bardzo szerokiej klasy modeli uwzględniającej złożenie procesów zmienności oraz efekt dźwigni. Ponadto metody symulacji z podrozdziału 3.7 umożliwiają zastosowanie jako rozkładów stacjonarnych uogólnionych rozkładów odwrotnych Gaussa oraz temperowane rozkłady stabilne. Użycie schematu Eulera umożliwia zastosowanie iterowanej filtracji także w przypadkach, gdy proces Ornsteina-Uhlenbecka jest określony poprzez wybór prowadzącego procesu Lévy'ego ukrytego w tle. Do obliczeń zastosowano program R, w tym pakiet POMP (King i in., 2019) stanowiący podstawę obliczeń algorytmu iterowanej filtracji. W tym celu napisano w języku programowania R oraz częściowo w języku programowania C skrypty umożliwiające użycie algorytmu iterowanej filtracji dla przedstawionych modeli przestrzeni stanów oraz generujące kolejne generacje cząsteczek dla przedstawionych procesów wariancji chwilowej.



Rysunek 3.7: Trajektorie oszacowań parametrów λ , α , ν dla dziesięciu 10 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych, każda po $J = 500$ iteracji (a) oraz oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów λ , α , ν wyznaczone przez filtr cząsteczkowy (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty) (b). Zastosowano podstawowy model BNS o wariancji chwilowej ze stacjonarnym rozkładem gamma z parametrami $\alpha = 12,5$, $\nu = 6,25$, $\lambda = 0,01$. Przyjęto

$$\mu = \beta = 0.$$

Źródło: opracowanie własne.

Modelowanie zmienności finansowych szeregów czasowych

4.1 Uwagi wstępne

W niniejszym rozdziale przedstawione zostaną przykłady zastosowania omówionych w rozdziale drugim niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka do modelowania zmienności dla trzech szeregów czasowych pochodzących z polskiego rynku finansowego: notowań Warszawskiego Indeksu Giełdowego (WIG), notowań dla indeksu WIG20 – indeksu dwudziestu największych spółek akcyjnych notowanych na warszawskiej Giełdzie Papierów Wartościowych oraz kursu USD/PLN (kursu dolara amerykańskiego wyrażonego w złotych).

W przypadku szeregu czasowych indeksu WIG oraz kursu USD/PLN są to dane dzienne z okresu od 6 września 2013 do 8 września 2017. Dane pochodzą z serwisu internetowego Stooq.pl (Stooq.pl, 2019). Modelowaniem objęte zostały dzienne (procentowe) logarytmiczne stopy zmian:

$$y_n = 100 \ln \left(\frac{x_n}{x_{n-1}} \right), \quad (4.1)$$

gdzie x_n jest ceną instrumentu finansowego (poziomem indeksu, notowaniem kursu walutowego). Są to stopy zwrotu z hipotetycznych jednodniowych inwestycji w dany walor finansowy przy kapitalizacji ciągłej i zerowych kosztach transakcji (Pajor, 2003, str. 20). Zdefiniowana w ten sposób logarytmiczna stopa zwrotu jest

dyskretyzacją procesu logarytmu cen występującego w modelach matematyki finansowej i pozwalają na obliczenie zwrotu z inwestycji wielodniowych w postaci sumy dziennych stóp zwrotu.

W przypadku szeregu czasowego indeksu WIG20 wykorzystano dane śróddzienne, 1 minutowe z okresu od 22 lutego 2017 do 21 lutego 2018. Dane pochodzą z serwisu internetowego Santander Biuro Maklerskie (Santander, 2019). W tym wypadku przedmiotem modelowania będą realizacje estymatora wariancji zrealizowanej wyznaczonego na podstawie zwrotów 5-minutowych. Wcześniejsze badania dotyczące rynku polskiego wskazują, że wartości wariancji zrealizowanej wyznaczone na podstawie zwrotów o częstotliwości od 5 do 10 minut dają najlepszy kompromis pomiędzy zbieżnością do wariancji kwadratowej procesu logarytmicznych cen wraz ze wzrostem częstotliwości pobierania obserwacji, a pojawiającą się dla bardzo wysokiej częstotliwości szumem mikrostruktury rynku (por. Kliber (2013, str. 105), Płucienik (2008)).

W rozdziale zostaną wykorzystane cztery wersje niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka: podstawowy model (definicja 2.1, oznaczenie: BNS), model ze złożeniem dwóch procesów (definicja 2.2, oznaczenie: BNS2), model z efektem dźwigni (definicja 2.3, oznaczenie: BNSL) oraz ze złożeniem dwóch procesów oraz z efektem dźwigni (definicja 2.4, oznaczenie: BNSL2). Powyższe modele są zagnieżdżone. Model BNS można traktować jako szczególny przypadek zarówno modelu BNS2 jak i BNSL, przyjmując przestrzeń parametrów modelu BNS jako podprzestrzeń bardziej ogólnych przestrzeni parametrów BNS2 i BNSL. Te dwa modele stanowią natomiast szczególne przypadki modelu BNS2L. W skrócie zależności na przestrzeniach parametrów można zapisać następująco: $BNS \subset BNSL \subset BNS2L$ oraz $BNS \subset BNS2 \subset BNS2L$.

Spośród możliwych rozkładów samorozkładalnych (por. podrozdział 2.5) jako rozkłady stacjonarne procesu wariancji chwilowej wybrane zostały uogólnione odwrotne rozkłady Gaussa. W szczególności dwa rozkłady tej klasy są szczególnie ważna: gamma i odwrotny Gaussa. Po pierwsze, jeżeli wariancja aktualna (dyskretyzacja wariancji chwilowej) miałaby rozkład odpowiednio gamma, odwrotny Gaussa, uogólniony odwrotny Gaussa, to zgodnie z twierdzeniem 1.6 logarytmiczne stopy

zwrotów miałyby bezwarunkowy rozkład VG, NIG oraz uogólniony hiperboliczny. Wszystkie trzy rozkłady mają bardzo dobre dopasowanie do empirycznie obserwowanych stóp zwrotu. Choć rozkład zmienności aktualnej jest nieznan, to część autorów¹ uważa że może być dostatecznie dobrze odwzorowany rozkładem stacjonarnym procesu wariancji chwilowej. Po drugie, Barndorff-Nielsen i Shephard (2003) analitycznie wykazali, że jeżeli wariancja chwilowa ma rozkład gamma lub odwrotny Gaussa, to wariancja scałkowana również ma rozkład bardzo zbliżony do odpowiednio rozkładu gamma lub odwrotnego Gaussa (por. twierdzenie 1.8 oraz wzór 2.17). Zależność ta nie jest prawdziwa dla wszystkich rozkładów stacjonarnych, na przykład dla rozkładu log-normalnego wariancja scałkowana nie ma podobnego rozkładu jak wariancja chwilowa. Własności te pozwalają przed dokonaniem estymacji modelu stochastycznej zmienności wstępnie ocenić, czy zastosowanie procesu wariancji chwilowej o danym rozkładzie jest w danym przypadku odpowiednie. Po trzecie, dla rozkładów stacjonarnych procesu wariancji chwilowej gamma i odwrotnego Gaussa zostały wyznaczone wzory na wycenę opcji europejskich w niegaussowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka². Ponadto, jedynie rozkłady odwrotny Gaussa oraz gamma spośród uogólnionych rozkładów Gaussa są rozkładami zamkniętymi na sploty, co umożliwia konstrukcje złożonego procesu zmienności.

Do estymacji parametrów modeli stochastycznej zmienności zostanie wykorzystana przedstawiona w podrozdziale 3.8 metoda iterowanej filtracji.

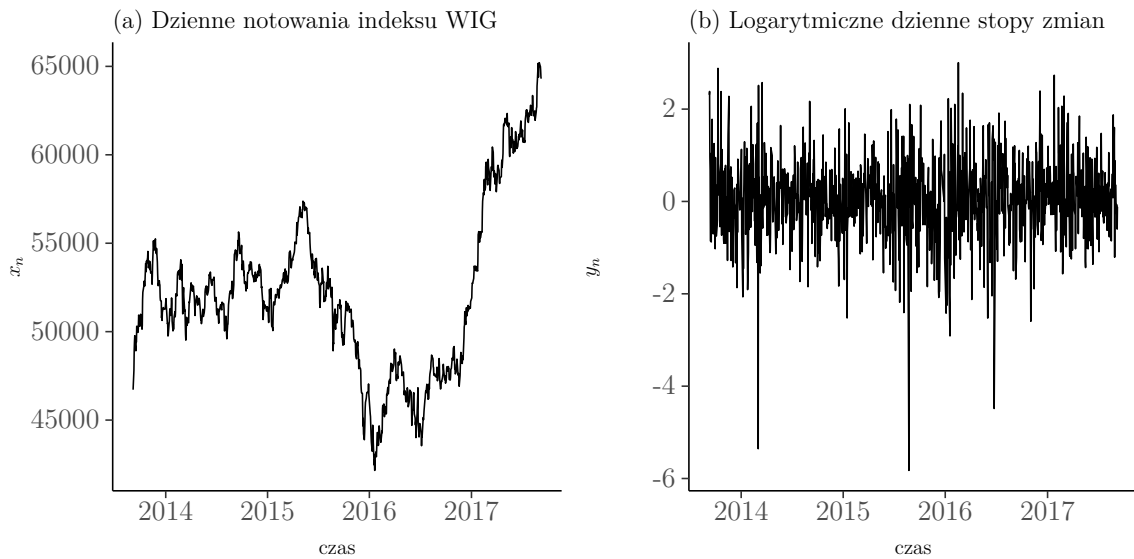
4.2 Charakterystyka danych empirycznych

Pierwszym badanym szeregiem czasowym jest 10001 kolejnych notowań indeksu Warszawskiego Indeksu Giełdowego (WIG) z okresu od 6 września 2013 do 8 września 2017. Na podstawie wyznaczono 1000 dziennych zwrotów logarytmicznych. Na rysunku 4.1 (a) i (b) przedstawiono odpowiednio notowania indeksu WIG i dzienne zwroty logarytmiczne w badanym okresie. Zauważmy, że logarytmiczne stopy zwrotu charakteryzują się grupowaniem zmienności: po dużych ruchach ceny występują za-

¹Na przykład Benth (2011) oraz Lindberg (2008).

²Wzory te zostały przedstawione w pracy Nicolato i Venardos (2003)

zwyczaj duże ruchy, a po małych małych.



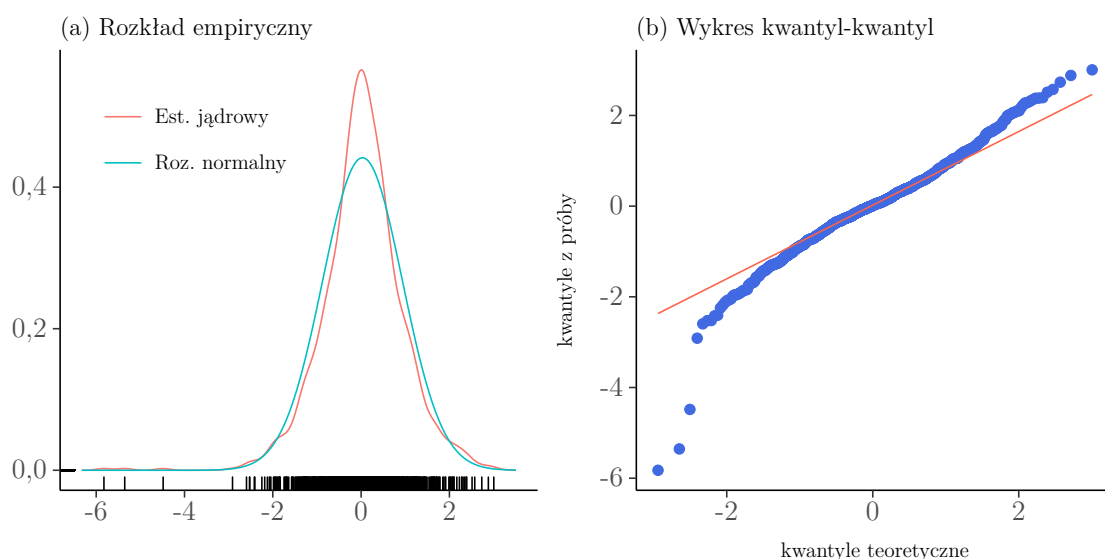
Rysunek 4.1: Dzielne notowania indeksu WIG (a) oraz logarytmiczne dzielne stopy zmian indeksu WIG (b) z okresu od 6 września 2013 do 8 września 2017.

Źródło: opracowanie własne.

Na podstawie rysunku 4.2 (a) można zauważyć, że oszacowanie funkcji gęstości otrzymane za pomocą estymatora jądrowego³ ma kształt bardziej skoncentrowany wokół wartości średniej niż w przypadku rozkładu normalnego oraz lewostronną asymetrię rozkładu empirycznego. Ponadto rysunek 4.2 (b) kwantyl-kwantyl (*qqplot*) wskazuje, że ogony rozkładu empirycznego znacznie różnią się od ogonów rozkładu normalnego. Potwierdzają to podstawowe charakterystyki próbkowe logarytmicznych dziennych stóp zmian indeksu WIG przedstawione w tablicy 4.1. Wartość kurtozy wskazuje na leptokurtyczność rozkładu, natomiast skośność na lewostronną asymetrię rozkładu. Wartości tych dwóch statystyk odbiegają od typowych wartości dla rozkładu normalnego (wynoszących odpowiednio 3 i 0). Leptokurtoza oraz asymetria spadków i wzrostów (spadki są co do wartości bezwzględnej większe niż wzrosty) są to typowe własności szeregów notowań akcji i indeksów giełdowych (Kliber, 2013, str. 50). Testy normalności Shapiro-Wilka oraz Jarque - Bera odrzucają

³Zastosowano jądro gaussowskie, parametr wygładzany wyznaczony zgodnie z regułą Silvermana (Silverman, 1986).

hipotezę o normalności rozkładu logarytmicznych stóp zmian indeksu WIG z p-wartościami testów poniżej 0,001.



Rysunek 4.2: Rozkład empiryczny dziennych logarytmicznych stóp zmian indeksu WIG wyznaczony na podstawie estymatora jądrowego oraz dopasowany wykres gęstości rozkładu normalnego (a) oraz wykres kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian indeksu WIG (b). Dane z okresu od 6 września 2013 do 8 września 2017.

Źródło: opracowanie własne.

Tablica 4.1: Statystyki opisowe dla logarytmicznych dziennych stóp zmian indeksu WIG z okresu od 6 września 2013 do 8 września 2017.

Liczba obserwacji	Wartość średnia	Odchylenie standardowe	Skośność	Kurtoza
1000	0,0320	0,9036	-0,5218	6,6328
Minimum	Kwartył pierwszy	Mediana	Kwartył trzeci	Maksimum
-5,8250	-0,4482	0,0429	0,5418	3,0052

Źródło: opracowanie własne.

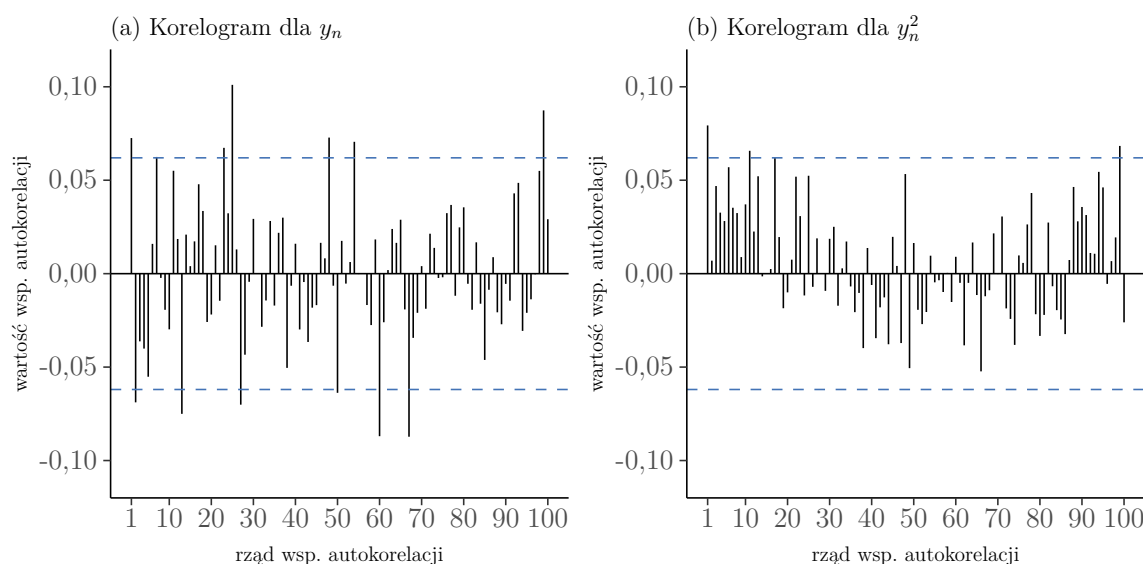
Korelogram kwadratów stóp zwrotu przedstawione na rysunku 4.3 (b) pokazuje, że nawet dla bardzo dużych opóźnień pojawiają się statystycznie istotne wartości

współczynnika korelacji⁴. Mniej typowe jest pojawienie się ujemnej korelacji. Również autokorelacja stóp zwrotu nie zanika szybko, tak jak wskazują typowe obserwacje zwrotów dla finansowych szeregów czasowych.

Tablica 4.2: Statystyki i p-wartości testów normalności dla zwrotów logarytmicznych indeksu WiG za okres od 6 września 2013 do 8 września 2017.

Test	Wartość statystyki testowej	P-wartość testu
Shapiro-Wilka	0,9668	0,0000
Jarque-Bera	595,4	0,0000

Źródło: opracowanie własne.

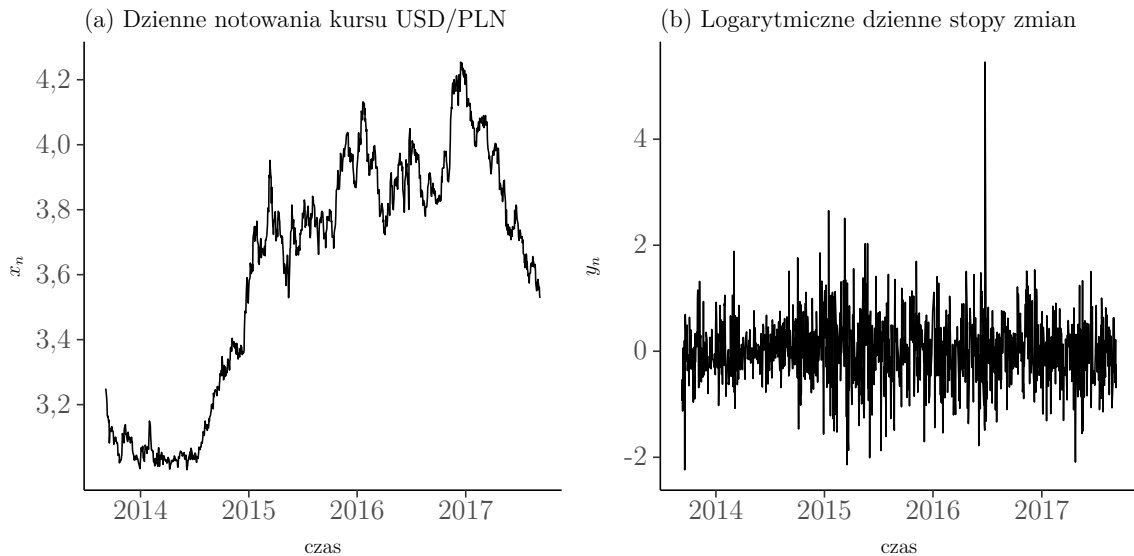


Rysunek 4.3: Korelogram dla dziennych logarytmicznych stóp zmian indeksu WIG (a) oraz dla kwadratów dziennych logarytmicznych stóp zmian indeksu WIG (b). Dane z okresu od 6 września 2013 do 8 września 2017.

Źródło: opracowanie własne.

⁴Linie przerywane wyznaczają obszary odrzucenia na poziomie istotności 0,05 hipotezy zerowej o zerowym współczynniku autokorelacji przy danym rzędzie opóźnienia. Linie te zostały wyznaczone przy założeniu, że szereg czasowy stanowi realizację ścisłego białego szumu (por. (Brockwell i in., 1991, roz. 7.2) oraz (Kwiatkowski, 2013, str. 199)).

Drugim badanym szeregiem czasowym jest 1035 kolejnych notowań kursu USD/PLN z okresu od 6 września 2013 do 8 września 2017. Na ich podstawie wyznaczono 1034 logarytmicznych dziennych stóp zmian (logarytmicznych stóp zwrotu w jednodniowe inwestycje w dolar amerykański). Na rysunku 4.4 (a) i (b) przedstawiono odpowiednio kurs i dzienne zwroty logarytmiczne w badanym okresie. Podobnie jak w przypadku kursu WIG widać efekt grupowania zmienności.

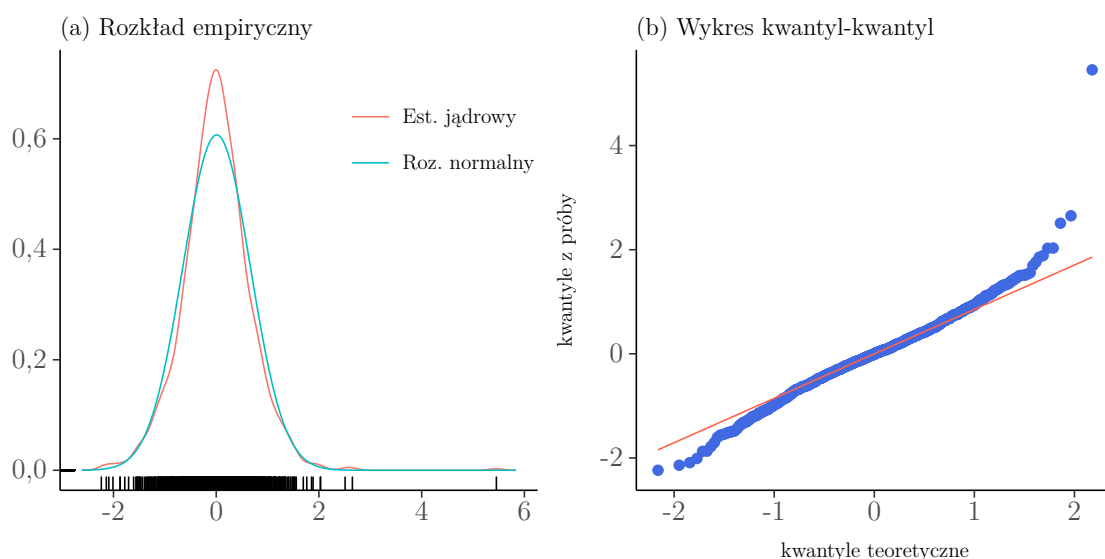


Rysunek 4.4: Dzielne wartości kursu USD/PLN (a) oraz logarytmiczne dzienne stopy zmian kursu USD/PLN (b) z okresu od 6 września 2013 do 8 września 2017.

Źródło: opracowanie własne.

Z rysunku 4.5 (a) wynika, że oszacowanie gęstości rozkładu ma kształt bardziej skoncentrowany wokół wartości średniej niż w przypadku rozkładu normalnego oraz prawostronną asymetrię wynikającą z wystąpienia pojedynczych bardzo dużych zwrotów. Wykres kwantyl-quantyl (rysunek 4.5 (b)) również wskazuje na istotne różnice empirycznego rozkładu od rozkładu normalnego zwłaszcza w prawym ogonie rozkładu. Podstawowe statystyki wyznaczone dla logarytmicznych dziennych stóp zmian kursu USD/PLN przedstawione w tabelicy 4.3 potwierdzają leptokurtozę i prawostronną asymetrię rozkładu. W przypadku kursów walutowych brak asymetrii spadków i wzrostów jest typowy. Testy normalności Shapiro-Wilka oraz Jarque - Bera podobnie jak w przypadku kursu WIG odrzucają hipotezę o normalności roz-

kładu logarytmicznych stóp zmian kursu USD/PLN z p-wartościami testów poniżej 0,001.



Rysunek 4.5: Rozkład empiryczny dziennych logarytmicznych stóp zmian kursu USD/PLN wyznaczony na podstawie estymatora jądrowego i dopasowany wykres gęstości rozkładu normalnego (a) oraz wykres kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian kursu USD/PLN (b). Dane z okresu od 6 września 2013 do 8 września 2017.

Źródło: opracowanie własne.

Tablica 4.3: Statystyki opisowe dla logarytmicznych dziennych stóp zmian kursu USD/PLN. Dane z okresu od 6 września 2013 do 8 września 2017.

Liczba obserwacji	Wartość średnia	Odchylenie standardowe	Skośność	Kurtoza
1034	0,0080	0,6572	0,5572	7,9791
Minimum	Kwartył pierwszy	Mediana	Kwartył trzeci	Maksimum
-2,2363	-0,3717	0,0053	0,3848	5,4542

Źródło: opracowanie własne.

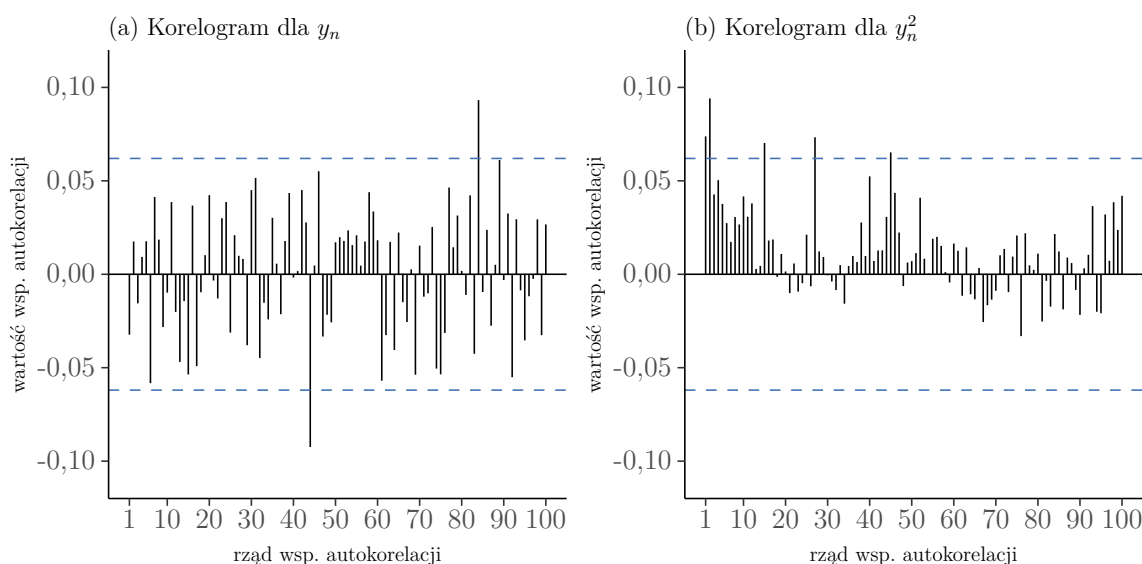
Korelogramy stóp i kwadratów stóp zwrotu (rysunek 4.6) również nie są typowe dla finansowych szeregów czasowych. Po pierwsze w przypadku autokorelacji stóp

Tablica 4.4: Statystyki i p-wartości testów normalności dla zwrotów logarytmicznych kursu USD/PLN za okres od 6 września 2013 do 8 września 2017.

Test	Wartość statystyki testowej	P-wartość testu
Shapiro-Wilka	0,96813	0,0000
Jarque-Bera	1121,8	0,0000

Źródło: opracowanie własne.

zwrotów pojawiają się istotne korelacje dla wysokich rzędów opóźnień. Po drugie pojawiają się ujemna (choć nieistotna statystycznie) korelacja dla kwadratów stóp zwrotu.



Rysunek 4.6: Korelogram dla dziennych logarytmicznych stóp zmian kursu USD/PLN (a) oraz dla kwadratów dziennych logarytmicznych stóp zmian kursu USD/PLN (b).

Dane z okresu od 6 września 2013 do 8 września 2017.

Źródło: opracowanie własne.

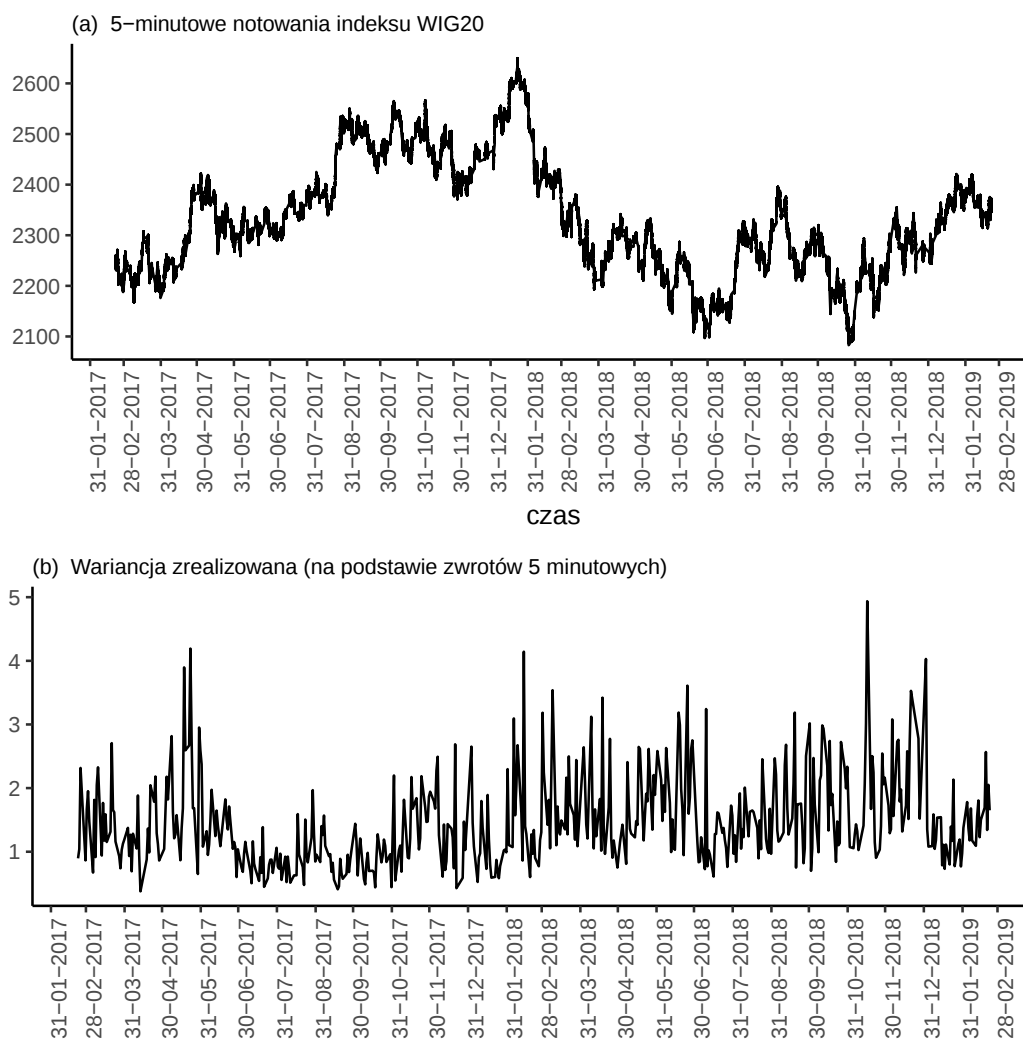
Trzeci szereg czasowy to 240018 1-minutowych notowań kursu WIG20 z okresu od 22 lutego 2017 do 21 lutego 2019. Na tej podstawie wyznaczono 499 oszacowań estymatora wariancji zrealizowanej na podstawie 5-minutowych zwrotów śróddziennych. Wykorzystanie wariancji zrealizowanej zmniejsza długość szeregu czasowego

blisko 481-razy. Na rysunku 4.7 przedstawiono 5-minutowe notowania kursu WIG20 oraz wyznaczone wartości wariancji zrealizowanej. Na rysunku 4.8 widać jeszcze wyraźniejszą niż w przypadku danych dziennych koncentrację oszacowania gęstości rozkładu empirycznego w pobliżu wartości średniej i lewostronną asymetrię. Potwierdzają to wyniki statystyk opisowych wyznaczone dla 5-minutowych logarytmicznych stóp zmian indeksu WIG20 przedstawione w tablicy 4.5 statystyki opisowe. Jest to zgodne z obserwowaną dla większości szeregów czasowych agregacyjną własnością szeregów czasowych: wraz ze wzrostem jednostki skali czasowej, względem której obliczane są zwroty, rozkład zwrotów upodabnia się do rozkładu normalnego (Kliber, 2013, str. 51).

Tablica 4.5: Statystyki opisowe dla 5-minutowych logarytmicznych stóp zmian indeksu WIG20.

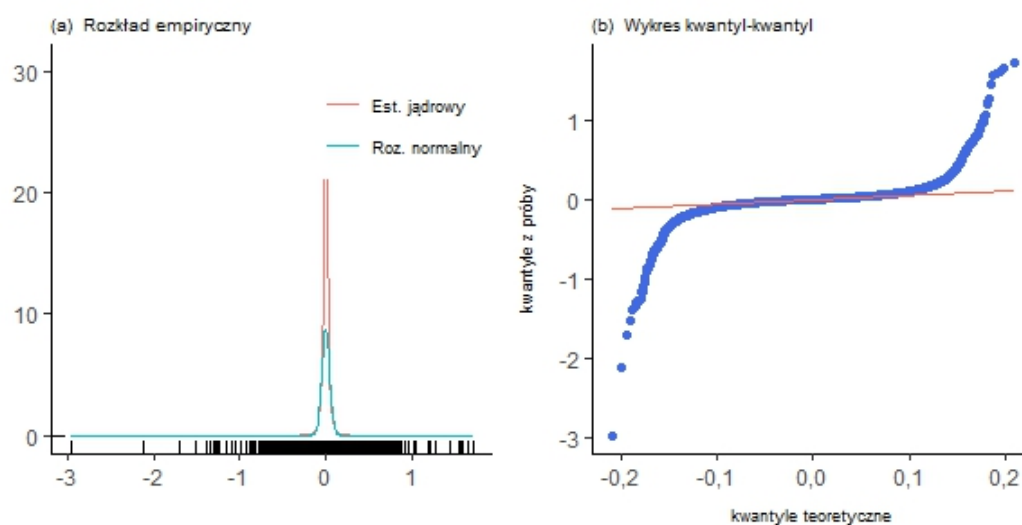
Liczba obserwacji	Wartość średnia	Odchylenie standardowe	Skośność	Kurtoza
48402	0,0000	0,0450	-0,4620	225,6326
Minimum	Kwartył pierwszy	Mediana	Kwartył trzeci	Maksimum
-2,9413	-0,0178	0,0000	0,0176	1,8524

Źródło: opracowanie własne.



Rysunek 4.7: 5-minutowe notowania indeksu WIG20 (a) oraz wariancja zrealizowana indeksu WIG20 (b) z okresu od 22 lutego 2017 do 21 lutego 2019.

Źródło: opracowanie własne.



Rysunek 4.8: Rozkład empiryczny 5-minutowych logarytmicznych stóp zmian indeksu WIG20 wyznaczony na podstawie estymatora jądrowego oraz dopasowany wykres gęstości rozkładu normalnego (a), wykres kwantyl-kwantyl dla 5-minutowych logarytmicznych stóp zmian indeksu WIG20 (b). Na podstawie danych z okresu od 22 lutego 2017 do 21 lutego 2019.

Źródło: opracowanie własne.

4.3 Wybór rozkładu stacjonarnego procesu wariancji chwilowej

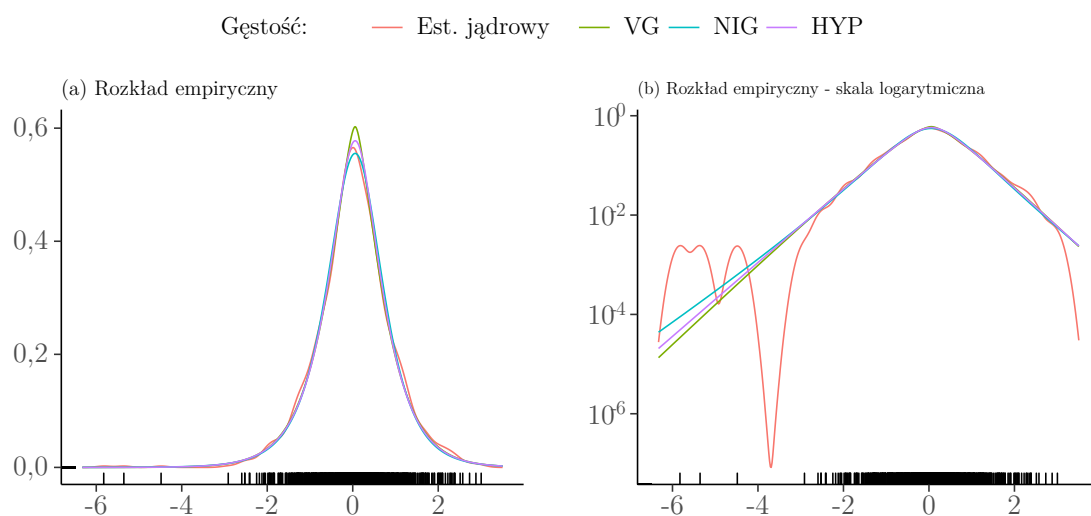
Jako rodzinę rozkładów procesu wariancji chwilowej zostały wybrane uogólnione rozkłady odwrotne Gaussa. Na rysunkach 4.9, 4.10 oraz 4.11 przedstawiono dopasowanie rozkładów VG, NIG oraz uogólnionego hiperbolicznego do odpowiednio dziennych stóp zwrotów indeksu WIG, kursu USD/PLN oraz WIG20. W tym ostatnim przypadku wyniki zostały przedstawione dla danych dziennych, ponieważ wariancja zrealizowana jest oszacowaniem dziennej wariancji aktualnej. Wszystkie trzy rozkłady mają bardzo dobre dopasowanie do danych środkowej części rozkładu. Nieco gorsze dopasowanie można zauważyć w ogonach rozkładu (skala logarytmiczna o podstawie równej 10). Wykresy kwantyl-kwantyl dla rozkładów uogólnionego rozkładu hiperbolicznego VG i NIG są praktycznie nierozróżnialne (rysunki 4.12, 4.13 oraz 4.14 przedstawiają wyniki dla rozkładów VG i NIG). Wszystkie rozkłady poza pojedynczymi, skrajnie nietypowymi obserwacjami mają dobre dopasowanie do obserwowanych empirycznie logarytmicznych stóp zwrotów.

W tablicy 4.6 przedstawiono wyniki statystyki Kołmogorowa-Smirnowa (por. Domański (1990, str. 64))

$$KS = \sup_x |\hat{F}(x) - F(x)| \quad (4.2)$$

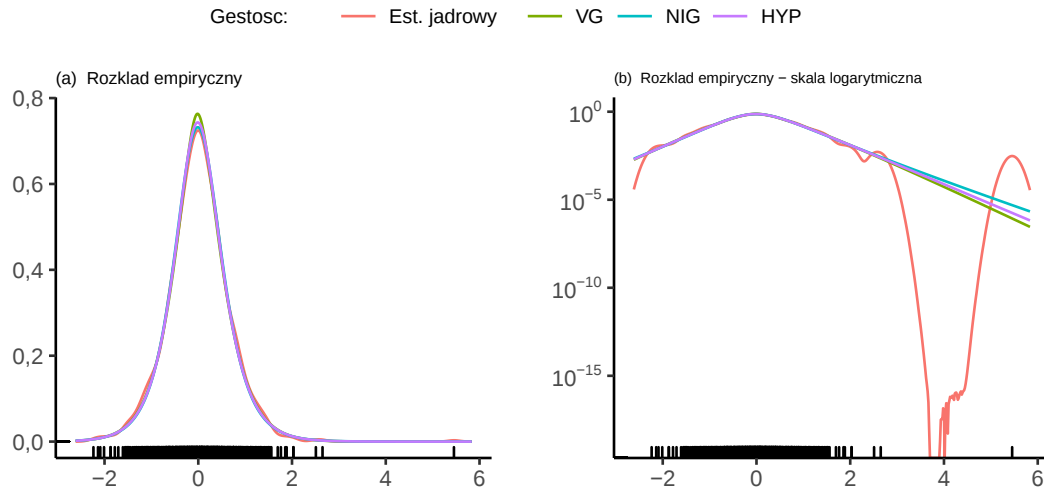
oraz p-wartości testu Kołmogorowa-Smirnowa dla trzech rozkładów: VG (*variance gamma*), NIG (normalny odwrotny Gaussa), HYP (uogólniony hiperboliczny) dla trzech rozważanych szeregów dziennych logarytmicznych stóp zwrotu. Wynik testów w żadnym z przypadków nie pozwalają na odrzucenie hipotezy o zgodności z rozkładem teoretycznym. Potwierdza to dobre dopasowanie do obserwowanych empirycznie rozkładów.

Rozkłady VG i NIG są specjalnymi przypadkami uogólnionego rozkładu hiperbolicznego przy spełnieniu pewnych restrykcji na przestrzeni parametrów tego rozkładu. Można zatem sprawdzić istotność ograniczeń parametrów za pomocą testu ilorazu wiarygodności dla modeli zagnieżdżonych (Domański i in., 2014, str. 279).



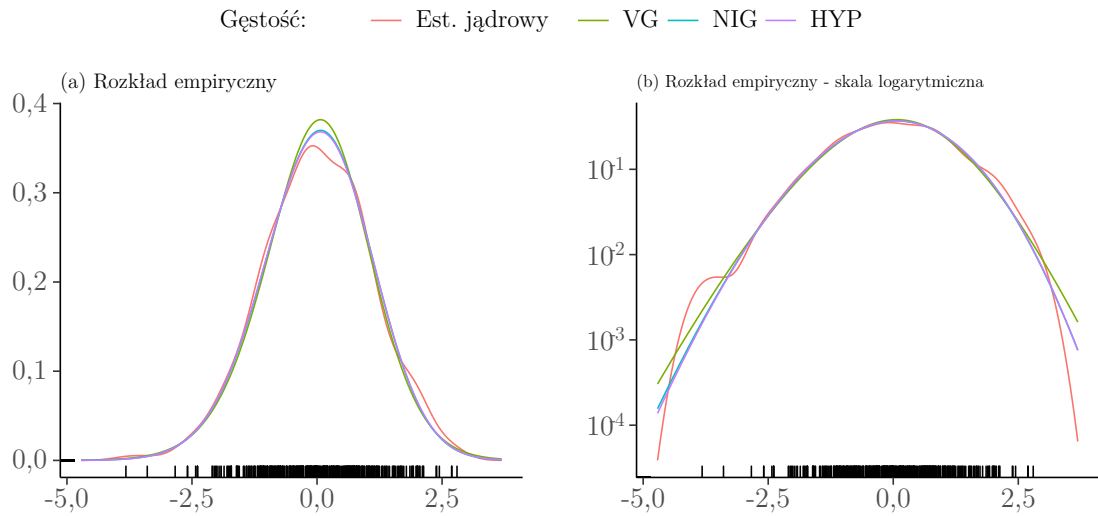
Rysunek 4.9: Rozkład empiryczny dziennych logarytmicznych stóp zmian indeksu WIG oraz dopasowane wykresy gęstości rozkładu VG, NIG, uogólnionego hiperbolicznego na skali zwykłej (a) oraz logarytmicznej (b).

Źródło: opracowanie własne.



Rysunek 4.10: Rozkład empiryczny dziennych logarytmicznych stóp zmian kursu USD/PLN oraz dopasowane wykresy gęstości rozkładu VG, NIG, uogólnionego hiperbolicznego na skali zwykłej (a) oraz logarytmicznej (b).

Źródło: opracowanie własne.



Rysunek 4.11: Rozkład empiryczny dziennych logarytmicznych stóp zmian indeksu WIG20 oraz dopasowane wykresy gęstości rozkładu VG, NIG, uogólnionego hiperbolicznego na skali zwykłej (a) oraz logarytmicznej (b).

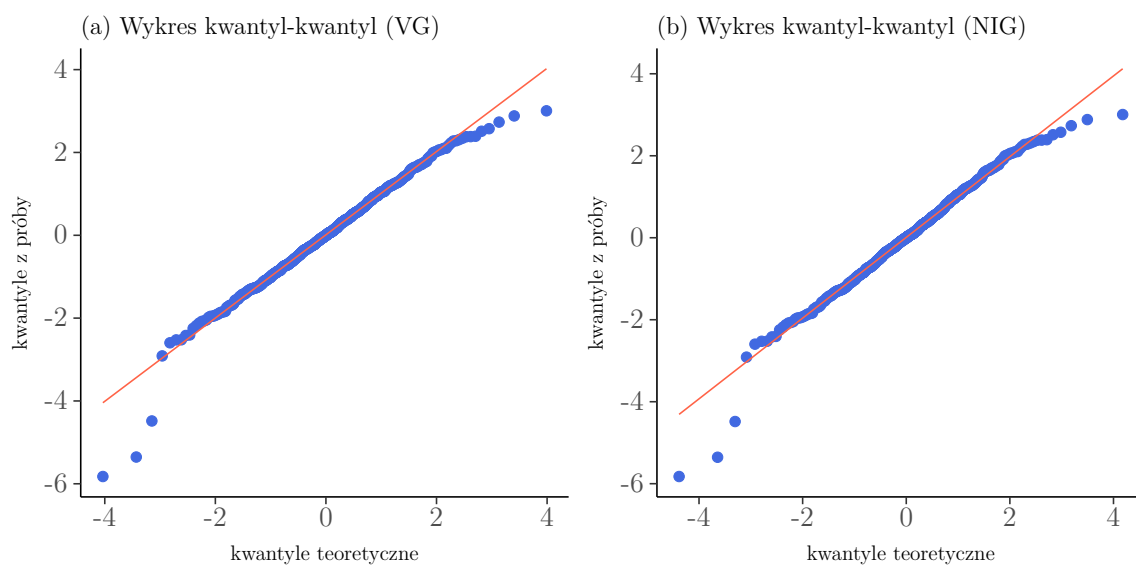
Źródło: opracowanie własne.

Odpowiednia statystyka testu przyjmuje wówczas postać

$$d = 2 \left(\sup_{\theta \in \Theta} \ell(\theta) - \sup_{\theta \in \Theta_0} \ell(\theta) \right), \quad (4.3)$$

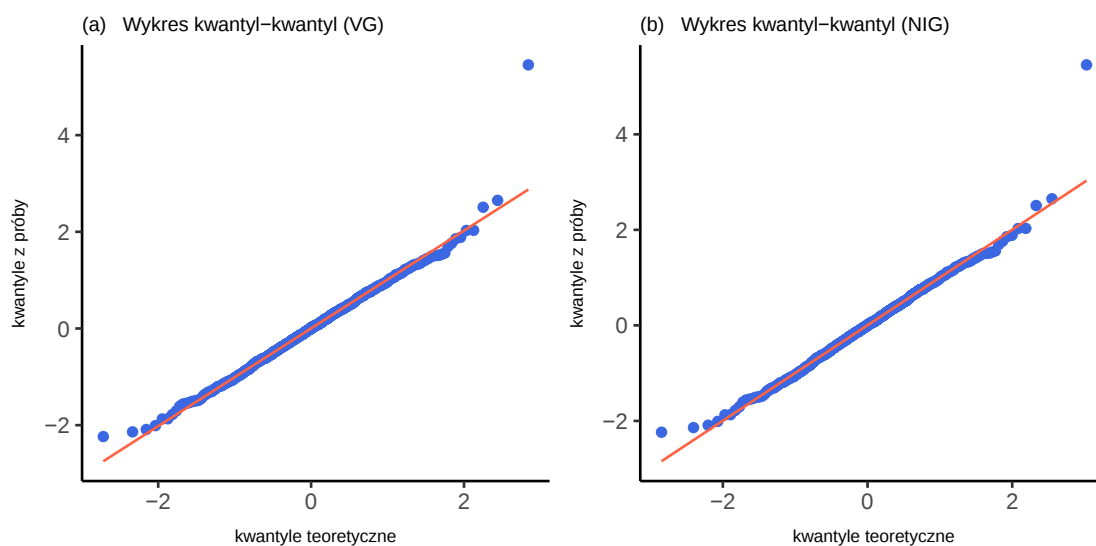
gdzie $\ell(\theta)$ to logarytm funkcji wiarygodności. Jeżeli próba rzeczywiście pochodzi z modelu z ograniczeniami Θ_0 , statystyka d ma asymptotycznie rozkład χ_k^2 (k stopniami swobody), gdzie k to liczba ograniczeń.

W 4.7 przedstawiono statystyki oraz p-wartości testu ilorazu wiarygodności dla rozkładów VG i NIG jako specjalnych przypadków uogólnionego rozkładu hiperbolicznego. W każdym z rozważanych przypadków brak podstaw do odrzucenia hipotez o istotności nałożonych restrykcji. Można zatem ograniczyć rozważania co do wyboru rozkładu stacjonarnego procesu wariancji chwilowej do dwóch rozkładów: gamma i odwrotnego Gaussa. Pozwala to na zastosowanie modeli uwzględniających złożenie procesów zmienności. Choć oba rozkłady należą do rodziny uogólnionych rozkładów odwrotnych Gaussa, to nieguassowskie procesy Ornsteina-Uhlenbecka o takich rozkładach stacjonarnych różnią się istotnie. W przypadku procesu o rozkła-



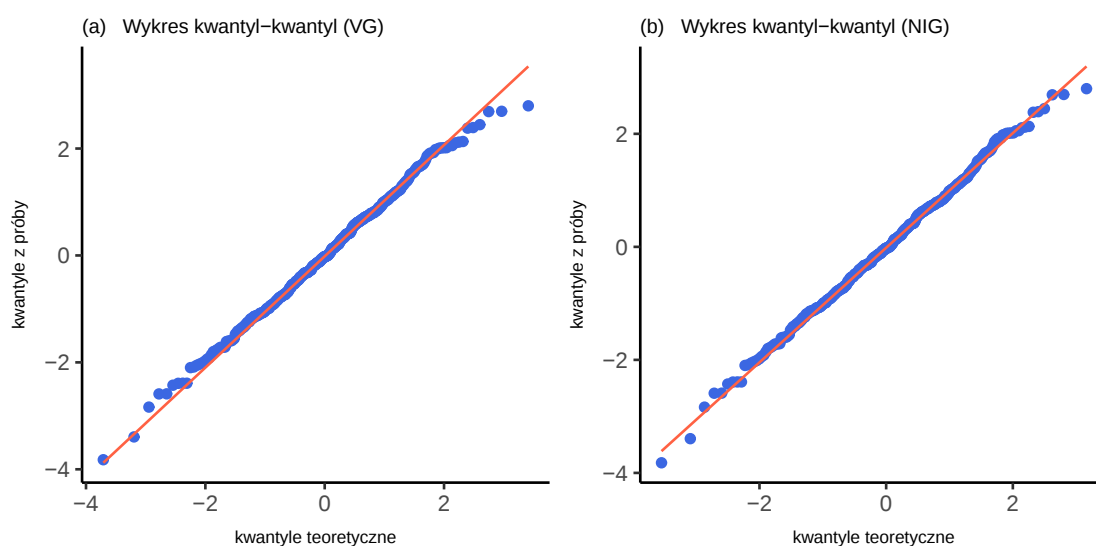
Rysunek 4.12: Wykresy kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian indeksu WIG względem rozkładu VG (c) i NIG (d).

Źródło: opracowanie własne.



Rysunek 4.13: Wykresy kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian kursu USD/PLN względem rozkładu VG (a) i NIG (b).

Źródło: opracowanie własne.



Rysunek 4.14: Wykresy kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian indeksu WIG względem rozkładu VG (c) i NIG (d).

Źródło: opracowanie własne.

dzie stacjonarnym gamma odpowiadający mu proces prowadzący Lévy'ego BDLP jest procesem o skończonej ilości skoków w jednostce czasu, a w przypadku odwrotnego Gaussa proces BDLP ma nieskończenie wiele skoków w jednostce czasu. Własności te przechodzą na proces wariancji chwilowej.

Tablica 4.6: Statystyki Kołmogorowa-Smirnowa (w nawiasach p-wartości testu Kołmogorowa-Smirnowa) dla trzech rozkładów: VG (*variance gamma*), NIG (normalny odwrotny Gaussa), HYP (uogólniony hiperboliczny).

Rozkład:	Szereg dziennych zwrotów:		
	WIG	USD/PLN	WIG20
VG	0,0150 (0,9777)	0,0101 (0,9999)	0,0106 (0,9998)
NIG	0,0188 (0,8683)	0,0096 (0,9999)	0,0125 (0,9975)
HYP	0,0156 (0,9675)	0,0094 (0,9999)	0,0117 (0,9991)

Źródło: opracowanie własne.

Tablica 4.7: Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla rozkładów VG (*variance gamma*) oraz NIG jako specjalnych przypadków uogólnionego rozkładu hiperbolicznego.

Rozkład:	Szereg dziennych zwrotów:		
	WIG	USD/PLN	WIG20
VG	0,6803 (0,5595)	0,8479 (0,3571)	0,3767 (0,5394)
NIG	0,1723 (0,6123)	0,3401 (0,4095)	0,0793 (0,7782)

Źródło: opracowanie własne.

4.4 Wyniki estymacji dla indeksu WIG

W podrozdziale przedstawione zostaną wyniki badania zmienności indeksu WIG z zastosowaniem niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka. Wykorzystano dwa rozkłady stacjonarne procesu wariancji chwilowej gamma i odwrotny Gaussa dla czterech specyfikacji niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka przedstawionych na początku rozdziału. Do obliczeń wykorzystano metodę iterowanej filtracji. Dla każdego z ośmiu modeli zastosowano pięć powtórzeń algorytmu iterowanej filtracji, każda o innym losowym wektorze wartości początkowych. Wybrano następujące ustawienia algorytmu iterowanej filtracji: $J = 200$ iteracji, $K = 1000$ cząsteczek, wielowymiarowy rozkład normalny z diagonalną macierzą wariancji kowariancji jako rozkład zaburzeń na przestrzeni parametrów, ciąg geometryczny jako ciąg schładzający o wartości początkowej ciągu 0,02, a iloraz w ciągu geometrycznym został dobrany tak, aby dla połowy iteracji otrzymać połowę wartości początkowej.

Do wyboru najlepszego modelu zastosowano kryterium informacyjne Akaikego (Akaike, 1973) (*Akaike information criterion*, AIC)⁵ postaci

$$AIC = -2\ell(\hat{\theta}) - 2q, \quad (4.4)$$

gdzie q to liczba szacowanych parametrów.

W tablicy 4.8 przedstawiono wyniki uzyskane dla rozkładu gamma, natomiast w tablicy 4.9 dla rozkładu odwrotnego Gaussa. Jako oszacowanie wartość wektora parametrów wybrano to spośród pięciu powtórzeń algorytmu, dla którego oszacowanie wartości logarytmu funkcji wiarygodności było największe. Oszacowanie logarytmu funkcji wiarygodności zostało wyznaczone na podstawie dziesięciu powtórzeń algorytmu filtru cząsteczkowego, jako średnia z oszacowań danych wzorem 3.75, przy czym

⁵W literaturze przedmiotu rozważane są także inne kryteria informacyjne (por. Doman i Doman (2009, str. 57)) m.in. Bayesowskie kryterium informacyjne Schwartza (Schwarz, 1978) (*Bayesian Information Criterion*, BIC), Kryterium informacyjne Hannan-Quinna (Hannan i Quinn, 1979) (*Hannan-Quinn information criterion*, HQC). Wybór kryterium jest kwestią subiektywną. Po szczególne kryteria różnią się „karą” za wprowadzenie kolejnych parametrów do modelu. Kryterium AIC jest jednym z najczęściej stosowanych.

algorytm był powtarzany dziesięć razy, aby otrzymać oszacowanie błędu standardowego estymatora logarytmu funkcji wiarygodności.

Interpretując wyniki należy jednak mieć na uwadze, że wartości logarytmu funkcji wiarygodności wykorzystywane do kryterium informacyjnego oraz do testów są wyznaczone za pomocą filtrów cząsteczkowych. Wiąże się to z dwoma implikacjami. Po pierwsze filtry cząsteczkowe dostarczają nieobciążonego i zgodnego oszacowania funkcji wiarygodności, ale obciążonego oszacowania logarytmu funkcji wiarygodności. W tej pracy wykorzystano powszechnie stosowaną korektę na redukcję obciążenia (wzór 3.75). Po drugie oszacowania są wyznaczone z pewnym błędem Monte Carlo, którego oceną jest podana wartość błędu standardowego estymatora logarytmu funkcji wiarygodności.

W tablicach 4.8 i 4.9 jako błędy standardowe oszacowań parametrów podano wartości wyznaczone za pomocą numerycznego szacowania macierzy informacyjnej Fishera według procedury opisanej w rozdziale 3. Ta sama metoda wyznaczania błędów standardowych została zastosowana w pracy Bretó (2014).

Tablica 4.8: Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej gamma wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.

Model	μ	β	α	ν	λ_1	λ_2	ρ_1	ρ_2	w_1	loglik	AIC
BNS Ga-OU	0,089 (0,0154)	0,0266 (0,0256)	3,8725 (0,3363)	3,0007 (0,2544)	0,0814 (0,0285)	-	-	-	-	-1272,748 (0,1539)	2550,496
BNSL Ga-OU	0,0884 (0,0231)	0,0226 (0,0393)	4,8746 (0,3071)	2,9445 (0,2295)	0,1019 (0,0175)	-	-2,4016 (0,3304)	-	-	-1266,538 (0,4290)	2537,076
BNS2 Ga-OU	0,09427 (0,0252)	0,0246 (0,0363)	3,9214 (0,7169)	2,2604 (0,4147)	0,0382 (0,0184)	0,6465 (0,3974)	-	-	0,00837 (0,0568)	-1267,935 (0,8382)	2542,87
BNS2L Ga-OU	0,10443 (0,0314)	0,0278 (0,0684)	3,6824 (0,3548)	2,8341 (0,2848)	0,1058 (0,1716)	0,2660 (0,0679)	-5,0210 (0,5156)	-0,0681 (0,4333)	0,0841 (0,0576)	-1260,374 (0,6898)	2529,748

Źródło: opracowanie własne.

Tablica 4.9: Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej odwrotnym Gaussa wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC.

W nawiasach podano błędy standardowe oszacowań.

Model	μ	β	γ	δ	λ_1	λ_2	ρ_1	ρ_2	w_1	loglik	AIC
BNS IG-OU	0,0804 (0,0265)	0,0304 (0,0355)	1,6178 (0,2404)	1,5077 (0,1905)	0,0487 (0,0385)	-	-	-	-	-1274,815 (0,6244)	2554,63
BNSL IG-OU	0,0718 (0,0262)	0,0181 (0,0359)	1,6203 (0,1768)	1,3766 (0,1366)	0,0881 (0,0251)	-	-0,0898 (0,3923)	-	-	-1270,561 (0,9147)	2545,122
BNS2 IG-OU	0,1206 (0,0262)	0,0109 (0,0376)	1,5578 (0,6594)	1,1376 (0,2175)	0,0188 (0,0106)	0,8277 (0,1697)	-	-	0,0274 (0,0514)	-1271,416 (0,4192)	2549,832
BNS2L IG-OU	0,1108 (0,0212)	0,0106 (0,0897)	1,2334 (0,6125)	1,4793 (0,1733)	0,1157 (0,0264)	0,5475 (0,1362)	-0,1232 (0,2373)	-0,0568 (0,1849)	0,0855 (0,0506)	-1266,4 (0,5431)	2539,8

Źródło: opracowanie własne.

Największą wartość kryterium informacyjnego AIC spośród ośmiu rozważanych modeli uzyskał model z efektem dźwigni oraz złożeniem procesów zmienności dla rozkładu stacjonarnego procesu wariancji chwilowej gamma (BNSL2 Ga-OU). Można zauważyć, że modele o rozkładzie stacjonarnym procesu wariancji gamma mają większe wartości funkcji wiarygodności od ich odpowiedników z rozkładem stacjonarnym odwrotnym Gaussa.

We wszystkich oszacowanych modelach parametr dryfu μ przyjmuje wartości bliskie 0,1, co wskazuje na nieznaczną tendencję do wzrostu kursu WIG w czasie. Jest to zgodne z otrzymaną dodatnią wartością średniej z logarytmicznych stóp zmian kursu WIG przedstawionymi w tablicy 4.1. Parametr β (interpretowany jako premia za ryzyko) we wszystkich modelach przyjmuje wartość dodatnią, ale bliską zera. Dodatnia wartość jest zgodna z teorią i dotychczasowymi badaniami empirycznymi opartymi na niegaussowskich modelach stochastycznej zmienności (por. Taufer i in. (2011)). W modelach ze złożeniem dla obu rozkładów stacjonarnych można zauważyć, że procesy o mniejszej wartości parametru λ mają znacznie mniejszą wagę od procesów z większą wartością parametru λ . Parametry ρ_1 i ρ_2 zgodnie z teorią przyjmują wartości ujemne. Uwagę zwraca szczególnie duża (co do wartości bezwzględnej) wartość parametru ρ_1 dla modelu BNSL2 Ga-OU. Jest to zgodne z wynikami empirycznymi uzyskanymi dla modelu modelu BNSL2 Ga-OU dla indeksu S&P500 przez Griffin i Steel (2006). Parametry mierzące siłę efektu dźwigni w odróżnieniu od ich odpowiedników w modelach opartych o dyskretyzację procesów SV, w których logarytm zmienność chwilowej jest procesem Ornsteina-Uhlenbecka oraz modeli należących do klasy CEV nie przyjmuje wartości tylko z przedziału $[-1, 1]$.

W tablicy 4.10 przedstawiono oszacowania wartości oczekiwanej $\xi = \mathbb{E}(\sigma^2(t))$ i odchylenia standardowego $\omega = \sqrt{\text{Var}(\sigma^2(t))}$ wariancji chwilowej dla wszystkich ośmiu modeli. Wartości uzyskano zgodnie ze wzorami odpowiednio $\xi = \nu/\alpha$, $\omega = \sqrt{\nu}/\alpha$ dla rozkładu gamma oraz $\xi = \delta/\gamma$, $\omega = \sqrt{\delta/\gamma^3}$ dla rozkładu odwrotnego Gaussa. Wartości uzyskane dla obu parametrów ξ i ω są większe dla rozkładu odwrotnego Gaussa niż dla rozkładu gamma.

Na rysunku 4.15 przedstawiono trajektorie oszacowań uzyskane metodą iterowanej filtracji dla poszczególnych składowych wektora $\theta = [\mu, \beta, \alpha, \nu, \lambda_1, \lambda_2, \rho_1, \rho_2, w_1]^T$

Tablica 4.10: Wyniki estymacji wartości średniej (parametr ξ) oraz odchylenia standardowego (parametr ω) rozkładu stacjonarnego procesu wariancji chwilowej dla ośmiu rozważanych modeli stochastycznej zmienności z podziałem na rozkłady stacjonarne: gamma (Ga-OU) oraz odwrotnym Gaussa (IG-OU).

Model	Ga-OU		Model	IG-OU	
	ξ	ω		ξ	ω
BNS Ga-OU	0,7697	0,4572	BNS IG-OU	0,9321	0,5967
BNSL Ga-OU	0,6041	0,3521	BNSL IG-OU	0,8496	0,5688
BNS2 Ga-OU	0,5764	0,3834	BNS2 IG-OU	0,7302	0,5486
BNS2L Ga-OU	0,7696	0,4572	BNS2L IG-OU	1,1994	0,8879

Źródło: opracowanie własne.

względem kolejnych iteracji algorytmu dla modelu BNSL2 Ga-OU, dla którego wartość kryterium AIC była najmniejsza. Na rysunku 4.16 przedstawiono jednowymiarowe ciecia wiarygodności dla tego modelu. Wykresy diagnostyczne 4.15 i 4.16 potwierdzają uzyskanie zbieżności algorytmu iterowanej filtracji dla modelu BNSL2 Ga-OU.

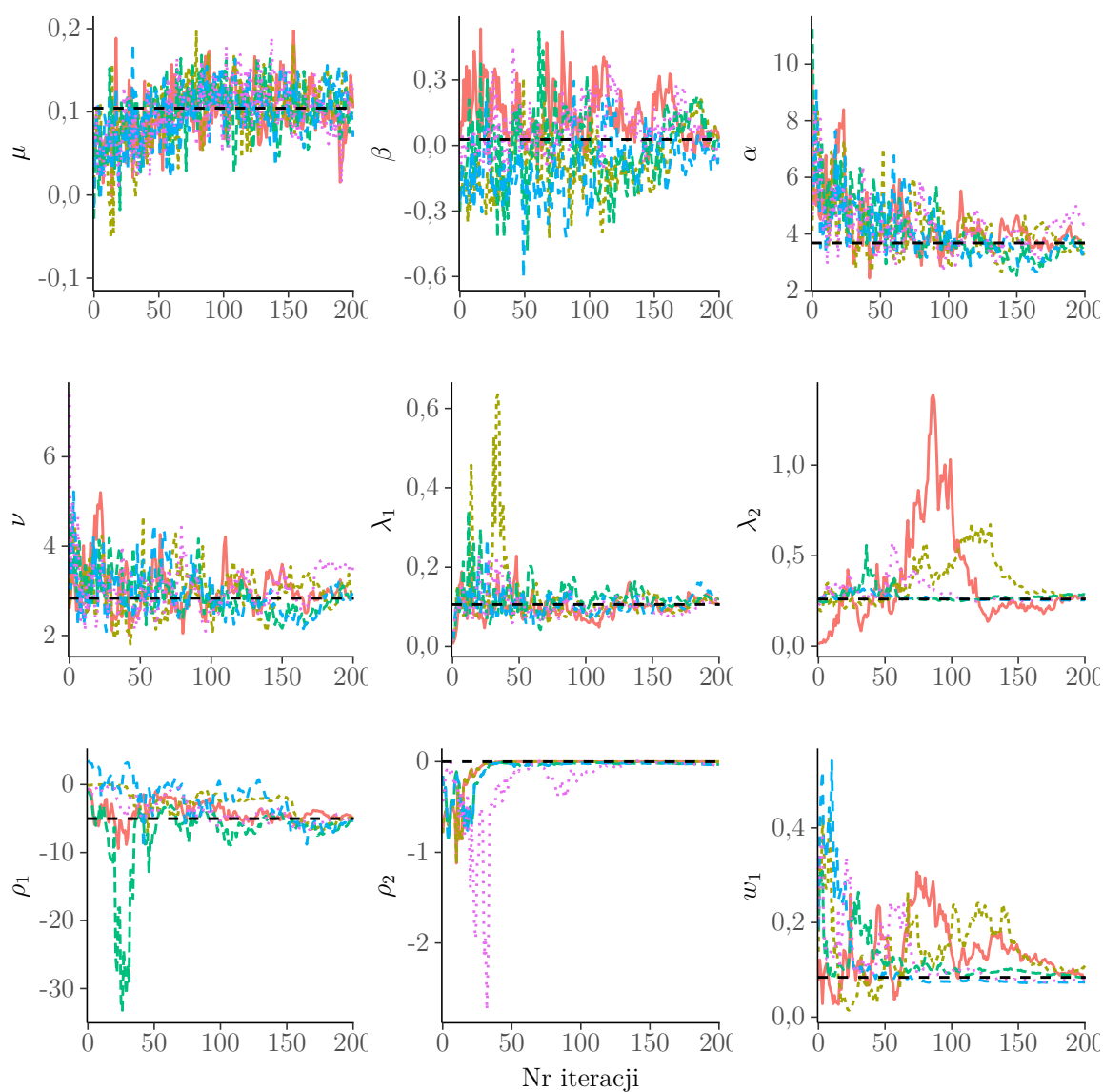
Modele BNS2 oraz BNSL można traktować jako modele z restrykcjami na przestrzeni parametrów dla modelu BNS2L. Można zatem rozważać następujące pary hipotez

$$\begin{aligned} H_0 : \rho_1 = \rho_2 = 0 \\ H_1 : \sim H_0 \end{aligned} \quad (4.5)$$

oraz

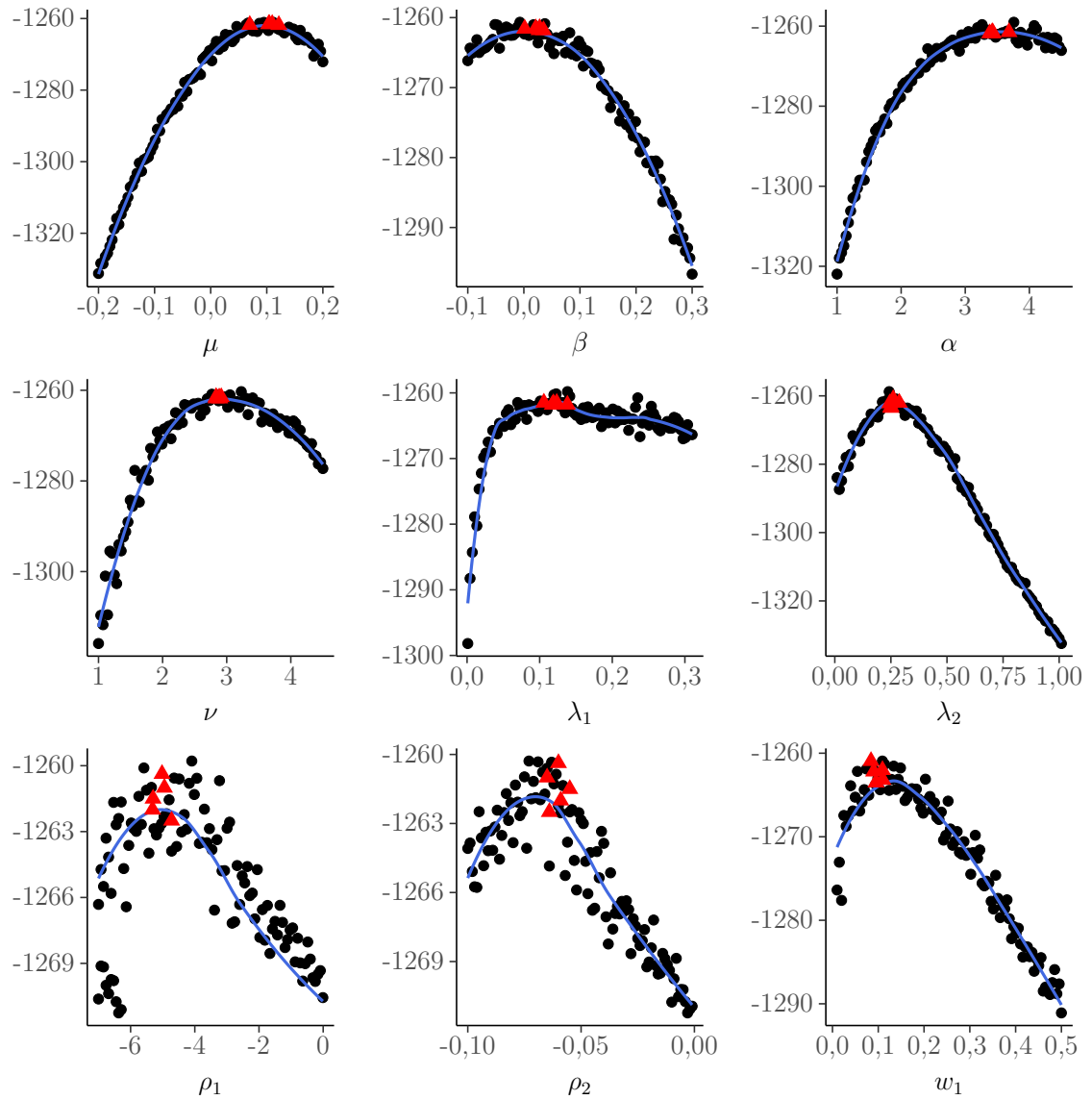
$$\begin{aligned} H_0 : \rho_2 = \lambda_2 = w_1 = 0 \\ H_1 : \sim H_0 \end{aligned} \quad (4.6)$$

W parze hipotez (4.5), hipoteza H_0 zakłada, że prawdziwe są restrykcje upraszczające model BNS2L do modelu BNS2, natomiast w przypadku pary hipotez (4.6)



Rysunek 4.15: Trajektorie oszacowań parametrów względem kolejnych iteracji algorytmu dla 5 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych w modelu ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU), każda po $J = 200$ iteracji.

Źródło: opracowanie własne.



Rysunek 4.16: Oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów wyznaczone za pomocą filtra cząsteczkowego (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty). Model ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU).

Źródło: opracowanie własne.

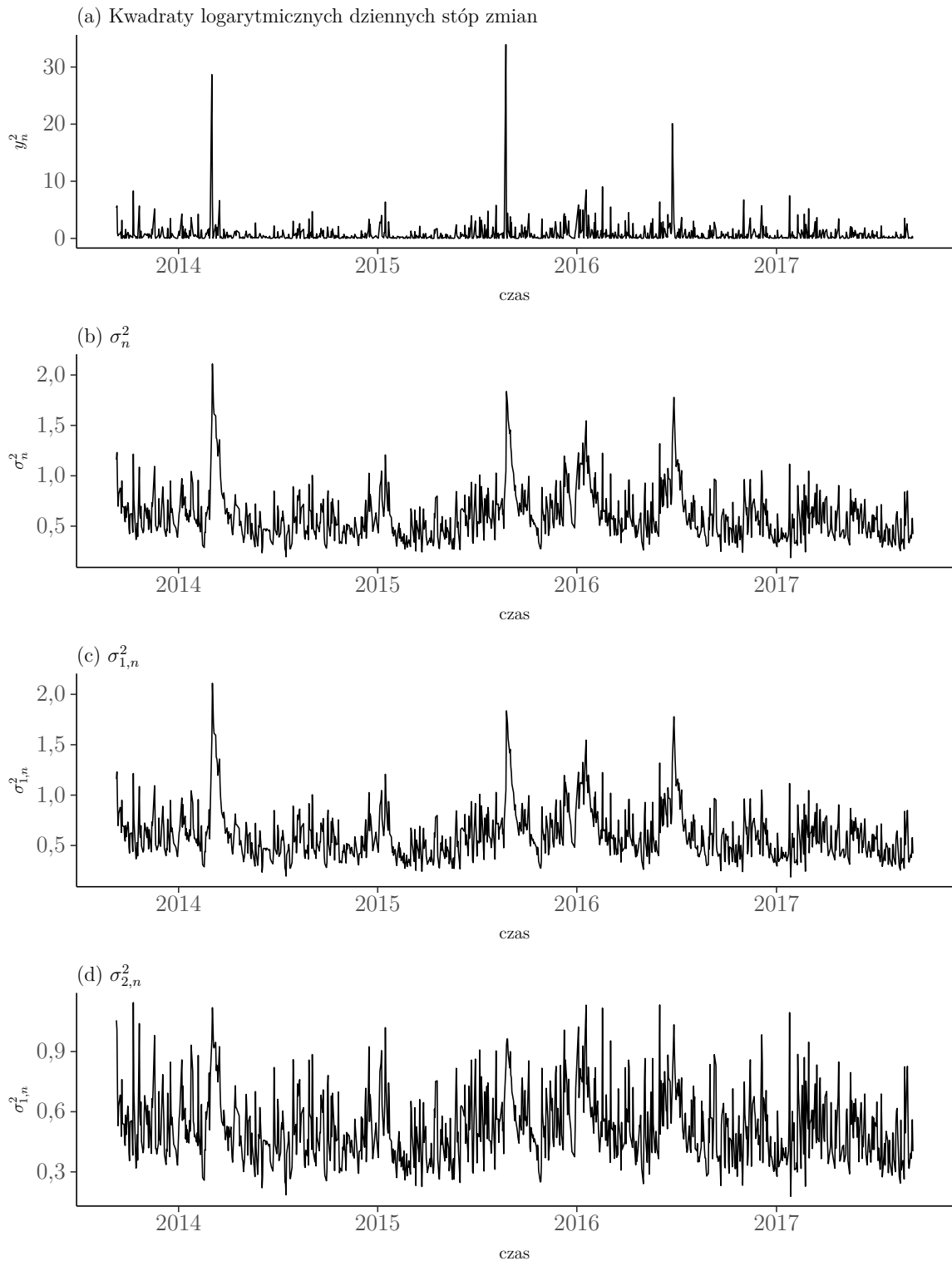
do modelu BNSL. Można sprawdzić istotność ograniczeń parametrów za pomocą testu ilorazu wiarygodności. W obu przypadkach przy założeniu prawdziwości hipotezy H_0 statystyka testowa ma rozkład χ^2 o odpowiednio dwóch i trzech stopniach swobody. W tablicy 4.11 przedstawiono wartości statyk testowych oraz p-wartości testów dla obu par hipotez i dla obu rozważanych rozkładów stacjonarnych procesu wariancji chwilowej. Tylko w przypadku rozkładu odwrotnego Gaussa i pary hipotez (4.6) można mieć wątpliwości, czy odrzucić hipotezę H_0 . W pozostałych przypadkach p-wartości testu wskazują na odrzucenie hipotezy H_0 dla typowych poziomów istotności testów. Wskazuje to, że zarówno złożenie procesów zmienności jak i efekt dźwigni należy uwzględnić w modelu.

Tablica 4.11: Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla par hipotez zagnieżdżonych.

Para hipotez	Rozkład stacjonarny wariancji chwilowej	
	Gamma	Odwrotny Gaussa
$H_0 : \rho_1 = \rho_2 = 0$ $H_1 : \sim H_0$	15,122 (0,0005)	9,832 (0,0073)
$H_0 : \rho_2 = \lambda_2 = w_1 = 0$ $H_1 : \sim H_0$	12,328 (0,0063)	6,122 (0,0435)

Źródło: opracowanie własne.

Na rysunku 4.17 przedstawiono oszacowanie wariancji aktualnej otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) oraz jej dwóch składowych w porównaniu do wartości kwadratów logarytmicznych stóp zwrotu. Jako model zmienności wybrano model ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym procesu zmienności gamma (BNS2L Ga-OU). Jako parametry modelu przyjęto oszacowaniu uzyskane metodą iterowanej filtracji przedstawione w tablicy 4.8. Zmiany w wahliwości kwadratów logarytmicznych stóp zwrotu są wyraźnie odzwierciedlone w zmianach wariancji aktualnej.



Rysunek 4.17: Kwadraty dziennych logarytmicznych stóp zmian indeksu WIG (a) oraz oszacowanie wariancji aktualnej w modelu ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU) otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) dla wartości parametrów przedstawionych w tabelicy 4.8 (b) wraz ze składowymi wariancjami aktualnej (c) i (d).

Źródło: opracowanie własne.

4.5 Wyniki estymacji dla kursu USD/PLN

W badaniu zmienności kursu USD/PLN z zastosowaniem niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka ponownie wykorzystano dwa rozkłady stacjonarne procesu wariancji chwilowej: gamma i odwrotny Gaussa dla czterech specyfikacji niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka przedstawionych na początku rozdziału. Do obliczeń wykorzystano metodę iterowanej filtracji z dokładnie takimi samymi ustawieniami algorytmu jak w przypadku badania dla indeksu WIG.

W tablicy 4.12 przedstawiono wyniki estymacji parametrów dla modeli o rozkładzie stacjonarnym procesu wariancji chwilowej gamma, a w tablicy o rozkładzie odwrotnym Gaussa 4.13. Dla wszystkich modeli parametr dryfu μ przyjmuje wartość bliską zera, co jest zgodne z wartością średnią logarytmicznych dziennych zmian kursu USD/PLN w badanym okresie. Parametr β dla wszystkich modeli przyjmuje wartość bliską zera. W modelach ze złożeniem procesów zmienności dla obu rozkładów stacjonarnych procesy o mniejszej wartości parametru persystencji λ mają mniejszą wagę od procesów z większą wartością parametru λ . Parametry ρ_1 i ρ_2 zgodnie z teorią przyjmują wartości ujemne, ale wartości są duże bliższe zera niż w przypadku badania dla kursu indeksu WIG. Jest to zgodne z wynikami uzyskanymi dla modeli stochastycznej zmienności z czasem dyskretnym: efekt dźwigni przejawia się silniej w przypadku modelowania zwrotów z akcji oraz indeksów, a słabiej w przypadku kursów walutowych (por. Pajor (2003)).

W tablicy 4.14 przedstawiono oszacowania wartości oczekiwanej (parametr ξ) i odchylenia standardowego (parametr ω) procesu wariancji chwilowej. Oszacowania wartości oczekiwanej są podobne, natomiast odchylenia standardowego wyższe dla rozkładu stacjonarnego wariancji chwilowej odwrotnego Gaussa.

Tablica 4.12: Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej gamma wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC.

W nawiasach podano błędy standardowe oszacowań.

Model	μ	β	α	ν	λ_1	λ_2	ρ_1	ρ_2	w_1	loglik	AIC
BNS Ga-OU	0,0322 (0,0185)	0,0016 (0,0484)	5,7266 (0,6065)	2,6392 (0,2422)	0,0736 (0,0194)	-	-	-	-	-975,853 (0,1344)	1956,7062
BNSL Ga-OU	0,0489 (0,0184)	-0,0092 (0,0484)	6,4053 (0,6738)	2,6603 (0,2359)	0,0824 (0,0208)	-	-0,5949 (0,4904)	-	-	-975,633 (0,1903)	1957,2662
BNS2 Ga-OU	0,0024 (0,0185)	0,0136 (0,0483)	5,8442 (0,2781)	2,0211 (0,0183)	0,0592 (0,0232)	0,8571 (0,1295)	-	-	0,0855 (0,0246)	-973,292 (0,4657)	1952,5846
BNS2L Ga-OU	0,0671 (0,0222)	0,0096 (0,0617)	6,1575 (0,4667)	2,0033 (0,0848)	0,0553 (0,0184)	0,4371 (0,0183)	-0,0061 (0,0828)	-0,1808 (0,5828)	0,0889 (0,0219)	-973,104 (0,1951)	1952,2008

Źródło: opracowanie własne.

Tablica 4.13: Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej odwrotnym Gaussa wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.

Model	μ	β	γ	δ	λ_1	λ_2	ρ_1	ρ_2	w_1	loglik	AIC
BNS IG-OU	0,0013 (0,0182)	0,0179 (0,0476)	1,675 (0,1811)	0,7784 (0,0561)	0,0907 (0,0294)	-	-	-	-	-974,285 (0,1983)	1953,574
BNSL IG-OU	0,0043 (0,0185)	-0,0078 (0,0181)	1,2663 (0,2328)	0,7386 (0,0757)	0,0828 (0,0124)	-	-0,3093 (0,3297)	-	-	-974,2725 (0,1907)	1954,545
BNS2 IG-OU	0,0086 (0,0187)	0,0115 (0,0488)	1,7263 (0,0201)	0,7856 (0,3777)	0,0726 (0,0131)	0,4985 (0,3776)	-	-	0,1739 (0,0208)	-970,4859 (0,1981)	1947,971
BNS2L IG-OU	0,0015 (0,0054)	0,0288 (0,4283)	1,0685 (0,0254)	0,6021 (0,0747)	0,0265 (0,0187)	0,4801 (0,0187)	-0,0068 (0,0033)	-0,2287 (0,0362)	0,1044 (0,0297)	-970,4057 (0,2910)	1949,811

Źródło: opracowanie własne.

Tablica 4.14: Wyniki estymacji wartości średniej (parametr ξ) oraz odchylenia standardowego (parametr ω) rozkładu stacjonarnego procesu wariancji chwilowej dla ośmiu rozważanych modeli stochastycznej zmienności z podziałem na rozkłady stacjonarne: gamma (Ga-OU) oraz odwrotnym Gaussa (IG-OU).

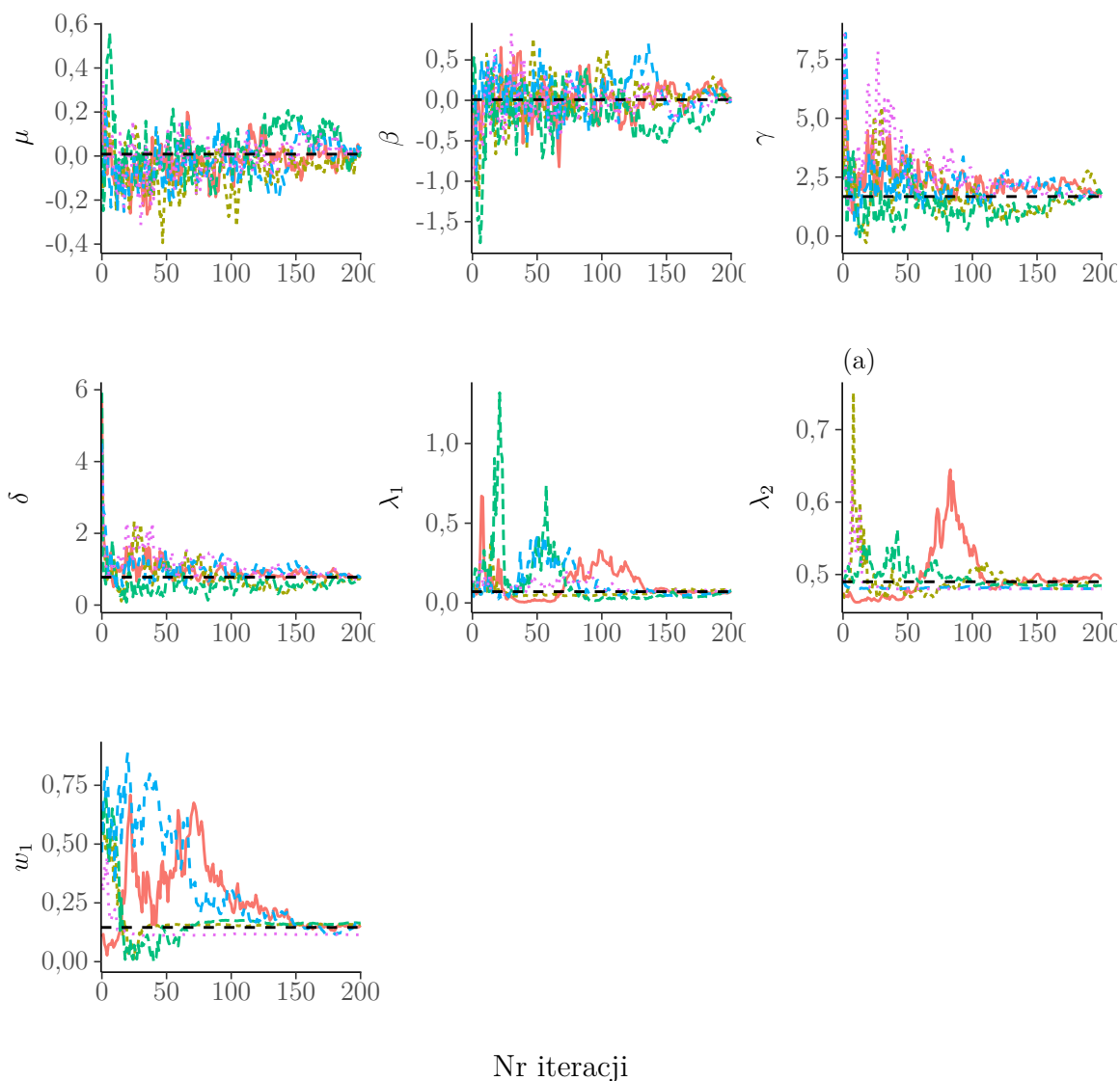
Model	Ga-OU		model	IG-OU	
	ξ	ω		ξ	ω
BNS Ga-OU	0,4609	0,2837	BNS IG-OU	0,4647	0,407
BNSL Ga-OU	0,4153	0,2546	BNSL IG-OU	0,5833	0,6031
BNS2 Ga-OU	0,3458	0,2433	BNS2 IG-OU	0,4551	0,3908
BNS2L Ga-OU	0,3253	0,2299	BNS2L IG-OU	0,5635	0,7025

Źródło: opracowanie własne.

Spośród wszystkich rozważanych modeli najmniejszą wartość kryterium AIC uzyskał model ze złożeniem procesów zmienności z rozkładem stacjonarnym odwrotnym Gaussa (BNS2 IG-OU). Modele z rozkładem stacjonarnym procesu wariancji chwilowej odwrotnym Gaussa mają nieco większe wartości funkcji wiarygodności niż z rozkładem gamma. Dla obu rozkładów uwzględnienie efektu dźwigni tylko nieznacznie zwiększa oszacowanie logarytmu funkcji wiarygodności. Większy wzrost oszacowania logarytmu funkcji wiarygodności można uzyskać poprzez uwzględnienie drugiego procesu zmienności.

Na rysunku 4.18 przedstawiono trajektorie oszacowań poszczególnych składowych wektora $\theta = [\mu, \beta, \gamma, \delta, \lambda_1, \lambda_2, w_1]^T$ względem kolejnych iteracji algorytmu iterowanej filtracji, a na rysunku 4.19 jednowymiarowe ciecia wiarygodności dla modelu BNS2 IG-OU. Wykresy diagnostyczne 4.18 i 4.19 potwierdzają uzyskanie zbieżności algorytmu iterowanej filtracji dla modelu BNS2 IG-OU.

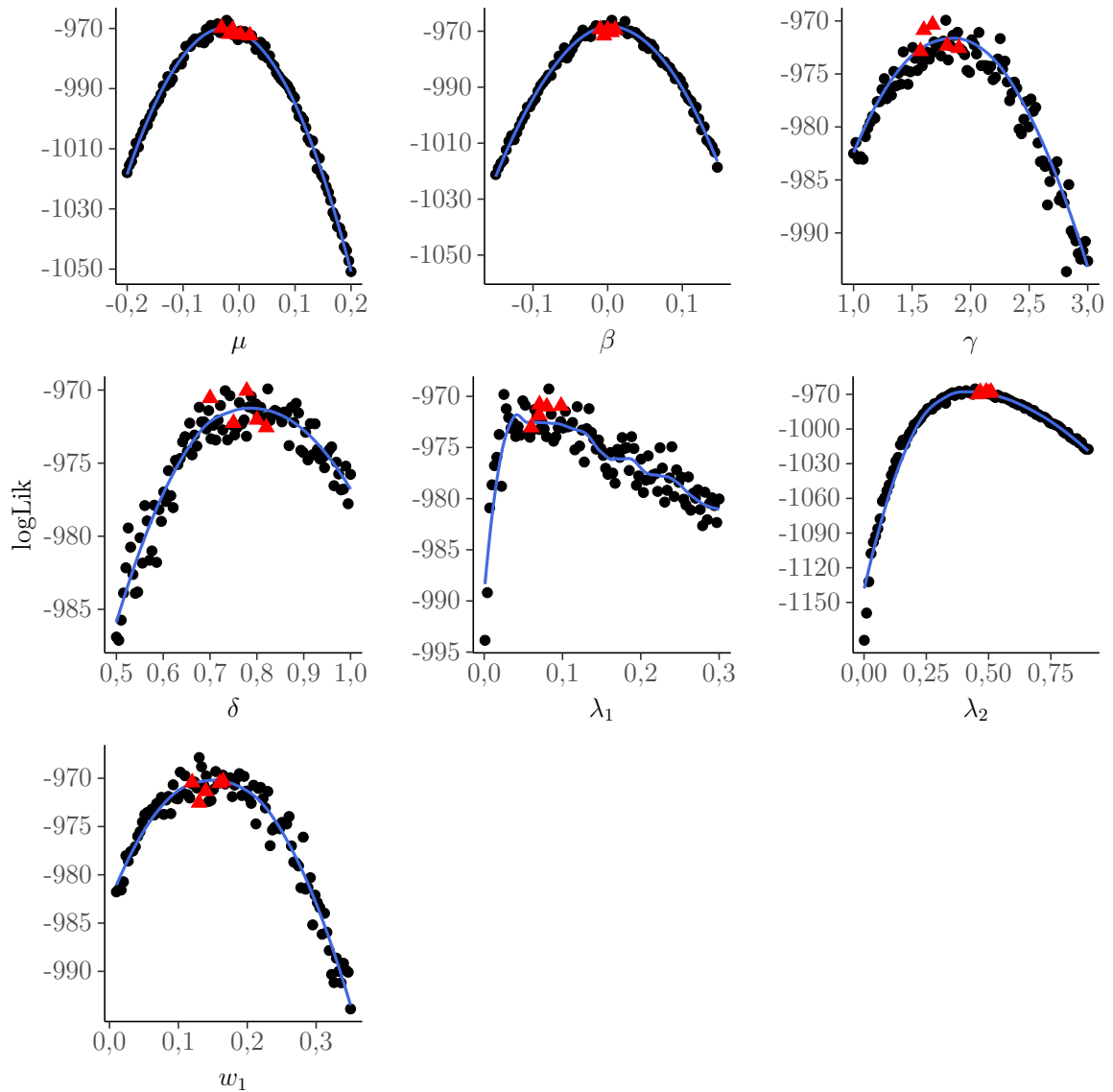
W tablicy 4.15 przedstawiono wyniki testów ilorazu wiarygodności dla czterech par hipotez dla modeli zagnieżdżonych. Oprócz dwóch rozważanych wcześniej (por.



Rysunek 4.18: Trajektorie oszacowań parametrów dla 5 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych w modelu ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU), każda po

$J = 200$ iteracji.

Źródło: opracowanie własne.



Rysunek 4.19: Oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów wyznaczone za pomocą filtru cząsteczkowego (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty). Model ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU).

Źródło: opracowanie własne.

(4.5) wzór oraz (4.6)) podano również wyniki dla par hipotez

$$\begin{aligned} H_0 : w_1 = \lambda_2 = 0 \\ H_1 : \sim H_0 \end{aligned} \quad (4.7)$$

oraz

$$\begin{aligned} H_0 : \rho_1 = 0 \\ H_1 : \sim H_0 \end{aligned} \quad (4.8)$$

W parze hipotez (4.7), przy prawdziwości hipotezy H_0 restrykcje umożliwiają uprościć model BNS2 do modelu BNS, natomiast w przypadku pary hipotez (4.8) model BNS2L do modelu BNSL. Wynik przedstawione w tablicy 4.15 potwierdzają spostrzeżenia, że efekt dźwigni można pominąć zarówno w modelu ze złożeniem procesów jak i bez dla obu rozkładów stacjonarnych (pary hipotez (4.5) oraz (4.8)). Natomiast więcej wątpliwości wzbudza uwzględnienie złożenia procesów zmienności. Dla obu rozkładów stacjonarnych procesu wariancji rozstrzygnięcie, czy należy odrzucić hipotezę o prawdziwości restrykcji zależy od wyboru poziomu istotności testu, przy czym p-wartość jest mniejsza w przypadku rozkładu odwrotnego Gaussa.

Tablica 4.15: Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla par hipotez zagnieżdżonych.

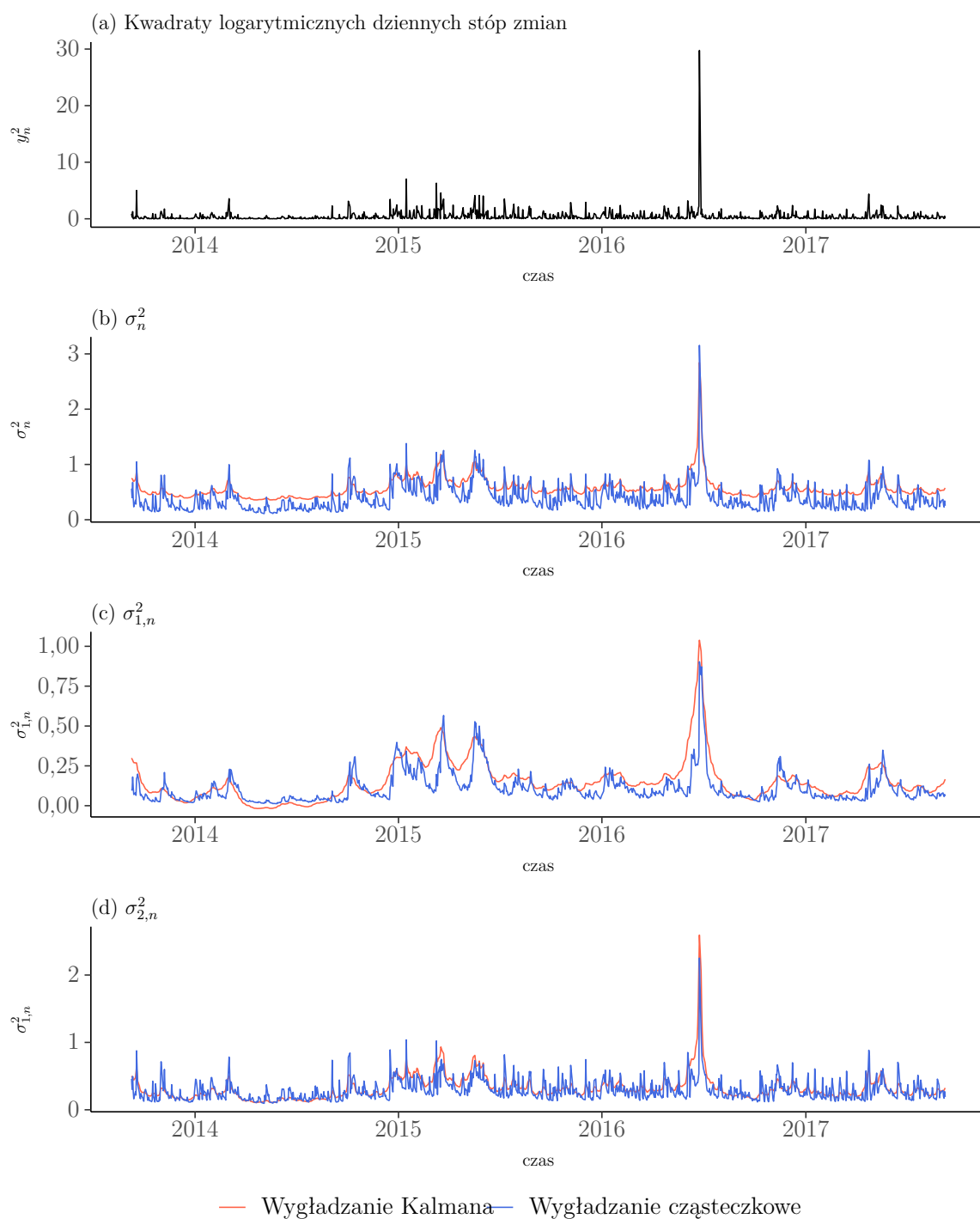
Para hipotez	Rozkład stacjonarny wariancji chwilowej	
	Gamma	Odwrotny Gaussa
$H_0 : \rho_1 = \rho_2 = 0$ $H_1 : \sim H_0$	0,384 (0,8254)	0,16 (0,9229)
$H_0 : \rho_2 = \lambda_2 = w_1 = 0$ $H_1 : \sim H_0$	5,065 (0,1671)	7,734 (0,5179)
$H_0 : w_1 = \lambda_2 = 0$ $H_1 : \sim H_0$	5,122 (0,0772)	7,599 (0,0224)
$H_0 : \rho_1 = 0$ $H_1 : \sim H_0$	0,441 (0,5071)	0,025 (0,8734)

Źródło: opracowanie własne.

Otrzymane wyniki metodą iterowanej filtracji wskazują, że w przypadku kursu USD/PLN można zastosować estymację metodą quasi-największej wiarygodności w

oparciu o filtr Kalmana. Po pierwsze, oszacowanie parametru β są bliskie zera. Po drugie, testy statystyczne nie odrzucają hipotez zakładających, że parametry ρ_1 , ρ_2 są równe 0. Jest to zgodne z uwagą z pracy Barndorff-Nielsen i Shephard (2001b), że w przypadku rynku walutowego przyjęcie założenia, że $\beta = \rho = 0$ może być odpowiednie. W tablicy 4.16 przedstawiono wyniki estymacji metodą quasi-największej wiarygodności w oparciu o filtr Kalmana. W nawiasach podano błędy standardowe wyznaczone numerycznie na podstawie wzorów asymptotycznych (por. (3.27)). Zastosowano modele przestrzeni stanów 3.1 oraz 3.2. Modele przestrzeni stanów dla filtru Kalmana nie rozróżniają rozkładu stacjonarnego procesu wariancji chwilowej, ponieważ są oparte jedynie na wartości oczekiwanej, odchyleniu standardowym i funkcji autokorelacji procesu wariancji chwilowej. Otrzymane wyniki są porównywalne do otrzymanych metodą iterowanej filtracji. Oszacowania parametru dryfu, wartości oczekiwane, odchylenia standardowego oraz wagi są nieco wyższe w przypadku metody quasi-największej wiarygodności. Natomiast oszacowania parametrów λ_1 i λ_2 są zbliżone.

Na rysunku 4.20 przedstawiono oszacowanie wariancji aktualnej dla modelu BNS2 IG-OU otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) oraz jej dwóch składowych w porównaniu do wartości kwadratów logarytmicznych stóp zmian kursu USD/PLN. Jako parametry modelu przyjęto oszacowaniu uzyskane metodą iterowanej filtracji przedstawione w tablicy 4.13. Dla porównania przedstawiono również wyniki wygładzania Kalmana dla modelu ze złożeniem dwóch procesów zmienności na podstawie oszacowań parametrów przedstawionych w tablicy 4.16. Oszacowania wariancji aktualnej wyznaczone za pomocą filtru Kalmana są wyższe niż wygładzania Kalmana. Może to się wiązać z wyższym oszacowaniem parametru ξ otrzymanym za pomocą metody quasi-największej wiarygodności.



Rysunek 4.20: Kwadraty dziennych logarytmicznych stóp zmian kursu USD/PLN (a) oraz oszacowanie wariancji aktualnej (b) wraz ze składowymi (c) i (d) w modelu ze złożeniem dwóch procesów zmienności o rozkładzie odwrotnym Gaussa (BNS2 IG-OU) otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) oraz wygładzania Kalmana.

Źródło: opracowanie własne.

Tablica 4.16: Wyniki estymacji parametrów modeli stochastycznej zmienności wyznaczone metoda quasi-największej wiarygodności. W nawiasach podano błędy standardowe oszacowań.

Model	μ	ξ	ω	λ_1	λ_2	w_1	loglik
BNS	0,0804 (0,0165)	0,6428 (0,0186)	0,4994 (0,0368)	0,0628 (0,0335)	-	-	-2618,953
BNS2	0,0853 (0,0516)	0,7528 (0,0493)	0,3969 (0,0743)	0,0795 (0,0735)	0,5833 (0,1274)	0,4012 (0,0153)	-1898,15

Źródło: opracowanie własne.

4.6 Wyniki estymacji dla indeksu WIG20

Dla indeksu WIG20 zastosowano jako zmienną obserwacji oszacowania wariancji zrealizowanej. Pozwala to znacznie zmniejszyć liczbę obserwacji w porównaniu do zastosowania bezpośrednio zwrotów śróddziennych, przy jednoczesnym wykorzystaniu informacji w nich zawartych.

Jako model przestrzeni stanów zastosowano reprezentację 3.7 dla modelu stochastycznej zmienności bez złożenia procesów zmienności oraz reprezentację 3.8 dla modelu ze złożeniem dwóch procesów zmienności. Wybrano dwa rozkłady stacjonarne procesu wariancji chwilowej gamma i odwrotny Gaussa. Wykorzystanie wariancji zrealizowanej umożliwia jedynie estymację parametrów rozkładu wariancji chwilowej oraz parametru persystencji λ , bez możliwości wyznaczenia parametrów dryfu czy premii za ryzyko. Tablica 4.17 przedstawia wyniki estymacji parametrów uzyskane dla rozkładu stacjonarnego gamma oraz odwrotnego Gaussa. Porównując wyniki dla obu rozkładów stacjonarnych procesu wariancji chwilowej można zauważyć zbliżone oszacowania parametrów λ_1 i λ_2 . Większe wartości logarytmu funkcji wiarygodności uzyskały model z rozkładem stacjonarnym odwrotnym Gaussa. Dla tego rozkładu zastosowanie złożenia procesów zmienności daje większy wzrost wartości logarytmu funkcji wiarygodności niż w przypadku rozkładu gamma. Kryterium informacyjnego AIC wskazuje jako najlepszy model ze złożeniem procesów zmienności i rozkładem stacjonarnym odwrotnym Gaussa. Dla tego modelu stochastycznej zmienności na rysunku 4.21 przedstawiono trajektorie oszacowań poszcze-

gólnych składowych względem kolejnych iteracji algorytmu iterowanej filtracji, a na rysunku 4.22 oszacowania jednowymiarowych cięć funkcji wiarygodności. Wszystkie pięć oszacowań mieszczą się w jednej jednostce logarytmu funkcji wiarygodności. Przedstawione rysunki potwierdza zbieżność algorytmu.

W tablicy 4.18 zestawiono uzyskane oszacowania wartości oczekiwanej (parametr ξ) i odchylenia standardowego (parametr ω) procesu wariancji chwilowej dla rozpatrywanych modeli. Wartości uzyskane dla rozkładu gamma są bardzo do siebie zbliżone. Natomiast wartości uzyskana dla rozkładu odwrotnego Gaussa dużo bardziej zależą od uwzględnienia drugiego procesu zmienności. W modelu ze złożeniem procesów zmienności pojawia się znaczny spadek oszacowań obu parametrów w porównaniu do modelu bez złożenia.

Tablica 4.17: Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej gamma i odwrotnym Gaussa wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.

Model	λ_1	λ_2	α	ν	γ	δ	w_1	loglik	AIC
BNS Ga-OU	0,0175 (0,0051)	-	5,9617 (0,5132)	5,3741 (0,6512)	-	-	-	-595,03 (0,2314)	-1196,06
BNS2 Ga-OU	0,0135 (0,0059)	0,3697 (0,1975)	5,6697 (0,4328)	5,1783 (0,5423)	-	-	0,2312 (0,0512)	-592,04 (0,2137)	-1194,08
BNS IG-OU	0,0147 (0,0069)	-	-	-	4,2774 (0,6113)	5,0196 (0,6113)	-	-594,05 (0,2176)	-1194,11
BNS2 IG-OU	0,0124 (0,0066)	0,4906 (0,2066)	-	-	5,4027 (0,7189)	4,2381 (0,6035)	0,1886 (0,0442)	-590,01 (0,2145)	-1190,02

Źródło: opracowanie własne.

Tablica 4.18: Wyniki estymacji wartości średniej (parametr ξ) oraz odchylenia standardowego (parametr ω) rozkładu stacjonarnego procesu wariancji chwilowej dla ośmiu rozważanych modeli stochastycznej zmienności z podziałem na rozkłady stacjonarne: gamma (Ga-OU) oraz odwrotnym Gaussa (IG-OU).

Model	Ga-OU		Model	IG-OU	
	ξ	ω		ξ	ω
BNS Ga-OU	0,9014	0,3888	BNS IG-OU	1,1735	0,2532
BNS2 Ga-OU	0,9133	0,4013	BNS2 IG-OU	0,7844	0,1639

Źródło: opracowanie własne.

Model BNS można traktować jako model z restrykcjami na przestrzeni parametrów dla modelu BNS2. Można zatem ponownie rozważać parę hipotez (4.7). Statystyki i p-wartości testu ilorazu wiarygodności przedstawia tablica 4.19. W przypadku rozkładu odwrotnego Gaussa wyniki testu wskazują na odrzucenie hipotezy H_0 na poziomie istotności 0,05. Natomiast p-wartość testu dla rozkładu gamma jest na granicy typowego poziomu istotności testu. Wyniki zatem mniej jednoznacznie wskazują na istotność uwzględnienia złożenia procesów zmienności.

W przypadku, gdy zmienną pomiaru jest wariancja zrealizowana można również zastosować filtr Kalmana i estymować parametry metodą quasi-największej wiarygodności. Odpowiednie modele przestrzeni stanu (3.7 dla modelu bez złożenia procesu oraz 3.4 dla modelu ze złożeniem) oparte są na wartości oczekiwanej, odchyleniu standardowym procesu i funkcji autokorelacji wariancji chwilowej. Nie zależą od wyboru rozkładu stacjonarnego procesu wariancji chwilowej. Wyniki estymacji przedstawia tablica 4.20. Porównując te wyniki z oszacowaniami uzyskanymi metodą iterowanej filtracji można zauważyć, że oszacowania wartości oczekiwanej procesu wariancji chwilowej są większe natomiast odchylenia standardowego mniejsze. Największa różnica jest w oszacowaniu wagi pierwszego składnika złożenia. Te uzyskane metodą quasi-największej wiarygodności jest znacznie większe niż w przypadku iterowanej filtracji. Natomiast wyniki estymacji parametrów λ_1 oraz λ_2 są

zbliżone.

Na rysunku 4.23 przedstawiono oszacowania wariancji aktualnej otrzymane za pomocą wygładzania cząsteczkowego (dla $K=1000$ cząsteczek, $L = 20$) oraz wygładzania Kalmana. Przyjęto oszacowania uzyskane odpowiednio algorytmem iterowanej filtracji (dla modelu BNS2 IG-OU) oraz metodą quasi-największej wiarygodności. Oszacowania wariancji aktualnej mimo różnic w przyjętych wartościach parametrów oraz różnych metod estymacji są bardzo zbliżone. Natomiast oszacowania poszczególnych składowych różnią się istotnie. Wpływa na to głównie duża różnica w oszacowaniach wagi pierwszej składowej.

Tablica 4.19: Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla pary hipotez zagnieżdżonych.

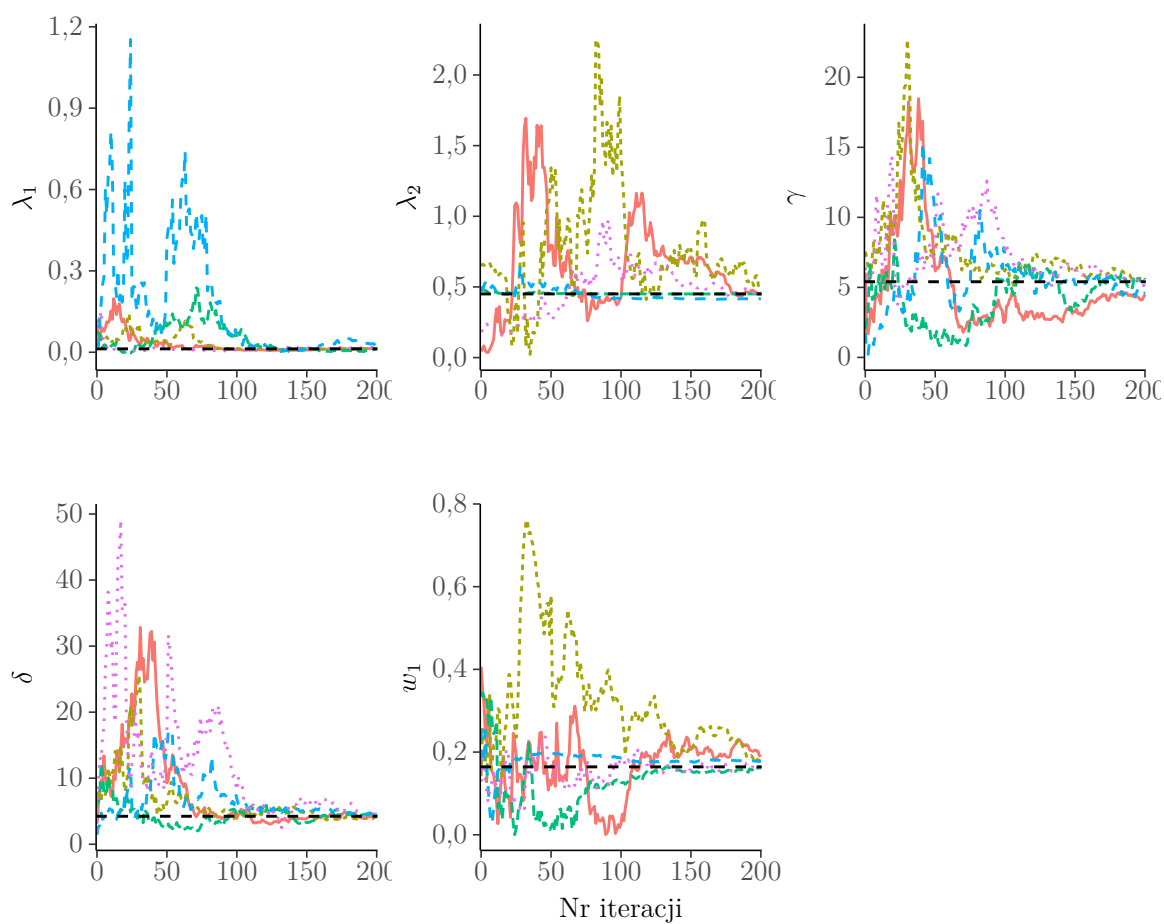
Para hipotez	Rozkład stacjonarny wariancji chwilowej	
	Gamma	Odwrotny Gaussa
$H_0 : w_1 = \lambda_2 = 0$	5,9801	8,0866
$H_1 : \sim H_0$	(0,0502)	(0,0175)

Źródło: opracowanie własne.

Tablica 4.20: Wyniki estymacji parametrów modeli stochastycznej zmienności wyznaczone metoda quasi-największej wiarygodności. W nawiasach podano błędy standardowe oszacowań.

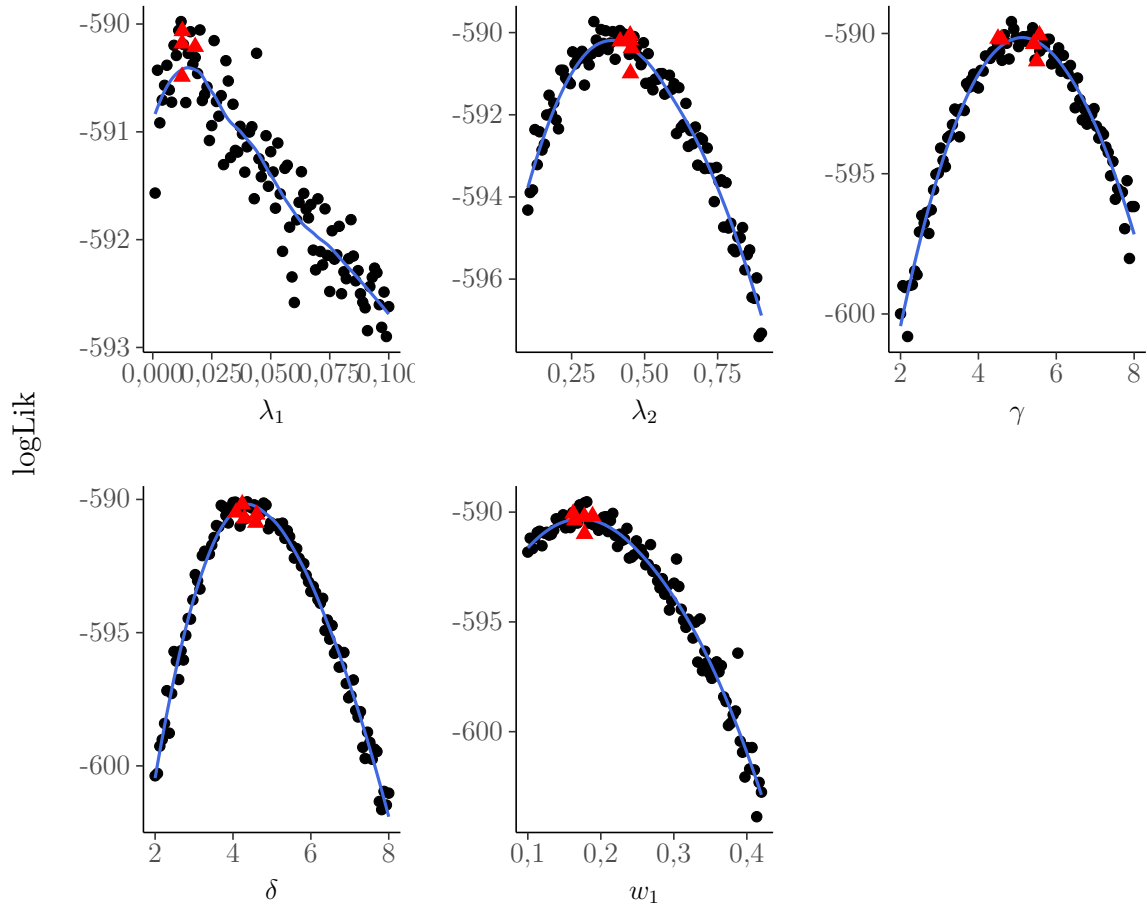
Model	ξ	ω	λ_1	λ_2	w_1	loglik
BNS	1,3714 (0,2791)	0,146 (0,0741)	0,0864 (0,0035)	-	-	-1309,477
BNS2	1,1142 (0,3419)	0,1327 (0,1098)	0,0821 (0,0308)	0,5128 (0,2279)	0,3314 (0,1241)	-1289,12

Źródło: opracowanie własne.



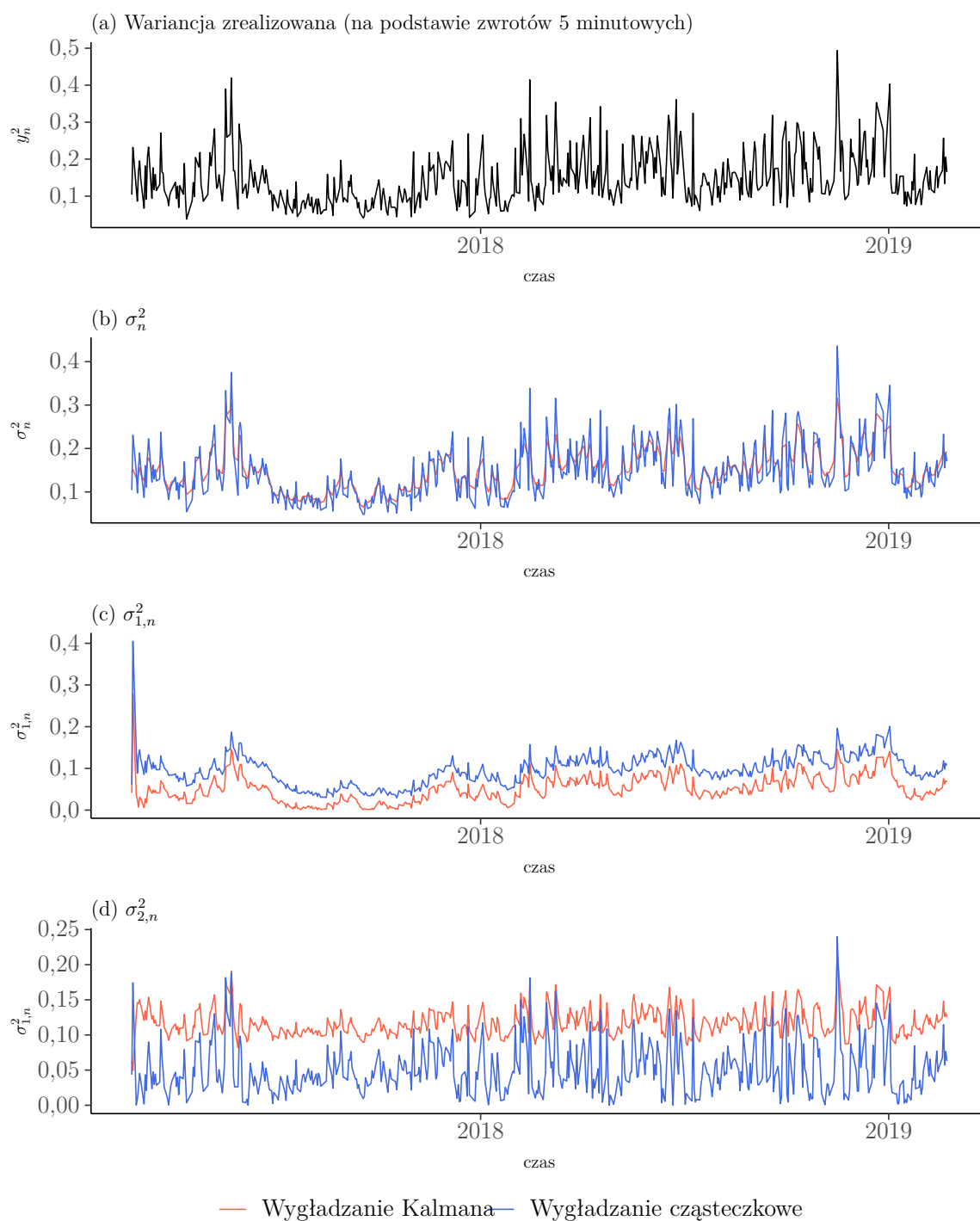
Rysunek 4.21: Trajektorie oszacowań parametrów dla 5 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych w modelu ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU), każda po $J = 200$ iteracji.

Źródło: opracowanie własne.



Rysunek 4.22: Oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów wyznaczone za pomocą filtra cząsteczkowego (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty). Model ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU).

Źródło: opracowanie własne.



Rysunek 4.23: Wariancja zrealizowana (a) oraz oszacowanie wariancji aktualnej (b) wraz ze składowymi (c) i (d) w model u ze złożeniem dwóch procesów zmienności o rozkładzie odwrotnym Gaussa (BNS2 IG-OU) otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) oraz wygładzania Kalmana.

Źródło: opracowanie własne.

Zakończenie

Głównym celem niniejszej pracy było przedstawienie niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka oraz opracowanie i zastosowanie metod opartych na filtrach Kalmana i filtrach cząsteczkowych do estymacji wariancji aktualnej i parametrów tych modeli. W ramach wymienionego celu głównego rozważane były trzy cele cząstkowe.

W rozdziale drugim zrealizowano pierwszy cel cząstkowy rozprawy – omówione zostały niegaussowskie modele stochastycznej zmienności typu Ornsteina-Uhlenbecka. Przedstawiono szczegółowo własności podstawowego modelu zaproponowanego przez Barndorff-Nielsen i Shephard (2001b), a następnie rozszerzenia tego modelu pozwalające na uchwycenie zależności długookresowych w procesie zmienności oraz korelacji pomiędzy procesem logarytmicznych cen i procesem zmienności, co pozwala na modelowanie efektu dźwigni. Zaprezentowano także powiązania modelu z estymatorami wariancji zrealizowanej.

Rozdział trzeci podejmuje drugi cel cząstkowy – zagadanie estymacji zmienności i parametrów modeli niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka za pomocą filtru Kalmana i filtrów cząsteczkowych. Przedstawiono liniowe modele przestrzeni stanów, które pozwalają wykorzystać filtr Kalmana do estymacji wariancji aktualnej oraz zastosować metodę quasi-największej wiarygodności do estymacji parametrów zarówno w przypadku, gdy zmienną pomiaru są zwroty logarytmiczne jak i wariancja zrealizowana. Ograniczeniem stosowania filtrów Kalmana jest brak możliwości przedstawienia niegaussowskich modeli stochastycznej zmienności w postaci liniowej przestrzeni stanów w przypadku, gdy parametr

β jest różny od 0. Ponadto estymacja metodą quasi-największej wiarygodności wymaga przyjęcia upraszczającego, ale nieprawdziwego założenia o normalnym rozkładzie zaburzeń. W takim przypadku estymatory pozostają zgodne, ale ich własności w małych próbach nie są znane. Następnie omówione zostały filtry cząsteczkowe jako narzędzie estymacji w przypadku nieliniowych niegaussowskich modeli przestrzeni stanów. Przedstawiono odpowiednie modele przestrzeni stanów pozwalające na estymację wariancji aktualnej w ogólnej postaci modelu stochastycznej zmienności zawierającym złożenie procesów zmienności i uwzględniającym efekt dźwigni. Zastosowanie filtrów cząsteczkowych wymaga opracowania metod symulacji kolejnych generacji cząsteczek. W kolejnym podrozdziale szczegółowo przeanalizowano to zagadnienie dla różnych rozkładów stacjonarnych procesu Ornsteina-Uhlenbecka. Rozważaniom teoretycznym towarzyszyło opracowanie w języku programowania R odpowiednich skryptów. Autorską propozycją jest wykorzystanie do estymacji parametrów metody iterowanej filtracji, która pozwala na wykorzystanie filtrów cząsteczkowych do estymacji parametrów. Zaprezentowano teoretyczne podstawy iterowanej filtracji oraz metody diagnostyczne pozwalające zweryfikować, czy udało się osiągnąć zbieżność algorytmu do wartości estymatora metody największej wiarygodności.

Przedstawione nieliniowe niegaussowskie modele przestrzeni stanów dla modeli ze złożeniem procesów zmienności i uwzględniające efekt dźwigni wraz z przedstawionymi metodami symulacji kolejnych generacji cząsteczek dla różnych rozkładów stacjonarnych procesów wariancji chwilowej w połączeniu z algorytmem iterowanej filtracji pozwalają na estymację parametrów modeli w przypadkach, które do tej pory były niedostępne w klasycznym wnioskowaniu. Jest to także istotne rozszerzenie propozycji przedstawionych w pracach Szczepocki (2019a) oraz Szczepocki (2019b), w których zastosowano iterowaną filtrację jedynie w przypadku procesu wariancji chwilowej o rozkładzie stacjonarnym gamma bez złożenia procesów zmienności i uwzględnienia efektu dźwigni.

Trzeciemu celowi cząstkowemu poświęcony został czwarty rozdział rozprawy. Przedstawiono wyniki estymacji zmienności i parametrów niegaussowskich modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka dla trzech szeregów czasowych pochodzących z polskiego rynku finansowego: notowań Warszawskiego Indeksu

Giełdowego (WIG), indeksu dwudziestu największych spółek akcyjnych notowanych na warszawskiej Giełdzie Papierów Wartościowych (WIG20) oraz kursu USD/PLN. Należy zwrócić uwagę, że są to jedne z pierwszych prac dotyczące zastosowania niegaussowskich modeli stochastycznej zmienności do analizy finansowych szeregów pochodzących z polskiego rynku finansowego. Wcześniej w pracy Szczepocki (2018) przedstawiono wyniki estymacji metodą quasi-największej wiarygodności dla kursu EUR/PLN (kursu Euro wyrażonego w złotych).

W przypadku szeregów czasowych indeksu WIG i kursu dolara w analizie wykorzystano dane dzienne, natomiast w przypadku indeksu WIG20 dane śróddzienne. Punktem wyjścia była statystyczna analiza szeregów czasowych, która pozwoliła wybrać dwa rozkłady jako potencjalne rozkłady stacjonarne procesu wariancji chwilowej: gamma i odwrotny Gaussa. W dwóch przypadkach: kursu WIG kryterium informacyjne Akaike AIC wskazało, że nieco lepiej odzwierciedlają dane modele stochastycznej zmienności o rozkładzie stacjonarnym gamma, a w przypadku – kursu USD/PLN oraz indeksu WIG20 o rozkładzie odwrotnym Gaussa. Na podstawie kryterium informacyjne Akaike AIC oraz testów statystycznych okazało się, że najlepszym modelem w przypadku kursu indeksu WIG jest model ze złożeniem dwóch procesów zmienności oraz uwzględniający efekt dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU). Natomiast w przypadku kursu USD/PLN oraz WIG20 modele ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa.

Przeprowadzonych w pracy rozważania teoretycznych i wyniki badań empirycznych pozwalają na pozytywną weryfikację wszystkich trzech hipotez. Niegaussowskie modele stochastycznej zmienności typu Ornsteina-Uhlenbecka pozwalają dobrze odzwierciedlać obserwowane własności szeregów czasowych: grupowanie zmienności, brak autokorelacji stóp zwrotu i powoli malejąca korelacja kwadratów stóp zwrotu, efekt dźwigni oraz bezwarunkowe rozkłady stóp zwrotu zbliżone do obserwowanych empirycznie. Dla przedstawionych szeregów czasowych pochodzących z polskiego rynku finansowego udało się dobrać rozkłady wariancji chwilowej. Przedstawione wyniki badań empirycznych w szczególności wykresy diagnostyczne wskazują, że udało osiągnąć się osiągnąć zbieżność algorytmu do wartości estymatora metody

największej wiarygodności.

Celem dalszych prac jest porównanie jakości oszacowań zmienności otrzymanych za pomocą niegaussowskich modeli stochastycznych z modelami GARCH i innymi modelami SV. W literaturze przedmiotu można znaleźć bardzo wiele specyfikacji modeli GARCH i modeli SV. W celu porównania wyników należałoby przeprowadzić bardzo rozbudowane badanie porównawcze. Ponadto musiałoby uwzględnić kontekst zagadnienia (prognoza zmienności, szacowanie ryzyka, wycena instrumentów pochodnych itp.). Realizacja tak rozbudowanego przedsięwzięcia jest celem przyszłych badań autora.

Celem dalszych badań są również wielowymiarowe modele stochastycznej zmienności. Pigorsch i Stelzer (2009) zaproponowali uogólnienie rozważanych w tej pracy nieujemnych procesów Ornsteina-Uhlenbecka do wielowymiarowych dodatnio określonych (*positive semidefinite*) procesów Ornsteina-Uhlenbecka. Barndorff-Nielsen i Stelzer (2013) rozszerzyli wielowymiarowe dodatnio określone procesy Ornsteina-Uhlenbecka na przypadek złożenia skończonej ilości takich procesów. Modele te ciągle mają jednak głównie znaczenie teoretyczne. Mało jest prac dotyczących estymacji. W jednej z nielicznych, Raknerud i Skare (2010) zaproponowali użycia wykorzystywanych w tej pracy filtrów Kalmana i estymację metodą quasi-największej wiarygodności.

Przedmiotem dalszych prac może być również zastosowanie iterowanej filtracji estymacji parametrów innych modeli stochastycznej m.in. Hestona, czy modeli, w którym logarytm procesu zmienności jest gaussowskim procesem Ornsteina-Uhlenbecka.

Rozkłady bezwarunkowe logarytmicznych stóp zwrotu

W niniejszej pracy wykorzystano następujące rozkłady do modelowania bezwarunkowych rozkładów stóp zwrotu: uogólniony rozkład hiperboliczny oraz dwa jego specjalne przypadki: rozkład normalny odwrotny Gaussa NIG (*Normal Inverse Gaussian*) oraz VG (*Variance Gamma*)¹.

1. Uogólniony rozkład hiperboliczny

Uogólniony rozkład hiperboliczny z parametrami $\mu \in \mathbb{R}$, $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$, $\delta \in \mathbb{R}$, $\nu \in \mathbb{R}$ ma gęstość daną wzorem

$$f_{GH}(x, \mu, \alpha, \beta, \delta, \nu) = a(x, \alpha, \beta, \delta, \nu) \left(\delta^2 + (x - \mu)^2 \right)^{(\nu-1/2)/2} \\ \times K_{\nu-1/2} \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right) \exp(\beta(x - \mu)),$$

$$a(x, \alpha, \beta, \delta, \nu) = \frac{(\alpha^2 - \beta^2)^{\nu/2}}{\sqrt{2\pi} \alpha^{\nu-1/2} \delta^\nu K_\nu(\delta \sqrt{\alpha^2 - \beta^2})}$$

gdzie K_ν jest zmodyfikowaną funkcją Bessela trzeciego rodzaju oraz:

$$\delta \geq 0, \quad |\beta| < \alpha \quad \text{jeśli} \quad \nu > 0$$

¹W literaturze przedmiotu rozważa się także inne rozkłady logarytmicznych stóp zwrotu, m.in.: α -stabilne (Mandelbrot, 1963), *Meixner* (Schoutens i Teugels, 2001), CGMY (Carr i in., 2002), uogólnione rozkład wartości ekstremalnych (*Generalized extreme value distribution*) (Makhwiting i in., 2014).

$$\delta > 0, \quad |\beta| < \alpha \quad \text{jeśli} \quad \nu = 0$$

$$\delta > 0, \quad |\beta| \leq \alpha \quad \text{jeśli} \quad \nu < 0.$$

Parametr μ jest parametrem położenia, β asymetrii, a δ skali.

Specjalnymi przypadkami uogólnionego rozkładu hiperbolicznego są rozkłady normalny odwrotny Gaussa NIG: $GH(\mu, \alpha, \beta, \delta, -1/2) = NIG(\mu, \alpha, \beta, \delta)$ oraz VG (dla $\nu = \sigma^2/\nu, \alpha = \sqrt{2/\nu + (\theta^2/\sigma^4)}, \beta = \theta/\sigma^2$ i $\delta \rightarrow 0$).

Uogólniony rozkład hiperboliczny jest rozkładem będącym normalną mieszaniną średnio-wariacyjną, dla którego rozkładem miksującym jest uogólniony rozkład odwrotny Gaussa (por. twierdzenie 1.6).

Funkcja charakterystyczna uogólnionego rozkładu hiperbolicznego przyjmuje postać

$$\varphi_{GH}(\zeta; \mu, \alpha, \beta, \delta, \nu) = e^{i\mu\zeta} \left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + i\zeta)^2} \right)^{\nu/2} \frac{K_\nu(\sqrt{\delta(\alpha^2 - (\beta + i\zeta)^2)})}{K_\nu(\delta\sqrt{\alpha^2 - \beta^2})}$$

Wartość oczekiwana i wariancja zmiennej losowej X o uogólnionym rozkładzie hiperbolicznym są równe odpowiednio

$$\mathbb{E}X = \mu + \frac{\beta\delta}{\alpha^2 - \beta^2} \frac{K_{\nu+1}(\delta\sqrt{\alpha^2 - \beta^2})}{K_\nu(\delta\sqrt{\alpha^2 - \beta^2})}$$

oraz

$$\begin{aligned} \text{Var}(X) = & \delta^2 \left(\frac{K_{\nu+1}(\delta\sqrt{\alpha^2 - \beta^2})}{\delta\sqrt{\alpha^2 - \beta^2} K_\nu(\delta\sqrt{\alpha^2 - \beta^2})} \right. \\ & \left. + \frac{\beta^2}{\alpha^2 - \beta^2} \left(\frac{K_{\nu+2}(\delta\sqrt{\alpha^2 - \beta^2})}{K_\nu(\delta\sqrt{\alpha^2 - \beta^2})} - \frac{K_{\nu+1}^2(\delta\sqrt{\alpha^2 - \beta^2})}{K_\nu^2(\delta\sqrt{\alpha^2 - \beta^2})} \right) \right). \end{aligned}$$

Uogólniony rozkład hiperboliczny jest rozkładem nieskończenie podzielny. Można z nim związać następujący proces Lévy'ego.

Definicja A.1.

Proces Lévy'ego $(L(t))_{t \geq 0}$, dla którego przyrosty $L(t+s) - L(t)$ (dla dowolnych $t \geq 0$ oraz $s > 0$) mają rozkład o funkcji charakterystycznej $(\varphi_{GH}(\zeta; \delta, \alpha, \beta))^s$ nazywamy uogólnionym procesem hiperbolicznym.

2. Rozkład normalny odwrotny Gaussa NIG

Rozkład normalny odwrotny Gaussa z parametrami $\mu \in \mathbb{R}$, $\alpha > 0$, $-\alpha < \beta < \alpha$, $\delta > 0$ ma gęstość daną wzorem

$$f_{NIG}(x; \mu, \alpha, \beta, \delta) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right) \frac{K_1(\alpha\sqrt{\delta^2 + (x - \mu)^2})}{\sqrt{\delta^2 + (x - \mu)^2}}.$$

Parametr μ jest parametrem położenia, β asymetrii, δ skali, a α „ciężkości” ogonów rozkładu. Rozkład NIG jest mieszaniną średnio-wariacyjną, dla którego rozkładem miksującym jest rozkład odwrotny Gaussa (por. twierdzenie 1.6). Jako pierwszy dowód przedstawił Barndorff-Nielsen (1978).

Funkcja charakterystyczna rozkładu $NIG(\mu, \alpha, \beta, \delta)$ przyjmuje postać

$$\varphi_{NIG}(\zeta; \mu, \alpha, \beta, \delta) = \exp\left(\mu i\zeta - \delta\left(\sqrt{\alpha^2 - (\beta + i\zeta)^2} - \sqrt{\alpha^2 - \beta^2}\right)\right).$$

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie $NIG(\mu, \alpha, \beta, \delta)$ są równe odpowiednio

$$\mathbb{E}X = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} \quad \text{oraz} \quad \text{Var}(X) = \frac{\alpha^2\delta}{(\alpha^2 - \beta^2)^{3/2}}.$$

Jako specjalny przypadek uogólnionego rozkładu hiperbolicznego jest również rozkładem nieskończenie podzielny i można związać z nim proces Lévy’ego.

Definicja A.2.

Proces Lévy’ego $(L(t))_{t \geq 0}$, dla którego przyrosty $L(t+s) - L(t)$ (dla dowolnych $t \geq 0$ oraz $s > 0$) mają rozkład $NIG(\mu, \alpha, \beta, s\delta)$ nazywamy procesem normalnym odwrotnym Gaussa NIG.

Zgodnie z twierdzeniem 1.8 proces NIG można przedstawić jako proces Wienera ze „zmianą czasu”, w którym procesem podporządkowanym jest proces odwrotny Gaussa.

3. Rozkład VG

Rozkład VG z parametrami $\mu \in \mathbb{R}$, $\sigma > 0$, $\nu > 0$, $\theta \in \mathbb{R}$ ma gęstość daną wzorem

$$f_{VG}(x; \mu, \sigma, \nu, \theta) = \frac{1}{\sigma\sqrt{\pi}\Gamma(\nu/2)} e^{\frac{\theta(x-\mu)}{\sigma^2}} \left(\frac{|x-\mu|}{2\sqrt{\theta^2 + \sigma^2}}\right)^{\frac{\nu-1}{2}} K_{\frac{\nu-1}{2}}\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|x-\mu|\right)$$

Rozkład VG jest mieszaniną średnio-wariacyjną, dla którego rozkładem miksującym jest rozkład gamma (por. twierdzenie 1.6).

Funkcja charakterystyczna rozkładu $VG(\mu, \sigma, \nu, \theta)$ przyjmuje postać

$$\varphi_{VG}(\zeta; \mu, \sigma, \nu, \theta) = e^{i\mu\zeta} \left(1 - i\zeta\theta\nu + \frac{1}{2}\sigma^2\nu\zeta^2 \right)^{-1/\nu}.$$

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie $VG(\mu, \sigma, \nu, \theta)$ są równe odpowiednio

$$\mathbb{E}X = \mu + \theta \quad \text{oraz} \quad \text{Var}(X) = \sigma^2 + \nu\theta^2.$$

Podobnie jak rozkład NIG, rozkład VG jako specjalny przypadek uogólnionego rozkładu hiperbolicznego jest rozkładem nieskończenie podzielny i można związać z nim proces Lévy'ego.

Definicja A.3.

Proces Lévy'ego $(L(t))_{t \geq 0}$, dla którego przyrosty $L(t+s) - L(t)$ (dla dowolnych $t \geq 0$ oraz $s > 0$) mają rozkład $VG(\mu, \sigma\sqrt{s}, \nu/s, s\theta)$ nazywamy procesem VG.

Proces VG można przedstawić jako proces Wienera ze „zmianą czasu”, w którym procesem podporządkowanym jest proces gamma (por. twierdzenie 1.8).

Symulacja niegaussowskich procesów Ornsteina-Uhlenbecka

Kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \ \eta_{2n}]^T$ dla procesu Ga-OU:

```
eta.vec<-function(lambda,alpha,ni,Delta){  
  
  ci=0  
  licznik=0  
  eta1=0  
  eta2=0  
  while(ci<1){  
    licznik=licznik+1  
    ci<-ci+rexp(1, rate = lambda*Delta*ni)  
    ri=runif(1)  
    if(ci>1) break  
    eta1=eta1+log(1/ci)*exp(lambda*Delta*ri)  
    eta2=eta2+log(1/ci)  
  }  
  if(licznik==1){  
    eta1=0  
    eta2=0  
  } else{  
    eta1=1/alpha*exp(-lambda*Delta)*eta1  
    eta2=1/alpha*eta2  
  }  
}
```

```

}
  return(c(eta1,eta2))
}

```

Kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ dla procesu GIG-OU:

```

eta.vec<-function(J,lambda,Delta,delta,gamma,nu){

g<-function(x,gamma){
  2/(x*pi^2)/(besselJ(sqrt(x), nu=abs(gamma))^2+
             bessely(sqrt(x), nu=abs(gamma))^2)
}

integrant<-function(zeta,z,gamma,nu){
  exp(-z*zeta/(2*delta^2))*g(zeta,gamma)
}

inverse<-function(z,gamma,nu,alpha){
  (1/2*(integrate(integrant,0,Inf,z,gamma,nu)$value)+max(0,nu))*exp(-gamma^2*z/2)
}

xx<-c(seq(0.0000001,.01,length.out = 100),seq(.01,1,length.out = 10))

dt = data.table(xx, val=xx)
dt=dt[order(-xx)]
setkey(dt, xx)

  eta<-c(0,0)
  ri=runif(J)
  ai=rexp(J)
  ai=cumsum(ai)
  temp=0
  for(j in 1:J){
    temp=xx[length(xx)+1-dt[J(ai[j]/lambda/Delta), roll = "nearest", which = TRUE]]
    eta[1]=eta[1]+temp*exp(lambda*Delta*ri[j])
    eta[2]=eta[2]+temp
  }
  return(c(eta[1]*exp(-lambda*Delta),eta[2]))
}

```

```
}
```

Kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ dla procesu TS-OU:

```
eta.vec<-function(J,lambda,Delta, kappa,nu,alpha){

  eta<-c(0,0)
  ri=runif(J)
  ai=rexp(J)
  ai=cumsum(ai)

  W<-function(x){
    A=nu*kappa*2^kappa/gamma(1-kappa)
    B=kappa
    C=(alpha^(1/kappa))/2
    (A/x)^(1/B)*exp(-lambertW(C/B*(A/x)^(1/B)))
  }

  temp=0
  for(j in 1:J){
    temp= W((ai[j]/(lambda*Delta)))
    eta[1]=eta[1]+temp*exp(lambda*Delta*ri[j])
    eta[2]=eta[2]+temp
  }
  return(c(eta[1]*exp(-lambda*Delta),eta[2]))
}
```

Kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ dla procesu IG-OU:

```
eta.vec<-function(J,lambda,Delta,delta,gamma){

  eta<-c(0,0)
  ri=runif(J)
  ai=rexp(J)
  ai=cumsum(ai)

  for(j in 1:J){
```

```

    temp=1/(gamma^2)*lambertW(delta^2*gamma^2/(2*pi*(ai[j]/(lambda*Delta))))
    eta[1]=eta[1]+temp*exp(lambda*Delta*ri[j])
    eta[2]=eta[2]+temp
  }
  return(c(eta[1]*exp(-lambda*Delta),eta[2]))
}

```

Na rysunku (3.6) przedstawiono symulację trajektorii procesu IG-OU za pomocą powyższego kodu dla $J=100$.

Kod w języku programowania R zwracający wektor $\eta_n = [\eta_{1n} \quad \eta_{2n}]^T$ dla procesu OU-Ga:

```

eta.vec<-function(hbar,lambda,Delta,nu,alpha){

  eta<-c(0,0)
  partition=seq(0,Delta,by = hbar)
  J=length(partition)
  temp=0
  for(j in 1:J){
    temp=rgamma(1,shape=hbar*nu,rate=alpha)
    eta[1]=eta[1]+temp*exp(lambda*Delta*partition[j])
    eta[2]=eta[2]+temp
  }
  return(c(eta[1]*exp(-lambda*Delta),eta[2]))
}

```

Symulacja procesu logarytmicznych cen

Symulacja procesu logarytmicznych cen w niegaussowskich modelach stochastycznej zmienności typu Ornsteina-Uhlenbecka

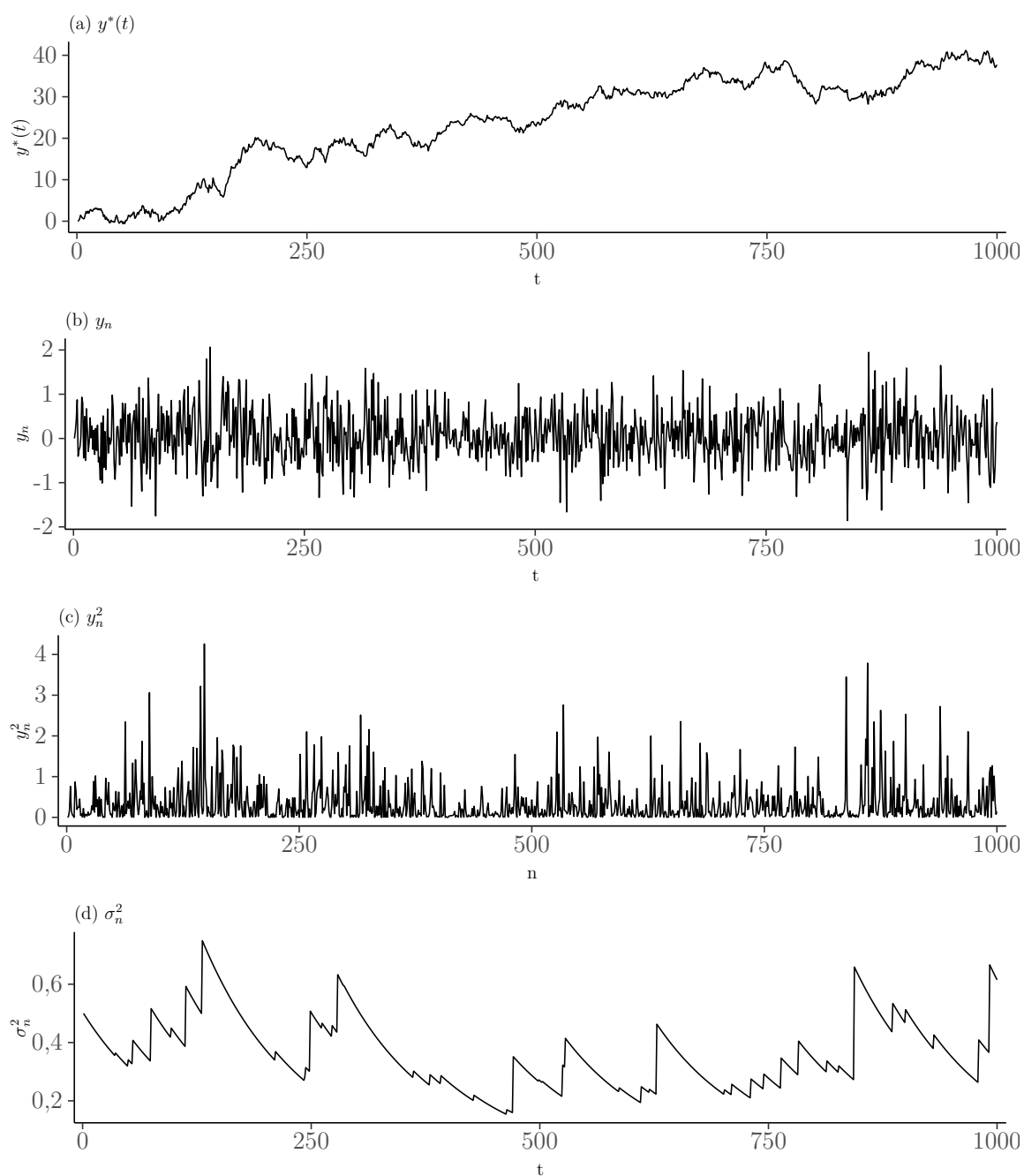
Na podstawie symulacji procesu wariancji chwilowej wraz z procesem BDLP dla momentów czasu $t = \Delta n$, dla $n = 0, 1, 2, \dots$ można dokonać symulacji procesu logarytmu cen. W szczególności dla modelu BNS z efektem dźwigni postaci

$$\begin{aligned} dY &= (\mu + \beta\sigma^2(t)) dt + \sigma(t)dW(t) + \rho d\bar{Z}(\lambda t) \quad Y(0) = 0, \\ d\sigma^2(t) &= -\lambda\sigma^2(t)dt + dZ(\lambda t), \quad \sigma^2(0) > 0. \end{aligned}$$

można otrzymać symulację procesu logarytmu cen za pomocą dyskretyzacji

$$\begin{aligned} Y(\Delta n) &= Y(\Delta(n-1)) + (\mu + \beta\sigma^2(\Delta n)) \Delta + \sqrt{\sigma^2(\Delta n)} \Delta Y_n + \\ &+ \rho (Z(\lambda \Delta n) - Z(\lambda \Delta(n-1)) - \lambda \Delta \xi) \quad Y(0) = 0, \end{aligned} \tag{C.1}$$

gdzie ξ jest wartością oczekiwaną rozkładu stacjonarnego procesu wariancji chwilowej, $Y_n \sim N(0, 1)$ dla $t = \Delta n$, dla $n = 0, 1, 2, \dots$. Na rysunku (C.1) przedstawiono symulację procesów logarytmów cen z efektem dźwigni na podstawie procesu wariancji z rozkładem stacjonarnym gamma.



Rysunek C.1: Symulacja trajektorii procesu logarytmu cen (a) oraz odpowiadające tej trajektorii stopy zwrotu (b), kwadraty stóp zwrotu (c), proces wariancji aktualnej (d). Wykorzystano proces wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,5$ i wariancji $\omega = 0,2$. Pozostałe parametry $\mu = 0$, $\beta = 0,05$, $\rho = -0,1$.

Źródło: opracowanie własne.

Bibliografia

- Akaikei H. (1973), Information theory and an extension of maximum likelihood principle, *Proc. 2nd Int. Symp. on Information Theory*, s. 267–281.
- Albrecher H., Predota M. (2004), On Asian option pricing for NIG Lévy processes, *Journal of Computational and Applied Mathematics*, 172 (1), s. 153–168.
- Alizadeh S., Brandt M. W., Diebold F. X. (2002), Range-based estimation of stochastic volatility models, *The Journal of Finance*, 57 (3), s. 1047–1091.
- Andersen T. G., Benzoni L., Lund J. (2002), An empirical investigation of continuous-time equity return models, *The Journal of Finance*, 57 (3), s. 1239–1284.
- Andersen T. G., Bollerslev T. (1998), Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, s. 885–905.
- Andersen T. G., Bollerslev T., Diebold F. X. (2001), The distribution of realized stock return volatility, *Journal of Financial Economics*, 61 (1), s. 43–76.
- Andersen T. G., Bollerslev T., Diebold F. X., Labys P. (1999), Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian, *Multinational Finance Journal*, 4, s. 159–179.
- Andrieu C., Doucet A., Holenstein R. (2010), Particle Markov Chain Monte Carlo Methods, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72 (3), s. 269–342.

- Applebaum D. (2009), *Lévy processes and stochastic calculus*, Cambridge University Press.
- Barndorff-Nielsen O. E. (1978), Hyperbolic distributions and distributions on hyperbolae, *Scandinavian Journal of statistics*, s. 151–157.
- Barndorff-Nielsen O. E. (1997), Processes of Normal Inverse Gaussian type, *Finance and Stochastics*, 2 (1), s. 41–68.
- Barndorff-Nielsen O. E., Cox D. R., *Inference and asymptotics*, Chapman & Hall London.
- Barndorff-Nielsen O. E., Hansen P. R., Lunde A., Shephard N. (2008), Designing realized kernels to measure the ex post variation of equity prices in the presence of noise, *Econometrica*, 76 (6), s. 1481–1536.
- Barndorff-Nielsen O. E., Shephard N. (1998), Aggregation and Model Construction for Volatility Models, *Department of Mathematical Sciences*.
- Barndorff-Nielsen O. E., Shephard N. (1999), Incorporation of a leverage effect in a stochastic volatility model, *Lévy Processes: Theory and Applications*, s. 161–189.
- Barndorff-Nielsen O. E., Shephard N. (2001), Normal modified stable processes, *Theory of Probability and Mathematical Statistics*, 1, s. 19–29.
- Barndorff-Nielsen O. E., Shephard N. (2001a), How accurate is the asymptotic approximation to the distribution of realised volatility?, (No. 2001-W16). Economics Group, Nuffield College, University of Oxford.
- Barndorff-Nielsen O. E., Shephard N. (2001b), Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63 (2), s. 167–241.
- Barndorff-Nielsen O. E., Shephard N. (2002), Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64 (2), s. 253–280.

- Barndorff-Nielsen O. E., Shephard N. (2003), Integrated OU Processes and Non-Gaussian OU-based Stochastic Volatility Models, *Scandinavian Journal of Statistics*, 30 (2), s. 277–295.
- Barndorff-Nielsen O. E., Shephard N. (2004), Power and bipower variation with stochastic volatility and jumps, *Journal of Financial Econometrics*, 2 (1), s. 1–37.
- Barndorff-Nielsen O. E., Stelzer R. (2013), The multivariate supOU stochastic volatility model, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 23 (2), s. 275–296.
- Beaulieu J. M., Jhwaeng D. C., Boettiger C., O’Meara B. C. (2012), Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution, *Evolution: International Journal of Organic Evolution*, 66 (8), s. 2369–2383.
- Behme A. (2011), *Generalized Ornstein-Uhlenbeck Processes and Extensions*, praca doktorska, TU Braunschweig.
- Benth F. E. (2011), The stochastic volatility model of Barndorff-Nielsen and Shephard in commodity markets, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 21 (4), s. 595–625.
- Benth F. E., Eriksson M., Westgaard S (2017), Stochastic Volatility Modeling of Emission Allowances Futures Prices in the European Union Emission Trading System Market, *Real Options in Energy and Commodity Markets*, s. 63–115.
- Benth F. E., Groth M., Kufakunesu R. (2007), Valuing Volatility and Variance Swaps for a Non-Gaussian Ornstein–Uhlenbeck Stochastic Volatility Model, *Applied Mathematical Finance*, 14 (4), s. 347–363.
- Benth F. E., Karlsen K. H., Reikvam K. (2003), Merton’s portfolio optimization problem in a Black and Scholes market with non-Gaussian stochastic volatility of Ornstein-Uhlenbeck type, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 13 (2), s. 215–244.

- Bhadra A., Ionides, E. L., Laneri K., Pascual M., Bouma M., Dhiman R. C. (2011), Malaria in Northwest India: Data analysis via partially observed stochastic differential equation models driven by Lévy noise, *Journal of the American Statistical Association*, 106 (494), s. 440–451.
- Black F. (1976), Studies of Stock Price Volatility Changes, *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section*, s. 177–181.
- Blackwood J. C., Streicker D. G., Altizer S., Rohani P. (2013), Resolving the roles of immunity, pathogenesis, and immigration for rabies persistence in vampire bats, *Proceedings of the National Academy of Sciences*, 110 (51), s. 20837–20842.
- Blæsild P. (1978), The shape of the generalized inverse Gaussian and hyperbolic distributions, Department of Theoretical Statistics, Inst., Univ.
- Blomberg S. P. (2017), Beyond Brownian motion and the Ornstein-Uhlenbeck process: Stochastic diffusion models for the evolution of quantitative characters, *bioRxiv*, s. 67–363.
- Bondesson L. (1982), On simulation from infinitely divisible distributions, *Advances in Applied Probability*, 14 (4), s. 855–869.
- Bretó C. (2007), *Statistical Inference for Nonlinear Dynamical Systems*, praca doktorska, University of Michigan, <https://ionides.github.io/531w18/12/notes12.html> (dostęp: 27.05.2019).
- Bretó C. (2014), On idiosyncratic stochasticity of financial leverage effects, *Statistics & Probability Letters*, 91, s. 20 – 26.
- Breymann W., Lüthi D. (2013), ghyp: A package on generalized hyperbolic distributions, *Manual for R Package ghyp*, https://cran.r-project.org/web/packages/ghyp/vignettes/Generalized_Hyperbolic_Distribution.pdf (dostęp: 27.05.2019).
- Broadie M., Kaya, Ö. (2006), Exact simulation of stochastic volatility and other affine jump diffusion processes, *Operations Research*, 54 (2), 217–231.

- Brockwell P. J., Davis R. A., Fienberg S. E. (1991), *Time Series: Theory and Methods: Theory and Methods*, Springer Science & Business Media.
- Brzozowska-Rup K., Dawidowicz A. L. (2009), Metoda filtru cząsteczkowego, *Matematyka Stosowana: matematyka dla społeczeństwa*, 10 (51), s. 69–107.
- Byrd R. H., Lu P., Nocedal J., Zhu C. (1995), A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16, s. 1190–1208.
- Caines P.E. (1988), *Linear Stochastic Systems*, New York: John Wiley and Sons, Inc.
- Cappé O., Moulines E., Rydén T. (2009), *Inference in Hidden Markov Models*, Springer Series in Statistics.
- Carr P., Geman H., Madan D. B., Yor M. (2002), The fine structure of asset returns: An empirical investigation, *The Journal of Business*, 75 (2), s. 305–332.
- Carr P., Geman H., Madan D. B., Yor M. (2007), Self-decomposability and option pricing, *Mathematical Finance*, 17 (1) s. 31–57.
- Carr P., Madan D. B. (1998), Option valuation using the fast Fourier transform, *Journal of Computational Finance*, 2 (4), s. 61–73.
- Carr P., Wu L. (2008), Variance risk premiums, *The Review of Financial Studies*, 22 (3), s. 1311–1341.
- Carvalho C. M., Lopes H. F. (2007), Simulation-based sequential analysis of Markov switching stochastic volatility models, *Computational Statistics & Data Analysis*, 51 (9), s. 4526–4542.
- Chesney M., Scott L. (1989), Pricing European currency options: A comparison of the modified Black-Scholes model and a random variance model, *Journal of Financial and Quantitative Analysis*, 24 (3) s. 267–284.
- Chib S., Nardari, F., Shephard, N. (2002), Markov chain Monte Carlo methods for stochastic volatility models, *Journal of Econometrics*, 108 (2), s. 281–316.

- Chinczyn A. Y. (1938), *Limit laws of sums of independent random variables*, ONTI, Moskwa.
- Christie A. A. (1982), The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of Financial Economics*, 10 (4), s. 407–432.
- Clark P. K. (2973), A subordinated stochastic process model with finite variance for speculative prices, *Econometrica: Journal of the Econometric Society*, 41 (1), s. 135–155.
- Cohen J. B., Black F., Scholes M. (1972), The valuation of option contracts and a test of market efficiency, *The Journal of Finance*, 27 (2), s. 399–417.
- Cont R. (2005), Long range dependence in financial markets, w: *Fractals in Engineering*, Springer, London.
- Cont R. (2007), Volatility Clustering in Financial Markets: Empirical Facts and Agent-Based Models, w: *Long Memory in Economics*, Springer, London.
- Cont R., Tankov P. (1004), *Financial Modelling with Jump Processes*, Chapman and Hall.
- Cox J. (1975), *Notes on option pricing I: Constant elasticity of variance diffusions*, Stanford University, Graduate School of Business.
- Cox J. C., Ingersoll E., Ross S. (1985), A Theory of the Term Structure of Interest Rates, 53 (2), s. 385–407.
- Creal D. (2012), A Survey of Sequential Monte Carlo Methods for Economics and Finance, *Econometric Reviews*, 31 (3), s. 245–296.
- Crisan D., Doucet A. (2002), A Survey of Convergence Results on Particle Filtering Methods for Practitioners, *IEEE Transactions on Signal Processing*, 50 (3), s. 736–746.
- De Jong P. (1991), The diffuse Kalman filter, *The Annals of Statistics*, 19 (2), s. 1073–1083.

- Del Moral P. (1996), Non-linear filtering: interacting particle resolution, *Markov Processes and Related Fields*, 2 (4), s. 555–581.
- Diebold F. X. (1988), *Empirical modeling of exchange rate dynamics*, New York: Springer-Verlag.
- Doman M., Doman R. (2009), *Modelowanie zmienności i ryzyka: metody ekonometrii finansowej*, Wolters Kluwer.
- Domański C. (1990), *Testy statystyczne*, Państwowe Wydaw. Ekonomiczne.
- Domański C., Pekasiewicz D., Baszczyńska A., Witaszczyk A. (2014), *Testy statystyczne w procesie podejmowania decyzji*, Wydawnictwo Uniwersytetu Łódzkiego.
- Domański C., Pruska K. (2000), *Nieklasyczne metody statystyczne*, Polskie Wydawnictwo Ekonomiczne.
- Doucet A., De Freitas N., Gordon N. (2001), An introduction to sequential Monte Carlo methods, w: *Sequential Monte Carlo methods in practice*, New York: Springer-Verlag.
- Doucet A., Jacob P. E., Rubenthaler S. (2013), Derivative-free estimation of the score vector and observed information matrix with application to state-space models, arXiv preprint arXiv:1304.5768, <https://arxiv.org/pdf/1304.5768.pdf> (dostęp: 27.05.2019).
- Drózd K. (2017), Estymacja zmiennych stanu i parametrów układu dwumasowego przy pomocy rozmytych filtrów Kalmana, praca doktorska, Politechnika Poznańska, Wydział Elektryczny Instytut Automatyki i Inżynierii Informatycznej.
- Dziechciarz J. (1993), Modele ekonometryczne z filtrami Kalmana, *Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, 658.
- Eberlein E., Keller U. (1995), Hyperbolic distributions in finance, *Bernoulli*, 1 (3), s. 281–299.

- Eraker B., Johannes M., Polson N. (2003), The impact of jumps in volatility and returns, *The Journal of Finance*, 58 (3), s. 1269–1300.
- Fama E. F. (1965), The behavior of stock-market prices, *The Journal of Business*, 38 (1), s. 34–105.
- Feller W. (1977), *Wstęp do rachunku prawdopodobieństwa*, Polskie Wydawnictwo Naukowe.
- Frühwirth-Schnatter S. and Sögner L. (2009), Bayesian estimation of stochastic volatility models based on OU processes with marginal Gamma law, *Annals of the Institute of Statistical Mathematics*, 61 (1), s. 159–179.
- Gander M. P. S., Stephens D. A. (2007a), Simulation and inference for stochastic volatility models driven by Lévy processes, *Biometrika*, 94 (3), s. 627–646.
- Gander M. P. S., Stephens D. A. (2007b), Stochastic volatility modelling in continuous time with general marginal distributions: Inference, prediction and model selection, *Journal of Statistical Planning and Inference*, 137 (10), s. 3068–3081.
- Ghosh D. (1989), Maximum likelihood estimation of the dynamic shock-error model, *Journal of Econometrics*, 41 (1), s. 121–143.
- Gierej K. (2008), Długa pamięć w szeregach stóp zwrotu i w szeregach zmienności indeksów giełdowych, *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania/Uniwersytet Szczeciński*, 10, s. 434–445.
- Glasserman P., Kim K.-K. (2011), Gamma expansion of the Heston stochastic volatility model, *Finance and Stochastics*, 15 (2), s. 267–296.
- Gleim A., Pigorsch, C. (2013), Approximate Bayesian computation with indirect summary statistics, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.665.5503&rep=rep1&type=pdf> (dostęp: 27.05.2019).
- Glosten L. R., Jagannathan R., Runkle, D. E. (1993), On the relation between the expected value and the volatility of the nominal excess return on stocks, *The Journal of Finance*, 48 (5), s. 1779–1801.

- Good I. J. (1953), The population frequencies of species and the estimation of population parameters, *Biometrika*, 40 (3-4), s. 237–264.
- Gopikrishnan P., Plerou V., Liu Y., Amaral L. A., Gabaix X., Stanley, H. E. (2000), Scaling and correlation in financial time series, *Physica A: Statistical Mechanics and its Applications*, 287 (3-4), s. 362–373.
- Gordon N. J., Salmond D. J., Smith A. F. M. (1993), Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings F (Radar and Signal Processing)*, 140 (2), s. 107–113.
- Gourieroux C., Monfort A., Renault E. (1993), Indirect inference, *Journal of Applied Econometrics*, 8 (1), s. 85–118.
- Grewal M. S., Andrews A. P. (2010), Applications of Kalman filtering in aerospace 1960 to the present (historical perspectives), *IEEE Control Systems Magazine*, 30 (3), s. 69–78.
- Griffin J. E., Steel, M. F. J. (2006), Inference with non-Gaussian Ornstein–Uhlenbeck processes for stochastic volatility, *Journal of Econometrics*, 134 (2), s. 605–644.
- Griffin J. E., Steel, M. F. J. (2010), Bayesian inference with stochastic volatility models using continuous superpositions of non-Gaussian Ornstein–Uhlenbeck processes, *Computational Statistics & Data Analysis*, 54 (11), s. 2594–2608.
- Grzenda W. (2015), The advantages of bayesian methods over classical methods in the context of credible intervals, *Information Systems in Management*, 4 (1), s. 53–63.
- Grzenda W. (2016), Modelowanie bayesowskie: teoria i przykłady zastosowań, Oficyna Wydawnicza Szkoła Główna Handlowa, Warszawa.
- Grzesiak S. (1995), Równania filtru Kalmana w modelowaniu ekonometrycznym, *Przegląd Statystyczny*, 42 (1).
- Habtemicael S., SenGupta I. (2016), Pricing variance and volatility swaps for Barndorff-Nielsen and Shephard process driven financial markets, *International Journal of Financial Engineering*, 3 (4), s. 1–35.

- Halgreen C. (1979), Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions, *Probability Theory and Related Fields*, 47 (1), s. 13–27.
- Hamilton J. D. (1994), State-space models, w: *Handbook of econometrics*, 4, s. 3039–3080.
- Hannan E. J., Quinn B. G. (1979), The determination of the order of an autoregression, *Journal of the Royal Statistical Society: Series B (Methodological)*, 41 (2), s. 190–195.
- Harvey A. (1990), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press.
- Harvey A., Ruiz E., Shephard N. (1994), Multivariate stochastic variance models, *The Review of Economic Studies*, 61 (2), s. 247–264.
- Hendry D. F., Massmann M. (2007), Co-breaking: Recent advances and a synopsis of the literature, *Journal of Business & Economic Statistics*, 25 (1), s. 33–51.
- Heston S. L. (1993), A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies*, 6 (2), s. 327–343.
- Hougaard P. (1986), Survival models for heterogeneous populations derived from stable distributions, *Biometrika*, 73 (2), s. 387–396.
- Hubalek F., Posedel P. (2011), Joint analysis and estimation of stock prices and trading volume in Barndorff-Nielsen and Shephard stochastic volatility models, *Quantitative Finance*, 11 (6), s. 917–932.
- Hubalek F., Sgarra C. (2009), On the Esscher transforms and other equivalent martingale measures for Barndorff-Nielsen and Shephard stochastic volatility models with jumps, *Stochastic Processes and their Applications*, 119 (7), s. 2137–2157.

- Hull J., White A. (1987), The pricing of options on assets with stochastic volatilities, *The Journal of Finance*, 42 (2), s. 281–300.
- Iacus S. M. (2009), *Simulation and inference for stochastic differential equations: with R examples*, Springer Science & Business Media.
- Iacus S. M. (2009), *Option pricing and estimation of financial models with R*, John Wiley & Sons.
- Ionides E. L. (2005), Maximum smoothed likelihood estimation, *Statistica Sinica*, s. 1003–1014.
- Ionides E. L. (2019), *Practical likelihood-based inference for POMP models*, <https://ionides.github.io/531w18/12/notes12.html> (dostęp: 27.05.2019).
- Ionides E. L., Bhadra A., Atchadé Y., King A. (2011), Iterated filtering, *The Annals of Statistics*, 39 (3), s. 1776–1802.
- Ionides E. L., Bretó C., King A. A. (2006), Inference for nonlinear dynamical systems, *Proceedings of the National Academy of Sciences*, 103 (49), s. 18438–18443.
- Ionides E. L., Nguyen D., Atchadé Y., Stoev S., King A. A. (2015), Inference for dynamic and latent variable models via iterated, perturbed Bayes maps, *Proceedings of the National Academy of Sciences*, 112 (3), s. 719–724.
- Irony T. Z., Singpurwalla N. D. (1997), Non-informative priors do not exist A dialogue with José M. Bernardo, *Journal of Statistical Planning and Inference*, 65 (1), s. 159–177.
- Issaka A., SenGupta I. (2017), Analysis of variance based instruments for Ornstein–Uhlenbeck type models: swap and price index, *Annals of Finance*, 13 (4), s. 401–434.
- Jacquier E., Polson N. G., Rossi P. (1999), Stochastic volatility: Univariate and multivariate extensions, CIRANO Working Papers, <https://cirano.qc.ca/files/publications/99s-26.pdf> (dostęp: 27.05.2019).

- Jakubowski J., Palczewski A., Rutkowski M., Stettner Ł. (2003), *Matematyka finansowa. Instrumenty pochodne*, WNT.
- Jakubowski J., Sztencel R. (2001), *Wstęp do teorii prawdopodobieństwa*, Script.
- James L. F., Kim D., Zhang Z. (2013), Exact simulation pricing with Gamma processes and their extensions, arXiv preprint arXiv:1310.6526 <https://arxiv.org/pdf/1310.6526.pdf> (dostęp: 27.05.2019).
- James L. F., Müller G., Zhang Z. (2018), Stochastic volatility models based on OU-Gamma time change: theory and estimation, *Journal of Business & Economic Statistics*, 36 (1), s. 75–87.
- Janicki A., Izydorczyk A. (2001), *Komputerowe metody w modelowaniu stochastycznym: modele w finansach, technice i biologii, algorytmy numeryczne i statystyczne, symulacja i wizualizacja zjawisk losowych, autorski pakiet komputerowy SDE-Solver*, WNT.
- Kaarakka T. (2015), *Fractional Ornstein-Uhlenbeck Processes*, Julkaisu-Tampere University of Technology.
- Kalman R. E. (1961), A new approach to linear filtering and prediction problems, *Journal of Basic Engineering*, 82 (1), s. 35–45.
- Kalman R. E., Bucy R. S. (1961), New Results in Linear Filtering and Prediction Theory, *Journal of Basic Engineering*, 83 (1), s. 95–108.
- Kantas N., Doucet A., Singh S. S., Maciejowski J., Chopin N. (2015), On particle methods for parameter estimation in state-space models, *Statistical Science*, 30 (3), s. 328–351.
- Karandikar R. L., Rao B.V. (2014), On quadratic variation of martingales, *Proceedings-Mathematical Sciences*, 124 (3), s. 457–469.
- Kim S., Shephard N., Chib S. (1998), Stochastic volatility: likelihood inference and comparison with ARCH models, *The Review of Economic Studies*, 65 (3), s. 361–393.

- King A. A., Ionides E. L., Bretó C. M., S. P. Ellner, M. J. Ferrari, B. E. Kendall, Lavine M., Nguyen D., Reuman D. C., Wearing H., Wood S. N. (2019), *pomp: Statistical Inference for Partially Observed Markov Processes*, (R package, version 2.1), <https://cran.r-project.org/web/packages/pomp/index.html> (dostęp: 27.05.2019).
- Kitagawa G. (1996), Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *Journal of Computational and Graphical Statistics*, 5 (1), s. 1–25.
- Kitagawa G. (2014), Computational aspects of sequential Monte Carlo filter and smoother, *Annals of the Institute of Statistical Mathematics*, 66 (3), s. 443–471.
- Kliber P. (2013), *Zastosowanie procesów dyfuzji ze skokami do modelowania polskiego rynku finansowego*, Wydawnictwo Uniwersytetu Ekonomicznego, Poznań.
- Kong A., Liu J. S., Wong W. H. (1994), Sequential imputations and Bayesian missing data problems, *Journal of the American Statistical Association*, 89 (425), s. 278–288.
- Kwiatkowski Ł. (2013), *Bayesowskie modele SV z przełączaniem Markowa w analizie zmienności na rynkach finansowych*, Zeszyty Naukowe/Uniwersytet Ekonomiczny w Krakowie. Seria Specjalna, Monografie.
- Latała R. (2011), *Wstęp do analizy stochastycznej*, Uniwersytet Warszawski, Wydział Matematyki, Informatyki i Mechaniki.
- Lebaron B. (2001), Stochastic volatility as a simple generator of apparent financial power laws and long memory, *Quantitative Finance*, 1 (6), s. 621–631.
- Lehmann E. L. (1991), *Teoria estymacji punktowej*, Wydawnictwo Naukowe PWN.
- Lele S. R., Dennis B., Lutscher F. (2007), Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods, *Ecology Letters*, 10 (7), s. 551–563.

- Lévy P. (1937), L'addition des variables aléatoires définies sur une circonférence, *Bulletin de la Société mathématique de France*, 67, s. 1–41.
- Li L., Linetsky V. (2015), *Discretely monitored first passage problems and barrier options: an eigenfunction expansion approach*, *Finance and Stochastics*, 19 (4), s. 941–977.
- Lindberg C. (2008), The estimation of the Barndorff-Nielsen and Shephard model from daily data based on measures of trading intensity, *Applied Stochastic Models in Business and Industry*, 24 (4), s. 277–289.
- Lindström E., Ionides E. L., Frydendall J., Madsen H. (2012), Efficient iterated filtering, *IFAC Proceedings Volumes*, 45 (16), s. 1785–1790.
- Liu J. S. (1996), Metropolized independent sampling with comparisons to rejection sampling and importance sampling, *Statistics and Computing*, 6 (2), s. 113–119.
- Liu J. S., Chen R. (1995), Blind deconvolution via sequential imputations, *Journal of the American Statistical Association*, 90 (430), s. 567–576.
- Liu J., West M. (2001), Combined parameter and state estimation in simulation-based filtering, w: *Sequential Monte Carlo methods in practice*, Springer, s. 197–223.
- Liu L. Y., Patton A. J., Sheppard K., Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes, *Journal of Econometrics*, 187 (1), s. 293–311.
- Madan D. B., Carr P., Chang E. C. (1998), The variance gamma process and option pricing, *Review of Finance*, 2 (1), s. 79–105.
- Madan D. B., Seneta E. (1990), The variance gamma (VG) model for share market returns, *Journal of Business*, s. 511–524.
- Maddala G. S. (2008), *Ekonometria*, Wydawnictwo Naukowe PWN.

- Makhwiting M. R., Sigauke C., Lesaoana M. (2014), Modelling tail behavior of returns using the generalized extreme value distribution, *Economics, Management and Financial Markets*, 9 (1), s. 41-57.
- Malik S., Pitt M. K. (2011), Particle filters for continuous likelihood evaluation and maximisation, *Journal of Econometrics*, 165 (2), s. 190–209.
- Mandelbrot B. (1963), The variation of certain speculative prices, *The Journal of Business*, 36 (4), s. 394–419.
- Marcus M. B. (1987), Xi-radial Processes and Random Fourier Series, *American Mathematical Soc.*, s. 68–368.
- Meddahi N., Renault E. (1998), Aggregations and marginalization of GARCH and stochastic volatility models, Université de Montréal. Département de sciences économiques.
- Melino A., Turnbull S. M. (1990), Pricing foreign currency options with stochastic volatility, *Journal of Econometrics*, 45 (1-2), s. 239–265.
- Michael J. R., Schucany W. R., Haas R. W. (1976), Generating random variates using transformations with multiple roots, 30 (2), s. 88–90.
- Mikosch T., Starica C. (2000), Is it really long memory we see in financial returns, *Extremes and Integrated Risk Management*, 12, s. 149–168.
- Miszczak W. (2006), *Wielowymiarowe szeregi czasowe*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.
- Nelson D. B. (1990), ARCH models as diffusion approximations, *Journal of Econometrics*, 45 (1-2), s. 7–38.
- Nelson D. B. (1990), Conditional heteroskedasticity in asset returns: A new approach, *Econometrica: Journal of the Econometric Society*, 59 (2), s. 347-370.
- Nguyen D. (2016), Another look at Bayes map iterated filtering, *Statistics & Probability Letters*, 118, s. 32–36.

- Nguyen D., Ionides E. L. (2017), A second-order iterated smoothing algorithm, *Statistics and Computing*, 27 (6), s. 1677–1692.
- Nicolato E., Venardos E. (2003), Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 13 (4), s. 445–466.
- Pagan A. (1980), Some identification and estimation results for regression models with stochastically varying coefficients, *Journal of Econometrics*, 13 (3), s. 341–363.
- Pajor A. (2003), Procesy zmienności stochastycznej SV w bayesowskiej analizie finansowych szeregów czasowych, *Zeszyty Naukowe/Uniwersytet Ekonomiczny w Krakowie. Seria Specjalna, Monografie: prace doktorskie/Akademia Ekonomiczna w Krakowie*.
- Papapantoleon A. (2008), An introduction to Lévy processes with applications in finance, arXiv preprint arXiv:0804.0482 <https://arxiv.org/pdf/0804.0482.pdf> (dostęp: 27.05.2019).
- Pigorsch C., Stelzer R. (2009), A multivariate Ornstein-Uhlenbeck type stochastic volatility model, https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.050/people/stelzer/PigorschStelzer2009.pdf (dostęp: 27.05.2019).
- Piontek K. (2000), Modelowanie i prognozowanie zmienności instrumentów finansowych, praca doktorska, *Akademia Ekonomiczna we Wrocławiu*, http://www.fire.ue.wroc.pl/pracownicy/KPiontek_doktorat.pdf (dostęp: 27.05.2019).
- Piontek K. (2003), Wycena opcji w modelu uwzględniającym efekt AR-GARCH, *Prace Naukowe AE we Wrocławiu*, 990, s. 331–336.
- Piontek K. (2004), Zastosowanie modeli klasy ARCH do opisu własności szeregu stóp zwrotu indeksu WIG, *Prace Naukowe AE we Wrocławiu*, 1021, s. 152–169.
- Pitt M. K., Shephard N. (1999), Filtering via simulation: Auxiliary particle filters, *Journal of the American Statistical Association*, 94 (446), s. 590–599.

- Pluciennik P. (2008), Badanie finansowych szeregów czasowych za pomocą modeli z czasem ciągłym, praca doktorska, Uniwersytet im. Adama Mickiewicza w Poznaniu, http://www.wbc.poznan.pl/Content/113203/Pluciennik_Piotr-rozpraw_doktorska.pdf (dostęp: 27.05.2019).
- Prause K. (1999), *The generalized hyperbolic model: Estimation, financial derivatives and risk measures*, <https://d-nb.info/961152192/34> (dostęp: 27.05.2019).
- Protter P. E. (2005), *Stochastic integration and differential equations*, Springer.
- Raknerud A., Skare Ø., Multivariate stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck processes: A quasi-likelihood approach, Discussion Papers No. 614, Statistics Norway, Research Department, <https://www.ssb.no/a/publikasjoner/pdf/DP/dp614.pdf> (dostęp: 27.05.2019).
- Roberts G. O., Papaspiliopoulos O. Dellaportas P. (2004), Bayesian inference for non-Gaussian Ornstein–Uhlenbeck stochastic volatility processes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66 (2), s. 369–393.
- Rogers L. C. G., Williams D. (1996), *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, Cambridge University Press.
- Rosenberg B. (1973), Random coefficients models: the analysis of a cross section of time series by stochastically convergent parameter regression, *Annals of Economic and Social Measurement*, 2 (4), s. 399–428.
- Rosiński J. (1990), On series representations of infinitely divisible random vectors, *The Annals of Probability*, 18, s. 405–430.
- Rosiński J. (2001), Dyskusja do artykułu Barndorff-Nielsen i Shepharda, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63 (2), s. 167–241.
- Roy M., Bouma M. J., Ionides E. L., Dhiman R. C., Pascual M. (2013), The potential elimination of *Plasmodium vivax* malaria by relapse treatment: insights from a transmission model and surveillance data from NW India, *PLoS neglected tropical diseases*, 7 (1).

- Santander Biuro Maklerskie (2019), <https://bm.santander.pl/inwestowanie-online/inwestor-online/narzedzia-inwestycyjne/pliki-at/pliki-at.html> (dostęp: 27.05.2019).
- Sato K.-I. (1999), Lévy processes and infinitely divisible distributions, Cambridge University Press.
- Sato K.-I. (2014), Stochastic integrals with respect to Lévy processes and infinitely divisible distributions, *Sugaku Expositions*, 27 (1), s. 161–181.
- Schöbel R., Zhu J. (1999), Stochastic volatility with an Ornstein–Uhlenbeck process: an extension, *Review of Finance*, 3 (1), s. 23–46.
- Schountens W. (2003), Lévy Processes in Finance: Pricing financial derivatives, Wiley.
- Schountens W., Teugels J. L. (2003), *Meixner processes in finance*, <https://www.eurandom.tue.nl/reports/2001/002-report.pdf> (dostęp: 27.05.2019).
- Schwarz G. (1978), Estimating the dimension of a model, *The Annals of Statistics*, 6 (2), s. 461–464.
- Sentana E. (1995), Quadratic ARCH models, *The Review of Economic Studies*, 62 (4), s. 639–661.
- Shephard N., Andersen T. G. (2009), Stochastic volatility: origins and overview, w: *Handbook of Financial Time Series*, Springer, s. 233–254.
- Shi Q., Yang X. (2014), Pricing Asian options in a stochastic volatility model with jumps, *Applied Mathematics and Computation*, 228, s. 411–422.
- Silverman B. W. (1986), *Density estimation for statistics and data analysis*, London: Chapman & Hall/CRC.
- Skrzypek S. (1986), Zastosowanie filtracji kalmanowskiej do estymacji łącznej w modelach procesów gospodarczych, *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie*, 222.
- Sobczyk K. (1996), Stochastyczne równania różniczkowe, WNT, Warszawa.

- Sørensen M. (2000), Prediction-based estimating functions, *The Econometrics Journal*, 3 (2), s. 123–147.
- Stein E. M., Stein, J. C. (1991), Stock price distributions with stochastic volatility: an analytic approach, *The Review of Financial Studies*, 4 (4), s. 727–752.
- Serwis internetowy Stooq.pl, <https://stooq.pl> (dostęp: 27.05.2019).
- Szczepocki P. (2018), Zastosowanie filtru Kalmana do modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka, *Folia Oeconomica*, 4 (337), s. 183–201.
- Szczepocki P. (2019a), Application of iterated filtering to stochastic volatility models based on non-gaussian Ornstein-Uhlenbeck process, *Statistics in Transition New Series* (w trakcie recenzji).
- Szczepocki P. (2019b), Zastosowanie iterowanej filtracji do estymacji parametrów wariancji chwilowej w ramach niegaussowskich procesów stochastycznej zmienności typu Ornsteina-Uhlenbecka, *Przegląd Statystyczny* (przyjęty do druku).
- Tauber E., Leonenko N., Bee M. (2011), Characteristic function estimation of Ornstein–Uhlenbeck-based stochastic volatility models, *Computational Statistics & Data Analysis*, 55 (8), s. 2525–2539.
- Taylor S. J. (1982), Financial returns modelled by the product of two stochastic processes—a study of the daily sugar prices 1961–75, *Time Series Analysis: Theory and Practice*, 1, s. 203–226.
- Trabs M. (2014), Calibration of self-decomposable Lévy models, *Bernoulli*, 20 (1), s. 109–140.
- Tsay R. S. (2005), *Analysis of financial time series*, John Wiley & Sons.
- Tweedie M. C. K. (1984), An index which distinguishes between some important exponential families, *Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee International conference*, s. 579–604.
- Uhlenbeck G. E., Ornstein L. S. (1930), On the theory of the Brownian motion, *Physical Review*, 36 (5), s. 823–841.

- Valdivieso L., Schoutens W., Tuerlinckx F. (2009), Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type, *Statistical Inference for Stochastic Processes*, 12 (1), s. 1–19.
- Vasicek O. (1977), An equilibrium characterization of the term structure, *Journal of Financial Economics*, 5 (2), s. 177–188.
- Vaswani N. (2008), Particle filtering for large-dimensional state spaces with multimodal observation likelihoods, *IEEE Transactions on Signal Processing*, 56 (10), s. 4583–4597.
- Walker A. M. (1969), On the Asymptotic Behaviour of Posterior Distributions, *Journal of the Royal Statistical Society: Series B (Methodological)*, 31 (1), s. 80–88.
- White H. (1982), Maximum likelihood estimation of misspecified models, *Econometrica: Journal of the Econometric Society*, 50 (1), s. 1–25.
- Wiggins J. B. (1987), Option values under stochastic volatility: Theory and empirical estimates, *Journal of Financial Economics*, 19 (2), s. 351–372.
- Wolfe S. J. (1982), On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$, *Stochastic Processes and Their Applications*, 12 (3), s. 301–312.
- Wong E. (1976), Procesy stochastyczne w teorii informacji i układach dynamicznych, Wydawnictwa Naukowo-Techniczne.
- Zanini E., Nemeth C. (2012), Parameter estimation with Particle Filtering Algorithms, *STOR-i Internship 2012, Lancaster University*.
- Zhang L., Mykland P. A., Aït-Sahalia Y. (2005), A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, 100 (472), s. 1394–1411.
- Zhang S. (2011), Exact simulation of tempered stable Ornstein–Uhlenbeck processes, *Journal of Statistical Computation and Simulation*, 81 (11), s. 1533–1544.

Zhu, J. (2008), A simple and exact simulation approach to Heston model, <http://dx.doi.org/10.2139/ssrn.1153950> (dostęp: 27.05.2019).

Żółtowska E. (1997), *Funkcje produkcji, teoria, estymacja, zastosowanie*, Wydawnictwo Uniwersytetu Łódzkiego.

Spis rysunków

1.1	Trajektoria standardowego procesu Wienera.	26
1.2	Trajektoria procesu Poissona z parametrem intensywności $\lambda = 5$ (a) oraz odpowiadająca temu procesowi trajektoria skompensowanego procesu Poissona (b).	29
1.3	Trajektoria złożonego procesu Poissona z parametrem intensywności $\lambda = 2$ ze skokami o rozkładzie normalnym standardowym (a) oraz od- powiadająca temu procesowi trajektoria skompensowanego złożonego procesu Poissona (b).	29
1.4	Trajektoria złożonego procesu Poissona z parametrem intensywno- ści $\lambda = 0,1$ ze skokami o rozkładzie normalnym standardowym (a) oraz odpowiadająca temu procesowi trajektoria procesu skoków wraz zaznaczonym zbiorem $[0, 2] \times [0, 5, 1]$ (b).	31
1.5	Gęstości rozkładu gamma (a), gęstości Lévy'eg procesu gamma (b) oraz symulacje trajektorii procesu gamma (c) odpowiadające trzem rozkładom gamma $Ga(\nu, \alpha)$	35
1.6	Rozkład $IG(1,1/2)$ jako moment dojścia do poziomu $\delta = 1$ procesu Wienera z dryfem $\gamma = 1/2$ (por. podrozdział 2.5) oraz gęstości roz- kładu IG (b), gęstości Lévy'eg procesu IG (b) oraz trajektorie procesu IG (d) odpowiadające trzem rozkładom odwrotnym Gaussa $IG(\delta, \gamma)$	37

- 1.7 Trajektorie trzech różnych procesów Ornsteina-Uhlenbecka o tych samych parametrach $\mu = 0, \lambda = 0,01, \sigma = 0,05$, ale różnych wartościach początkowych. 44
- 2.1 Trajektoria procesu wariancji chwilowej o rozkładzie brzegowym gamma z parametrem kształtu $\nu = 1$ oraz parametrem skali $1/\alpha = 1/3$ (a) oraz odpowiadający mu proces Lévy'ego ukryty w tle (BDLP) z parametrem $\lambda = 0,01$ (b). 54
- 2.2 Trajektoria procesu wariancji scałkowanej odpowiadającej procesowi wariancji chwilowej przedstawionemu na rysunku 2.1. 59
- 2.3 Trajektoria procesu wariancji chwilowej o rozkładzie brzegowym gamma z parametrem kształtu $\nu = 2$, parametrem skali $1/\alpha = 1/10$, parametrem $\lambda = 0,01$ (a) oraz rozkład empiryczny wariancji aktualnej: estymator jądrowy gęstości (kolor czerwony) oraz gęstość rozkładu gamma (kolor niebieski) z parametrem kształtu $\nu = 2$ oraz parametrem skali $1/\alpha = 1/10$ (b). 63
- 2.4 Zależność pomiędzy parametrem autoregresyjnym AR(1) a parametrem średniej ruchomej MA(1) odpowiednio dla σ_n^2 (a) i y_n^2 (b), przy czym zaznaczono punkt dla $\Delta = 1$ i $\lambda = 0,1$ 66
- 2.5 Funkcje autokorelacji dla jednego procesu zmienności o parametrze $\lambda = 1$ (kolor czerwony), dwóch procesów o parametrach $\lambda = 1$ oraz $\lambda = 0,1$ (kolor zielony) i trzech procesów zmienności o parametrach $\lambda = 1, \lambda = 0,1$ oraz $\lambda = 0,01$ (kolor niebieski). Dla uproszczenia zapisu przyjęto $\Delta = 1$ 70
- 2.6 Wariancja aktualna σ_n (linia czerwona) oraz wariancja zrealizowana $\{y\}_n$ (symbol +) dla $M = 1, 12, 48, 288$. Zastosowano proces wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,2$ i wariancji $\omega^2 = 0,02$ oraz parametrze $\lambda = 0,01$ 85

- 2.7 Błąd wariancji zrealizowanej oryginalny (a) oraz po transformacji logarytmicznej (b) dla symulacji przedstawionej na rysunku 2.6. Linie wyznaczają błąd standardowy przemnożony przez 1,96 (linie kwantylowe dla prawdopodobieństwa 0,95 przy założeniu, że doszło do zbieżności do rozkładu normalnego). Prawy segment (c) przedstawia wykresy kwantyl-kwantyl dla wystandaryzowanego błędu oryginalnego (kolor czerwony) i po transformacji logarytmicznej (kolor turkusowy). Górny panel przedstawia wyniki dla $M=24$, dolny dla $M=288$ 89
- 3.1 Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru Kalmana (czerwona linia) i wygładzania Kalmana (linia niebieska). 107
- 3.2 Trajektoria wariancji aktualnej σ_n^2 (a) powstałej ze złożenia dwóch wariancji aktualnych $\sigma_{1,n}^2$ (b) i $\sigma_{2,n}^2$ (c) oraz ich oszacowania za pomocą filtru Kalmana (czerwona linia) i wygładzania Kalmana (linia niebieska). 110
- 3.3 Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru Kalmana (czerwona linia) i wygładzania Kalmana (linia niebieska) dla $M = 24, 288$. Punktami + oznaczono oszacowania wariancji zrealizowanej. 115
- 3.4 Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru cząsteczkowego (czerwona linia) i wygładzania cząsteczkowego (linia niebieska) dla $K = 1000$ i $L = 20$ (a) oraz efektywna liczebność próby ESS (b). 128
- 3.5 Trajektoria wariancji aktualnej σ_n (linia czarna) oraz jej oszacowania za pomocą filtru cząsteczkowego (czerwona linia) i wygładzania cząsteczkowego (linia niebieska) dla $M = 12, 288$. Punktami + oznaczono oszacowania wariancji zrealizowanej. 131
- 3.6 Trajektoria procesu wariancji chwilowej o rozkładzie stacjonarnym odwrotnym Gaussa z parametrami $\delta = 1$ i $\gamma = 2$ (a) oraz odpowiadający mu proces Lévy'ego ukrytym w tle (BDLP) z parametrem $\lambda = 0,01$ (b). 137

- 3.7 Trajektorie oszacowań parametrów λ , α , ν dla dziesięciu 10 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych, każda po $J = 500$ iteracji (a) oraz oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów λ , α , ν wyznaczone przez filtr cząsteczkowy (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty) (b). Zastosowano podstawowy model BNS o wariancji chwilowej ze stacjonarnym rozkładem gamma z parametrami $\alpha = 12, 5$, $\nu = 6, 25$, $\lambda = 0, 01$. Przyjęto $\mu = \beta = 0$. . . 157
- 4.1 Dienne notowania indeksu WIG (a) oraz logarytmiczne dzienne stopy zmian indeksu WIG (b) z okresu od 6 września 2013 do 8 września 2017. 161
- 4.2 Rozkład empiryczny dziennych logarytmicznych stóp zmian indeksu WIG wyznaczony na podstawie estymatora jądrowego oraz dopasowany wykres gęstości rozkładu normalnego (a) oraz wykres kwantylkwantyl dla dziennych logarytmicznych stóp zmian indeksu WIG (b). Dane z okresu od 6 września 2013 do 8 września 2017. 162
- 4.3 Korelogram dla dziennych logarytmicznych stóp zmian indeksu WIG (a) oraz dla kwadratów dziennych logarytmicznych stóp zmian indeksu WIG (b). Dane z okresu od 6 września 2013 do 8 września 2017. 163
- 4.4 Dienne wartości kursu USD/PLN (a) oraz logarytmiczne dzienne stopy zmian kursu USD/PLN (b) z okresu od 6 września 2013 do 8 września 2017. 164
- 4.5 Rozkład empiryczny dziennych logarytmicznych stóp zmian kursu USD/PLN wyznaczony na podstawie estymatora jądrowego i dopasowany wykres gęstości rozkładu normalnego (a) oraz wykres kwantylkwantyl dla dziennych logarytmicznych stóp zmian kursu USD/PLN (b). Dane z okresu od 6 września 2013 do 8 września 2017. 165

4.6	Korelogram dla dziennych logarytmicznych stóp zmian kursu USD/PLN (a) oraz dla kwadratów dziennych logarytmicznych stóp zmian kursu USD/PLN (b). Dane z okresu od 6 września 2013 do 8 września 2017.	166
4.7	5-minutowe notowania indeksu WIG20 (a) oraz wariancja zrealizowana indeksu WIG20 (b) z okresu od 22 lutego 2017 do 21 lutego 2019.	168
4.8	Rozkład empiryczny 5-minutowych logarytmicznych stóp zmian indeksu WIG20 wyznaczony na podstawie estymatora jądrowego oraz dopasowany wykres gęstości rozkładu normalnego (a), wykres kwantyl-kwantyl dla 5-minutowych logarytmicznych stóp zmian indeksu WIG20 (b). Na podstawie danych z okresu od 22 lutego 2017 do 21 lutego 2019.	169
4.9	Rozkład empiryczny dziennych logarytmicznych stóp zmian indeksu WIG oraz dopasowane wykresy gęstości rozkładu VG, NIG, uogólnionego hiperbolicznego na skali zwykłej (a) oraz logarytmicznej (b).	171
4.10	Rozkład empiryczny dziennych logarytmicznych stóp zmian kursu USD/PLN oraz dopasowane wykresy gęstości rozkładu VG, NIG, uogólnionego hiperbolicznego na skali zwykłej (a) oraz logarytmicznej (b).	171
4.11	Rozkład empiryczny dziennych logarytmicznych stóp zmian indeksu WIG20 oraz dopasowane wykresy gęstości rozkładu VG, NIG, uogólnionego hiperbolicznego na skali zwykłej (a) oraz logarytmicznej (b).	172
4.12	Wykresy kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian indeksu WIG względem rozkładu VG (c) i NIG (d).	173
4.13	Wykresy kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian kursu USD/PLN względem rozkładu VG (a) i NIG (b).	173
4.14	Wykresy kwantyl-kwantyl dla dziennych logarytmicznych stóp zmian indeksu WIG względem rozkładu VG (c) i NIG (d).	174

- 4.15 Trajektorie oszacowań parametrów względem kolejnych iteracji algorytmu dla 5 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych w modelu ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU), każda po $J = 200$ iteracji. 182
- 4.16 Oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów wyznaczone za pomocą filtru cząsteczkowego (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty). Model ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU). 183
- 4.17 Kwadraty dziennych logarytmicznych stóp zmian indeksu WIG (a) oraz oszacowanie wariancji aktualnej w modelu ze złożeniem dwóch procesów zmienności i efektem dźwigni o rozkładzie stacjonarnym gamma (BNS2L Ga-OU) otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) dla wartości parametrów przedstawionych w tablicy 4.8 (b) wraz ze składowymi wariancji aktualnej (c) i (d). 185
- 4.18 Trajektorie oszacowań parametrów dla 5 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych w modelu ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU), każda po $J = 200$ iteracji. 190

- 4.19 Oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów wyznaczone za pomocą filtru cząsteczkowego (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty). Model ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU). 191
- 4.20 Kwadraty dziennych logarytmicznych stóp zmian kursu USD/PLN (a) oraz oszacowanie wariancji aktualnej (b) wraz ze składowymi (c) i (d) w modelu ze złożeniem dwóch procesów zmienności o rozkładzie odwrotnym Gaussa (BNS2 IG-OU) otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek, $L = 20$) oraz wygładzania Kalmana. 194
- 4.21 Trajektorie oszacowań parametrów dla 5 realizacji iterowanej filtracji o różnych, losowo wybranych punktach startowych w modelu ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU), każda po $J = 200$ iteracji. 200
- 4.22 Oszacowania jednowymiarowych cięć funkcji wiarygodności względem parametrów wyznaczone za pomocą filtru cząsteczkowego (czarne koła) przy użyciu 1000 cząsteczek, interpolacja logarytmu funkcji wiarygodności lokalnymi wielomianami drugiego stopnia (linia niebieska) oraz oszacowania logarytmu funkcji wiarygodności otrzymane przy użyciu ostatniej iteracji algorytmu iterowanej filtracji (czerwone trójkąty). Model ze złożeniem dwóch procesów zmienności o rozkładzie stacjonarnym odwrotnym Gaussa (BNS2 IG-OU). 201
- 4.23 Wariancja zrealizowana (a) oraz oszacowanie wariancji aktualnej (b) wraz ze składowymi (c) i (d) w modelu u ze złożeniem dwóch procesów zmienności o rozkładzie odwrotnym Gaussa (BNS2 IG-OU) otrzymane za pomocą wygładzania cząsteczkowego ($K = 1000$ cząsteczek) oraz wygładzania Kalmana. 202

- C.1 Symulacja trajektorii procesu logarytmu cen (a) oraz odpowiadające tej trajektorii stopy zwrotu (b), kwadraty stóp zwrotu (c), proces wariancji aktualnej (d). Wykorzystano proces wariancji chwilowej o rozkładzie stacjonarnym gamma o wartości oczekiwanej $\xi = 0,5$ i wariancji $\omega = 0,2$. Pozostałe parametry $\mu = 0$, $\beta = 0,05$, $\rho = -0,1$.

Źródło: opracowanie własne. 216

Spis tablic

4.1	Statystyki opisowe dla logarytmicznych dziennych stóp zmian indeksu WIG z okresu od 6 września 2013 do 8 września 2017.	162
4.2	Statystyki i p-wartości testów normalności dla zwrotów logarytmicznych indeksu WiG za okres od 6 września 2013 do 8 września 2017. .	163
4.3	Statystyki opisowe dla logarytmicznych dziennych stóp zmian kursu USD/PLN. Dane z okresu od 6 września 2013 do 8 września 2017. . .	165
4.4	Statystyki i p-wartości testów normalności dla zwrotów logarytmicznych kursu USD/PLN za okres od 6 września 2013 do 8 września 2017.	166
4.5	Statystyki opisowe dla 5-minutowych logarytmicznych stóp zmian indeksu WIG20.	167
4.6	Statystyki Kołmogorowa-Smirnowa (w nawiasach p-wartości testu Kołmogorowa-Smirnowa) dla trzech rozkładów: VG (<i>variance gamma</i>), NIG (normalny odwrotny Gaussa), HYP (uogólniony hiperboliczny).	175
4.7	Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla rozkładów VG (<i>variance gamma</i>) oraz NIG jako specjalnych przypadków uogólnionego rozkładu hiperbolicznego.	175

4.8	Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej gamma wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.	178
4.9	Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej odwrotnym Gaussa wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.	179
4.10	Wyniki estymacji wartości średniej (parametr ξ) oraz odchylenia standardowego (parametr ω) rozkładu stacjonarnego procesu wariancji chwilowej dla ośmiu rozważanych modeli stochastycznej zmienności z podziałem na rozkłady stacjonarne: gamma (Ga-OU) oraz odwrotnym Gaussa (IG-OU).	181
4.11	Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla par hipotez zagnieżdżonych.	184
4.12	Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej gamma wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.	187
4.13	Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej odwrotnym Gaussa wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.	188

4.14	Wyniki estymacji wartości średniej (parametr ξ) oraz odchylenia standardowego (parametr ω) rozkładu stacjonarnego procesu wariancji chwilowej dla ośmiu rozważanych modeli stochastycznej zmienności z podziałem na rozkłady stacjonarne: gamma (Ga-OU) oraz odwrotnym Gaussa (IG-OU).	189
4.15	Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla par hipotez zagnieżdżonych.	192
4.16	Wyniki estymacji parametrów modeli stochastycznej zmienności wyznaczone metoda quasi-największej wiarygodności. W nawiasach podano błędy standardowe oszacowań.	195
4.17	Wyniki estymacji parametrów modeli stochastycznej zmienności z rozkładem stacjonarnym procesu wariancji chwilowej gamma i odwrotnym Gaussa wraz z uzyskanym oszacowaniem logarytmu funkcji wiarygodności oraz wartością kryterium informacyjnego AIC. W nawiasach podano błędy standardowe oszacowań.	197
4.18	Wyniki estymacji wartości średniej (parametr ξ) oraz odchylenia standardowego (parametr ω) rozkładu stacjonarnego procesu wariancji chwilowej dla ośmiu rozważanych modeli stochastycznej zmienności z podziałem na rozkłady stacjonarne: gamma (Ga-OU) oraz odwrotnym Gaussa (IG-OU).	198
4.19	Statystyki testu ilorazu wiarygodności (w nawiasach p-wartości testu) dla pary hipotez zagnieżdżonych.	199
4.20	Wyniki estymacji parametrów modeli stochastycznej zmienności wyznaczone metoda quasi-największej wiarygodności. W nawiasach podano błędy standardowe oszacowań.	199