

Polska Akademia Nauk  
Instytut Podstawowych Problemów Techniki

Rozprawa doktorska

**Skierowane liczby rozmyte  
w modelowaniu i symulacji  
finansowych szeregów czasowych**

**mgr Adam Marszałek**

Promotor: prof. dr hab. inż. Tadeusz Burczyński  
Promotor pomocniczy: dr inż. Michał Bereta

Kraków 2017

Pragnę serdecznie podziękować wszystkim, których pomoc przyczyniła się do powstania niniejszej rozprawy, w szczególności promotorowi, prof. dr hab. inż. Tadeuszowi Burczyńskiemu oraz promotorowi pomocniczemu, dr inż. Michałowi Berecie za wiele cennych rad i wskazówek oraz poświęcony czas i opiekę naukową. Koleżankom i kolegom z Instytutu Informatyki Politechniki Krakowskiej za inspiracje i twórcze dyskusje.

Żonie Ewelinie za cierpliwość i wsparcie.

# Spis treści

<b>1. Wstęp</b>	<b>1</b>
1.1. Wprowadzenie . . . . .	1
1.2. Cel i teza rozprawy . . . . .	4
1.3. Przegląd treści rozprawy . . . . .	5
<b>2. Elementy teorii zbiorów rozmytych</b>	<b>7</b>
2.1. Zbiory rozmyte . . . . .	7
2.1.1. Definicja zbioru rozmytego . . . . .	8
2.1.2. Podstawowe pojęcia związane ze zbiorami rozmytymi . . . . .	8
2.1.3. Podstawowe klasy funkcji przynależności . . . . .	10
2.1.4. Podstawowe operacje na zbiorach rozmytych . . . . .	14
2.2. Liczby rozmyte . . . . .	15
2.2.1. Definicja liczby rozmytej . . . . .	15
2.2.2. Arytmetyka liczb rozmytych . . . . .	16
2.2.3. Liczby typu L-R . . . . .	18
<b>3. Skierowane liczby rozmyte</b>	<b>21</b>
3.1. Niedoskonałości liczb rozmytych . . . . .	21
3.2. Definicja skierowanej liczby rozmytej . . . . .	22
3.3. Algebra skierowanych liczb rozmytych . . . . .	25
3.3.1. Operacje na skierowanych liczbach rozmytych . . . . .	25
3.3.2. Reprezentacja numeryczna . . . . .	27
3.3.3. Algebra Banacha . . . . .	28
3.4. Operatory wyostrzania . . . . .	29

3.5. Liczby niewłaściwe . . . . .	30
<b>4. Reprezentacja danych za pomocą skierowanych liczb rozmytych</b>	<b>32</b>
4.1. Finansowe szeregi czasowe wysokiej częstotliwości . . . . .	32
4.2. Skierowane świece rozmyte . . . . .	35
4.2.1. Skierowane świece rozmyte z funkcjami ustalonego typu . . . . .	36
4.2.2. Empiryczne skierowane świece rozmyte . . . . .	37
4.3. Przykłady skierowanych świec rozmytych . . . . .	38
4.3.1. Trapezoidalna skierowana świeca rozmyta . . . . .	38
4.3.2. Gaussowska skierowana świeca rozmyta . . . . .	40
4.3.3. Empiryczna skierowana świeca rozmyta . . . . .	41
4.3.4. Szeregi czasowe o identycznych świecach japońskich . . . . .	42
4.4. Dodatkowe informacje o notowaniach . . . . .	43
4.4.1. Wolumen transakcji . . . . .	43
4.4.2. Spread transakcyjny . . . . .	44
<b>5. Modele rozmytych szeregów czasowych</b>	<b>46</b>
5.1. Rozmyte szeregi czasowe . . . . .	46
5.1.1. Podejście oparte na rozmytym wnioskowaniu . . . . .	47
5.1.2. Podejście oparte na rozmytej regresji . . . . .	48
5.1.3. Podejście oparte na modelach Boxa i Jenkinsa . . . . .	50
5.2. Skierowane rozmyte szeregi czasowe . . . . .	52
5.2.1. Skierowany rozmyty model regresji liniowej . . . . .	53
5.2.2. Skierowany rozmyty model autoregresyjny . . . . .	60
<b>6. Symulacje rozmytych szeregów czasowych</b>	<b>65</b>
6.1. Rozmyte zmienne losowe . . . . .	66
6.1.1. Model Kwakernaaka . . . . .	66
6.1.2. Model Puri i Ralescu . . . . .	67
6.1.3. Model Liu i Liu . . . . .	69
6.2. Skierowane rozmyte zmienne losowe . . . . .	70
6.3. Skierowane rozmyte liczby pseudolosowe . . . . .	77

---

<b>7. Przykłady zastosowań</b>	<b>84</b>
7.1. Wycena opcji europejskiej na indeks WIG20 . . . . .	84
7.1.1. Dopasowanie modelu SRAR(p) do notowań . . . . .	85
7.1.2. Analiza reszt modelu . . . . .	90
7.1.3. Wycena opcji europejskiej . . . . .	93
7.2. Dywersyfikacja portfela akcji . . . . .	97
7.2.1. Rozmyte stopy zwrotu . . . . .	98
7.2.2. Rozmyta dywersyfikacja portfela akcji . . . . .	100
7.2.3. Wyniki empiryczne . . . . .	100
7.3. Analiza techniczna rynków finansowych . . . . .	110
7.3.1. Rozmyte wskaźniki analizy technicznej . . . . .	110
7.3.2. Skierowana świeca rozmyta jako wskaźnik . . . . .	112
<b>8. Podsumowanie i wnioski</b>	<b>114</b>
<b>Bibliografia</b>	<b>116</b>
<b>Spis rysunków</b>	<b>124</b>
<b>Spis tabel</b>	<b>127</b>
<b>Streszczenie</b>	<b>128</b>
<b>Abstract</b>	<b>129</b>

# Rozdział 1

## Wstęp

### 1.1. Wprowadzenie

Obecnie w bardzo wielu zagadnieniach rzeczywistych parametry, dane wejściowe wykorzystywane w modelowaniu jak również sam opis modelowanego zjawiska obarczone są różnego rodzaju niepewnością. Nie inaczej sytuacja wygląda w przypadku modelowania finansowych szeregów czasowych jakimi są notowania giełdowe. Niepewność danych giełdowych wynika m.in. stąd, że bardzo duża liczba informacji znajduje odzwierciedlenie na rynku finansowym. Poza aspektami bezpośrednio związanymi z danym instrumentem finansowym (np. kondycja notowanej spółki) są to również czynniki takie jak stan gospodarki, sytuacja polityczna, oczekiwania inwestorów, katastrofy oraz wiele innych niemożliwych do przewidzenia zdarzeń. Ponadto, w ekonometrii do konstruowania modeli finansowych szeregów czasowych najczęściej wykorzystywane są dzienne ceny zamknięcia, pomija się w ten sposób setki danych z całego dnia notowań. Wobec przypadku modelowania danych giełdowych można zatem mówić nie tylko o danych obarczonych niepewnością ale również o niepełnej informacji.

Stąd, coraz większą popularnością cieszą się metody wchodzące w skład tzw. *obliczeń miękkich* (ang. *soft computing*). Jednym z popularniejszych podejść jest podejście rozmyte, w którym niepewność można opisać za pomocą zbiorów i liczb rozmytych. Teoria zbiorów

rozmytych [99, 14] rozszerza pojęcie czysto ostrego zbioru poprzez przypisanie elementom zbioru stopni przynależności tak, aby przejście ze stanu nienależenia do stanu należenia do zbioru było stopniowe, a nie nagłe.

Logika rozmyta znalazła bardzo wiele zastosowanie m.in. w zagadnieniach sterowania z wykorzystaniem zmiennych lingwistycznych [53, 80]. Podejście to wykorzystano również do modelowania regresji nieliniowej poprzez zastosowanie regułowych neuronowo-rozmytych systemów wnioskujących, takich jak system ANFIS (Adaptive Neuro-Fuzzy Inference System) [1, 25]. W obszarze modelowania szeregów czasowych pionierskie prace na temat *rozmytej regresji liniowej* (ang. *fuzzy linear regression*) wprowadzili Tanaka, Uejima i Asai [81]. Zdefiniowali oni model rozmyty jako funkcję liniową, której parametrami są zbiory rozmyte. W przeciwieństwie do konwencjonalnej regresji liniowej, gdzie różnice między obserwacjami i przewidywanymi wartościami traktowane są jako błędy obserwacji, różnice traktowane są jako rozmytość/niepewność w systemie. Pojęcie *rozmytego szeregu czasowego* (ang. *fuzzy time series*) zostało wprowadzone przez Songa i Chissoma [75, 76, 77] oraz Chena [6, 7]. Rozważane przez nich modele rozmytych szeregów czasowych konstruowane są w oparciu o pojęcie relacji rozmytej między przeszłymi wartościami szeregu czasowego. Ozawa wraz ze współpracownikami [64], skonstruowali rozmyty model autoregresyjny w postaci klasycznego równania autoregresji o współczynnikach rzeczywistych, gdzie wartości szeregu czasowego zostały rozmyte za pomocą liczb rozmytych. Natomiast Tseng et al. [86] skonstruowali rozmyty model ARIMA, gdzie współczynniki klasycznego modelu ARIMA zostały zastąpione liczbami rozmytymi.

W modelowaniu finansowych szeregów czasowych wyróżnić można dwa równie ważne aspekty: losowość i niedokładności. Losowość odnosi się do zmienności stochastycznej wszystkich możliwych stanów modelu. Natomiast niedokładność może wynikać z niepełnej wiedzy na temat modelowanego zjawiska. Te dwa pojęcia zostały połączone ze sobą za pomocą koncepcji *rozmytej zmiennej losowej* (ang. *fuzzy random variable*). Obecnie w literaturze pojawiają się trzy definicje rozmytych zmiennych losowych. Pierwsza wprowadzona przez Kwakernaaka [43, 44], który przedstawił rozmytą zmienną losową jako

niejasne postrzeganie ostrej ale nie obserwowalnej zmiennej losowej. Druga zaproponowana przez Puri i Ralescu [70], którzy rozmyte zmienne losowe traktują jak zbiory losowe oraz trzecia zaproponowana przez Liu i Liu [46, 47], którzy koncepcję rozmytej zmiennej losowej oparli na koncepcji miary wiarygodności.

Wszystkie wymienione powyżej modele i definicje odnoszą się do liczb rozmytych zdefiniowanych zgodnie z klasyczną teorią zbiorów rozmytych zaproponowaną przez Zadeha [99], a także do ogólnie akceptowalnej teorii liczb rozmytych, którą zaproponowali Dubois i Prade [14]. Arytmetyka dla tych liczb rozmytych definiuje podstawowe operacje matematyczne na liczbach rozmytych poprzez rozszerzenie tych operacji z liczb rzeczywistych. Sposób ich rozszerzania podaje tzw. zasada rozszerzania Zadeha. Stosowanie tej zasady stwarza parę ograniczeń, a mianowicie bez względu na to jakie operacje arytmetyczne wykonujemy na liczbach rozmytych, nośnik liczby wynikowej powiększa się. Nośniki nazywane są czasem przedziałami nieprecyzyjności. W efekcie może się okazać, że po wykonaniu wielu działań nośnik wynikowej liczby będzie tak szeroki, że informacja, którą dana liczba będzie przenosić, stanie się bezużyteczna.

W latach 2001–2003 Kosiński wraz ze współpracownikami zaproponował nowe ujęcie liczby rozmytej, eliminujące m.in. problem rozszerzania się nośnika [33, 34, 35, 36]. Zdefiniowali oni liczbę rozmytą jako uporządkowaną parę funkcji ciągłych prowadzących z przedziału  $[0, 1]$  w zbiór liczb rzeczywistych i nazwali ją *skierowaną liczbą rozmytą* (ang. *ordered fuzzy number*). Taka definicja pozwoliła im na zdefiniowanie arytmetyki liczb rozmytych, w której istnieją elementy neutralne dla takich działań jak dodawanie i mnożenie oraz możliwe jest rozwiązywanie w niej równań algebraicznych. Dzięki dobrze zdefiniowanej arytmetyce skierowanych liczb rozmytych, możliwe staje się definiowanie w pełni rozmytych modeli przeróżnych zjawisk w postaci prostych modeli równaniowych, w których zmienne objaśniane, objaśniające jak i współczynniki są skierowanymi liczbami rozmytymi. To właśnie te cechy oraz względnie prosta algorytmizacja działań zdecydowała o wykorzystaniu skierowanych liczb rozmytych w zagadnieniach opisywanych w niniejszej rozprawie.



## 1.2. Cel i teza rozprawy

Celem głównym rozprawy jest opracowanie metod, modeli i algorytmów umożliwiających modelowanie i symulację finansowych szeregów czasowych z uwzględnieniem niepewności związanej z notowaniami giełdowymi wysokiej częstotliwości (wyższej niż dzienna). Do celów modelowania niepewności zdecydowano się wykorzystać model skierowanych liczb rozmytych, który dzięki dobrze zdefiniowanej arytmetyce, zgodnej z arytmetyką liczb rzeczywistych, umożliwia definiowanie w pełni rozmytych modeli stochastycznych w postaci klasycznych równań.

W ramach realizacji celu głównego zostały zrealizowane zadania cząstkowe opisane w niniejszej rozprawie:

- Reprezentacja danych giełdowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych (rozmycie danych giełdowych), w taki sposób, aby zmniejszyć rozmiar danych zachowując przy tym jak najwięcej informacji zawartych w notowaniach giełdowych.
- Opracowanie metodologii w pełni rozmytych stochastycznych modeli szeregów czasowych w postaci klasycznych równań, np. takich, jak modele regresji liniowej, autoregresji itp..
- Rozwinięcie modelu skierowanych liczb rozmytych o pojęcie rozmytej zmiennej losowej dla tego typu liczb.
- Opracowanie algorytmu generowania skierowanych liczb pseudolosowych o rozkładzie normalnym.
- Adaptacja metod estymacji parametrów dla skonstruowanych modeli oraz praktyczne zastosowanie ich w zagadnieniach modelowania finansowego.

Zrealizowanie powyżej wymienionych zadań wymagało od autora wiedzy zarówno z zakresu nauk matematycznych jak i informatycznych, m.in. rachunku prawdopodobieństwa, procesów stochastycznych, analizy szeregów czasowych, programowania, metod numerycznych, symulacji komputerowej, a przede wszystkim teorii zbiorów i liczb rozmytych oraz modelowania rozmytego.

Z dokonanego przez autora przeglądu literaturowego wynika, że tego typu podejście do modelowania finansowych szeregów czasowych oraz taki zakres rozprawy doktorskiej, przedstawiony został po raz pierwszy, a częściowe wyniki badań zostały opublikowane w pracach własnych [58, 59, 60].

Określone powyżej cele i zadania doprowadziły do sformułowania następujących tez rozprawy doktorskiej:

**Teza 1.** Możliwe jest przedstawienie danych finansowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych, zachowując przy tym więcej informacji o danym szeregu czasowym (więcej niż w przypadku klasycznych reprezentacji notowań giełdowych takich jak: ceny zamknięcia, słupki OHLC czy świece japońskie).

**Teza 2.** Skonstruowane na bazie skierowanych liczb rozmytych, w pełni rozmyte modele szeregów czasowych, pozwalają na efektywne modelowanie i symulację finansowych szeregów czasowych oraz możliwe jest wykorzystanie ich w praktycznych zagadnieniach inżynierii finansowej, takich jak np. wycena instrumentów finansowych.

### 1.3. Przegląd treści rozprawy

Niniejsza rozprawa składa się z ośmiu rozdziałów. Rozdział pierwszy stanowi wprowadzenie do zagadnień, które są przedmiotem niniejszej rozprawy. Przedstawiono w nim również cele oraz tezy rozprawy.

W rozdziale drugim omówione zostały pojęcia i definicje dotyczące klasycznej koncepcji zbiorów rozmytych zaproponowanej przez Zadeha. Przytoczone zostały także pojęcia dotyczące szczególnej klasy zbiorów rozmytych, a mianowicie liczb rozmytych i ich arytmetyki.

Rozdział trzeci w całości poświęcony jest skierowanym liczbom rozmytym zaproponowanym przez Kosińskiego. Przedstawiono w nim algebrę skierowanych liczb rozmytych, która umożliwia posługiwanie się liczbami rozmytymi w sposób podobny do rachunku na liczbach rzeczywistych.

W rozdziale czwartym omówiono zaproponowaną przez autora koncepcję reprezentowania danych finansowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych. Skierowana liczba rozmyta, utożsamiona z danymi, nazwana jest skierowaną świecą rozmytą. Przedstawiono zarówno ogólną ideę, jak i metody konstrukcji przykładowych klas skierowanych świec rozmytych.

Rozdział piąty poświęcony jest modelom rozmytych szeregów czasowych. W pierwszej części przytoczono znane w literaturze podejścia do modelowania rozmytych szeregów czasowych. W drugiej części omówiono nowe podejście do konstruowania modeli rozmytych szeregów czasowych z wykorzystaniem skierowanych liczb rozmytych. Aspekty takie jak estymacja współczynników modelu oraz predykcja, zostały przedstawione na przykładzie rozmytych modeli regresji liniowej i autoregresji.

W rozdziale szóstym przedstawiono zagadnienie symulacji rozmytych szeregów czasowych, które sprowadzono do zagadnienia symulacji rozmytych zmiennych losowych. W pierwszej części rozdziału przytoczono trzy najbardziej cytowane w literaturze koncepcje rozmytych zmiennych losowych, które odnoszą się do klasycznego pojęcia zbioru rozmytego. Druga część rozdziału poświęcona jest zaproponowanej przez autora koncepcji rozmytej zmiennej losowej dla skierowanych liczb rozmytych. W zaproponowanym podejściu skierowane rozmyte zmienne losowe definiowane są jako uporządkowane pary procesów stochastycznych. Dla nowej definicji wprowadzone zostały pojęcia takie jak wartość oczekiwana oraz wariancja, zarówno o wartościach rozmytych jak i ostrych. W ostatniej części tego rozdziału przedstawiono definicję oraz algorytm generowania skierowanych rozmytych liczb pseudolosowych o rozkładzie normalnym.

W rozdziale siódmym omówione koncepcje z rozdziałów od czwartego do szóstego zastosowane zostały z powodzeniem w kilku wybranych zagadnieniach praktycznych, takich jak wycena instrumentów finansowych, dywersyfikacja portfela akcji oraz analiza techniczna rynków finansowych.

Podsumowanie i wnioski przedstawiono w rozdziale ósmym.

## Rozdział 2

# Elementy teorii zbiorów rozmytych

W niniejszym rozdziale zaprezentowane zostaną podstawowe pojęcia i definicje dotyczące teorii zbiorów rozmytych. Omówione zostaną także pojęcia i definicje dotyczące szczególnej klasy tych zbiorów, a mianowicie liczb rozmytych. W obu przypadkach ograniczono się jedynie do zakresu pojęć ułatwiających zrozumienie dalszych części pracy.

### 2.1. Zbiory rozmyte

W wielu zagadnieniach, klasyczne założenie, że elementy mogą należeć bądź nie należeć do zbioru, jest zbyt restrykcyjne. Najwyraźniej zaobserwować to można na przykładzie modelowania pojęć języka naturalnego. Jednym z popularniejszych tego typu zbiorów jest „zbiór osób o wysokim wzroście”, dla którego powstaje naturalne pytanie o granicę między osobą wysoką a nie wysoką. Przykładowo, jeśli granicą tą będzie 170 cm, osoba o wzroście 171 cm będzie do tego zbioru należeć, a o wzroście 169 cm już nie. Wynika to z „ostrego” przejścia między przynależnością, a nie przynależnością elementów do danego zbioru. Mimo to, porównując te dwie osoby ze sobą można by powiedzieć, że są one podobnego wzrostu. W teorii *zbiorów rozmytych* (ang. *fuzzy sets*), dopuszcza się możliwość częściowej przynależności elementu do zbioru. W ten sposób zarówno osoba o wzroście 171 cm, jak i 169 cm będzie należeć do rozważanego zbioru, jednak każda z nich w innym stopniu.

### 2.1.1. Definicja zbioru rozmytego

Pojęcie zbioru rozmytego podał po raz pierwszy L. A. Zadeh w artykule „*Fuzzy sets*” opublikowanym w czasopiśmie *Information and Control* w 1965 roku [99]. Podobnie jak klasyczne zbiory mogą być opisane za pomocą funkcji charakterystycznych, tak zbiory rozmyte zostały opisane za pomocą tzw. *funkcji przynależności* (ang. *membership functions*). Funkcja charakterystyczna może przyjmować jedynie wartość zero (element nie należy do zbioru) lub jeden (element należy do zbioru). Funkcja przynależności natomiast, przypisuje każdemu elementowi zbioru liczbę rzeczywistą z przedziału  $[0, 1]$ , określającą stopień przynależności elementu do danego zbioru.

Niech dana będzie przestrzeń  $\mathcal{X}$ , niepusta, skończona lub nie. Aby zatem określić zbiór rozmyty na uniwersum  $\mathcal{X}$ , należy każdemu elementowi z przestrzeni  $\mathcal{X}$  przypisać wartość funkcji przynależności. Formalnie, wielu autorów zbiorów rozmyty określa zgodnie z poniższą definicją [11, 29, 45, 99].

**Definicja 2.1.** *Zbiorem rozmytym*  $A$  na uniwersum  $\mathcal{X}$  nazywa się zbiór uporządkowanych par

$$A = \{(x, \mu_A(x)) : x \in \mathcal{X}\}, \quad \text{gdzie } \mu_A(x) : \mathcal{X} \rightarrow [0, 1]. \quad (2.1)$$

Funkcja  $\mu_A$  nazywana jest funkcją przynależności zbioru rozmytego  $A$ , przypisującą każdemu elementowi  $x \in \mathcal{X}$  stopień jego przynależności do zbioru rozmytego  $A$ . Graficznie zbiór rozmyty  $A$  przedstawia się jako wykres jego funkcji przynależności (por. rys. 2.1).

### 2.1.2. Podstawowe pojęcia związane ze zbiorami rozmytymi

Poniżej przytoczonych zostanie kilka pojęć istotnych w teorii zbiorów rozmytych [48, 49]. Interpretacja graficzna niektórych z nich przedstawiona została na rys. 2.1.

- Zbiory klasyczne (dystrybutywne) w dalszej części pracy nazywane będą *zbiorami ostrymi* (ang. *crisp set*).

- *Nośnikiem* (ang. *support*) zbioru rozmytego  $A$  nazywa się ostry zbiór tych elementów uniwersum, dla których funkcja przynależności jest dodatnia

$$\text{supp}(A) = \{x \in \mathcal{X} : \mu_A(x) > 0\}. \quad (2.2)$$

Ponadto, mówi się, że zbiór rozmyty jest *pusty*, gdy jego nośnik jest zbiorem pustym.

- *Jądrem* (ang. *kernel*) zbioru rozmytego  $A$  nazywa się ostry zbiór tych elementów uniwersum, które w pełni należą do zbioru  $A$  (tzn. należą do  $A$  w stopniu 1)

$$\text{ker}(A) = \{x \in \mathcal{X} : \mu_A(x) = 1\}. \quad (2.3)$$

- Liczbę będącą supremum wartości funkcji przynależności zbioru rozmytego  $A$  nazywa się *wysokością* (ang. *height*) tego zbioru

$$\text{hgt}(A) = \sup_{x \in \mathcal{X}} \mu_A(x). \quad (2.4)$$

- Zbiór rozmyty nazywa się *normalnym*, gdy jego wysokość wynosi 1. Każdy niepusty zbiór rozmyty, który nie jest normalny, można znormalizować, dzieląc wartości funkcji przynależności przez wysokość tego zbioru.
- $\alpha$ -*przekrojem* (ang.  $\alpha$ -*cut*) zbioru rozmytego  $A$  nazywa się ostry zbiór tych elementów, dla których funkcja przynależności przyjmuje wartość większą lub równą  $\alpha$ . Zbiór ten oznacza się przez  $A_\alpha$  i formalnie zapisuje jako

$$A_\alpha = \{x \in \mathcal{X} : \mu_A(x) \geq \alpha\}, \quad (2.5)$$

czyli jest to ostry zbiór o następującej funkcji charakterystycznej

$$\chi_{A_\alpha}(x) = \begin{cases} 1, & \mu_A(x) \geq \alpha, \\ 0, & \mu_A(x) < \alpha. \end{cases} \quad (2.6)$$

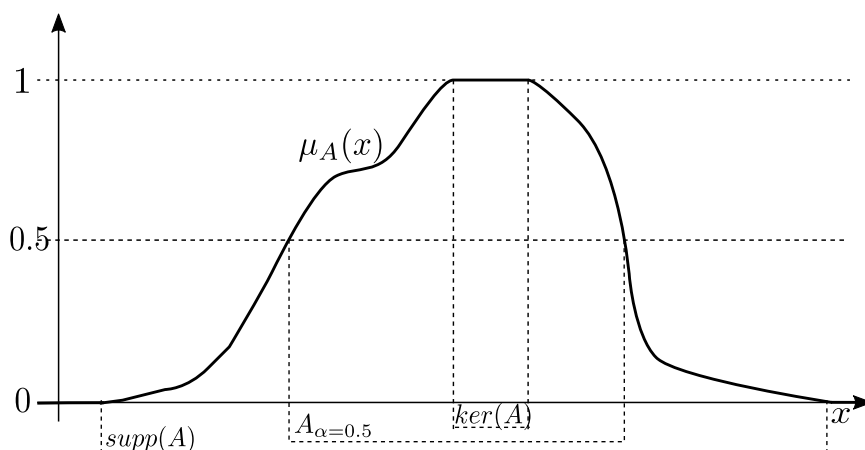
- Zbiór rozmyty  $A$  jest *wypukły* (ang. *convex*) wtedy i tylko wtedy, gdy dla dowolnych  $x_1, x_2 \in \mathcal{X}$  i dla dowolnej liczby  $\lambda \in [0, 1]$  spełniona jest nierówność postaci

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \min(\mu_A(x_1), \mu_A(x_2)). \quad (2.7)$$

- Zbiór rozmyty  $A$  jest *wklęsły* (ang. *concave*) wtedy i tylko wtedy, gdy dla dowolnych  $x_1, x_2 \in \mathcal{X}$  i dla dowolnej liczby  $\lambda \in [0, 1]$  spełniona jest nierówność postaci

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \leq \max(\mu_A(x_1), \mu_A(x_2)). \quad (2.8)$$

- Funkcje przynależności zbiorów wypukłych nazywa się wypukłymi, a wklęsłych funkcjami wklęsłymi. Należy zaznaczyć, że tak pojmowane wypukłość i wklęsłość funkcji przynależności zbiorów rozmytych są istotnie różne od tak samo nazwanych własności analitycznych funkcji.



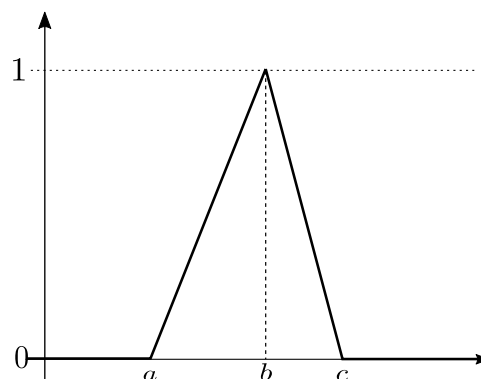
Rysunek 2.1: Nośnik, jądro i  $\alpha$ -przekrój wypukłego zbioru rozmytego  $A$ .

### 2.1.3. Podstawowe klasy funkcji przynależności

Funkcja przynależności należy do najważniejszych pojęć teorii zbiorów rozmytych, gdyż każdy zbiór rozmyty jest jednoznacznie opisany przez swoją funkcję przynależności. W teorii, funkcje przynależności mogą przyjmować dowolny kształt, a ich wartości mogą być interpretowane na różne sposoby, np. jako stopień podobieństwa, preferencji lub niepewności. W praktycznych zastosowaniach często korzysta się z kilku rodzajów funkcji przynależności określonych w zbiorze liczb rzeczywistych. Poniżej wymienione zostaną najczęściej występujące klasy funkcji przynależności [49]:

- *Trójkątna*

$$\mu_A(x; a, b, c) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x \leq b, \\ \frac{c-x}{c-b}, & b < x \leq c, \\ 0, & x > c, \end{cases} \quad (2.9)$$

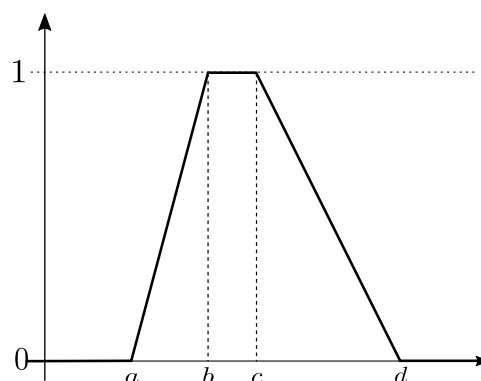


Rysunek 2.2: Trójkątna f. przynależności.

gdzie liczby  $a \leq b \leq c$  są parametrami.

- *Trapezowa*

$$\mu_A(x; a, b, c, d) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x \leq b, \\ 1, & b < x \leq c, \\ \frac{d-x}{d-c}, & c < x \leq d, \\ 0, & x > d, \end{cases} \quad (2.10)$$



Rysunek 2.3: Trapezowa f. przynależności.

gdzie liczby  $a \leq b \leq c \leq d$  są parametrami. Dla  $b = c$  otrzymuje się trójkątną funkcję przynależności. Dla  $a = b$  oraz  $c = d$  uzyskuje się *prostokątną* lub inaczej *przedziałową* funkcję przynależności.

- *Gaussowska*

$$\mu_A(x; m, \sigma) = \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad (2.11)$$

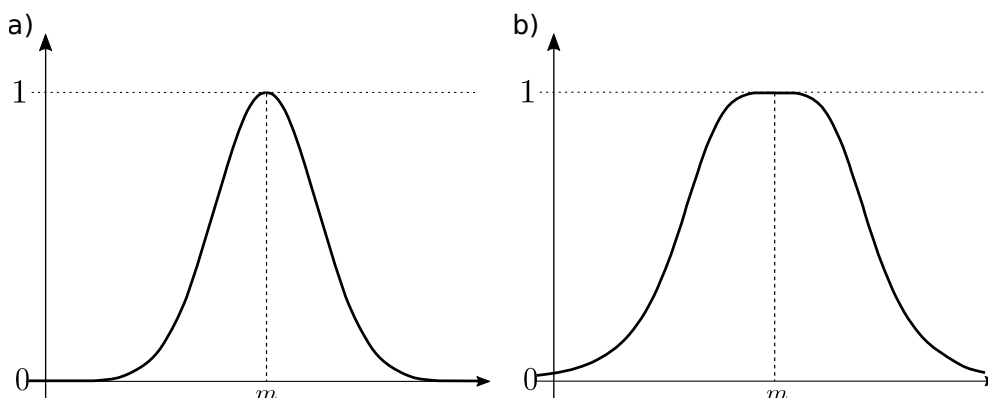
gdzie liczby  $m, \sigma$  są parametrami. Dla  $x = m$  powyższa funkcja przyjmuje wartość 1. Parametr  $\sigma > 0$  określa szerokość zbioru rozmytego (por. rys. 2.4a).



- *Uogólniona dzwonowa*

$$\mu_A(x; m, \sigma, \gamma) = \frac{1}{1 + \left|\frac{x-m}{\sigma}\right|^{2\gamma}}, \quad (2.12)$$

gdzie liczby  $m, \sigma, \gamma$  są parametrami. Dla  $x = m$  powyższa funkcja przyjmuje wartość 1. Parametry  $\sigma > 0, \gamma > 0$  określają szerokość zbioru rozmytego i nachylenie jego zboczy (por. rys. 2.4b).

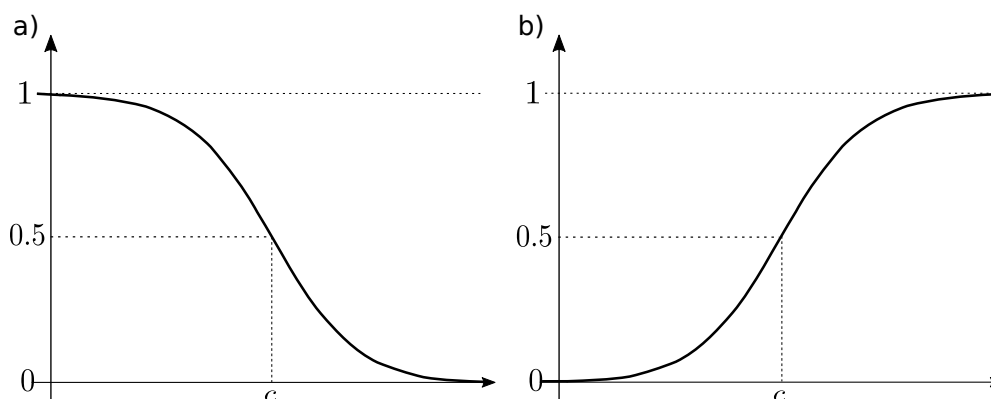


Rysunek 2.4: Funkcje przynależności: a) gaussowska, b) uogólniona dzwonowa.

- *Sigmoidalna*

$$\mu_A(x; c, \beta) = \frac{1}{1 + \exp(-\beta(x - c))}, \quad (2.13)$$

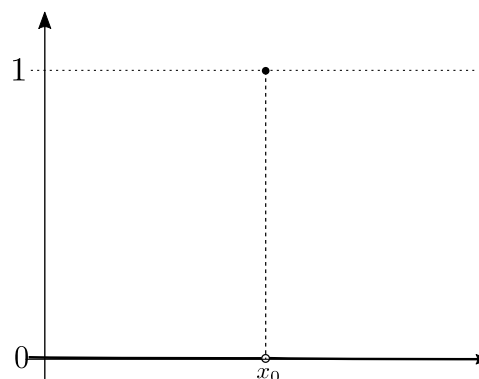
gdzie liczby  $c, \beta$  są parametrami. Parametr  $c$  określa punkt, dla którego stopień przynależności wynosi  $\frac{1}{2}$ , a parametr  $\beta$  wpływa na nachylenie funkcji przynależności (por. rys. 2.5).



Rysunek 2.5: Sigmoidalna funkcja przynależności: a)  $\beta < 0$ , b)  $\beta > 0$ .

- *Rozmyty singleton*

$$\mu_A(x; x_0) = \begin{cases} 1, & x = x_0, \\ 0, & x \neq x_0, \end{cases} \quad (2.14)$$

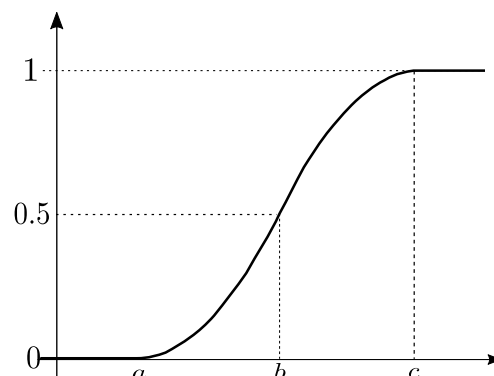


Rysunek 2.6: Rozmyty singleton.

gdzie liczba  $x_0$  jest parametrem określającym położenie singletonu. Rozmyty singleton jest przykładem zbioru, dla którego jądro jest równoważne nośnikowi i składa się z jednego elementu.

- *Funkcja typu S*

$$\mu_A(x; a, c) = \begin{cases} 0, & x \leq a, \\ 2\left(\frac{x-a}{c-a}\right)^2, & a < x \leq b, \\ 1 - 2\left(\frac{x-c}{c-a}\right)^2, & b < x \leq c, \\ 1, & x > c, \end{cases} \quad (2.15)$$



Rysunek 2.7: Funkcja przynależności typu S.

gdzie liczby  $a, c$  są parametrami. Wartość  $b = \frac{a+c}{2}$  jest punktem, dla którego funkcja przyjmuje wartość  $\frac{1}{2}$ . Funkcje przynależności tej klasy często są oznaczane jako  $S(x; a, c)$ .

- *Funkcja klasy Z*

Funkcja klasy  $Z$  jest definiowana za pomocą funkcji klasy  $S$  następująco:

$$\mu_A(x; a, c) = Z(x; a, c) = 1 - S(x; a, c). \quad (2.16)$$

- *Funkcja klasy  $\Pi$*

Funkcję klasy  $\Pi$  również definiowana jest za pomocą funkcji klasy  $S$  w następujący sposób:

$$\mu_A(x; a, b) = \Pi(x; a, b) = \begin{cases} S(x; b - a, b), & x \leq b, \\ 1 - S(x; b, a + b), & x > b, \end{cases} \quad (2.17)$$

gdzie liczby  $a, b$  są parametrami. Parametr  $a$  określa szerokość zbioru rozmytego, a dla  $x = b$  funkcja przyjmuje wartość 1.

#### 2.1.4. Podstawowe operacje na zbiorach rozmytych

Podobnie jak dla zbiorów ostrych, dla zbiorów rozmytych można zdefiniować trzy klasyczne operacje mnogościowe [99]:

- *Dopełnieniem* zbioru rozmytego  $A$  na przestrzeni  $\mathcal{X}$  jest zbiór rozmyty  $\neg A$  określony funkcją przynależności

$$\mu_{\neg A}(x) = 1 - \mu_A(x) \quad \text{dla każdego } x \in \mathcal{X}. \quad (2.18)$$

- *Sumą* zbiorów rozmytych  $A$  i  $B$  na tej samej przestrzeni  $\mathcal{X}$  jest zbiór rozmyty  $A \cup B$  określony funkcją przynależności

$$\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = \max(\mu_A(x), \mu_B(x)) \quad \text{dla każdego } x \in \mathcal{X}. \quad (2.19)$$

- *Iloczynem* zbiorów rozmytych  $A$  i  $B$  na tej samej przestrzeni  $\mathcal{X}$  jest zbiór rozmyty  $A \cap B$  określony funkcją przynależności

$$\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \min(\mu_A(x), \mu_B(x)) \quad \text{dla każdego } x \in \mathcal{X}. \quad (2.20)$$

Tak zdefiniowane przez Zadeha operacje mnogościowe posiadają własności przemienności, łączności i rozdzielności, a ponadto zachodzą również prawa de Morgana oraz absorpcji. Nie są spełnione natomiast prawa dopełnienia. Dzięki prawu łączności, sumę i iloczyn zbiorów rozmytych można zdefiniować również dla skończonej liczby argumentów [48]. Warto zaznaczyć, że nie są to jedyne definicje tych operacji. Standardowe wymagania dla operacji sumy i iloczynu zbiorów rozmytych znane są pod nazwą dualnych norm

trójkątnych (odpowiednio  $T$ -normy i  $S$ -normy). W literaturze zaproponowano wiele różnych par norm trójkątnych [11, 29, 45, 72, 97].

## 2.2. Liczby rozmyte

Jak piszą Czogała i Pedrycz, jednym z ciekawszych aspektów teorii zbiorów rozmytych jest koncepcja różnego rodzaju obliczeń na liczbach nieprecyzyjnych (przybliżonych) [11]. Przykładowo, mogą to być zbiory rozmyte opisujące stwierdzenia języka naturalnego dotyczące liczb (przedziałów liczbowych), na przykład stwierdzenia typu „temperatura wynosi około  $23^{\circ}\text{C}$ ”, na których występuje konieczność wykonywania obliczeń arytmetycznych. Wobec czego w teorii zbiorów rozmytych wyszczególniono pewną klasę zbiorów rozmytych nazwanych *liczbami rozmytymi* (ang. *fuzzy numbers*) [11, 14, 29, 45, 72, 84].

### 2.2.1. Definicja liczby rozmytej

W literaturze nie ma jednej ścisłej definicji liczby rozmytej. Niektórzy z autorów definicją liczby rozmytej obejmują zarówno zbiory o jądrze jednoelementowym (właściwe liczby rozmyte), jak i zbiory o jądrze w postaci przedziału (płaskie liczby rozmyte lub inaczej przedziały rozmyte). Inni oba te pojęcia definiują osobno. W niniejszej pracy przyjęto definicję liczby rozmytej w następującej postaci:

**Definicja 2.2.** Zbiór rozmyty określony na przestrzeni liczb rzeczywistych nazywa się *liczbą rozmytą* jeśli spełnia on następujące warunki:

1. Zbiór ten jest normalny.
2. Zbiór ten jest wypukły.
3. Nośnik tego zbioru jest przedziałem.
4. Funkcja przynależności tego zbioru jest przedziałami ciągła.

Zbiory rozmyte spełniające powyższe warunki nazywane są również *wypukłymi liczbami rozmytymi*. Tak przyjęta definicja obejmuje zarówno właściwe liczby rozmyte, jak

i przedziały rozmyte. Ponadto, dopuszcza zbiory jedno punktowe (czyli np. liczby rzeczywiste reprezentowane przez funkcje charakterystyczne), a także zbiory rozmyte o nieograniczonym nośniku (np. zbiory o funkcji przynależności klasy  $S$  czy  $Z$ ), opisujące pojęcia typu „5 lub więcej” albo „mniej niż  $-5$ ”. Liczby rozmyte można podzielić na liczby dodatnie, ujemne i mieszane. Liczbę rozmytą nazywa się dodatnią (ujemną), gdy jej funkcja przynależności jest zerowa dla ujemnych (dodatnich) argumentów. Liczbę rozmytą nazywa się mieszaną jeżeli nie jest ani dodatnia, ani ujemna.

### 2.2.2. Arytmetyka liczb rozmytych

Podstawowe operacje arytmetyczne na liczbach rozmytych można zdefiniować korzystając z tzw. *zasady rozszerzania Zadeha* (ang. *Zadeh's extension principle*), opracowanej przez Zadeha [100, 102], a później rozwiniętej przez Yagera [96]. Zasada ta pozwala na rozszerzenie dziedziny funkcji matematycznych z tradycyjnych zbiorów lub ich elementów do zbiorów rozmytych. Ogólniej, dla danych dwóch zbiorów rozmytych  $A$  i  $B$ , określonych na dwóch przestrzeniach, odpowiednio  $\mathcal{X}$  i  $\mathcal{Y}$ . Dla danej funkcji  $f: \mathcal{X} \rightarrow \mathcal{Y}$ , zbiór  $A$  wyznacza poprzez funkcję  $f$  zbiór rozmyty  $B$  w  $\mathcal{Y}$ , dany przez

$$\mu_B(y) = \sup_{x \in f^{-1}(\{y\})} \mu_A(x) \quad \text{dla } y \in \mathcal{Y}. \quad (2.21)$$

W przypadku operacji (funkcji) dwuargumentowej, gdy  $F: \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$ , to zbiór  $C$  w  $\mathcal{Z}$  dany jest przez

$$\mu_C(z) = \sup_{\{(x,y): F(x,y)=z\}} \min(\mu_A(x), \mu_B(y)) \quad \text{dla } z \in \mathcal{Z}. \quad (2.22)$$

Stąd, korzystając z zasady rozszerzania, dla liczb rozmytych  $A$  i  $B$  będących argumentami definiuje się operacje arytmetyczne w następujący sposób:

- dodawanie

$$\mu_{A+B}(z) = \sup_{\{(x,y): x+y=z\}} \min(\mu_A(x), \mu_B(y)) \quad \text{dla } z \in \mathcal{Z}, \quad (2.23)$$

- odejmowanie

$$\mu_{A-B}(z) = \sup_{\{(x,y): x-y=z\}} \min(\mu_A(x), \mu_B(y)) \quad \text{dla } z \in \mathcal{Z}, \quad (2.24)$$

- mnożenie

$$\mu_{A \cdot B}(z) = \sup_{\{(x,y): x \cdot y=z\}} \min(\mu_A(x), \mu_B(y)) \quad \text{dla } z \in \mathcal{Z}, \quad (2.25)$$

- dzielenie

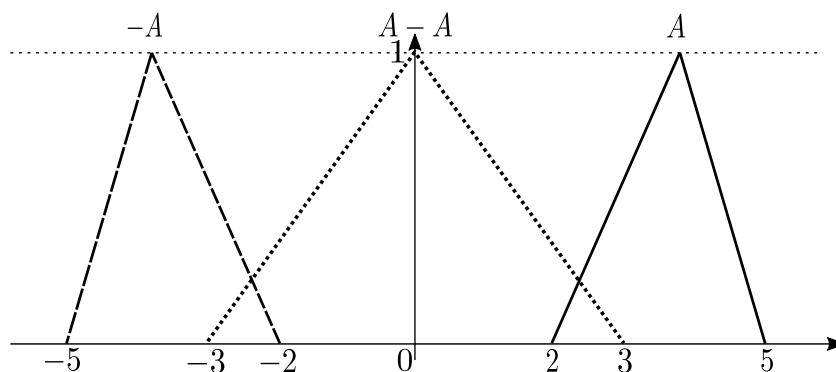
$$\mu_{A/B}(z) = \sup_{\{(x,y): x/y=z\}} \min(\mu_A(x), \mu_B(y)) \quad \text{dla } z \in \mathcal{Z}, \quad (2.26)$$

Ponadto, traktując liczby rzeczywiste 0 oraz 1 jako liczby rozmyte i wstawiając odpowiednio do (2.24) i (2.26), uzyskuje się

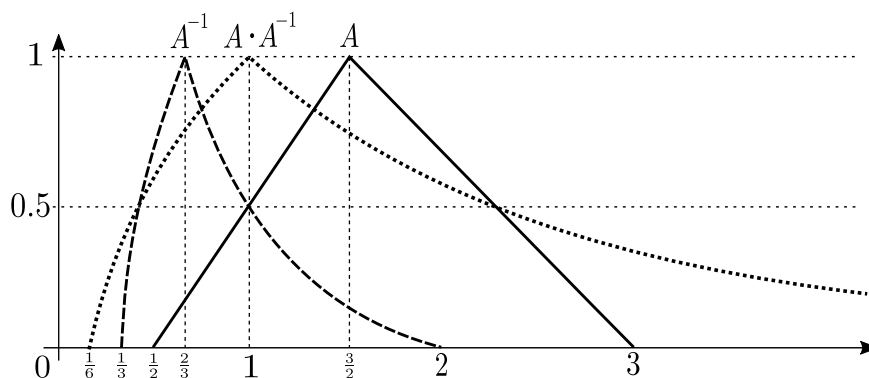
$$\mu_{0-A}(z) = \mu_{-A}(z) = \mu_A(-z), \quad (2.27)$$

$$\mu_{1/B}(z) = \mu_{B^{-1}}(z) = \mu_B(1/z). \quad (2.28)$$

Powyżej zdefiniowane dodawanie liczb rozmytych jest przemienne i łączne oraz zachodzą równości  $-(A+B) = (-A)+(-B)$ ,  $A \pm 0 = A$ . Nie istnieje natomiast element przeciwny dla dodawania (czyli  $A+(-A) \neq 0$ ) (por. rys. 2.8). Podobnie mnożenie jest łączne i przemienne, zachodzą równości  $(-A) \cdot B = -(A \cdot B)$ ,  $A/B = A \cdot B^{-1}$ ,  $A \cdot 1 = A$ . Nie istnieje element odwrotny dla mnożenia (czyli  $A \cdot A^{-1} \neq 1$ ) (por. rys. 2.9), a rozdzielność  $A \cdot (B+C) = (A \cdot B) + (A \cdot C)$  zachodzi jedynie dla jednakowych znaków  $B$  i  $C$  [48, 49].



Rysunek 2.8: Suma liczby rozmytej i liczby przeciwnej do niej.



Rysunek 2.9: Iloczyn liczby rozmytej i liczby odwrotnej do niej.

### 2.2.3. Liczby typu L-R

Przytoczone wzory na działania na liczbach rozmytych zawierają operator supremum, co powoduje, że policzenie każdego stopnia przynależności liczby wynikowej wymaga wykonania nieskończenie wielu operacji na stopniach przynależności argumentów. Wobec czego, nie są one łatwo algorytmizowalne, nawet jeśli nośniki nieskończone zastąpimy dostatecznie dokładnymi nośnikami skończonymi. Podobne problemy napotkamy definiując działania na liczbach rozmytych w języku  $\alpha$ -przekrojów funkcji przynależności [21].

Dubois i Prade dla ułatwienia rachunków zaproponowali ograniczoną klasę funkcji przynależności, umiejscawiając w niej tzw. *liczby typu L-R*, z dwoma funkcjami kształtu  $L$  i  $R$  [14].

**Definicja 2.3.** *Przedziałem rozmytym typu L-R* (ang. *L-R fuzzy interval*) nazywamy zbiór rozmyty  $A$  określony na przestrzeni liczb rzeczywistych, taki że

$$\mu_A(x) = \begin{cases} L\left(\frac{m-x}{\alpha}\right) & \text{dla } x < m \\ 1 & \text{dla } m \leq x \leq n \\ R\left(\frac{x-n}{\beta}\right) & \text{dla } x > n \end{cases} \quad (2.29)$$

gdzie  $\alpha, \beta, m, n \in \mathbb{R}$ , takie, że  $\alpha, \beta > 0, m < n$  oraz  $L$  i  $R$ , ustalone funkcje bazowe.

Przez funkcję bazową rozumie się każdą funkcję rzeczywistą  $f: \mathbb{R}^+ \rightarrow [0, 1)$ , słabo malejącą, przedziałami ciągłą oraz zbieżną do wartości 1 dla argumentów dążących do 0. Przykładami funkcji bazowych są następujące trzy rodziny funkcji, dla ustalonego rzeczywistego  $p > 0$ :

- $f(x) = e^{-x^p}$
- $f(x) = \frac{1}{1+x^p}$
- $f(x) = \max(0, 1 - x^p)$

Przedział rozmyty typu  $L$ - $R$  zapisuje się w postaci  $(m, n, \alpha, \beta)_{LR}$ . Jeżeli  $n = m$  to przedział  $L$ - $R$  staje się *liczbą rozmytą typu  $L$ - $R$* , którą zapisuje się w postaci  $(m, \alpha, \beta)_{LR}$ . Ponadto, mówi się, że dwie liczby typu  $L$ - $R$  należą do tej samej klasy, gdy mają te same funkcje bazowe  $L$  i  $R$ . Natomiast, należą do klas przeciwnych, gdy jedna ma funkcje bazowe  $L$ , i  $R$ , a druga przeciwnie - funkcje  $R$  i  $L$ . W pozostałych przypadkach należą one do różnych klas. Tak zdefiniowane przedziały (liczby) rozmyte są liczbami rozmytymi zgodnie z definicją 2.2.

Działania dodawania i odejmowania na liczbach typu  $L$ - $R$  sprowadzają się do prostych operacji na ich parametrach (o ile funkcje bazowe tych liczb są tej samej klasy). Niech  $A = (m_A, \alpha_A, \beta_A)_{LR}$ ,  $B = (m_B, \alpha_B, \beta_B)_{LR}$  to liczby tej samej klasy, wówczas:

- $-A = -(m_A, \alpha_A, \beta_A)_{LR} = (-m_A, \beta_A, \alpha_A)_{RL}$
- $A + B = (m_A, \alpha_A, \beta_A)_{LR} + (m_B, \alpha_B, \beta_B)_{LR} = (m_A + m_B, \alpha_A + \alpha_B, \beta_A + \beta_B)_{LR}$
- $A - B = (m_A, \alpha_A, \beta_A)_{LR} - (m_B, \alpha_B, \beta_B)_{RL} = (m_A - m_B, \alpha_A + \beta_B, \beta_A + \alpha_B)_{LR}$

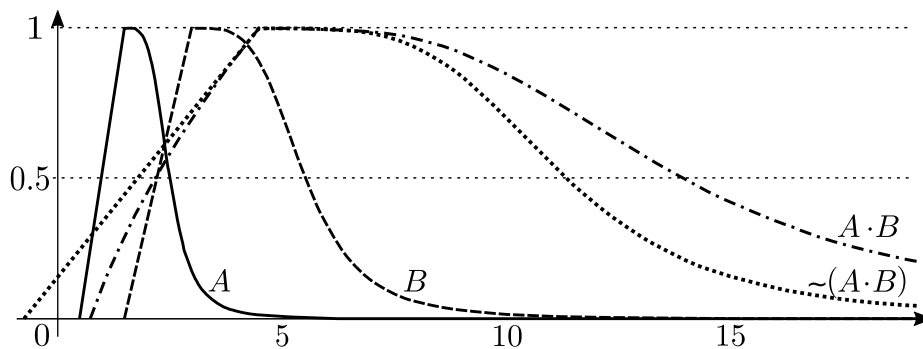
Mnożenie i dzielenie liczb typu  $L$ - $R$ , podobnie jak odwrotność takiej liczby, nie są liczbami tego typu  $L$ - $R$ . Można jednak zaproponować przybliżone wyniki tych operacji w postaci liczb typu  $L$ - $R$ . Wzory na mnożenie i dzielenie wyprowadza się dla czterech przypadków zależnych od znaków tych liczb. Poniżej podane zostały przykładowe przybliżenia tylko dla przypadku liczb dodatnich. Wzory dla pozostałych przypadków, a także inne przybliżenia można znaleźć w [48, 66].

- $A^{-1} = (m_A, \alpha_A, \beta_A)_{LR}^{-1} = \left( \frac{1}{m_A}, \frac{\beta_A}{m_A^2}, \frac{\alpha_A}{m_A^2} \right)_{RL}$



- $A \cdot B = (m_A, \alpha_A, \beta_A)_{LR} \cdot (m_B, \alpha_B, \beta_B)_{LR}$   
 $\approx (m_A \cdot m_B, m_B \cdot \alpha_A + m_A \cdot \alpha_B, m_B \cdot \beta_A + m_A \cdot \beta_B)_{LR}$
- $A : B = (m_A, \alpha_A, \beta_A)_{LR} : (m_B, \alpha_B, \beta_B)_{RL}$   
 $\approx \left( \frac{m_A}{m_B}, \frac{m_B \cdot \alpha_A + m_A \cdot \beta_B}{m_B^2}, \frac{m_B \cdot \beta_A + m_A \cdot \alpha_B}{m_B^2} \right)_{LR}$

Na rys. 2.10 przedstawiono iloczyn  $(A \cdot B)$  dwóch liczb rozmytych typu  $L$ - $R$ :  $A = (1.5, 1, 1)_{LR}$ ,  $B = (3, 1.5, 2.5)_{LR}$  oraz jego przybliżenie w postaci liczby typu  $L$ - $R$  ( $\sim (A \cdot B)$ ). W przedstawionym przykładzie, obie liczby składają się z funkcji bazowych postaci  $L(x) = \max(0, 1 - x)$  oraz  $R(x) = \frac{1}{1 + x^4}$ . Jak można zauważyć, różnica między rzeczywistym wynikiem mnożenia liczb  $A$  i  $B$  a jego przybliżeniem jest znacząca.



Rysunek 2.10: Iloczyn liczb rozmytych typu  $L$ - $R$ .

# Rozdział 3

## Skierowane liczby rozmyte

W rozdziale tym omówiony zostanie model skierowanych liczb rozmytych, który odbiega nieco od koncepcji liczb rozmytych przedstawionych w poprzednim rozdziale. Przedstawiona zostanie algebra skierowanych liczb rozmytych, która umożliwia posługiwanie się pojęciami rozmytymi, ilościowo nieostrymi, w sposób podobny do rachunku na liczbach rzeczywistych. To właśnie ta cecha, oraz względnie prosta algorytmizacja działań, zdecydowała, o wykorzystaniu skierowanych liczb rozmytych w zagadnieniach opisywanych w niniejszej rozprawie.

### 3.1. Niedoskonałości liczb rozmytych

Liczby rozmyte przedstawione w rozdziale 2 wraz z ich arytmetyką opartą na zasadzie rozszerzania Zadeha, obarczone są pewnymi niedoskonałościami.

**Błędy zaokrągleń.** Aproksymacja wyników mnożenia i dzielenia liczb rozmytych, prowadzi w naturalny sposób do błędów zaokrągleń, które w miarę obliczeń kumulują się, czasem bez możliwości kontroli [89, 90].

**Rozszerzanie nośnika.** Bez względu na to jakie działania wykonuje się na liczbach rozmytych, nośnik liczb wynikowych ulega powiększeniu. Nośniki nazywane są czasem przedziałami nieprecyzyjności. W konsekwencji, ich powiększanie się nazywane jest niekiedy zwiększaniem nieprecyzyjności lub zwiększaniem rozmytości liczby rozmytej. W efekcie

może się okazać, że po wykonaniu wielu działań, nośnik wynikowej liczby będzie tak szeroki, że informacja, którą dana liczba będzie przenosić, stanie się mniej użyteczna [37].

**Brak spójności z arytmetyką liczb rzeczywistych.** Mimo, że dla dowolnej liczby rozmytej  $A$  zachodzi równość  $A + 0 = A$ , gdzie  $0$  jest liczbą rzeczywistą reprezentowaną przez funkcję charakterystyczną, to jednak nie zachodzi warunek na istnienie elementów odwrotnych (tj. dla każdej liczby rozmytej  $A$  istnieje liczba rozmyta  $T$  taka, że  $A + T = 0 = T + A$ ). Przyjmując za  $T$  liczbę  $(-A)$  zgodnie z wzorem (2.27), wynikiem działania  $A + (-A)$  jest zero rozmyte, a nie zero rzeczywiste (por. rys. 2.8). Analogicznie sytuacja wygląda dla elementu neutralnego i elementów odwrotnych dla mnożenia liczb rozmytych. Kolejną cechą niezgodną z intuicją i działaniami na liczbach rzeczywistych są operacje na prostych równaniach. Dla dowolnych liczb rzeczywistej  $a, b, c$  prawdziwa jest implikacja  $(a + b = c) \Rightarrow (c - b = a \wedge c - a = b)$ . Zależność ta na ogół nie zachodzi dla liczb rozmytych zdefiniowanych w poprzednim rozdziale, co prowadzi do braku możliwości rozwiązania prostego równania  $A + X = C$ . Co więcej, jeżeli nośnik liczby  $C$  jest mniejszy niż nośniki liczby  $A$ , to taka liczba  $X$  nie istnieje [37, 67].

W literaturze istnieją różne podejścia, zwykle dość złożone, które pozwalają na obejście niektórych problemów z działaniami na liczbach rozmytych [73, 32]. Model skierowanych liczb rozmytych jest względnie prostą w implementacji alternatywą dla tych rozwiązań.

## 3.2. Definicja skierowanej liczby rozmytej

W. Kosiński, wraz ze współpracownikami, postawili sobie za cel stworzenie nowego modelu liczb rozmytych, pozbawionego wspomnianych wcześniej niedogodności. Zaproponowane przez nich na przełomie lat 2002–2003 uogólnienie liczby rozmytej, nazwane *skierowaną liczbą rozmytą*<sup>1</sup> (ang. *ordered fuzzy number*) prowadzi do istnienia wymaganych elementów neutralnych, możliwości rozwiązywania równań oraz daje możliwość kontrolowania rozmycia wyników działań. Nowy model spełnia nie tylko aksjomaty pierścienia,

<sup>1</sup>Po śmierci W. Kosińskiego, aby uhonorować jego wkład w rozwój i popularyzację skierowanych liczb rozmytych, zaproponowano zmianę nazwy na *liczby rozmyte Kosińskiego* (ang. *Kosinski's fuzzy numbers*) [68, 69].

ale i algebry, czego nie można powiedzieć o liczbach rozmytych zdefiniowanych z użyciem pojęcia zbioru rozmytego według Zadeha. Rezultaty swoich prac przedstawili oni w wielu publikacjach m.in. [34, 35, 36, 37, 38].

**Definicja 3.1.** *Skierowaną liczbę rozmytą (SLR)* definiuje się jako uporządkowaną parę ciągłych funkcji rzeczywistych określonych na przedziale  $[0, 1]$ , tj.

$$A = (f_A, g_A), \quad \text{gdzie } f_A, g_A: [0, 1] \rightarrow \mathbb{R} \text{ są funkcjami ciągłymi.} \quad (3.1)$$

Funkcje  $f_A$  i  $g_A$  nazywane są odpowiednio częścią *up* i częścią *down* skierowanej liczby rozmytej  $A$ . Z ciągłości obu części wynika, że ich obrazy są ograniczonymi przedziałami<sup>2</sup>, które oznaczane są odpowiednio  $UP = (l_A, 1_A^-)$  i  $DOWN = (1_A^+, p_A)$ .

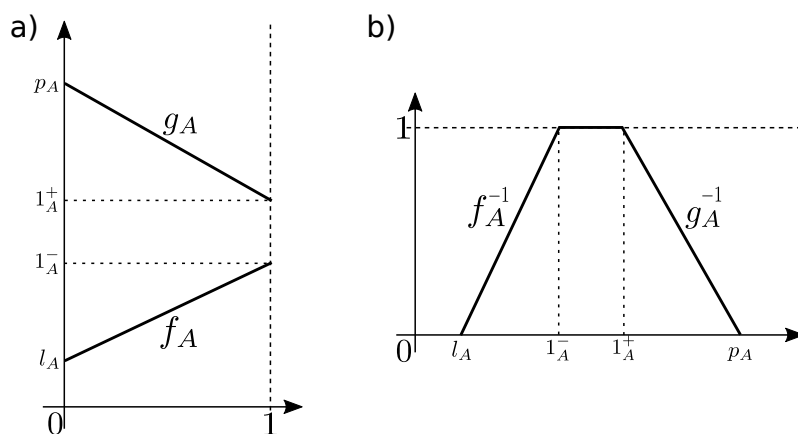
W definicji SLR wymagana jest jedynie ciągłość funkcji  $f_A$  oraz  $g_A$ , niemniej, jeżeli założy się, że są one ściśle monotoniczne, wówczas istnieją dla nich funkcje odwrotne  $f_A^{-1}$  i  $g_A^{-1}$  określone na przedziałach, odpowiednio  $UP$  i  $DOWN$ . Jeżeli dodatkowo założy się, że  $f_A$  jest rosnąca, a  $g_A$  malejąca, takie że  $f_A \leq g_A$ , wtedy dokładając do nich funkcję stałe równą 1 na przedziale  $[1_A^-, 1_A^+]$ , otrzymuje się razem z  $UP$  i  $DOWN$  jeden przedział, na którym można zdefiniować funkcję przynależności  $\mu_A$  zbioru rozmytego określonego na  $\mathbb{R}$  w następujący sposób

$$\mu(x) = \begin{cases} f_A^{-1}(x) & \text{dla } x \in UP, \\ g_A^{-1}(x) & \text{dla } x \in DOWN, \\ 1 & \text{dla } x \in [1_A^-, 1_A^+], \\ 0 & \text{w p.p.} \end{cases} \quad (3.2)$$

Tak zdefiniowana funkcja przynależności  $\mu_A$  definiuje wypukłą liczbę rozmytą rozumianą w klasycznym sensie. Zatem, można powiedzieć, że SLR stanowią rozszerzenie wypukłych liczb rozmytych, ponieważ każdą wypukłą liczbę rozmytą o ciągłej funkcji przynależności można przedstawić w postaci SLR. Graficzna interpretacja SLR oraz jej postać w klasycznym sensie przedstawiona została na rys. 3.1. Te skierowane liczby rozmyte, które da się przedstawić w postaci klasycznych liczby rozmytych, nazywane są

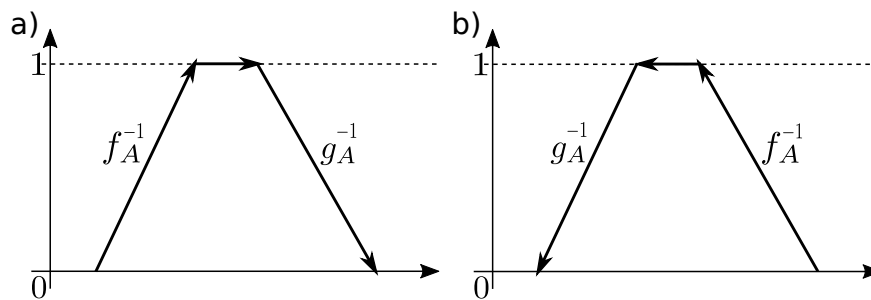
<sup>2</sup>Wniosek z tw. Weierstrassa o przyjmowaniu kresów.

właściwymi SLR. Ponadto, skierowane liczby rozmyte stanowią również rozszerzenie liczb rzeczywistych, bowiem każdą liczbę rzeczywistą reprezentuje SLR złożona z funkcji  $f_A$  i  $g_A$  stałych i równych sobie.



Rysunek 3.1: a) Skierowana liczba rozmyta, b) Skierowana liczba rozmyta jako klasyczna liczba rozmyta.

Skierowane liczby rozmyte wyposażone są w dodatkową własność nazywaną *orientacją* (*skierowaniem*). Własność ta wynika wprost z definicji skierowanej liczby rozmytej jako pary uporządkowanej, stąd liczba  $(f_A, g_A)$  nie jest tym samym co liczba  $(g_A, f_A)$ . Dla właściwych SLR, można wyróżnić dwa typy orientacji. Orientację *pozytywną* (*dodatnią*) wtedy, gdy kierunek skierowanej liczby rozmytej jest zgodny z kierunkiem osi  $OX$  liczb rzeczywistych oraz orientację *negatywną* (*ujemną*) wtedy, gdy kierunek skierowanej liczby rozmytej jest przeciwny do kierunku osi  $OX$  liczb rzeczywistych. Graficznie, orientacja oznaczana jest za pomocą strzałki ze zwrotem od funkcji  $f_A$  do  $g_A$ . Przykłady SLR o orientacji pozytywnej i negatywnej przedstawione zostały na rys. 3.2. Orientacja danej liczby rozmytej  $A$  ilustruje tylko położenie części up  $f_A$  względem części down  $g_A$ . Liczby rzeczywiste reprezentowane przez SLR w oczywisty sposób nie są zorientowane ani dodatnio ani ujemnie.



Rysunek 3.2: Skierowana liczba rozmyta zorientowana: a) pozytywnie, b) negatywnie.

### 3.3. Algebra skierowanych liczb rozmytych

#### 3.3.1. Operacje na skierowanych liczbach rozmytych

Skierowane liczby rozmyte to para funkcji rzeczywistych, wobec czego podstawowe operacje arytmetyczne zdefiniowane są w naturalny sposób jako operacje na tych funkcjach.

**Definicja 3.2.** Niech  $A = (f_A, g_A)$ ,  $B = (f_B, g_B)$  oraz  $C = (f_C, g_C)$  będą skierowanymi liczbami rozmytymi. Wówczas działania: *dodawania*  $C = A + B$ , *odejmowania*  $C = A - B$ , *mnożenia*  $C = A \cdot B$  oraz *dzielenia*  $C = A \div B$  zdefiniowane są następująco

$$f_C(x) = f_A(x) * f_B(x) \wedge g_C(x) = g_A(x) * g_B(x) \quad \text{dla każdego } x \in [0, 1], \quad (3.3)$$

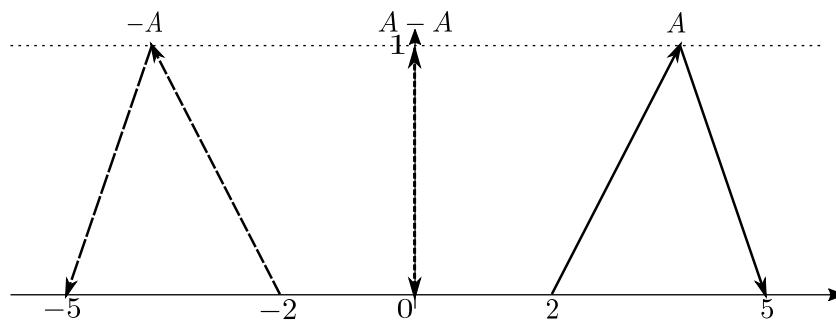
gdzie symbol  $*$  zastępuje odpowiednio  $+$ ,  $-$ ,  $\cdot$  lub  $\div$ , działanie dzielenia  $C = A \div B$  zdefiniowane jest przy założeniu, że liczba  $B$  nie zawiera zera (tj.  $|f_B| > 0$  oraz  $|g_B| > 0$ ).

**Definicja 3.3.** Działanie mnożenia skierowanej liczby rozmytej  $A = (f_A, g_A)$  przez skalar  $\lambda \in \mathbb{R}$  definiuje się następująco

$$C = \lambda \cdot A, \quad \text{gdzie } f_C(x) = \lambda \cdot f_A(x) \wedge g_C(x) = \lambda \cdot g_A(x) \quad \text{dla każdego } x \in [0, 1]. \quad (3.4)$$

Warto, zauważyć, że odejmowanie od liczby  $A$  liczby  $B$  jest tym samym, co dodanie do liczby  $A$  liczby przeciwnej do  $B$ , w postaci

$$(-B) = (-1) \cdot B = (-f_B, -g_B). \quad (3.5)$$



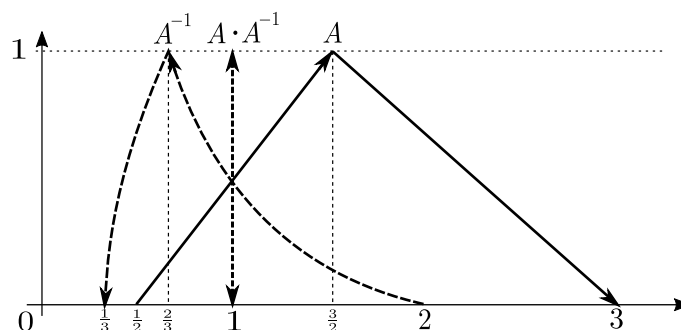
Rysunek 3.3: Suma skierowanej liczby rozmytej i liczby przeciwnej do niej.

W szczególności, w wyniku działania  $A + (-A)$  otrzymuje się liczbę o funkcjach  $f$  i  $g$  stale równych zero, a więc zero rzeczywiste (por. rys. 3.3). Oznacza to, że odejmowanie SLR jest niezgodne z zasadą rozszerzania Zadeha dla klasycznych liczb rozmytych. Liczbą przeciwną w sensie Zadeha do skierowanej liczby rozmytej  $A = (f_A, g_A)$  jest liczba  $\bar{A} = (-g_A, -f_A)$ . Wówczas suma  $A + \bar{A}$  daje w wyniku (na ogół) liczbę około zera (rozmyte zero). W arytmetyce SLR, liczba  $\bar{A}$  spełnia rolę *liczby komplementarnej* (ang. *complementary number*) do liczby  $A$ .

Dzielenie liczby  $A$  przez liczbę  $B$  jest tym samym co pomnożenie liczby  $A$  przez liczbę odwrotną do  $B$ , w postaci

$$B^{-1} = \frac{1}{B} = \left( \frac{1}{f_B}, \frac{1}{g_B} \right). \quad (3.6)$$

Mnożąc liczbę  $A$  przez jej odwrotność, otrzymujemy liczbę o funkcjach  $f$  i  $g$  stale równych jeden, a więc jedynekę rzeczywistą (por. rys. 3.4).



Rysunek 3.4: Iloczyn skierowanej liczby rozmytej i liczby odwrotnej do niej.

W analogiczny sposób jak podstawowe operacje arytmetyczne dla skierowanych liczb rozmytych można zdefiniować funkcje, takie jak, wartość bezwzględna, potęgowanie czy logarytmowanie [67].

**Definicja 3.4.** Niech  $A = (f_A, g_A)$ , będzie skierowaną liczbą rozmytą, wówczas wartością bezwzględną z liczby  $A$  jest skierowana liczba rozmyta  $|A| = C = (f_C, g_C)$  taka, że

$$f_C(x) = |f_A(x)| \wedge g_C(x) = |g_A(x)| \quad \text{dla każdego } x \in [0, 1]. \quad (3.7)$$

**Definicja 3.5.** Niech  $A = (f_A, g_A)$  oraz  $B = (f_B, g_B)$  będą skierowanymi liczbami rozmytymi, wówczas wynikiem podniesienia liczby  $A$  do potęgi  $B$  jest skierowana liczba rozmyta  $A^B = C = (f_C, g_C)$  taka, że

$$f_C(x) = f_A(x)^{f_B(x)} \wedge g_C(x) = g_A(x)^{g_B(x)} \quad \text{dla każdego } x \in [0, 1]. \quad (3.8)$$

**Definicja 3.6.** Niech  $A = (f_A, g_A)$  oraz  $B = (f_B, g_B)$  będą skierowanymi liczbami rozmytymi, wówczas logarytmem o podstawie  $B$  z liczby  $A$  jest skierowana liczba rozmyta  $\log_B A = C = (f_C, g_C)$  taka, że

$$f_C(x) = \log_{f_B(x)} f_A(x) \wedge g_C(x) = \log_{g_B(x)} g_A(x) \quad \text{dla każdego } x \in [0, 1]. \quad (3.9)$$

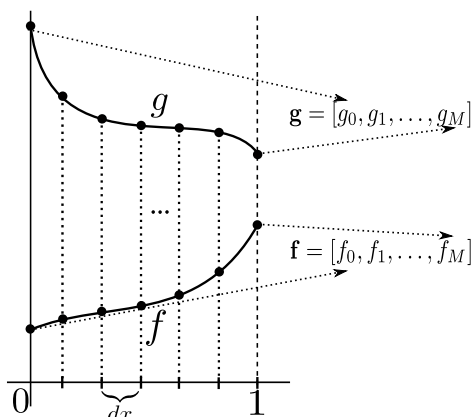
### 3.3.2. Reprezentacja numeryczna

Na potrzeby obliczeń numerycznych, w niniejszej rozprawie skierowana liczba rozmyta  $A = (f, g)$  reprezentowana jest poprzez uporządkowaną parę wektorów  $(\mathbf{f}, \mathbf{g})$ . Wektory  $\mathbf{f}$  i  $\mathbf{g}$  składają się z wartości odpowiednio funkcji  $f$  i  $g$  obliczonych dla  $(M + 1)$  punktów powstałych w wyniku równomiernej dyskretyzacji przedziału  $[0, 1]$  tj.

$$\begin{aligned} \mathbf{f} &= [f(0), f(dx), f(2 \cdot dx), \dots, f((M - 1) \cdot dx), f(1)], \\ \mathbf{g} &= [g(0), g(dx), g(2 \cdot dx), \dots, g((M - 1) \cdot dx), g(1)], \end{aligned} \quad dx = \frac{1}{M}. \quad (3.10)$$

Taka, dyskretna reprezentacja SLR pozwala na wygodną implementację operacji zdefiniowanych w poprzednim podrozdziale, których wynikiem mogą być SLR o funkcjach  $f$  i  $g$  przyjmujących dowolną postać. Sposób dyskretyzacji skierowanej liczby rozmytej przedstawiony jest na rys. 3.5.





Rysunek 3.5: Dyskretyzacja skierowanej liczby rozmytej.

### 3.3.3. Algebra Banacha

Niech  $\mathcal{OF}$  będzie przestrzenią wszystkich skierowanych liczb rozmytych. Mnożenie przez skalary oraz operacja dodawania, określone przez wzory (3.4), (3.3), prowadzą do struktury przestrzeni liniowej  $\mathcal{OF}$ . Ponadto zachodzi następujący fakt [37]:

**Lemat 3.1.** *Przestrzeń skierowanych liczb rozmytych  $\mathcal{OF}$  jest izomorficzna z liniową przestrzenią rzeczywistych, 2-wymiarowych wektorowych funkcji określonych na odcinku  $I = [0, 1]$ . Ponadto, jest to przestrzeń unormowana przez normę*

$$\|A\| = \max(\sup_{s \in I} |f_A(s)|, \sup_{s \in I} |g_A(s)|).$$

Stąd przestrzeń  $\mathcal{OF}$  może być utożsamiona z iloczynem kartezjańskim przestrzeni funkcji ciągłych określonych na przedziale  $[0, 1]$  (tj.  $\mathcal{C}^0([0, 1]) \times \mathcal{C}^0([0, 1])$ ). Dowód wynika bezpośrednio z definicji działań i zbioru  $\mathcal{OF}$ . Mnożenie elementów w przestrzeni  $\mathcal{OF}$ , które jest przemienne, w oczywisty sposób spełnia warunek rozdzielności względem dodawania i wprowadza w przestrzeni  $\mathcal{OF}$  strukturę pierścienia przemienne. W konsekwencji, przestrzeń  $\mathcal{OF}$  jest *algebrą Banacha* z jednością  $e$  jako parą funkcji ciągłych stale równych jeden.

Dzięki tak zdefiniowanej algebrze skierowanych liczb rozmytych możliwe staje się rozwiązywanie równań, które nie miały rozwiązań w klasycznych liczbach rozmytych. Ponadto możemy rozwiązywać zadania, w których część danych jest skierowanymi liczbami rozmytymi, a część liczbami rzeczywistymi.

### 3.4. Operatory wyostrzania

Operacje wyostrzania są podstawowymi operacjami występującymi w rozmytych systemach wnioskujących i rozmytych sterownikach. Jeśli konkluzja wynikająca z modelu rozmytego ma mieć praktyczne zastosowanie, nie może być ona zbiorem (liczbą) rozmytą, powinna być wartością liczbową. Można to uzyskać poprzez zastosowanie odpowiedniego *operatora wyostrzania* (ang. *defuzzification*), czyli odwzorowania, które zmiennej rozmytej przyporządkuje liczbę rzeczywistą. Dla klasycznych zbiorów (liczb) rozmytych zaproponowano dużą liczbę takich operatorów [11, 66].

Dla skierowanych liczb rozmytych operatory wyostrzania mogą być zdefiniowane za pomocą ciągłego, liniowego funkcjonału na przestrzeni  $\mathcal{OF}$  [37, 40]. Skoro przestrzeń  $\mathcal{OF}$  jest iloczynem kartezyjańskim znanej przestrzeni Banacha funkcji ciągłych na odcinku  $[0, 1]$ , korzystając z twierdzenia Banacha-Kakutaniego-Riesza, każdy liniowy, ciągły funkcjonał  $\phi$  jest jednoznacznie określony przez parę miar Radona  $(\nu_1, \nu_2)$  na przedziale  $[0, 1]$ , jako

$$\phi(A) = \phi(f_A, g_A) = \int_0^1 f_A(s)\nu_1(ds) + \int_0^1 g_A(s)\nu_2(ds). \quad (3.11)$$

Korzystając z tej formuły, można określić praktycznie nieskończenie wiele operatorów wyostrzania, m.in. odpowiedniki stosowanych operatorów dla klasycznych zbiorów rozmytych. W szczególności, znana metoda wyostrzania poprzez wyznaczenie pola powierzchni pod funkcją przynależności jest realizowana przez liniową kombinację pary miar Lebesgue'a na  $[0, 1]$ . Miarę Radona na  $[0, 1]$  można przedstawić w postaci całki Stieltjesa względem funkcji o wahanii ograniczonym, wówczas funkcjonał  $\phi$  przyjmuje postać

$$\phi(f_A, g_A) = \int_0^1 f_A(s)d\nu_1(s) + \int_0^1 g_A(s)d\nu_2(s), \quad (3.12)$$

gdzie  $\nu_1, \nu_2$  to funkcje o wahanii skończonym [37].

Należy zwrócić uwagę, że operatory wyostrzania nie sprowadzają się jedynie do metod liniowych, gdyż złożenie nieliniowej funkcji jednej czy wielu zmiennych rzeczywistych, z jednym czy wieloma liniowymi funkcjonałami na  $\mathcal{OF}$ , prowadzi do kolejnych (nieliniowych) operatorów wyostrzania. Co oznacza, że operatory wyostrzania dla skierowanych

liczb rozmytych można zdefiniować jako dowolną kombinację funkcyjałów postaci (3.12). Klasyczna metoda wyostrzania poprzez wyznaczanie *środku ciężkości* (ang. *center of gravity* (CoG)) realizowana jest w następujący sposób

$$CoG(A) = \frac{\phi_1(A)}{\phi_2(A)} = \frac{\int_0^1 \frac{1}{2}(f_A(s) + g_A(s))|f_A(s) - g_A(s)|ds}{\int_0^1 |f_A(s) - g_A(s)|ds}, \quad (3.13)$$

przy założeniu, że  $\int_0^1 |f_A(s) - g_A(s)|ds \neq 0$  [40].

### 3.5. Liczby niewłaściwe

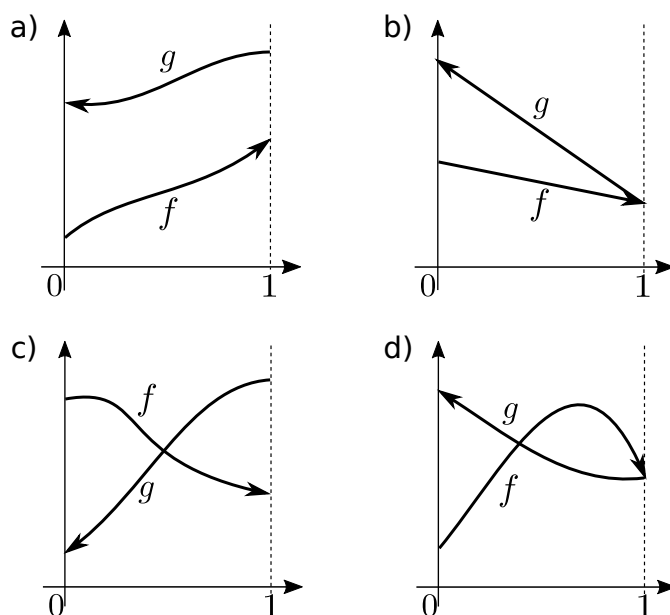
W zbiorze wszystkich skierowanych liczb rozmytych można wyróżnić trzy podzbiory:

- Właściwe skierowane liczby rozmyte, czyli takie, które za pomocą funkcji przynależności można przedstawić w postaci wypukłych liczb rozmytych.
- Liczby rzeczywiste, reprezentowane przez parę równych sobie funkcji stałych.
- Wszystkie pozostałe, nazywane *niewłaściwymi skierowanymi liczbami rozmytymi*.

Ze względu na ograniczenia co do funkcji  $f$  i  $g$ , pierwsze dwa zbiory stanowią stosunkowo niewielki podzbiór wszystkich skierowanych liczb rozmytych. Wobec czego, bardzo często wynikiem działań na liczbach skierowanych są liczby niewłaściwe (nawet, jeśli działania wykonywane są tylko na liczbach właściwych).

Na rys. 3.6 przedstawiono cztery przykłady liczb niewłaściwych. Mimo, że liczby te mogą nie mieć jasnej interpretacji, jak liczby w klasycznym modelu liczb rozmytych, to wciąż przechowują one pewną informację o rozmytej przynależności. Informacje te uzależnione są jednak od kontekstu, w którym dana liczba jest rozważana (zobacz interpretację SLR z rys. 7.10). Ich interpretacja opisywana jest za pomocą koncepcji *rozmytej obserwacji*, a niektóre z nich mogą być interpretowane w taki sposób, jak zbiory rozmyte 2 typu [37, 39, 67, 94].

W rozdziale 3.2 wprowadzono pojęcie orientacji SLR jako relacji między funkcją  $f$  i  $g$ . Dla właściwych SLR, wyszczególniono orientację pozytywną i negatywną. W przy-



Rysunek 3.6: Niewłaściwe skierowane liczby rozmyte.

padku liczb niewłaściwych określenie czy dana liczba posiada orientację pozytywną czy negatywną, nie zawsze jest oczywiste. Przyglądając się liczbie z rysunku 3.6b, podążając od funkcji  $f$  do  $g$ , można przypisać jej orientację negatywną, ze względu na malejący charakter funkcji  $f$ , jak również pozytywną, ze względu na rosnący charakter funkcji  $g$ .

W celu uniknięcia niejednoznaczności, do określenia orientacji skierowanej liczby rozmytej, w niniejszej pracy autor wprowadza następujący operator skierowania.

**Definicja 3.7.** Niech  $\mathcal{OF}$  będzie zbiorem wszystkich skierowanych liczb rozmytych. Operatorem skierowania nazywa się odwzorowanie  $Ord: \mathcal{OF} \rightarrow \{-1, 0, 1\}$ , takie że

$$Ord(A) = \begin{cases} 1 & \text{gdy } Defuzzy(A) > f_A(0), \\ 0 & \text{gdy } Defuzzy(A) = f_A(0) \text{ lub } Defuzzy(A) \text{ nie istnieje,} \\ -1 & \text{gdy } Defuzzy(A) < f_A(0), \end{cases} \quad (3.14)$$

gdzie  $Defuzzy(A)$  to ustalony operator wyostżenia liczby  $A = (f_A, g_A)$  np. operator środka ciężkości zdefiniowany za pomocą wzoru (3.13). Liczby  $-1, 1, 0$  oznaczają, odpowiednio, orientację negatywną, pozytywną oraz brak orientacji.

## Rozdział 4

# Reprezentacja danych za pomocą skierowanych liczb rozmytych

W pierwszej części tego rozdziału przedstawiona zostanie charakterystyka danych finansowych wysokiej częstotliwości i związane z nimi aspekty niepewności. Następnie omówiona zostanie zaproponowana przez autora koncepcja reprezentowania takich danych za pomocą skierowanych liczb rozmytych, zwanych dalej skierowanymi świecami rozmytymi. Przedstawiona zostanie ogólna idea, jak i metody konstrukcji przykładowych klas skierowanych świec rozmytych. Na końcu, przedstawione zostaną metody pozwalające na uwzględnienie dodatkowych informacji, takich jak np. volumen transakcji.

### 4.1. Finansowe szeregi czasowe wysokiej częstotliwości

Głównym przedmiotem badań przedstawionych w niniejszej rozprawie są szeregi czasowe instrumentów finansowych notowanych w systemie notowań ciągłych. Instrumentami takimi są na przykład akcje, instrumenty pochodne (opcje, kontrakty futures) czy też kursy par walutowych. Kurs instrumentu finansowego jest ceną kształtowaną na rynku na podstawie popytu i podaży generowanego przez różnego rodzaju inwestorów. W systemie notowań ciągłych, kursy notowań zmieniają się wielokrotnie w czasie trwania danej sesji giełdowej, a odpowiednie ceny generowane są wraz z napływem nowych zleceń od inwestorów [23].

Rynek finansowy to złożony, niestacjonarny, chaotyczny system dynamiczny, jednak nie jest to proces błędzenia losowego [50]. Tłumaczyć to zjawisko można m.in. tym, że bardzo duża liczba informacji znajduje odzwierciedlenie na rynku finansowym. Poza aspektami bezpośrednio związanymi z danym instrumentem finansowym (np. kondycja notowanej spółki), są to również czynniki takie jak stan gospodarki, sytuacja polityczna, oczekiwania inwestorów, katastrofy oraz wiele innych niemożliwych do przewidzenia zdarzeń. Teoria efektywności rynku mówi, że informacyjnie efektywny rynek kapitałowy charakteryzuje się tym, że ceny instrumentów finansowych występujących na nim w pełni i odpowiednio szybko odzwierciedlają wszystkie informacje ich dotyczące [5]. Formalnie, hipoteza efektywnego rynku została stworzona w 1970 roku przez Eugene E. Fama. W swojej pracy podzielił on teorię efektywnego rynku kapitałowego na trzy hipotezy: słabą, półsilną i silną [17].

- Hipoteza słaba zakłada, że obecne ceny instrumentów finansowych odzwierciedlają wszystkie historyczne informacje oraz dane cenowe.
- Hipoteza półsilna zakłada, że obecne ceny instrumentów finansowych odzwierciedlają wszystkie publicznie dostępne informacje, włączając w to dane historyczne, raporty finansowe, prognozy ekonomiczne, itp..
- Hipoteza silna zakłada, że obecne ceny instrumentów finansowych odzwierciedlają wszystkie dostępne informacje, zarówno publiczne, jak i niepubliczne.

W rzeczywistości, żadnego rynku kapitałowego nie można uznać za rynek w pełni efektywny, gdyż nie istniałyby wówczas strategie umożliwiające osiąganie ponadprzeciętnych zysków poprzez stosowanie analizy technicznej (hipoteza słaba), analizy fundamentalnej (hipoteza półsilna), informacji poufnych (hipoteza silna). Są jednak inwestorzy, którzy potrafią osiągać systematycznie ponadprzeciętne wyniki. Niemniej, istnieją liczne badania empiryczne w zakresie testowania efektywności rynków kapitałowych, które pozwalają na określenie, które hipotezy rynku efektywnego i na jakim poziomie istotności można przyjąć za prawdziwe dla danego rynku w badanym okresie. Nigdy jednak nie ma pewności, czy cena danego instrumentu finansowego, ustalona przez popyt i podaż, uwzględnia wszystkie

istotne informacje, wobec czego **naturalnym wydaje się być rozpatrywanie danych finansowych w kontekście danych obarczonych niepewnością.**

Dane zawierające najpełniejszą wiedzę o notowaniach danego instrumentu finansowego to ceny odpowiadające każdej pojedynczej transakcji zawartej na tym instrumencie. Są to równocześnie dane o najwyższej możliwej częstotliwości, które nazywane są *danymi ultra wysokiej częstotliwości* (ang. *ultra-high-frequency data*) lub kolokwialnie danymi *tick-by-tick*. W badaniach ekonometrycznych, za dane wysokiej częstotliwości uznaje się dane o częstotliwości wyższej niż dzienna, czyli tzw. dane *intraday*. Rozwój rynków finansowych, z równoległym rozwojem technologii komputerowych w zakresie zapisu, przesyłu i przechowywania danych spowodował, że obecnie transakcje na instrumentach finansowych mogą być dokonywane z częstością poniżej ułamków sekundy. Liczba dokonywanych transakcji zależy w dużym stopniu od płynności danego instrumentu finansowego. Średnia dzienna liczba notowań dla pary walutowej USD/EUR z łatwością przekracza 20 tys., a na rynkach wysoko rozwiniętych, o dużej kapitalizacji, jak na przykład giełda nowojorska (NYSE), może być ona dużo większa.

Analiza danych typu tick-by-tick jest bardzo utrudniona, m.in. ze względu na bardzo dużą liczbę obserwacji, nieregularne odstępy między obserwacjami, występowanie wzorców cenowych oraz długotrwałych zależności. Z różnych powodów mogą one również zawierać błędne obserwacje, brakujące dane czy też błędną kolejność danych [16, 98]. Wobec czego, ciągle najpopularniejszymi danymi używanymi przy konstrukcji modeli ekonometrycznych są dane o częstotliwości dziennej, najczęściej są to ceny zamknięcia z danego dnia.

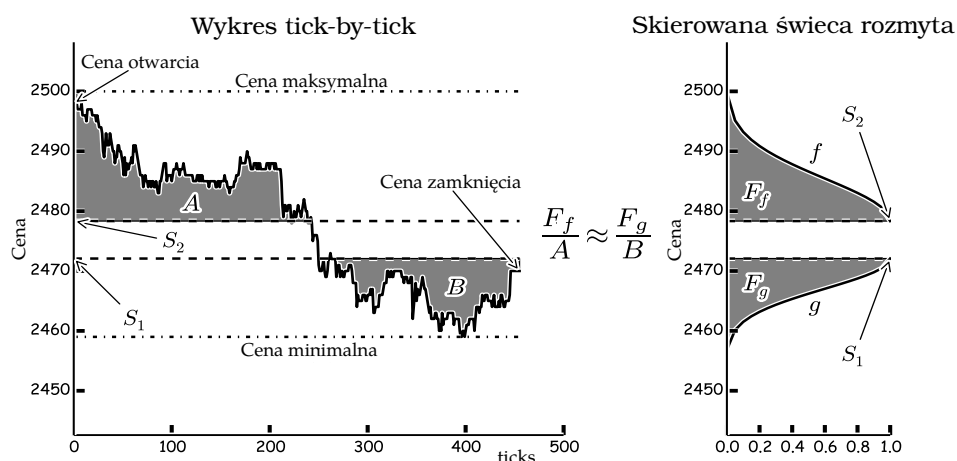
Również inwestorzy podejmujący decyzję na podstawie wykresów cenowych nie wykorzystują danych tick-by-tick, gdyż śledzenie takich wykresów jest po prostu niemożliwe. Korzystają oni z wykresów o niższej częstotliwości, konstruowanych najczęściej za pomocą *słupków OHLC* (ang. *Open-High-Low-Close bar*) lub *świec japońskich* (ang. *Japanese Candlestick*). Wykres świecowy jest najczęściej stosowanym wykresem przy analizie technicznej rynków finansowych. Na wykresie tym każda pojedyncza świeca jest prostą graficzną interpretacją tego co się dzieje z ceną w jednym ustalonym interwale czasowym.

Ilustruje ona poziomy czterech cen: otwarcia, zamknięcia, maksymalną i minimalną oraz dodatkowo pokazuje trend panujący na rynku [61, 63].

## 4.2. Skierowane świece rozmyte

W proponowanym przez autora podejściu do reprezentacji danych tick-by-tick, każdemu ustalonymu przedziałowi czasowemu odpowiada odpowiednio zdefiniowana skierowana liczba rozmyta, zwana dalej *skierowaną świecą rozmytą* - *SSR* (ang. *ordered fuzzy candlestick* - *OFC*). Koncepcja ta została zdefiniowana, a następnie rozszerzana w publikacjach autora [58, 59, 60].

Główna idea skierowanej świecy rozmytej przedstawiona jest na rys. 4.1. Polega ona na tym, że informacje o danym procesie notowań zawarte są w kształtach funkcji  $f$  i  $g$  skierowanej świecy rozmytej. W tym sensie, że kształty te nie zależą bezpośrednio od wartości notowań, ale od relacji pomiędzy odpowiednimi polami powierzchni (por. rys. 4.1). Ponadto, orientacja skierowanej świecy rozmytej powinna być zgodna z trendem panującym na rynku w rozważanym okresie. W dotychczasowych rozważaniach wyszczególniono dwa przypadki konstrukcji skierowanej świecy rozmytej.



Rysunek 4.1: Szkie koncepcji skierowanej świecy rozmytej.



### 4.2.1. Skierowane świece rozmyte z funkcjami ustalonego typu

Pierwszy z przypadków konstrukcji skierowanych świec rozmytych zakłada, że funkcje  $f$  i  $g$  są funkcjami z góry ustalonego typu, przy czym kształty tych funkcji mogą zależeć od dwóch parametrów. Wówczas skierowaną świecę rozmytą dla danego szeregu czasowego można zdefiniować w poniższy sposób.

**Definicja 4.1.** Niech dany będzie szereg czasowy  $\{X_t: t \in T\}$  dla  $T = \{1, 2, \dots, n\}$ . Skierowaną świecę rozmytą nazywamy skierowaną liczbę rozmytą  $C = (f, g)$  wtedy, gdy spełnione są warunki 1-4 (dla świecy wzrostowej) lub 1'-4' (dla świecy spadkowej).

1.  $X_1 \leq X_n$ ,
2. funkcja  $f: [0, 1] \rightarrow \mathbb{R}$  jest ciągła i rosnąca na  $[0, 1]$ ,
3. funkcja  $g: [0, 1] \rightarrow \mathbb{R}$  jest ciągła i malejąca na  $[0, 1]$ ,
4.  $S_1 < S_2$ ,  $f(0) = \min_{t \in T} X_t - C_1$ ,  $f(1) = S_1$ ,  $g(1) = S_2$  oraz wartość  $g(0)$  jest tak dobrana aby stosunki  $\frac{F_g}{A}$  i  $\frac{F_f}{B}$  były sobie równe,

- 1'.  $X_1 > X_n$ ,
- 2'. funkcja  $f: [0, 1] \rightarrow \mathbb{R}$  jest ciągła i malejąca na  $[0, 1]$ ,
- 3'. funkcja  $g: [0, 1] \rightarrow \mathbb{R}$  jest ciągła i rosnąca na  $[0, 1]$ ,
- 4'.  $S_1 < S_2$ ,  $f(0) = \max_{t \in T} X_t + C_2$ ,  $f(1) = S_2$ ,  $g(1) = S_1$  oraz wartość  $g(0)$  jest tak dobrana aby stosunki  $\frac{F_f}{A}$  i  $\frac{F_g}{B}$  były sobie równe,

gdzie  $S_1, S_2, C_1, C_2, A, B$  są liczbami rzeczywistymi wyznaczonymi na podstawie wartości szeregu czasowego  $X_t$ . Liczby  $F_f$  oraz  $F_g$  są odpowiednio polami po wykresami funkcji  $f^{-1}$  oraz  $g^{-1}$  (por. rys. 4.1).

Kształt skierowanej świecy rozmytej zdefiniowanej według powyższej definicji zależy nie tylko od przyjętego typu dla funkcji  $f$  i  $g$  ale w dużym stopniu również od sposobu wyznaczenia parametrów  $S_1, S_2, C_1, C_2, A$  oraz  $B$ .

**Parametry  $S_1, S_2$ :**

Parametry  $S_1, S_2$  takie, że  $S_1 < S_2$  wyznaczają przedział  $[1_A^-, 1_A^+]$  (centrum, jądro) skierowanej liczby rozmytej i powinny należeć do przedziału  $[\min_{t \in T} X_t, \max_{t \in T} X_t]$ . Parametry te mogą być wyznaczone m.in. za pomocą różnego rodzaju średnich z wartości szeregu czasowego  $X_t$ . Najczęściej przyjmowanymi przez autora wartościami dla  $S_1, S_2$  są odpowiednio minimum i maksimum ze średnich: arytmetycznej, ważonej liniowo oraz ważonej eksponencjalnie.

**Parametry  $C_1, C_2$ :**

Parametry  $C_1, C_2 \geq 0$  pozwalają dodatkowo rozszerzyć rozmycie (nośnik) skierowanej liczby rozmytej poza wartości minimalne i maksymalne szeregu czasowego. Parametry te mogą być wyznaczone m.in. za pomocą odchylenia standardowego lub zmienności szeregu czasowego  $X_t$ .

**Parametry  $A, B$ :**

Parametry  $A, B \geq 0$  odpowiadają za relację między funkcjami  $f$  i  $g$ . Parametry te mogą być wyznaczone za pomocą pól powierzchni obszarów między wartościami szeregu czasowego  $X_t$ , a parametrami  $S_1$  oraz  $S_2$  (por. rys. 4.1).

**4.2.2. Empiryczne skierowane świece rozmyte**

W drugim zaproponowanym przez autora przypadku konstrukcji skierowanych świec rozmytych brak jest założeń o kształtach funkcji  $f$  oraz  $g$ . Dla danego szeregu czasowego funkcje te wyznaczone są z wykorzystaniem metody, którą można by porównać do konstrukcji dystrybuant empirycznych.

**Definicja 4.2.** Niech dany będzie szereg czasowy  $\{X_t: t \in T\}$  dla  $T = \{1, 2, \dots, n\}$ . *Empiryczną skierowaną świecą rozmytą* nazywamy skierowaną liczbę rozmytą  $C = (f, g)$  wtedy, gdy funkcje  $f$  i  $g$  są dowolnymi ciągłymi aproksymacjami lub interpolacjami<sup>1</sup> dyskretnych funkcji odpowiednio  $\Psi_1$  i  $\Psi_2$  (dla świecy wzrostowej) lub  $\Psi_2$  i  $\Psi_1$  (dla świecy

<sup>1</sup>W swoich badaniach autor wykorzystuje interpolacje funkcjami liniowymi.

spadkowej). Ponadto, funkcje  $\Psi_1, \Psi_2$  są funkcjami dyskretnymi określonymi na przedziale  $[0, 1]$  z przyjętym krokiem  $dx = \frac{1}{M}$  ( $M + 1$  równo rozłożonych punktów) w następujący sposób

$$\Psi_1(k \cdot dx) = \begin{cases} Y_{\lfloor \frac{k}{dx} \rfloor}^{(1)} - \frac{M-k}{M} C_1 & \text{dla } k \in \{0, 1, \dots, M-1\}, \\ S_1 & \text{dla } k = M, \end{cases} \quad (4.1)$$

$$\Psi_2(k \cdot dx) = \begin{cases} Y_{K_2 - \lfloor \frac{k}{dx} \rfloor}^{(2)} + \frac{M-k}{M} C_2 & \text{dla } k \in \{0, 1, \dots, M-1\}, \\ S_2 & \text{dla } k = M, \end{cases} \quad (4.2)$$

gdzie wartości szeregu czasowego  $Y_t$  to posortowane rosnąco wartości szeregu  $X_t$ , a szeregi  $Y_t^{(1)}$  i  $Y_t^{(2)}$  są szeregami czasowymi powstałymi przez podział szeregu czasowego  $Y_t$  w poniższy sposób

$$Y_t^{(1)} = \{Y_i : i \in T \wedge Y_i \leq S_1\} \quad t \in \{0, 1, \dots, K_1\},$$

$$Y_t^{(2)} = \{Y_i : i \in T \wedge S_2 \leq Y_i\} \quad t \in \{0, 1, \dots, K_2\},$$

gdzie  $K_1$  i  $K_2$  to odpowiednio liczby obserwacji szeregów czasowych  $Y_t^{(1)}$  i  $Y_t^{(2)}$ . Natomiast,  $S_1, S_2, C_1, C_2$  są liczbami rzeczywistymi wyznaczonymi na podstawie wartości szeregu czasowego  $X_t$ , a ich znaczenie jest takie same jak w poprzednim podrozdziale.

### 4.3. Przykłady skierowanych świec rozmytych

W niniejszym podrozdziale przedstawione zostaną empiryczne przykłady konstrukcji skierowanych świec rozmytych.

#### 4.3.1. Trapezoidalna skierowana świeca rozmyta

Zakładając w definicji 4.1 postać liniową funkcji  $f$  oraz  $g$ , tj.

$$f(x) = (b_f - a_f)x + a_f \quad \text{oraz} \quad g(x) = (b_g - a_g)x + a_g \quad \text{dla } x \in [0, 1], \quad (4.3)$$

gdzie parametry  $a_f, a_g$  oraz  $b_f, b_g$  to wartości funkcji  $f$  i  $g$  odpowiednio dla  $x = 0$  oraz  $x = 1$ . Wówczas, skierowana świeca rozmyta  $C = (f, g)$  nazywana jest *trapezoidalną SŚR* (ang. *trapezoid OFC*), w szczególności, gdy  $b_f = b_g$  to świeca ta nazywana jest *trójkąt-  
ną SŚR* (ang. *triangular OFC*).

Zgodnie z definicją 4.1 dla parametrów funkcji liniowych  $f$  i  $g$  dla danego szeregu czasowego  $\{X_t: t \in T\}$  dla  $T = \{1, 2, \dots, n\}$  otrzymuje się następujące wzory.

Dla świecy wzrostowej ( $X_1 \leq X_n$ ):

$$a_f = \min_{t \in T} X_t - C_1 \quad \wedge \quad b_f = S_1 \quad (4.4)$$

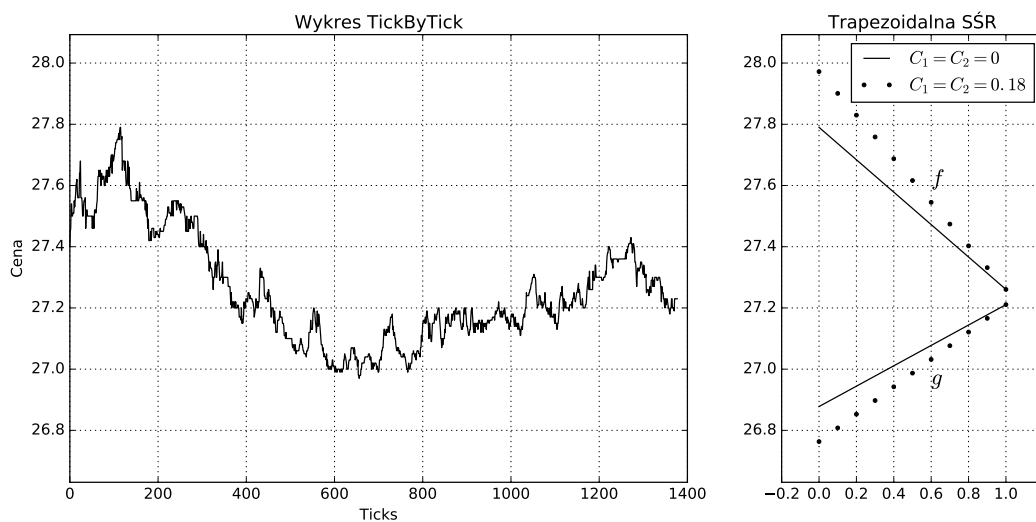
$$a_g = \frac{A}{B}(S_1 - \min_{t \in T} X_t + C_1) + S_2 \quad \wedge \quad b_g = S_2 \quad (4.5)$$

Dla świecy spadkowej ( $X_1 > X_n$ ):

$$a_f = \max_{t \in T} X_t + C_2 \quad \wedge \quad b_f = S_2 \quad (4.6)$$

$$a_g = \frac{B}{A}(S_2 - \max_{t \in T} X_t - C_2) + S_1 \quad \wedge \quad b_g = S_1 \quad (4.7)$$

Rys. 4.2 przedstawia wykres notowań spółki PKOBP z dnia 14 marca 2016 roku (1378 notowań) oraz realizację dwóch trapezoidalnych skierowanych świec rozmytych. Dla świec tych obliczono parametry  $S_1 = 27.21$ ,  $S_2 = 27.26$  (minimum i maksimum z trzech różnych średnich) oraz parametry  $A = 102.35$ ,  $B = 64.35$  (odpowiednie pola powierzchni). Świece różnią się natomiast parametrami  $C_1$  i  $C_2$ . W pierwszej przyjęto dla obu parametrów wartość zero, a w drugiej wartość odchylenia standardowego wynoszącego 0.18.



Rysunek 4.2: Przykłady trapezoidalnej skierowanej świecy rozmytej.

### 4.3.2. Gaussowska skierowana świeca rozmyta

Skierowana świeca rozmyta  $C = (f, g)$  nazywana jest *gaussowską SŚR* (ang. *gaussian OFC*) wówczas gdy funkcje  $f$  oraz  $g$  przyjmują poniższą postać.

$$f(x) = f(z) = \sigma_f \sqrt{-2 \ln(z)} + m_f \quad \text{oraz} \quad g(x) = g(z) = \sigma_g \sqrt{-2 \ln(z)} + m_g, \quad (4.8)$$

gdzie  $z = (1 - \alpha)x + \alpha$  dla  $x \in [0, 1]$ . Parametry  $m_f, m_g$  to wartości funkcji  $f$  oraz  $g$  dla  $x = 1$ , natomiast parametry  $\sigma_f, \sigma_g$  odpowiadają szerokość liczby rozmytej. Współczynnik  $\alpha > 0$  bliski zeru umożliwia określenie wartości dla  $x = 0$ .

Zgodnie z definicją 4.1 dla parametrów funkcji  $f$  i  $g$  dla danego szeregu czasowego  $\{X_t: t \in T\}$  dla  $T = \{1, 2, \dots, n\}$  otrzymuje się następujące wzory.

Dla świecy wzrostowej ( $X_1 \leq X_n$ ):

$$m_f = S_1 \quad \wedge \quad \sigma_f = \frac{\min_{t \in T} X_t - C_1 - S_1}{\sqrt{-2 \ln(\alpha)}} \quad (4.9)$$

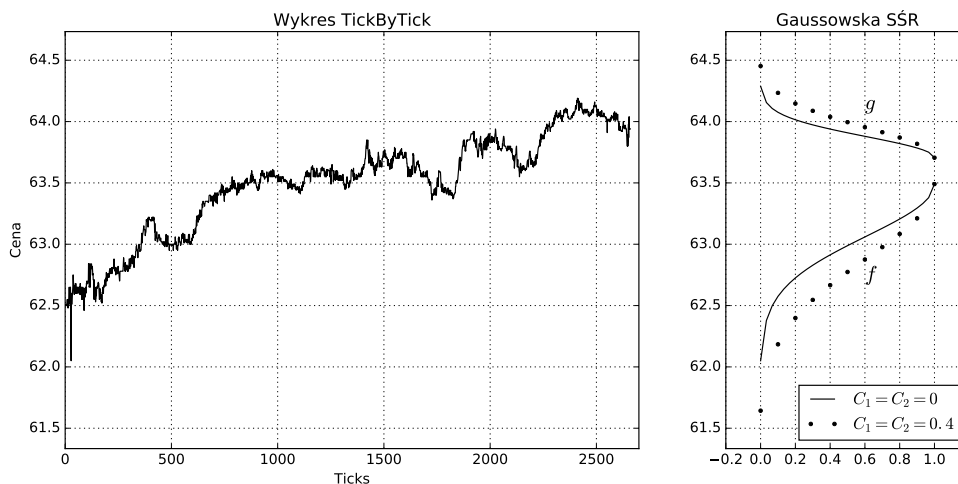
$$m_g = S_2 \quad \wedge \quad \sigma_g = -\frac{A}{B} \sigma_f \quad (4.10)$$

Dla świecy spadkowej ( $X_1 > X_n$ ):

$$m_f = S_2 \quad \wedge \quad \sigma_f = \frac{\max_{t \in T} X_t + C_2 - S_2}{\sqrt{-2 \ln(\alpha)}} \quad (4.11)$$

$$m_g = S_1 \quad \wedge \quad \sigma_g = -\frac{B}{A} \sigma_f \quad (4.12)$$

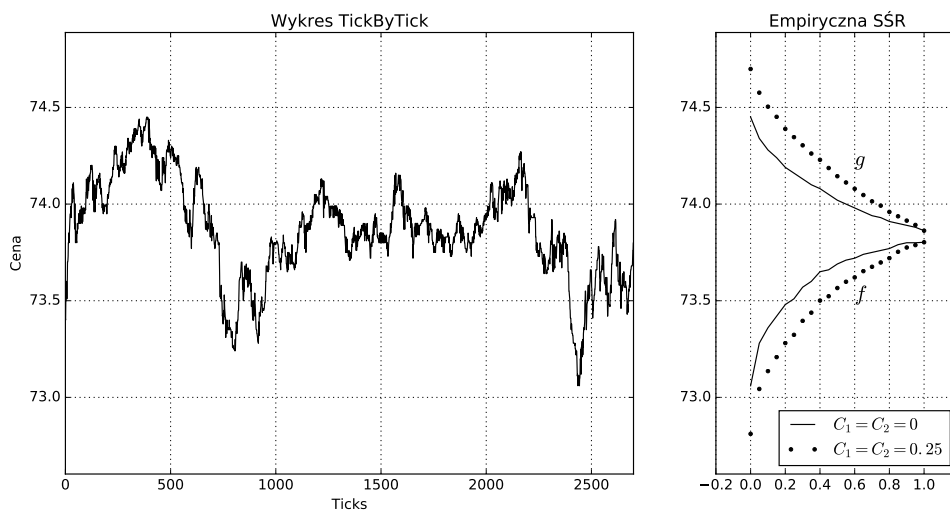
Na rys. 4.3 przedstawiono realizację dwóch gaussowskich skierowanych świec rozmytych dla notowań spółki KGHM z dnia 28 czerwca 2016 roku (2661 notowań). Dla świec tych obliczono parametry  $S_1 = 63.49$ ,  $S_2 = 63.71$  (minimum i maksimum z trzech różnych średnich) oraz parametry  $A = 168.73$ ,  $B = 417.20$  (odpowiednie pola powierzchni). Świece różnią się natomiast parametrami  $C_1$  i  $C_2$ . W pierwszej przyjęto dla obu parametrów wartość zero, a w drugiej wartość odchylenia standardowego wynoszącego 0.40.



Rysunek 4.3: Przykłady gaussowskiej skierowanej świecy rozmytej.

### 4.3.3. Empiryczna skierowana świeca rozmyta

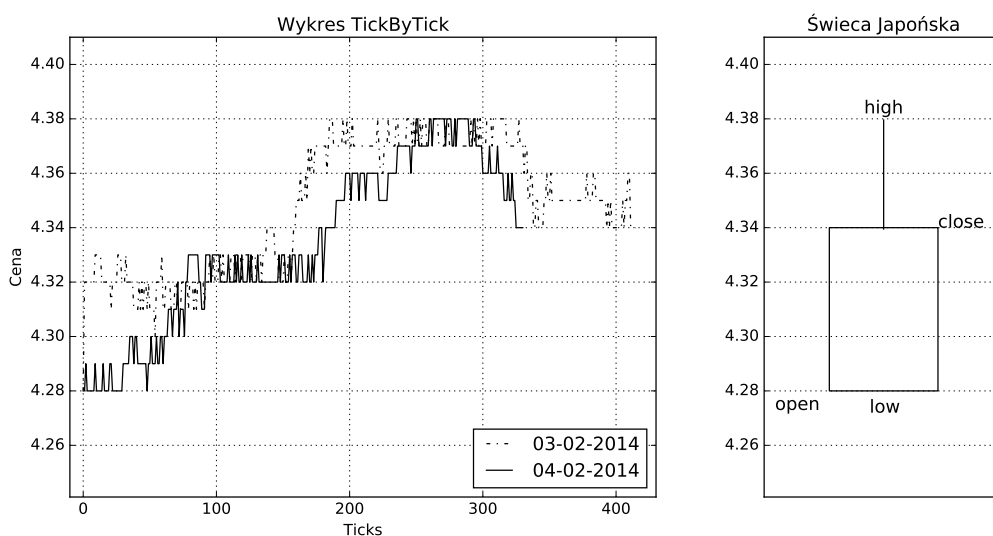
W kolejnym przykładzie skonstruowano dwie empiryczne świece rozmyte zgodnie z procedurą przedstawioną w definicji 4.2 dla szeregu czasowego notowań spółki KGHM z dnia 14 marca 2016 roku (2902 notowań). Dla świec tych przyjęto parametry  $M = 20$ ,  $S_1 = 73.80$ ,  $S_2 = 73.86$  (minimum i maksimum z trzech różnych średnich) oraz parametry  $C_1 = C_2 = 0$  w pierwszym przypadku, natomiast  $C_1 = C_2 = 0.25$  (odchylenie standardowe) w drugim przypadku. Realizację skonstruowanych świec przedstawia rys. 4.4.



Rysunek 4.4: Przykłady empirycznej skierowanej świecy rozmytej.

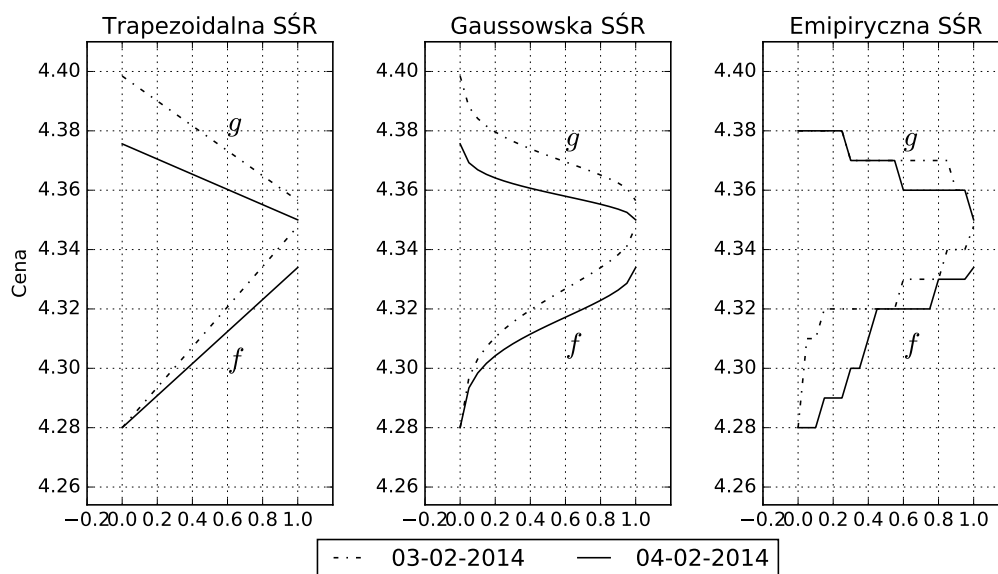
#### 4.3.4. Szeregi czasowe o identycznych świecach japońskich

W niniejszym przykładzie rozważono dwa szeregi czasowe notowań spółki TAURON-PE. Pierwszy z nich przedstawia notowania z dnia 3 lutego 2014 roku (416 notowań), drugi natomiast z dnia 4 lutego 2014 roku (331 notowań). Oba te szeregi charakteryzują się identycznymi cenami otwarcia, cenami maksymalnymi, cenami minimalnymi oraz cenami zamknięcia. Stąd skonstruowane na ich podstawie świece japońskie lub słupki OHLC są identyczne, mimo że przebiegi notowań są istotnie różne (por. rys. 4.5).



Rysunek 4.5: Dwa szeregi czasowe o identycznych świecach japońskich.

Na rys. 4.6 przedstawiono realizacje trzech rodzajów skierowanych świec rozmytych dla rozważanych szeregów czasowych. Dla notowań z dnia 3 lutego obliczono parametry  $S_1 = 4.348$ ,  $S_2 = 4.357$  (minimum i maksimum z trzech różnych średnich), parametry  $A = 2.67$ ,  $B = 4.33$  (odpowiednie pola powierzchni) oraz  $C_1 = C_2 = 0$ . Natomiast, dla notowań z dnia 4 lutego parametry te wyniosły  $S_1 = 4.334$ ,  $S_2 = 4.35$ ,  $A = 2.05$ ,  $B = 4.32$  oraz  $C_1 = C_2 = 0$ . Ponadto, dla świecy empirycznej przyjęto  $M = 20$ . W każdym z przypadków skonstruowane skierowane świece rozmyte są istotnie różne, co pozwala sądzić, że skierowane świece rozmyte zawierają efektywnie więcej informacji o danym szeregu czasowym niż świece japońskie czy słupki OHLC.



Rysunek 4.6: Realizacje SRS dla szeregów czasowych o identycznych świecach japońskich.

## 4.4. Dodatkowe informacje o notowaniach

Notowania, czyli ceny po których dokonano transakcji to nie jedyne wartości opisujące zachowanie danego instrumentu finansowego. W niniejszym podrozdziale zaprezentowano, jak w prosty sposób przy konstrukcji skierowanych świec rozmytych uwzględnić takie informacje jak wolumen czy spread.

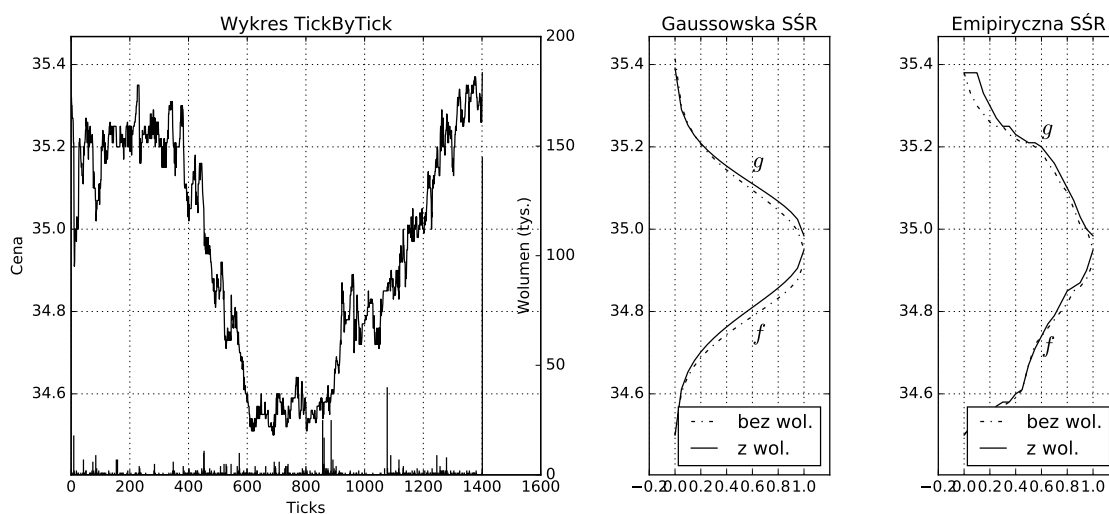
### 4.4.1. Wolumen transakcji

Wolumen to liczba instrumentów finansowych, które były w danym okresie przedmiotem transakcji i obok ceny stanowi podstawową informację o danej transakcji. W analizie technicznej wolumen wykorzystywany jest m.in. do potwierdzania siły trendu i formacji cenowych. Zmiana ceny, czy to w dół czy w górę z relatywnie większym wolumenem jest postrzegana silniej niż podobna zmiana z mniejszym wolumenem. Przyjmuje się także, że rosnący wolumen potwierdza rozwijający się na rynku trend. W trendzie wzrostowym wolumen powinien rosnać w miarę wzrostu cen i spadać podczas spadkowych korekt, w trendzie spadkowym powinno być na odwrót, wolumen rośnie, gdy ceny spadają, a podczas wzrostowych korekt wolumen spada [61].



W przypadku konstrukcji skierowanych świec rozmytych, wolumen transakcji może być uwzględniony w bardzo prosty sposób. Wystarczy, że parametry SŚR, takie jak  $S_1, S_2, A, B, C_1, C_2$  oraz funkcje  $\Psi_1, \Psi_2$  w przypadku świecy empirycznej, zostaną obliczone z wykorzystaniem funkcji gęstości stowarzyszonej z wolumenem. W praktyce, wystarczyłoby każdą cenę powtórzyć zgodnie z odpowiadającym jej wolumenem. Niemniej wartość wolumenu może sięgać milionów jednostek, co znacznie zwiększyłoby rozważany szereg czasowy. Dlatego, przed powieleniem cen, wartości wolumenu są liniowo skalowane do przedziału liczb całkowitych od 1 do 100.

Na rys. 4.7 przedstawiono notowania spółki PZU z dnia 1 kwietnia 2016 roku (1402 notowań) wraz z wolumenem transakcji oraz odpowiadające im realizacje gaussowskiej i empirycznej SŚR. Dla każdego rodzaju świecy przedstawiona została świeca z uwzględnieniem wolumenu jaki i bez niego.



Rysunek 4.7: Realizacje SŚR bez oraz z informacją o wolumenie.

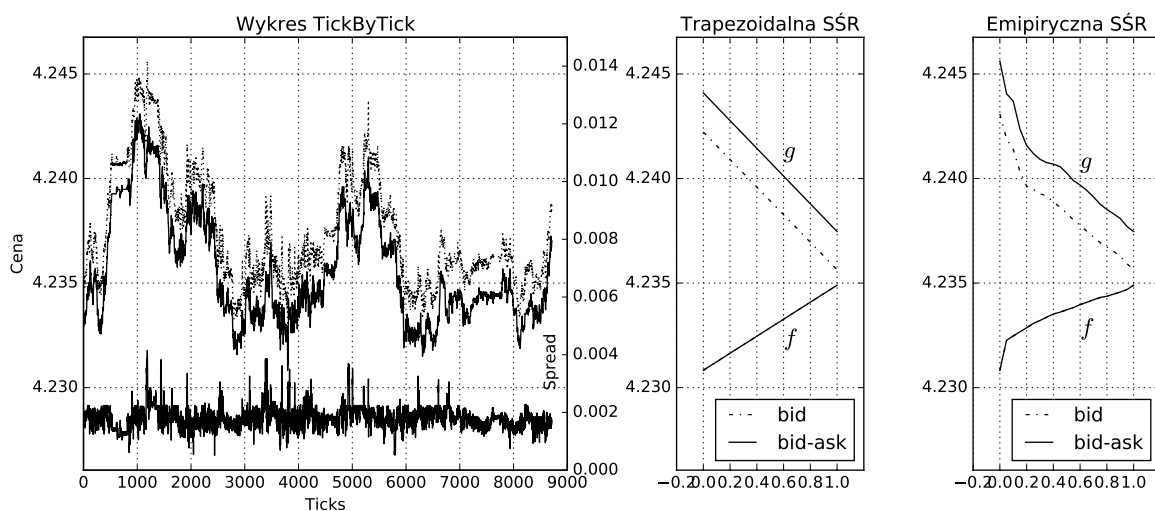
#### 4.4.2. Spread transakcyjny

Spread (ang. *bid-ask spread*) to różnica między najwyższą ceną jaką kupujący są w stanie zapłacić (cena bid), a najniższą ceną jaką sprzedawcy są w stanie zaakceptować (cena ask) za dany instrument finansowy. Na rynku pozagiełdowym, takim jak np. rynek walutowy, spready ustalane są wprost przez brokera, który odgrywa rolę animatora rynku. Oznacza to, że inwestor może w każdym momencie dokonać transakcji kupna lub sprze-

daży bez względu na to czy istnieje druga strona transakcji. Ryzyko znalezienia drugiej strony transakcji przejmuje na siebie animator, za co otrzymuje premię, właśnie w postaci spreadu. Z drugiej strony spread stanowi koszt dla inwestora. Wielkość spreadu bardzo często ulega zmianie w czasie. Związane jest to silnie z płynnością danego aktywu. Im większa płynność rynku tym spready niższe i odwrotnie. Na płynność wpływ mają m.in. warunki, w jakich działa rynek, jak i nadchodzące publikacje danych ekonomicznych.

Konstruując skierowane świece rozmyte, różnicę między cenami kupna i sprzedaży można uwzględnić obliczając parametry świec  $S_1, C_1, B$  oraz funkcję  $\Psi_1$  korzystając z szeregu czasowego cen kupna, natomiast parametry  $S_2, C_2, A$  oraz funkcję  $\Psi_2$  korzystając z szeregu czasowego cen sprzedaży.

Rys. 4.8 przedstawia wykresy cen kupna i sprzedaży notowań pary walutowej EUR/PLN z dnia 1 kwietnia 2016 roku od godziny 10:00 do godziny 14:00 UTC (4 godziny - 8713 notowań) oraz odpowiadające im realizacje trapezoidalnej i empirycznej SŚR. Dla każdego rodzaju świecy przedstawiona jest świeca z uwzględnieniem spreadu jak i bez niego (tylko ceny kupna).



Rysunek 4.8: Realizacje SŚR bez oraz z informacją o spreadzie.

# Rozdział 5

## Modele rozmytych szeregów czasowych

W niniejszym rozdziale przytoczone zostanie pojęcie rozmytego szeregu czasowego oraz główne podejścia do modelowania takich szeregów. W dalszej części omówione zostanie nowe podejście konstruowania modeli rozmytych szeregów czasowych z wykorzystaniem skierowanych liczb rozmytych. W szczególności omówione zostaną aspekty estymacji parametrów oraz predykcji w tak skonstruowanym rozmytym modelu regresji liniowej oraz modelu autoregresyjnym.

### 5.1. Rozmyte szeregi czasowe

W ogólnym rozumieniu szereg czasowy to ciąg informacji uporządkowanych w czasie, których pomiary wykonywane są w określonych odstępach. A bardziej formalnie, szereg czasowy  $\{x_1, x_2, \dots, x_n\}$  rozumiany jest jako wartości w chwilach  $t_1, t_2, \dots, t_n$  zmiennych losowych  $\xi_1, \xi_2, \dots, \xi_n$  należących do rodziny zmiennych losowych  $\{\xi_t : t \in T\}$ , którą nazywamy procesem stochastycznym. Tak, więc szereg czasowy to „ucięta realizacja” pewnego procesu stochastycznego (losowego) [93].

Wartości szeregu czasowego, pochodzące z pomiarów, w sposób naturalny mogą być obarczone różnymi błędami, jak również pewnymi aspektami niepewności charakterystycz-

nymi dla badanego zjawiska. Jednym ze sposobów modelowania niepewności jest użycie teorii zbiorów rozmytych. W niniejszej rozprawie jako *rozmyty szereg czasowy* (ang. *fuzzy time series*) rozumiemy ciąg uporządkowanych w czasie obserwacji przedstawionych w postaci zbiorów lub liczb rozmytych. W literaturze można wyróżnić trzy podejścia do konstrukcji modeli rozmytych szeregów czasowych. Podejścia te zostaną pokrótce przybliżone w kolejnych podrozdziałach.

### 5.1.1. Podejście oparte na rozmytym wnioskowaniu

W bardzo wielu pracach za prekursorów pojęcia rozmytych szeregów czasowych uważa się Songa i Chissoma, którzy w swoich pracach [75, 76, 77] zdefiniowali rozmyty szereg czasowy w następujący sposób.

**Definicja 5.1.** Niech  $Y(t)$  ( $t = \dots, 0, 1, 2, \dots$ ), podzbiór liczb rzeczywistych, będzie *przestrzenią rozważań* (ang. *universe of discourse*), na której zdefiniowane są zbiory rozmyte  $f_i(t)$  ( $i = 1, 2, \dots$ ). Niech  $F(t)$  będzie rodziną zbiorów rozmytych  $f_i(t)$  ( $i = 1, 2, \dots$ ). Wówczas  $F(t)$  nazywana jest *rozmytym szeregiem czasowym* nad  $Y(t)$  ( $t = \dots, 0, 1, 2, \dots$ ).

Przykładowo  $F(t)$  może być rozumiany jako pewna zmienna lingwistyczna (np. dzienne warunki pogodowe), która może przyjmować  $f_i(t)$  ( $i = 1, 2, \dots$ ) możliwych wartości lingwistycznych (np. zimno, umiarkowanie, gorąco itp.) zdefiniowanych za pomocą zbiorów rozmytych.

Song i Chissom zaproponowali również podejście do modelowania rozmytych szeregów czasowych oparte na pojęciu relacji rozmytej w następujący sposób.

**Definicja 5.2.** Załóżmy, że  $F(t)$  zależy tylko od  $F(t-1)$ , to znaczy, że  $F(t-1) \rightarrow F(t)$ . Wówczas relacja ta może być wyrażona jako rozmyte równanie relacyjne  $F(t) = F(t-1) \circ R(t, t-1)$ , gdzie  $R(t, t-1)$  jest rozmytą zależnością między  $F(t-1)$  i  $F(t)$ . Relacja ta nazywana jest *modelem pierwszego rzędu*  $F(t)$ .

W podobny sposób zdefiniowane zostały również modele wyższych rzędów. Przez następne lata wielu badaczy bardzo rozwinęło powyższą koncepcję rozmytych szeregów cza-

sowych, na szczególną uwagę zasługują prace Chena [6, 7, 8], który zaproponował prostą metodę obliczeniową pozwalającą uzyskać wyższą dokładność prognoz w modelach rozmytych szeregów czasowych. Dużo więcej na temat tego podejścia można przeczytać w publikacja przeglądowych [78, 79]. Do podejścia tego można również zaliczyć metody modelowania regresji nieliniowej poprzez zastosowanie regułowych neuronowo-rozmytych systemów wnioskujących, takich jak system ANFIS (Adaptive Neuro-Fuzzy Inference System) [1, 25].

### 5.1.2. Podejście oparte na rozmytej regresji

Podejście to oparte jest na modelu rozmytej *regresji liniowej* (ang. *fuzzy linear regression*), który zaproponowali Tanaka, Uejima i Asai [81]. Zdefiniowali oni model rozmyty jako funkcję liniową, której parametrami są zbiory rozmyte. W przeciwieństwie do konwencjonalnej funkcji regresji liniowej, gdzie różnice między obserwacjami i przewidywanymi wartościami traktowane są jako błędy obserwacji, różnice traktowane są jako rozmytość (niepewność) w systemie. Funkcje przynależności zbiorów rozmytych często określane są mianem *rozkładów możliwości* (ang. *possibility distributions*), dlatego też podejście to występuje w literaturze również pod nazwą *possibilistic regression analysis* [82, 83].

**Definicja 5.3.** *Modelem rozmytej regresji liniowej* nazywamy model postaci

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 \cdot x_1 + \cdots + \tilde{A}_n \cdot x_n = \tilde{A}_0 + \sum_{i=1}^n \tilde{A}_i \cdot x_i, \quad (5.1)$$

gdzie  $x_i \in \mathbb{R}$  to zmienne objaśniające, liczby ostre,  $\tilde{A}_i = (a_i, c_i)$  to współczynniki modelu, symetryczne trójkątne liczby rozmyte o funkcji przynależności  $\mu_{\tilde{A}_i}(x) = \max\{1 - \frac{|x-a_i|}{c_i}, 0\}$ ,  $a_i, c_i \in \mathbb{R}$  parametry liczby rozmytej,  $\tilde{Y}$  to zmienna objaśniana, liczba rozmyta. Składnik losowy modelu został zastąpiony poprzez rozmytość współczynników.

W celu estymacji parametrów modelu, minimalizowane jest rozmycie modelu, które odbywa się poprzez minimalizację całkowitego rozmycia współczynników modelu, przy warunku aby każdy objaśniany punkt danych zawierał się w estymowanym zbiorze roz-

mytym. Problem ten sprowadzony został do następującego problemu programowania liniowego.

Niech  $y_t \in \mathbb{R}$  będą obserwacjami pewnego badanego zjawiska, zależnymi od zmiennych  $\mathbf{x}_t \in \mathbb{R}^n$  dla  $t = 1, \dots, T$ . Wówczas, problem dopasowania modelu regresji rozmytej do danych sprowadza się do poniższego problemu programowania liniowego z ograniczeniami.

$$\begin{aligned} \tilde{Y}_t &= \tilde{A}_0 + \sum_{i=1}^n \tilde{A}_i \cdot x_{ti} = (a_0, c_0) + \sum_{i=1}^n (a_i, c_i)x_{ti}, \\ \sum_{t=1}^T c_0 + \sum_{t=1}^T \sum_{i=1}^n c_i |x_{ti}| &\rightarrow \min, \end{aligned} \quad (5.2)$$

przy ograniczeniach:

$$\begin{aligned} y_t &\leq a_0 + c_0 + \sum_{i=1}^n a_i x_{ti} + \sum_{i=1}^n c_i |x_{ti}| \quad (t = 1, \dots, T), \\ y_t &\geq a_0 - c_0 + \sum_{i=1}^n a_i x_{ti} - \sum_{i=1}^n c_i |x_{ti}| \quad (t = 1, \dots, T), \\ c_i &\geq 0 \quad (i = 0, 1, \dots, n), \end{aligned}$$

Innym podejściem do estymacji parametrów modelu jest wykorzystanie metody najmniejszych kwadratów. Wówczas w modelu uwzględniony jest również czynnik błędu, a model przyjmuje postać

$$\tilde{Y}_t = \tilde{A}_0 + \sum_{i=1}^n \tilde{A}_i \cdot x_{ti} + \tilde{e}_t \quad (t = 1, \dots, T), \quad (5.3)$$

gdzie błędy modelu wyrażają się wzorem

$$\tilde{e}_t = \tilde{Y}_t - \tilde{A}_0 - \sum_{i=1}^n \tilde{A}_i \cdot x_{ti}. \quad (5.4)$$

Warto zauważyć, że błędy  $\tilde{e}_t$  są liczbami rozmytymi, stąd utrudnione jest zastosowanie metody najmniejszych kwadratów w sposób bezpośredni. W 1988 roku Diamond jako pierwszy zaproponował użycie metody najmniejszych kwadratów z wykorzystaniem odpowiednio zdefiniowanej miary odległości między liczbami rozmytymi [12]. Zdefiniował on

metrykę  $L^2$  między trójkątnymi liczbami rozmytymi w następujący sposób

$$\begin{aligned} d(\tilde{A}_1, \tilde{A}_2)^2 &= d((m_1, l_1, r_1), (m_2, l_2, r_2))^2 \\ &= (m_1 - m_2)^2 + ((m_1 - l_1) - (m_2 - l_2))^2 + ((m_1 + r_1) - (m_2 + r_2))^2. \end{aligned} \quad (5.5)$$

Wówczas, estymacja modelu rozmytej regresji sprowadza się do minimalizacji błędu postaci

$$\sum_{t=1}^T d\left(\tilde{Y}_t, \tilde{A}_0 + \sum_{i=1}^n \tilde{A}_i \cdot x_{ti}\right)^2. \quad (5.6)$$

W niniejszym podrozdziale przedstawiono tylko podstawowe informacje na temat modelu regresji rozmytej. Przez ponad trzy dekady podejście to zostało rozwinięte i z powodzeniem zastosowane w wielu praktycznych problemach przez liczną grupę badaczy. Obszerny przegląd literatury w tym zakresie można znaleźć w publikacjach [30, 88]. Jako pierwszy do modelowania szeregów czasowych zastosował je Watada [91, 92].

### 5.1.3. Podejście oparte na modelach Boxa i Jenkinsa

Bazując na metodologii modelowania szeregów czasowych przedstawionej przez Boxa i Jenkinsa [3], Ozawa, wraz ze współbadaczami, zaproponowali *rozmyty model autoregresyjny* (ang. *fuzzy auto-regressive model*) [64] w postaci klasycznego równania autoregresyjnego o współczynnikach rzeczywistych, gdzie wartości szeregu czasowego zostały rozmyte za pomocą liczb rozmytych.

**Definicja 5.4.** *Rozmytym modelem autoregresyjnym rzędu  $p$  ( $FAR(p)$ ) nazywamy model postaci*

$$\left. \begin{aligned} \hat{\tilde{Y}}_t &= \phi_1 \cdot \tilde{Y}_{t-1} + \dots + \phi_p \cdot \tilde{Y}_{t-p} + \tilde{U} \\ \tilde{Y}_t &\subseteq \hat{\tilde{Y}}_t, \quad \hat{\tilde{Y}}_t = (\hat{\alpha}_t, \hat{\beta}_t, \hat{\delta}_t) \end{aligned} \right\} \quad (5.7)$$

gdzie  $\tilde{Y}_t$  ( $t = 1, \dots, T$ ) to trójkątne liczby rozmyte określone przez trzy parametry rzeczywiste  $\alpha_t \leq \beta_t \leq \delta_t$  tak, że  $\tilde{Y}_t = (\alpha_t, \beta_t, \delta_t)$ , relacja zawierania dla trójkątnych liczb rozmytych zdefiniowana jest następująco

$$\tilde{Y}_t \subseteq \tilde{Y}_s \Leftrightarrow \alpha_t \geq \alpha_s \wedge \delta_t \leq \delta_s. \quad (5.8)$$

Operacje arytmetyczne dla rozmytych liczb trójkątnych zdefiniowane są w poniższy sposób

$$\begin{aligned}\tilde{Y}_t + \tilde{Y}_s &= (\alpha_t + \alpha_s, \beta_t + \beta_s, \delta_t + \delta_s) \\ \tilde{Y}_t - \tilde{Y}_s &= (\alpha_t - \alpha_s, \beta_t - \beta_s, \delta_t - \delta_s) \\ p \cdot \tilde{Y}_t &= \begin{cases} (p\alpha_t, p\beta_t, p\delta_t), & p \geq 0, \\ (p\delta_t, p\beta_t, p\alpha_t), & p < 0, \end{cases} \quad p \in \mathbb{R}.\end{aligned}\tag{5.9}$$

Ponadto,  $\phi_i$  ( $i = 1, \dots, p$ ) to współczynniki modelu, liczby rzeczywiste, natomiast  $\tilde{U} = (u_\alpha, u_\beta, u_\delta)$  to trójkątna liczba rozmyta.

Rozmyty model autoregresyjny opisuje zależność między przyszłymi a przeszłymi wartościami rozmytego szeregu czasowego za pomocą rzeczywistych współczynników  $\phi_i$  oraz nieprecyzyjność systemu poprzez liczbę rozmytą  $\tilde{U}$ . Dopasowanie modelu do danych odbywa się podobnie jak w modelu rozmytej regresji poprzez minimalizację rozmycia modelu i sprowadza się do poniższego problemu programowania liniowego

$$\sum_{t=p+1}^T (\hat{\delta}_t - \hat{\alpha}_t) \rightarrow \min,$$

przy ograniczeniach:

$$\alpha_t \geq \hat{\alpha}_t, \tag{5.10}$$

$$\delta_t \leq \hat{\delta}_t,$$

$$(t = p + 1, p + 2, \dots, T),$$

$$u_\alpha \leq u_\delta.$$

W nieco odmienny sposób, rozwijając podejście zaproponowane przez Watade, bazujące na modelu regresji rozmytej, Tseng i inni [86] skonstruowali rozmyty model ARIMA oraz rozmyty sezonowy model ARIMA [87], gdzie współczynniki klasycznego modelu ARIMA zostały zastąpione liczbami rozmytymi.

**Definicja 5.5.** *Rozmytym, autoregresyjnym, zintegrowanym modelem średniej ruchomej rzędów  $p, d, q$  (ARIMA( $p, d, q$ )) nazywamy model postaci*

$$\hat{Z}_t = \tilde{A}_1 \cdot Z_{t-1} + \dots + \tilde{A}_p \cdot Z_{t-p} + \varepsilon - \tilde{A}_{p+1} \cdot \varepsilon_{t-1} - \dots - \tilde{A}_{p+q} \cdot \varepsilon_{t-q}, \tag{5.11}$$



gdzie  $Z_t = \nabla_d(Y_t - \mu)$  to szereg czasowy powstały przez zintegrowanie w stopniu  $d$  szeregu czasowego  $Y_t - \mu$  ( $t = 1, \dots, T$ ), gdzie  $\mu$  to wartość średnia  $Y_t$ ,  $\tilde{A}_i$  ( $i = 1, \dots, p + q$ ) to współczynniki modelu będące trójkątnymi liczbami rozmytymi. Ponadto  $\varepsilon_t$  ( $t = 1, \dots, T$ ) to niezależne zmienne losowe o identycznych rozkładach normalnych z wartością średnią zero i wariancją równą  $\sigma^2$ .

Szereg czasowy  $Y_t$  jest szeregiem czasowym o wartościach rzeczywistych, stąd badanie stacjonarności, jak również identyfikacja rzędów modelu odbywa się zgodnie z metodologią Boxa-Jenkinsa. Estymacja współczynników modelu wykonywana jest poprzez rozwiązanie odpowiednio skonstruowanego problemu programowania liniowego poprzez analogię do problemu (5.2).

## 5.2. Skierowane rozmyte szeregi czasowe

Konstruowanie modeli rozmytych w oparciu o klasyczną definicję zbiorów rozmytych i zasadę rozszerzania dla operacji arytmetycznych obarczone jest kilkoma wadami. Wady te wynikają wprost z niedoskonałości klasycznych liczb rozmytych (zobacz roz. 3.1). Mianowicie, wprowadzenie większej liczby współczynników modelu, a co za tym idzie większej liczby operacji, powoduje zwiększenie rozmycia zbioru wynikowego. Ponadto, w większości modeli wprowadza się uproszczenie, że wszystkie zbiory rozmyte posiadają funkcję przynależności tego samego typu (np. trójkątną, typu L-R) i aby uniknąć błędów zaokrągleń wynikających z przybliżenia liczby wynikowej do danego typu unika się operacji mnożenia i dzielenia liczb rozmytych.

Wykorzystanie skierowanych liczb rozmytych wraz z ich arytmetyką pozwala w prosty sposób na skonstruowanie w pełni rozmytych modeli pozbawionych powyższych niedoskonałości w postaci dowolnych równań. W pełni rozmytych, czyli takich gdzie zarówno współczynniki, dane wejściowe jak i również błędy modelu są skierowanymi liczbami rozmytymi. Ogólnie model taki można przedstawić w następujący sposób

$$\tilde{Y}_t = F(\tilde{A}_i; \tilde{X}_{ti}) + \tilde{\varepsilon}_t, \quad (5.12)$$

gdzie  $\tilde{Y}_t$  to zmienna objaśniana, SLR, ( $t = 1, \dots, T$ ),  $\tilde{X}_{ti}$  to zmienne objaśniające, SLR, ( $i = 1, \dots, n$ ),  $\tilde{A}_i$  to współczynniki modelu, SLR, natomiast  $\tilde{\varepsilon}_t$  to błąd modelu, również SLR. Funkcja  $F$  to dowolna zależność funkcyjna określona na skierowanych liczbach rozmytych.

Zagadnienia estymacji współczynników oraz predykcji w dużym stopniu zależą od postaci funkcji  $F$ , dlatego też zostaną one omówione w kolejnych podrozdziałach dla podstawowych modeli takich jak model regresji liniowej oraz model autoregresyjny.

### 5.2.1. Skierowany rozmyty model regresji liniowej

Przyjmując w równaniu (5.12) liniową względem współczynników postać funkcji  $F$  otrzymujemy *skierowane rozmyte równanie regresji liniowej* w następującej postaci

$$\tilde{Y}_t = \tilde{A}_0 + \sum_{i=1}^n (\tilde{A}_i \cdot \tilde{X}_{ti}) + \tilde{\varepsilon}_t, \quad (5.13)$$

gdzie błędy modelu mogą zostać wyrażone wzorem

$$\tilde{\varepsilon}_t = \tilde{Y}_t - \tilde{A}_0 - \sum_{i=1}^n (\tilde{A}_i \cdot \tilde{X}_{ti}). \quad (5.14)$$

**Estymacja:** Celem estymacji jest dopasowanie modelu regresji (współczynników  $\tilde{A}_i$ ) do zaobserwowanego zbioru danych empirycznych. Model jest tym lepiej dopasowany im mniejsza jest odległość wartości teoretycznych  $\hat{Y}_t = \tilde{A}_0 + \sum_{i=1}^n \tilde{A}_i \cdot \tilde{X}_{ti}$  od wartości zaobserwowanych  $\tilde{Y}_t$ . Celem zatem jest minimalizacja całkowitego błędu postaci

$$E = \sum_{t=1}^T d(\tilde{Y}_t, \hat{Y}_t), \quad (5.15)$$

gdzie  $d(\cdot, \cdot)$  oznacza miarę odległość między skierowanymi liczbami rozmytymi. W klasycznym modelu regresji liniowej dla liczb rzeczywistych najczęściej wybierana jest metryka indukowana przez normę  $l_2$  (metoda najmniejszych kwadratów). Skierowane liczby rozmyte  $A$  i  $B$  to z definicji uporządkowane pary funkcji ciągłych tj.  $A = (f_A, g_A)$ ,  $B = (f_B, g_B)$ . Wobec czego naturalnym wydaje się zaproponowanie metryki dla SLR wykorzystującej normę z przestrzeni  $L_2$ .

**Definicja 5.6.** Odległość między skierowanymi liczbami rozmytymi  $A = (f_A, g_A)$ ,  $B = (f_B, g_B)$  to metryka dana wzorem

$$d(A, B) = d((f_A, g_A), (f_B, g_B)) = \|f_A - f_B\|_2 + \|g_A - g_B\|_2, \quad (5.16)$$

gdzie  $\|\cdot\|_2$  to norma  $L_2$  w przestrzeni  $\mathcal{C}([0, 1])$  (funkcji ciągłych o wartościach rzeczywistych określonych na przedziale  $[0, 1]$ ) dana wzorem

$$\|f\|_2 = \left( \int_0^1 |f(x)|^2 dx \right)^{\frac{1}{2}}. \quad (5.17)$$

W modelu (5.13) optymalizowanymi współczynnikami są skierowane liczby rozmyte postaci  $\tilde{A}_i = (f_{\tilde{A}_i}, g_{\tilde{A}_i})$ , wobec czego zagadnienie estymacji można sprowadzić do rozwiązania zadania rachunku wariacyjnego tj. znalezienia funkcji  $f_{\tilde{A}_i}, g_{\tilde{A}_i}$  ( $i = 0, \dots, n$ ) minimalizujących funkcjonal postaci

$$E(f_{\tilde{A}_i}, g_{\tilde{A}_i}) = \int_0^1 \sum_{t=1}^T (f_{\tilde{Y}_t}(x) - f_{\hat{Y}_t}(x))^2 dx + \int_0^1 \sum_{t=1}^T (g_{\tilde{Y}_t}(x) - g_{\hat{Y}_t}(x))^2 dx \quad (5.18)$$

Poprzez dyskretyzację skierowanych liczb rozmytych (por. rys. 3.5) oraz zastosowanie metod całkowania numerycznego powyższy problem można rozwiązać wykorzystując szeroką gamę algorytmów optymalizacji.

Z drugiej strony, dzięki dobrze zdefiniowanej arytmetyce skierowanych liczb rozmytych, optymalne współczynniki modelu mogą być wyliczone, tak jak w przypadku klasycznym, poprzez operacje na postaci macierzowej modelu. Model (5.13) możemy zapisać jako

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{a} + \mathbf{e}, \quad (5.19)$$

gdzie:

$$\mathbf{y} = \begin{bmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \\ \vdots \\ \tilde{Y}_T \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \tilde{1} & \tilde{X}_{11} & \cdots & \tilde{X}_{1n} \\ \tilde{1} & \tilde{X}_{21} & \cdots & \tilde{X}_{2n} \\ \vdots & \vdots & & \vdots \\ \tilde{1} & \tilde{X}_{T1} & \cdots & \tilde{X}_{Tn} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} \tilde{A}_0 \\ \tilde{A}_1 \\ \vdots \\ \tilde{A}_n \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_T \end{bmatrix}. \quad (5.20)$$

Niech  $\tilde{S}$  będzie minimalizowaną sumą kwadratów reszt modelu, która, jak łatwo zauważyć, jest SLR. Wówczas korzystając z rachunku macierzowego otrzymujemy

$$\begin{aligned}\tilde{S} &= \sum_{t=1}^T \tilde{\varepsilon}_t^2 = \mathbf{e}^T \mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{a})^T (\mathbf{y} - \mathbf{X}\mathbf{a}) = \\ &= \dots = \mathbf{y}^T \mathbf{y} - 2\mathbf{a}^T \mathbf{X}^T \mathbf{y} + \mathbf{a}^T \mathbf{X}^T \mathbf{X} \mathbf{a}\end{aligned}\quad (5.21)$$

Różniczkując powyższe wyrażenie względem  $\mathbf{a}$  i przyrównując do zera otrzymujemy

$$\frac{\partial \tilde{S}}{\partial \mathbf{a}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \mathbf{a} = 0. \quad (5.22)$$

A po przekształceniu

$$\mathbf{X}^T \mathbf{X} \mathbf{a} = \mathbf{X}^T \mathbf{y}. \quad (5.23)$$

Przyjmując założenie jak w klasycznym modelu regresji, że macierz  $\mathbf{X}$  ma rząd równy liczbie współczynników ( $n + 1$ , pełny rząd kolumnowy). Wówczas macierz  $\mathbf{X}^T \mathbf{X}$  jest dodatnio określona, a zatem istnieje do niej macierz odwrotna  $(\mathbf{X}^T \mathbf{X})^{-1}$  i równanie (5.23) ma rozwiązanie postaci

$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (5.24)$$

Elementami macierzy  $\mathbf{X}$  są uporządkowane pary funkcji określonych na  $[0, 1]$ , wobec czego powyższe założenie musi być spełnione dla każdego  $x \in [0, 1]$ . Ponadto, ponieważ druga pochodna  $\tilde{S}$  jest równa  $2\mathbf{X}^T \mathbf{X}$  i zgodnie z założeniem jest dodatnio określona, zatem rozwiązanie (5.24) wyznacza minimum sumy kwadratów reszt  $\tilde{S}$ .

**Przykład:** Dla zilustrowania modelu rozmytej regresji liniowej opartego na arytmetyce skierowanych liczb rozmytych przeanalizowany zostanie prosty problem regresji liniowej jednej zmiennej dla danych Isibuchi [22]. Dane Isibuchi  $(x_t, y_t)$  wraz z losowo wygenero-

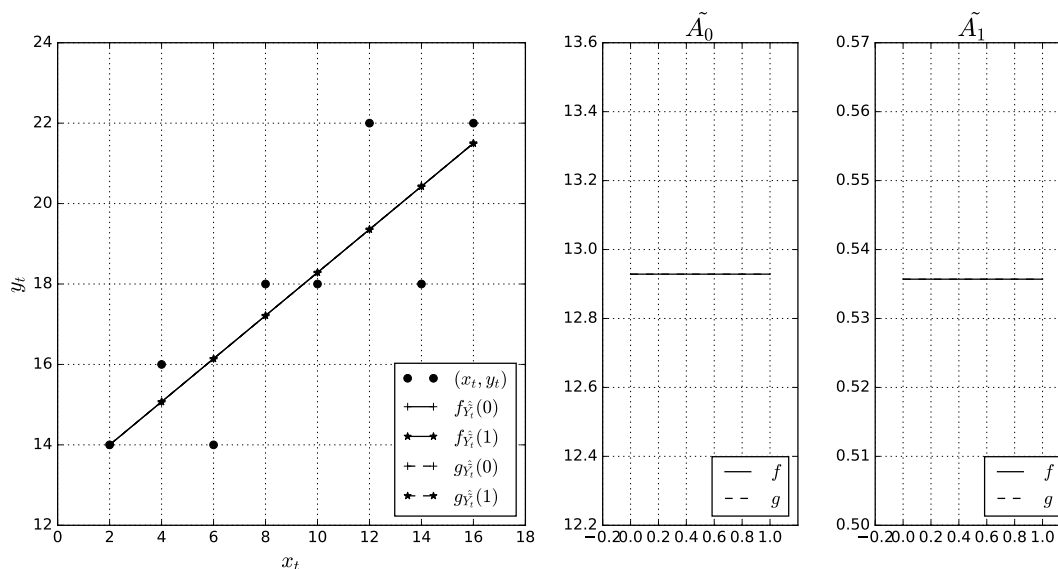
$t$	1	2	3	4	5	6	7	8
$x_t$	2	4	6	8	10	12	14	16
$y_t$	14	16	14	18	18	22	18	22
$er_{y_t}$	0.12417	0.45587	1.16176	0.70287	0.05801	1.90618	1.26043	0.14221
$er_{x_t}$	0.47001	0.11422	0.67202	0.16389	0.44385	0.13626	0.79932	0.35240

Tabela 5.1: Zbiór danych Ishibuchi z losowym błędem pomiarowym.

wanymi błędami dla  $y_t$  i  $x_t$  przedstawione są w tabeli 5.1. Dla danych tych model (5.13) przyjmuje postać

$$\tilde{Y}_t = \tilde{A}_0 + \tilde{A}_1 \cdot \tilde{X}_{t1} + \tilde{\varepsilon}_t. \quad (5.25)$$

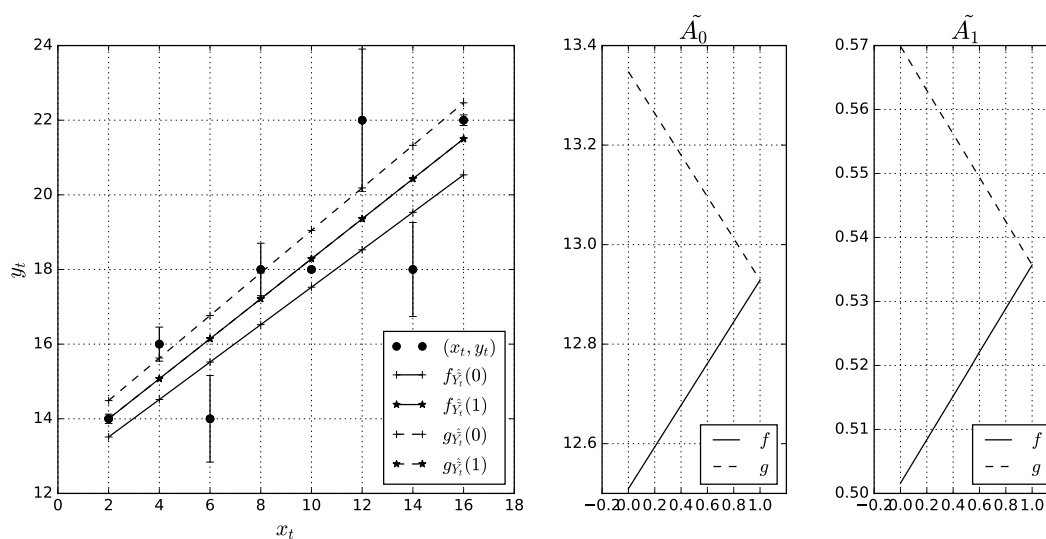
W pierwszej kolejności rozważony zostanie przypadek, w którym wszystkie zmienne modelu tj.  $\tilde{Y}_t, \tilde{X}_{t1}$  zostały rozmyte za pomocą skierowanych liczb rozmytych o funkcjach stałych i równych sobie (czyli są to po prostu liczby rzeczywiste przedstawione jak SLR, brak rozmycia). Wówczas, jak należałoby oczekiwać, w wyniku estymacji współczynników modelu również są SLR o funkcjach stałych i równych sobie. Co więcej, model jest tożsamy z klasycznym modelem regresji liniowej. Dopasowanie modelu rozmytej regresji do danych wraz z wykresami współczynników przedstawia rys. 5.1.



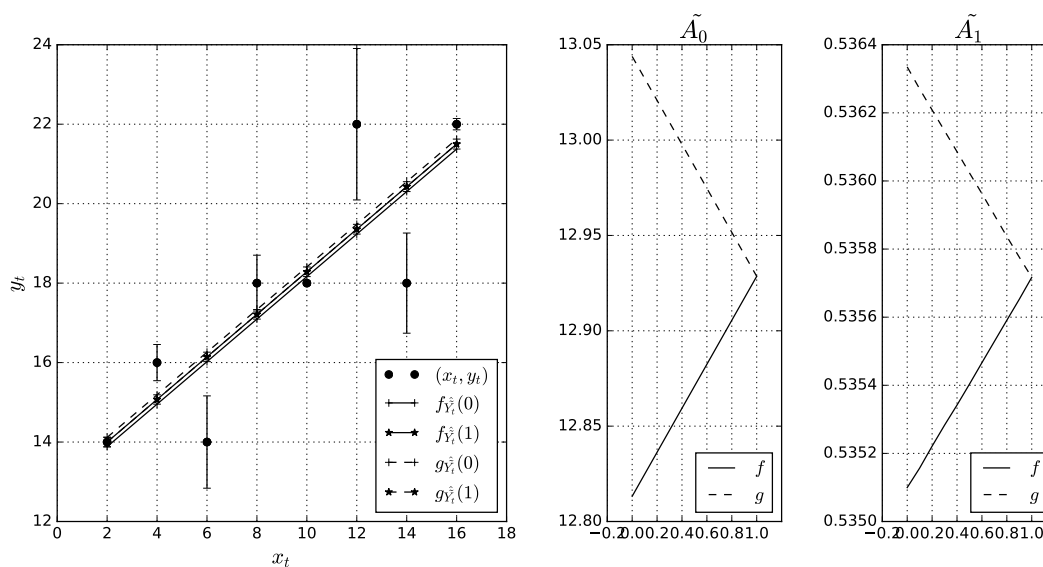
Rysunek 5.1: Skierowany rozmyty model regresji dla danych Ishibuchi - brak rozmycia.

W kolejnym przypadku, wartości zaobserwowane  $y_t$  zostały rozmyte za pomocą symetrycznych trójkątnych skierowanych liczb rozmytych zorientowanych dodatnio, tak że nośniki tych liczb wyznaczone są poprzez błędy pomiarowe  $er_{y_t}$ , czyli są przedziałami  $[y_t - er_{y_t}, y_t + er_{y_t}]$ . Wówczas otrzymane optymalne współczynniki są istotnie SLR. Dopasowanie modelu do danych wraz z wykresami współczynników przedstawia rys. 5.2. Warto zaznaczyć, że otrzymane w tym przypadku rezultaty są tożsame z modelem rozmytej regresji (5.3) przy zastosowaniu estymacji metodą najmniejszych kwadratów z metryką zaproponowaną przez Diamonda (5.5).

Jedną z cech odróżniających SLR od klasycznych liczb rozmytych jest skierowanie. Znaczenie i przypisanie odpowiedniego skierowania liczbie rozmytej w dużym stopniu zależy od kontekstu i charakterystyki rozważanego zagadnienia. Najczęściej jednak, skierowanie reprezentuje tendencję występującą w danych. W rozważonym poprzednio przypadku wprowadzona została modyfikacja polegająca na przypisaniu rozmywanym zmiennym  $y_t$  skierowania zgodnie z następującą zasadą. Jeżeli  $y_t < y_{t-1}$  to  $\tilde{Y}_t$  posiada orientację



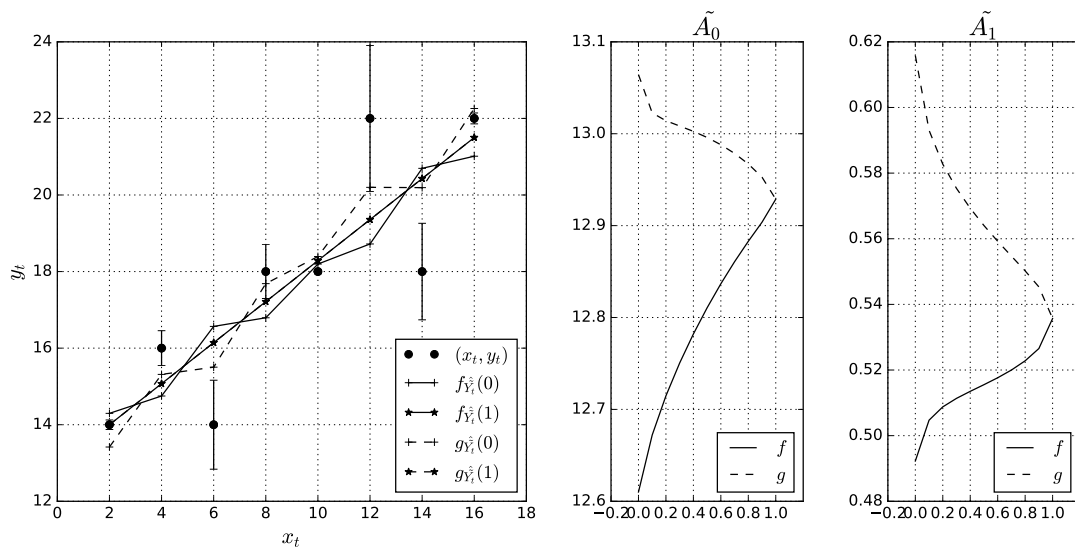
Rysunek 5.2: Skierowany rozmyty model regresji dla danych Ishibuchi - rozmycie  $y_t$ .



Rysunek 5.3: Skierowany rozmyty model regresji dla danych Ishibuchi - rozmycie  $y_t$  plus skierowanie.

ujemną, natomiast w przeciwnym przypadku orientację dodatnią. Dopasowanie modelu do danych wraz z wykresami współczynników przedstawia rys. 5.3. Jak można zauważyć, wprowadzenie liczb przeciwnie skierowanych wpłynęło na zmniejszenie rozmycia modelu poprzez zmniejszenie rozmycia współczynników. Jednak, podobnie jak poprzednio, wszystkie otrzymywane wartości  $\hat{Y}_t$  z modelu są SLR zorientowanymi dodatnio, a granice ich nośników wyznaczają linie proste.

Ostatnim rozważanym przypadkiem jest model, w którym uwzględniona została również niepewność związana ze zmiennymi objaśniającymi  $x_t$ . W modelu tym wartości zaobserwowane  $y_t$  rozmyte są jak w poprzednim przypadku, natomiast zmienne objaśniające  $x_t$  rozmyte są za pomocą gaussowskich skierowanych liczb rozmytych zorientowanych dodatnio o parametrach  $m_f = m_g = x_t$ ,  $\sigma_f = er_{x_t}$ ,  $\sigma_g = -er_{x_t}$  (zobacz wzory (4.8)). Dopasowanie modelu rozmytej regresji do danych wraz z wykresami współczynników przedstawia rys. 5.4. Tak skonstruowany model wciąż wyznacza pewną zależność liniową, jednak rozmycie linii regresji ulega dynamicznieszym zmianom. Ponadto, również orientacja otrzymanych liczb dla  $\hat{Y}_t$  ulega zmianie i jest zgodna z zadanymi orientacjami liczb  $\tilde{Y}_t$ .



Rysunek 5.4: Skierowany rozmyty model regresji dla danych Ishibuchi - rozmycie  $y_t$  plus skierowanie oraz rozmycie  $x_t$ .

**Predykcja:** Oszacowany na podstawie szeregów czasowych model regresji liniowej może zostać wykorzystany do celów prognozowania (predykcji), czyli wnioskowania w przyszłość na podstawie modelu. W klasycznym modelu regresji liniowej prognoza dla zmiennej  $Y_{t+s}$  (na chwilę  $t + s$ ) określana jest jako jej wartość oczekiwana tj.

$$\hat{Y}_{t+s} = \mathbb{E}(Y_{t+s}) = \mathbb{E} \left( A_0 + \sum_{i=1}^n (A_i \cdot X_{(t+s)i}) + \varepsilon_{t+s} \right). \quad (5.26)$$

Jedynym czynnikiem niedeterministycznym w równaniu (5.26) jest błąd  $\varepsilon_{t+s}$ , jednak z założeń klasycznego modelu regresji liniowej ma on wartość oczekiwaną równą zero dla każdego  $t$ . Wobec czego prognoza na chwilę  $t + s$  wyraża się wzorem

$$\hat{Y}_{t+s} = A_0 + \sum_{i=1}^n (A_i \cdot X_{(t+s)i}). \quad (5.27)$$

W analogiczny sposób można zdefiniować prognozę w rozmytym modelu regresji liniowej danym równaniem (5.13). Jedyny problem polega na tym, że czynnik błędu  $\tilde{\varepsilon}_t$  jest skierowaną liczbą rozmytą. Niemniej, może być ona potraktowana jako zmienna losowa przyjmująca wartości w zbiorze skierowanych liczb rozmytych, więcej na ten temat w rozdziale 6.2 poświęconym rozmytym zmiennym losowym. Na chwilę obecną przyjmijmy założenie, że  $\tilde{\varepsilon}_t$  to skierowana rozmyta zmienna losowa w tym sensie, że dla każdego  $x \in [0, 1]$ ,  $f_{\tilde{\varepsilon}_t}(x)$  i  $g_{\tilde{\varepsilon}_t}(x)$  są rzeczywistymi zmiennymi losowymi o wartości oczekiwanej równej zero. Wówczas, prognoza na chwilę  $t + s$  w modelu rozmytym wyraża się wzorem

$$\hat{\tilde{Y}}_{t+s} = \tilde{A}_0 + \sum_{i=1}^n (\tilde{A}_i \cdot \tilde{X}_{(t+s)i}). \quad (5.28)$$

Warto zwrócić uwagę, że  $\hat{\tilde{Y}}_{t+s}$  jest skierowaną liczbą rozmytą, co powala na wyróżnienie trzech typów prognoz:

- **prognoza punktowa:** wyznaczona poprzez wyostrzoną wartość liczby  $\hat{\tilde{Y}}_{t+s}$ ,
- **prognoza przedziałowa:** wyznaczona poprzez podzbiór nośnika liczby  $\hat{\tilde{Y}}_{t+s}$ ,
- **prognoza kierunkowa:** wyznaczona poprzez orientację liczby  $\hat{\tilde{Y}}_{t+s}$ .



### 5.2.2. Skierowany rozmyty model autoregresyjny

W rozważonym poprzednio modelu regresji liniowej struktura zjawiska uzależniona była od zbioru zmiennych objaśniających. Dla prognozowania niezbędna jest więc znajomość tych zmiennych dla okresu prognozowanego, co w praktyce nie zawsze jest możliwe. W takich sytuacjach można skorzystać z modeli opartych na analizie pojedynczego szeregu czasowego, których prostym przykładem jest model autoregresyjny.

*Klasyczny model autoregresyjny rzędu  $p$  ( $AR(p)$ )* to model, w którym obecna wartość szeregu czasowego zależy tylko od wartości tego samego szeregu w przeszłości oraz pewnego błędu losowego [85]. W pełni rozmytym, bazującym na skierowanych liczbach rozmytych uogólnieniem tego modelu jest *skierowany rozmyty model autoregresyjny rzędu  $p$  ( $SRAR(p)$ )* postaci

$$\tilde{Y}_t = \tilde{A}_0 + \sum_{i=1}^p (\tilde{A}_i \cdot \tilde{Y}_{t-i}) + \tilde{\varepsilon}_t, \quad (5.29)$$

gdzie  $\tilde{Y}_t$  to wartość szeregu czasowego w chwili  $t$ , SLR, ( $t = 0, \dots, T$ ),  $\tilde{Y}_{t-i}$  to wartości szeregu czasowego w przeszłości, zmienne opóźnione, SLR, ( $i = 1, \dots, p$ ),  $\tilde{A}_i$  to współczynniki modelu, SLR, natomiast  $\tilde{\varepsilon}_t$  to błąd modelu, również SLR.

**Stacjonarność:** Jednym z podstawowych pojęć stosowanych w analizie szeregów czasowych jest pojęcie stacjonarności szeregu czasowego. Intuicyjnie, szereg czasowy stacjonarny (stabilny) to szereg, którego własności (wartość oczekiwana, wariancja, itp.) nie zmieniają się wraz z upływem czasu [93]. W literaturze występuje kilka definicji stacjonarności, jednak najczęściej używane jest pojęcie stacjonarności słabej (kowariancyjnej, stacjonarności w szerszym sensie).

**Definicja 5.7.** Szereg czasowy  $Y_t$  nazywamy szeregiem *słabo stacjonarnym* (*stacjonarnym*), gdy:

$$\mathbb{E}(|Y_t|^2) = \sigma^2 < \infty \quad (5.30)$$

$$\mathbb{E}(Y_t) = \mu < \infty \quad (5.31)$$

$$\text{cov}(Y_t, Y_{t+h}) = \text{cov}(Y_s, Y_{s+h}) = \gamma_h \text{ dla dowolnych } t, s, h. \quad (5.32)$$

W ten sam sposób można podejść do stacjonarności skierowanych rozmytych szeregów czasowych.

**Definicja 5.8.** Skierowany rozmyty szereg czasowy  $\tilde{Y}_t$  nazywamy *stacjonarnym* jeżeli dla każdego ustalonego  $x \in [0, 1]$  szeregi czasowe  $f_{\tilde{Y}_t}(x)$  oraz  $g_{\tilde{Y}_t}(x)$  są stacjonarne.

Założenie stacjonarności badanego szeregu czasowego jest istotne, gdyż szacowanie modeli regresyjnych na podstawie niestacjonarnych szeregów czasowych często może prowadzić do tak zwanej *regresji pozornej* (ang. *spurious regression*), co może prowadzić do błędnych wyników wnioskowania statystycznego. Istnieje wiele metod sprowadzania niestacjonarnego szeregu czasowego do szeregu stacjonarnego, na przykład poprzez różnicowanie, odejmowanie wartości średniej lub trendu. W przypadku finansowych szeregów czasowych, najczęściej rozważa się szeregi czasowe będące szeregami prostych lub logarytmicznych stóp zwrotu, które w większości przypadków okazują się być szeregami stacjonarnymi. W przypadku skierowanych rozmytych szeregów czasowych można zastosować te same metody. Można jednak wykonać je na różnych etapach modelowania, przed albo po rozmyciu szeregu czasowego.

Ponieważ stacjonarność skierowanego rozmytego szeregu czasowego została sprawdzona do stacjonarności odpowiednich ostrych szeregów czasowych, to do testowania ich stacjonarności można wykorzystać dobrze znane testy stacjonarności, takie jak na przykład test pierwiastka jednostkowego Dickey-Fullera (DF) [15, 85]. Test ten polega na testowaniu, czy szereg czasowy  $Y_t$  da się zapisać w jednej z dwóch postaci

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t, \quad (5.33)$$

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t, \quad (5.34)$$

gdzie  $\varepsilon_t$  jest białym szumem, a weryfikowana jest hipoteza zerowa  $H_0: \phi_1 = 1$  (niestacjonarność) przeciw hipotezie alternatywnej  $H_a: \phi_1 < 1$  (stacjonarność). Do weryfikacji hipotezy zerowej wykorzystywana jest statystyka Dickey-Fullera postaci

$$DF = \frac{\hat{\phi}_1 - 1}{SE(\hat{\phi}_1)}, \quad (5.35)$$

gdzie  $\hat{\phi}_1$  jest estymatorem najmniejszych kwadratów parametru  $\phi_1$ , natomiast  $SE(\hat{\phi}_1)$  jest standardowym błędem estymatora  $\hat{\phi}_1$ . Okazuje się, że statystyka  $DF$  nie ma rozkładu t-Studenta i nie jest asymptotycznie zbieżna do żadnego ze znanych rozkładów. Statystyka  $DF$  ma tzw. rozkład Dickey-Fullera, dla którego wartości krytyczne, zależne od poziomu istotności testu, można znaleźć w literaturze m.in. w [4].

**Identyfikacja rzędu  $p$ :** Do wyboru rzędu  $p$  można posłużyć się procedurą Boxa i Jenkinsa [3] dla modelu autoregresyjnego, która polega na porównaniu *funkcji autokorelacji cząstkowej* (ang. *partial autocorrelation function (PACF)*) dla badanego szeregu z postacią tej funkcji dla teoretycznych modeli  $AR(p)$  o znanej liczbie parametrów. Autokorelacja cząstkowa mierzy zależność liniową jednej zmiennej po usunięciu innej zmiennej, która ma wpływ na obie zmienne [85].

Teoretyczny model  $AR(p)$  ma funkcję PACF równą zero dla opóźnień większych niż  $p$  (tzn. współczynniki autokorelacji cząstkowej dla opóźnień większych niż  $p$  są nieistotne). Tak, więc za  $p$  przyjmuje się maksymalne opóźnienie dla którego współczynnik autokorelacji jest istotny. W przypadku skierowanych rozmytych szeregów czasowych, procedurze tej można poddać każdy ze skonstruowanych szeregów dla ustalonego  $x \in [0, 1]$  i wybrać największy z otrzymanych rzędów lub szereg czasowy powstały przez wyostrenie rozmytego szeregu czasowego.

**Estymacja:** Oszacowanie wartości współczynników  $\tilde{A}_i$  w modelu (5.29) może odbywać się w taki sam sposób jak w modelu regresji z wykorzystaniem metody najmniejszych kwadratów. Wówczas, aby wyznaczyć optymalne współczynniki  $\tilde{A}_i = (f_{\tilde{A}_i}, g_{\tilde{A}_i})$  należy znaleźć funkcje  $f_{\tilde{A}_i}, g_{\tilde{A}_i}$  ( $i = 0, \dots, p$ ) minimalizujące funkcjonal

$$E(f_{\tilde{A}_i}, g_{\tilde{A}_i}) = \int_0^1 \sum_{t=p}^T (f_{\tilde{Y}_t}(x) - f_{\hat{Y}_t}(x))^2 dx + \int_0^1 \sum_{t=p}^T (g_{\tilde{Y}_t}(x) - g_{\hat{Y}_t}(x))^2 dx \quad (5.36)$$

gdzie  $\tilde{Y}_t$  to wartości zaobserwowane szeregu czasowego, natomiast  $\hat{Y}_t = \tilde{A}_0 + \sum_{i=1}^p \tilde{A}_i \cdot \tilde{Y}_{t-i}$  to wartości wyliczone z modelu.

Ponadto, zakładając, że  $\tilde{A}_0 = 0$  (proces scentrowany w zerze) to dzięki arytmetyce skierowanych liczb rozmytych do estymacji parametrów może być zastosowana metoda Yule'a-Walkera [31]. W metodzie tej optymalne wartości współczynników wyznaczone są z układu równań liniowych postaci

$$\begin{bmatrix} \hat{\gamma}_0 & \hat{\gamma}_1 & \cdots & \hat{\gamma}_{p-1} \\ \hat{\gamma}_1 & \hat{\gamma}_0 & \cdots & \hat{\gamma}_{p-2} \\ \vdots & \vdots & & \vdots \\ \hat{\gamma}_{p-1} & \hat{\gamma}_{p-2} & \cdots & \hat{\gamma}_0 \end{bmatrix} \cdot \begin{bmatrix} \tilde{A}_1 \\ \tilde{A}_2 \\ \vdots \\ \tilde{A}_p \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \vdots \\ \hat{\gamma}_p \end{bmatrix}, \quad (5.37)$$

gdzie  $\hat{\gamma}_k$  to wartości funkcji autokowariancji z próby danej wzorem

$$\hat{\gamma}_k = \hat{\gamma}(k) = \frac{1}{T} \sum_{t=k+1}^T (\tilde{Y}_t \cdot \tilde{Y}_{t-k}), \quad (k = 0, 1, \dots). \quad (5.38)$$

Ponieważ elementami macierzy w równaniu (5.37) są skierowane liczby rozmyte, warunkiem na to, aby układ ten posiadał dokładnie jedno rozwiązanie, jest aby  $f_{\hat{\gamma}_0}(x) > 0$  oraz  $g_{\hat{\gamma}_0}(x) > 0$  dla każdego  $x \in [0, 1]$ .

**Predykcja:** Inaczej niż w przypadku modelu regresji liniowej, prognoza wartości szeregu czasowego w chwili  $t + s$  ( $s$ -ta prognoza) w klasycznym modelu autoregresyjnym rzędu  $p$  definiowana jest jako warunkowa wartość oczekiwana względem filtracji  $\mathcal{F}_t$ , co zapisuje się jako

$$\hat{Y}_{t+s} = \mathbb{E}(Y_{t+s} | \mathcal{F}_t) = \mathbb{E} \left( A_0 + \sum_{i=1}^p (A_i \cdot Y_{t-i+s}) + \varepsilon_{t+s} | \mathcal{F}_t \right). \quad (5.39)$$

Nie wchodząc w formalizm matematyczny, filtracja  $\mathcal{F}_t$  oznacza wszystkie informacje na temat szeregu czasowego  $Y_t$  dostępne do chwili  $t$ . Korzystając z własności warunkowej wartości oczekiwanej [24] oraz zakładając, że warunkowa wartość oczekiwana  $\varepsilon_t$  jest równa zeru i nie zależy od filtracji  $\mathcal{F}_t$ , wówczas prognoza na jeden krok do przodu jest równa

$$\begin{aligned} \hat{Y}_{t+1} &= \mathbb{E}(Y_{t+1} | \mathcal{F}_t) = \mathbb{E} \left( A_0 + \sum_{i=1}^p (A_i \cdot Y_{t-i+1}) + \varepsilon_{t+1} | \mathcal{F}_t \right) \\ &= A_0 + \sum_{i=1}^p (A_i \cdot \mathbb{E}(Y_{t-i+1} | \mathcal{F}_t)) = A_0 + \sum_{i=1}^p (A_i \cdot Y_{t-i+1}). \end{aligned} \quad (5.40)$$

Prognoza na dwa kroki do przodu jest równa

$$\begin{aligned}
\hat{Y}_{t+2} &= \mathbb{E}(Y_{t+2}|\mathcal{F}_t) = \mathbb{E}\left(A_0 + \sum_{i=1}^p(A_i \cdot Y_{t-i+2}) + \varepsilon_{t+2}|\mathcal{F}_t\right) \\
&= A_0 + A_1 \cdot \mathbb{E}(Y_{t+1}|\mathcal{F}_t) + \sum_{i=2}^p(A_i \cdot Y_{t-i+2}) \\
&= A_0 + A_1 \cdot \hat{Y}_{t+1} + \sum_{i=2}^p(A_i \cdot Y_{t-i+2}).
\end{aligned} \tag{5.41}$$

Stąd  $s$ -ta prognoza dana jest wzorem

$$\begin{aligned}
\hat{Y}_{t+s} &= \mathbb{E}(Y_{t+s}|\mathcal{F}_t) = \mathbb{E}\left(A_0 + \sum_{i=1}^p(A_i \cdot Y_{t-i+s}) + \varepsilon_{t+s}|\mathcal{F}_t\right) \\
&= A_0 + \sum_{i=1}^p(A_i \cdot \hat{Y}_{t-i+s}).
\end{aligned} \tag{5.42}$$

Traktując w równaniach (5.39), (5.40), (5.41) oraz (5.42) wartości szeregu czasowego oraz współczynniki jako skierowane liczby rozmyte,  $s$ -ta prognoza w skierowanym rozmytym modelu autoregresyjnym wyraża się wzorem

$$\hat{\tilde{Y}}_{t+s} = \tilde{A}_0 + \sum_{i=1}^p(\tilde{A}_i \cdot \hat{\tilde{Y}}_{t-i+s}). \tag{5.43}$$

Podobnie jak w modelu rozmytej regresji, prognoza  $\hat{\tilde{Y}}_{t+s}$  jest skierowaną liczbą rozmytą, co pozwala na wyróżnienie trzech typów prognoz:

- **prognoza punktowa:** wyznaczona poprzez wyostrzoną wartość liczby  $\hat{\tilde{Y}}_{t+s}$ ,
- **prognoza przedziałowa:** wyznaczona poprzez podzbiór nośnika liczby  $\hat{\tilde{Y}}_{t+s}$ ,
- **prognoza kierunkowa:** wyznaczona poprzez orientację liczby  $\hat{\tilde{Y}}_{t+s}$ .

Przykład budowy krok po kroku skierowanego rozmytego modelu autoregresyjnego dla empirycznych finansowych szeregów czasowych przedstawiony jest w rozdziale 7.1.

## Rozdział 6

# Symulacje rozmytych szeregów czasowych

W wielu praktycznych zagadnieniach związanych z szeregami czasowymi, w szczególności takich, które zależą od przebiegu realizacji procesu, istnieje konieczność symulacji (najczęściej wielokrotnej) skonstruowanego modelu opisującego dane zjawisko w pewnym przedziale czasu. Symulacje stochastycznych modeli szeregów czasowych dokonywane są przy odpowiednio zadanych rozkładach prawdopodobieństwa dla zmiennych losowych występujących w modelu. W modelach skierowanych rozmytych szeregów czasowych, które zostały zaproponowane w poprzednim rozdziale (zobacz (5.12)) i mogą być traktowane jako uogólnienie klasycznych modeli, czynnikiem, który można rozważać jako niedeterministyczny jest błąd modelu  $\tilde{\varepsilon}_t$ . Ponieważ błąd ten wyrażony jest za pomocą skierowanej liczby rozmytej, istnieje konieczność określenia, jak rozumieć pojęcie skierowanej rozmytej zmiennej losowej i jak takie zmienne symulować.

W niniejszym rozdziale przytoczone zostaną trzy najbardziej cytowane w literaturze koncepcje rozmytych zmiennych losowych odnoszących się do klasycznego pojęcia zbioru rozmytego. Następnie wprowadzone i omówione zostanie pojęcie zmiennej losowej dla skierowanych liczb rozmytych wraz z pojęciami takimi jak wartość oczekiwana oraz wariancja. Na końcu przedstawiony zostanie przykład generowania takich zmiennych losowych.

## 6.1. Rozmyte zmienne losowe

### 6.1.1. Model Kwakernaaka

Określenie *rozmyta zmienna losowa* (ang. *fuzzy random variable*) zostało wprowadzona przez Kwakernaaka w 1978 roku [43, 44], który określa rozmytą zmienną losową jako niejasne postrzeżenie ostrej ale nieobserwowalnej zmiennej losowej. Jako przykład rozważmy zmienną losową opisującą wiek osób mijanych na ulicy. Wiek spotkanej osoby jest oczywiście ostrą zmienną losową o wartościach rzeczywistych dodatnich. Jest to jednak zmienna nieobserwowalna, gdyż nie pytamy każdej mijanej osoby o jej wiek. Postrzegamy ją natomiast jako zmienną losową o wartościach lingwistycznych opisanych zbiorami rozmytymi np. „osoba młoda”, „osoba w średnim wieku” lub „osoba w podeszłym wieku”.

Niech określona będzie klasyczna przestrzeń probabilistyczna  $(\Omega, \mathcal{A}, \Pr)$ , gdzie  $\Omega$  to przestrzeń zdarzeń elementarnych,  $\mathcal{A}$  to  $\sigma$ -algebra podzbiorów  $\Omega$ , natomiast  $\Pr$  to miara probabilistyczna określona na  $\Omega$  [24]. Niech  $\mathcal{F}(\mathbb{R})$  oznacza zbiór wszystkich wypukłych liczb rozmytych określonych na przestrzeni euklidesowej  $\mathbb{R}$ , których  $\alpha$ -przekroje dla każdego  $\alpha \in [0, 1]$  są zbiorami zwartymi. Wówczas rozmyta zmienna losowa zaproponowana przez Kwakernaaka, następnie ulepszona przez Kruse i Meyera [42] oraz zinterpretowana przez Gila [19], określona jest następującą definicją.

**Definicja 6.1.** *Rozmytą zmienną losową* nazywamy odwzorowanie  $\tilde{\xi}: \Omega \rightarrow \mathcal{F}(\mathbb{R})$  takie, że dla każdego  $\alpha \in [0, 1]$  oraz  $\omega \in \Omega$  odwzorowania

$$\inf \tilde{\xi}_\alpha: \Omega \rightarrow \mathbb{R} \text{ takie, że } \inf \tilde{\xi}_\alpha(\omega) = \inf (\tilde{\xi}(\omega))_\alpha, \quad (6.1)$$

$$\sup \tilde{\xi}_\alpha: \Omega \rightarrow \mathbb{R} \text{ takie, że } \sup \tilde{\xi}_\alpha(\omega) = \sup (\tilde{\xi}(\omega))_\alpha, \quad (6.2)$$

są klasycznymi zmiennymi losowymi o wartościach rzeczywistych (por. rys. 6.1a).

Dla tak zdefiniowanej rozmytej zmiennej losowej, poprzez zastosowanie zasady rozszerzania Zadeha, określone zostały własności takie jak wartość oczekiwana oraz wariancja rozmytej zmiennej losowej [41].

Niech  $\tilde{\xi}$  będzie rozmytą zmienną losową, natomiast  $\mathcal{U}_{\mathcal{A}}$  zbiorem wszystkich klasycznych  $\mathcal{A}$ -mierzalnych zmiennych losowych określonych na  $\Omega$ .

**Definicja 6.2.** Wartością oczekiwaną rozmytej zmiennej losowej  $\tilde{\xi}$  nazywamy zbiór rozmyty  $\mathbb{E}[\tilde{\xi}]$  taki, że

$$\mu_{\mathbb{E}[\tilde{\xi}]}(x) = \sup_{\{U \in \mathcal{U}_{\mathcal{A}}: \mathbb{E}[U]=x\}} \mu_{\tilde{\xi}}(U) \quad \text{dla } x \in \mathbb{R}. \quad (6.3)$$

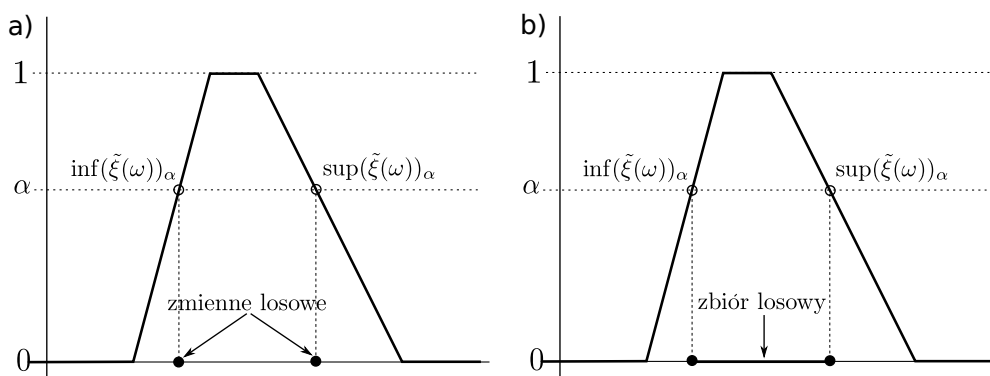
**Definicja 6.3.** Wariancją rozmytej zmiennej losowej  $\tilde{\xi}$  nazywamy zbiór rozmyty  $\text{Var}[\tilde{\xi}]$  taki, że

$$\mu_{\text{Var}[\tilde{\xi}]}(x) = \sup_{\{U \in \mathcal{U}_{\mathcal{A}}: \text{Var}[U]=x\}} \mu_{\tilde{\xi}}(U) \quad \text{dla } x \in \mathbb{R}. \quad (6.4)$$

### 6.1.2. Model Puri i Ralescu

Pojęcie rozmytej zmiennej losowej wprowadzone przez Kwakernaaka dotyczy jedynie liczb rozmytych określonych na  $\mathbb{R}$ . Niezadowoleni z tego faktu, Puri i Ralescu zaproponowali koncepcję rozmytej zmiennej losowej, której wartościami są rozmyte podzbiory przestrzeni  $\mathbb{R}^n$  (lub bardziej ogólnie, pewnej przestrzeni Banacha) [70]. Zdefiniowali oni rozmytą zmienną losową jako swego rodzaju rozmycie zbioru losowego, łącząc w ten sposób rozmyte zmienne losowe z dobrze rozwiniętą teorią *zbiorów losowych* (ang. *random sets* [57]).

Niech określona będzie przestrzeń probabilistyczna  $(\Omega, \mathcal{A}, \text{Pr})$ ,  $\mathcal{F}(\mathbb{R}^n)$  będzie zbiorem rozmytych podzbiorów  $\mathbb{R}^n$  o funkcjach przynależności  $\mu: \mathbb{R}^n \rightarrow [0, 1]$  oraz  $\mathcal{K}(\mathbb{R}^n)$  oznacza niepusty zwarty podzbiór  $\mathbb{R}^n$ .



Rysunek 6.1:  $\alpha$ -przekrój rozmytej zmiennej losowej: a) Kwakernaak, b) Puri i Ralescu.



**Definicja 6.4.** *Rozmytą zmienną losową* nazywamy odwzorowanie  $\tilde{\xi}: \Omega \rightarrow \mathcal{F}(\mathbb{R}^n)$  takie, że dla każdego  $\alpha \in [0, 1]$  odwzorowanie  $\tilde{\xi}_\alpha: \Omega \rightarrow \mathcal{K}(\mathbb{R}^n)$  (gdzie  $\tilde{\xi}_\alpha(\omega) = (\tilde{\xi}(\omega))_\alpha$  dla każdego  $\omega \in \Omega$ ) jest zbiorem losowym (por. rys. 6.1b).

Puri i Ralescu określili pojęcie wartości oczekiwanej dla rozmytej zmiennej losowej jako uogólnienie klasycznej wartości oczekiwanej z wykorzystaniem całki Aumanna [2].

**Definicja 6.5.** *Wartością oczekiwaną* rozmytej zmiennej losowej  $\tilde{\xi}$ , całkowalnej ( $\int_\Omega |\tilde{\xi}_0|^1 d\text{Pr} < \infty$ ) nazywamy zbiór rozmyty  $\mathbb{E}[\tilde{\xi}]$  określony na przestrzeni  $\mathbb{R}^n$  taki, że

$$(\mathbb{E}[\tilde{\xi}])_\alpha = \int_\Omega \tilde{\xi}_\alpha d\text{Pr} \text{ dla każdego } \alpha \in [0, 1], \quad (6.5)$$

gdzie  $\int_\Omega \tilde{\xi}_\alpha d\text{Pr}$  to całka Aumanna funkcji  $\tilde{\xi}_\alpha$  względem miary  $\text{Pr}$ .

W przypadku gdy rozważana jest przestrzeń  $\mathbb{R}$  oraz rozmyta zmienna losowa  $\tilde{\xi}: \Omega \rightarrow \mathcal{F}(\mathbb{R})$  jest całkowalna. Wówczas wartość oczekiwana  $\mathbb{E}[\tilde{\xi}]$  jest wyznaczona jednoznacznie i dla każdego  $\alpha \in [0, 1]$  dana jest poprzez zwarty przedział w przestrzeni  $\mathbb{R}$  [51]. Ponadto, przedziały te są postaci

$$\left[ \mathbb{E}[\inf \tilde{\xi}_\alpha], \mathbb{E}[\sup \tilde{\xi}_\alpha] \right] \text{ dla każdego } \alpha \in [0, 1]. \quad (6.6)$$

Wariancja natomiast określona została podobnie jak dla klasycznych zmiennych losowych jako miara rozrzutu wokół wartości oczekiwanej i została zdefiniowana jako liczba ostra [18].

**Definicja 6.6.** *Wariancją* rozmytej zmiennej losowej  $\tilde{\xi}$  nazywamy liczbę rzeczywistą  $\text{Var}[\tilde{\xi}]$  taką, że

$$\text{Var}[\tilde{\xi}] = \frac{1}{2} \int_0^1 \left[ \text{Var}[\tilde{\xi}_\alpha^-] + \text{Var}[\tilde{\xi}_\alpha^+] \right] d\alpha, \quad (6.7)$$

gdzie  $\tilde{\xi}_\alpha^-$  oraz  $\tilde{\xi}_\alpha^+$  to rzeczywiste zmienne losowe będące odpowiednio lewym i prawym końcem  $\alpha$ -przekroju  $\tilde{\xi}$  dla ustalonego  $\alpha$ .

### 6.1.3. Model Liu i Liu

W obu przytoczonych w poprzednich rozdziałach podejściach, rozmyta zmienna losowa definiowana jest jako mierzalna funkcja z przestrzeni probabilistycznej w zbiór wszystkich liczb rozmytych. W swoich pracach Liu i Liu narzucając odmienny warunek mierzalności funkcji, zaproponowali nową definicję rozmytej zmiennej losowej [46, 47].

Niech  $\mathcal{F}_v$  będzie zbiorem wszystkich *zmiennych rozmytych* (ang. *fuzzy variables* [62]) określonych na *przestrzeni możliwości*  $(\Theta, \mathcal{P}(\Theta), \text{Pos})$  (ang. *possibility space* [101]). Każdy element  $\tilde{a}$  należący do zbioru  $\mathcal{F}_v$  jest określony poprzez funkcję przynależności  $\mu_{\tilde{a}}$ , nazwaną również funkcją rozkładu możliwości. Niech określona będzie przestrzeń probabilistyczna  $(\Omega, \mathcal{A}, \text{Pr})$ .

**Definicja 6.7.** *Rozmytą zmienną losową* nazywamy odwzorowanie  $\tilde{\xi}: \Omega \rightarrow \mathcal{F}_v$  takie, że dla każdego domkniętego podzbioru przestrzeni  $\mathbb{R}$

$$\tilde{\xi}^*(C)(\omega) = \text{Pos}\{\tilde{\xi}(\omega) \in C\} = \sup_{x \in C} \mu_{\tilde{\xi}(\omega)}(x) \quad (6.8)$$

jest funkcją mierzalną  $\omega$ , gdzie  $\mu_{\tilde{\xi}(\omega)}$  jest funkcją przynależności zmiennej rozmytej  $\tilde{\xi}(\omega)$ .

Zmienne rozmyte określone są poprzez funkcję przynależności, której kształt uzależniony jest od przyjętego typu i parametrów. Na przykład dla trójkątnej zmiennej rozmytej, funkcja przynależności uzależniona jest od parametrów  $a, b$  oraz  $c$  (zobacz wzór (2.9)). W rozumieniu Liu i Liu, rozmyta zmienna losowa  $\tilde{\xi}$  jest trójkątną rozmytą zmienną losową, gdy dla każdego  $\omega$ ,  $\tilde{\xi}(\omega)$  jest trójkątną zmienną rozmytą, wyznaczoną przez parametry  $a(\omega), b(\omega), c(\omega)$ , gdzie  $a, b, c$  są zmiennymi losowymi określonymi na  $\Omega$ . Wobec czego, rozmyte zmienne losowe wg. Liu i Liu można rozumieć jako zbiory lub liczby rozmyte, których funkcje przynależności zależą od parametrów będących zmiennymi losowymi.

W odróżnieniu do poprzednich podejść, Liu i Liu określili wartość oczekiwaną oraz wariancję rozmytej zmiennej losowej jako liczbę rzeczywistą, a nie rozmytą. W tym celu, aby wyostrzyć zmienną rozmytą, zaproponowali operator wyostrzania, nazwany operatorem wartości oczekiwanej, w następujący sposób.

**Definicja 6.8.** Operatorem wartości oczekiwanej zmiennej rozmytej  $\tilde{a}$  nazywamy liczbę rzeczywistą  $E(\tilde{a})$  taką, że

$$E(\tilde{a}) = \int_0^{\infty} \text{Cr}\{\tilde{a} \geq r\} dr - \int_{-\infty}^0 \text{Cr}\{\tilde{a} \leq r\} dr, \quad (6.9)$$

pod warunkiem, że co najmniej jedna z całek jest skończona, gdzie  $\text{Cr}\{\tilde{a} \leq r\}$  jest *miarą wiarygodności* (ang. *credibility measure*) taką, że

$$\text{Cr}\{\tilde{a} \leq r\} = 1 - \text{Cr}\{\tilde{a} \geq r\} = \frac{1}{2} \left( \sup_{x \leq r} \mu_{\tilde{a}}(x) + 1 - \sup_{x \geq r} \mu_{\tilde{a}}(x) \right). \quad (6.10)$$

Liu i Liu wykazali, że jeżeli  $\tilde{\xi}$  jest rozmytą zmienną losową, to wartość wyostrzona  $E(\tilde{\xi})$  jest zmienną losową o wartościach rzeczywistych [47]. Fakt ten pozwala na określenie wartości oczekiwanej oraz wariancji rozmytej zmiennej losowej jako wartości oczekiwanej oraz wariancji klasycznej zmiennej losowej, której wartościami są wyostrzone wartości rozmytej zmiennej losowej.

**Definicja 6.9.** Wartością oczekiwaną rozmytej zmiennej losowej  $\tilde{\xi}$  określonej na przestrzeni probabilistycznej  $(\Omega, \mathcal{A}, \text{Pr})$  nazywamy liczbę rzeczywistą  $\mathbb{E}[\tilde{\xi}]$  taką, że

$$\begin{aligned} \mathbb{E}[\tilde{\xi}] &= \int_{\Omega} E(\tilde{\xi}(\omega)) \text{Pr}(d\omega) \\ &= \int_{\Omega} \left[ \int_0^{\infty} \text{Cr}\{\tilde{\xi}(\omega) \geq r\} dr - \int_{-\infty}^0 \text{Cr}\{\tilde{\xi}(\omega) \leq r\} dr \right] \text{Pr}(d\omega) \end{aligned} \quad (6.11)$$

**Definicja 6.10.** Niech  $\tilde{\xi}$  będzie rozmytą zmienną losową o skończonej wartości oczekiwanej  $\mathbb{E}[\tilde{\xi}]$ . Wówczas *wariancją* rozmytej zmiennej losowej  $\tilde{\xi}$  nazywamy liczbę rzeczywistą  $\text{Var}[\tilde{\xi}]$  będącą wartością oczekiwaną rozmytej zmiennej losowej  $(\tilde{\xi} - \mathbb{E}[\tilde{\xi}])^2$ . A więc

$$\text{Var}[\tilde{\xi}] = \mathbb{E}[(\tilde{\xi} - \mathbb{E}[\tilde{\xi}])^2]. \quad (6.12)$$

## 6.2. Skierowane rozmyte zmienne losowe

Z pośród trzech przytoczonych definicji rozmytej zmiennej losowej, najbardziej naturalną i najprostszą w adaptacji dla skierowanych liczb rozmytych wydaje się być definicja zaproponowana przez Kwakernaaka. W definicji tej dla każdego  $\alpha \in [0, 1]$  końce

$\alpha$ -przekrojów utożsamiane są ze zmiennymi losowymi o wartościach rzeczywistych. Przypomnijmy, że każda wypukła liczba rozmyta może być przedstawiona w postaci skierowanej liczby rozmytej (por. rys. 3.1) i wówczas końce  $\alpha$ -przekrojów odpowiadają wartością funkcji  $f(x)$  i  $g(x)$ , dla  $x = \alpha$ . Stąd, skierowana rozmyta zmienna losowa mogłaby być zdefiniowana jako odwzorowanie z przestrzeni probabilistycznej w zbiór wszystkich skierowanych liczb rozmytych, takie że dla każdego  $x \in [0, 1]$ ,  $f(x)$  oraz  $g(x)$  są zmiennymi losowymi o wartościach rzeczywistych. Jednakże, taka definicja jest bardzo ogólna i nie określa żadnych relacji między zmiennymi losowymi dla różnych wartości argumentu  $x$ . Z tego powodu, a także ponieważ skierowana liczba rozmyta to uporządkowana para funkcji ciągłych, w niniejszej rozprawie zaproponowana przez autora została definicja skierowanej rozmytej zmiennej losowej jako uporządkowanej pary ciągłych procesów stochastycznych.

W tym celu, w pierwszej kolejności przypomniane zostaną podstawowe definicje dotyczące procesów stochastycznych [65, 93] .

**Definicja 6.11.** *Procesem stochastycznym*, lub inaczej *procesem losowym*, nazywa się jednoparametrową rodzinę  $\{X_t: t \in \mathcal{T}\}$  zmiennych losowych określonych na wspólnej przestrzeni probabilistycznej  $(\Omega, \mathcal{A}, \Pr)$ , indeksowaną parametrem  $t$  ze zbioru indeksów  $\mathcal{T}$ .

Jeżeli zbiór  $\mathcal{T}$  jest zbiorem liczb naturalnych, to proces nazywany jest *dyskretnym procesem stochastycznym*. Natomiast, jeżeli  $\mathcal{T}$  jest podzbiorem zbioru  $[0, \infty] \subset \mathbb{R}$  to proces nazywany jest *ciągłym procesem stochastycznym*. Każdy proces stochastyczny może być rozważany jako funkcja dwóch argumentów  $t$  i  $\omega$ . W taki sposób, że dla każdego ustalonego  $t \in \mathcal{T}$  funkcja  $\omega \rightarrow X_t(\omega)$  jest zmienna losową. Natomiast, dla każdego ustalonego  $\omega \in \Omega$  funkcja  $t \rightarrow X_t(\omega)$  jest deterministyczną funkcją o wartościach rzeczywistych. Funkcja ta nazywana jest *realizacją* albo *trajektorią* procesu stochastycznego.

**Definicja 6.12.** Proces stochastyczny  $\{X_t: t \in \mathcal{T}\}$  nazywany jest *procesem drugiego rzędu*, jeżeli dla każdego  $t \in \mathcal{T}$  wartość oczekiwana zmiennej losowej  $X_t^2$  istnieje i jest skończona (tj.  $\mathbb{E}[X_t^2] < \infty$ ).

**Definicja 6.13.** *Wartością oczekiwaną* procesu stochastycznego drugiego rzędu  $\{X_t: t \in \mathcal{T}\}$  nazywamy funkcję  $m_X: \mathcal{T} \rightarrow \mathbb{R}$  taką, że dla każdego  $t \in \mathcal{T}$

$$m_x(t) = \mathbb{E}[X_t] = \int_{\Omega} X_t(\omega) \Pr(d\omega). \quad (6.13)$$

**Definicja 6.14.** *Wariancją* procesu stochastycznego drugiego rzędu  $\{X_t: t \in \mathcal{T}\}$  nazywamy funkcję  $var_X: \mathcal{T} \rightarrow \mathbb{R}$  taką, że dla każdego  $t \in \mathcal{T}$

$$var_x(t) = \text{Var}[X_t] = \mathbb{E}[(X_t - m_X(t))^2]. \quad (6.14)$$

Niech określona będzie przestrzeń probabilistyczna  $(\Omega, \mathcal{A}, \Pr)$  oraz niech  $\mathcal{O}\mathcal{F}$  będzie zbiorem wszystkich skierowanych liczb rozmytych.

**Definicja 6.15.** *Skierowaną rozmytą zmienną losową* nazywamy odwzorowanie  $\tilde{\xi}: \Omega \rightarrow \mathcal{O}\mathcal{F}$  takie, że  $\tilde{\xi}(\omega) = (f_{\tilde{\xi}}(t, \omega), g_{\tilde{\xi}}(t, \omega))$  jest uporządkowaną parą ciągłych procesów stochastycznych, gdzie  $f_{\tilde{\xi}}, g_{\tilde{\xi}}: \mathcal{T} \times \Omega \rightarrow \mathbb{R}$ ,  $\mathcal{T} = [0, 1] \subset \mathbb{R}$ . Dodatkowo zakładamy, że procesy te są procesami drugiego rzędu.

Jak można zauważyć, dla każdego ustalonego  $t \in [0, 1]$  funkcja  $t \rightarrow (f_{\tilde{\xi}}(t, \omega), g_{\tilde{\xi}}(t, \omega))$  jest zmienną losową o wartościach w  $\mathbb{R}^2$  (wektor losowy), natomiast dla każdego ustalonego  $\omega \in \Omega$  wartości funkcji  $\omega \rightarrow (f_{\tilde{\xi}}(t, \omega), g_{\tilde{\xi}}(t, \omega))$  to skierowane liczby rozmyte, których funkcje  $f$  oraz  $g$  wyznaczone są poprzez trajektorie procesów stochastycznych.

Dla tak zdefiniowanej skierowanej rozmytej zmiennej losowej, pojęcia takie jak wartość oczekiwana oraz wariancja mogą zostać zdefiniowane w następujący sposób.

**Definicja 6.16.** *Rozmytą wartośćią oczekiwaną* skierowanej rozmytej zmiennej losowej  $\tilde{\xi}: \Omega \rightarrow \mathcal{O}\mathcal{F}$  nazywamy skierowaną liczbę rozmytą  $\mathbb{E}[\tilde{\xi}]$  taką, że

$$\mathbb{E}[\tilde{\xi}] = (f_{m_{\tilde{\xi}}}, g_{m_{\tilde{\xi}}}), \quad (6.15)$$

gdzie funkcje  $f_{m_{\tilde{\xi}}}$  oraz  $g_{m_{\tilde{\xi}}}$  to wartości oczekiwane procesów stochastycznych  $f_{\tilde{\xi}}$  i  $g_{\tilde{\xi}}$ .

**Definicja 6.17.** *Rozmytą wariancją* skierowanej rozmytej zmiennej losowej  $\tilde{\xi}: \Omega \rightarrow \mathcal{O}\mathcal{F}$  nazywamy skierowaną liczbę rozmytą  $\text{Var}[\tilde{\xi}]$  taką, że

$$\text{Var}[\tilde{\xi}] = (f_{var_{\tilde{\xi}}}, g_{var_{\tilde{\xi}}}), \quad (6.16)$$

gdzie funkcje  $f_{var_{\tilde{\xi}}}$  oraz  $g_{var_{\tilde{\xi}}}$  to wariancje procesów stochastycznych  $f_{\tilde{\xi}}$  i  $g_{\tilde{\xi}}$ .

Warto zaznaczyć, że dzięki takim definicjom, utrzymane zostają podstawowe własności wartości oczekiwanej oraz wariancji.

**Twierdzenie 6.1.** *Zakładając, że istnieją rozmyte wartości oczekiwane  $\mathbb{E}[\tilde{\xi}]$  oraz  $\mathbb{E}[\tilde{\eta}]$  skierowanych rozmytych zmiennych losowych  $\tilde{\xi}$  i  $\tilde{\eta}$ . Wtedy*

(i) *Jeżeli  $\tilde{\xi} \geq 0$ , to  $\mathbb{E}[\tilde{\xi}] \geq 0$ .*

(ii)  $|\mathbb{E}[\tilde{\xi}]| \leq \mathbb{E}[|\tilde{\xi}|]$ .

(iii) *Dla  $a, b \in \mathbb{R}$  istnieje rozmyta wartość oczekiwana  $a\tilde{\xi} + b\tilde{\eta}$  i*

$$\mathbb{E}[a\tilde{\xi} + b\tilde{\eta}] = a\mathbb{E}[\tilde{\xi}] + b\mathbb{E}[\tilde{\eta}].$$

**Twierdzenie 6.2.** *Jeżeli  $\tilde{\xi}$  jest skierowaną rozmytą zmienną losową, dla której istnieje rozmyta wariancja  $\text{Var}[\tilde{\xi}]$ . Wtedy*

(i)  $\text{Var}[\tilde{\xi}] \geq 0$ .

(ii) *Dla  $c \in \mathbb{R}$   $\text{Var}[c\tilde{\xi}] = c^2 \text{Var}[\tilde{\xi}]$ .*

(iii) *Dla  $c \in \mathbb{R}$   $\text{Var}[\tilde{\xi} + c] = \text{Var}[\tilde{\xi}]$ .*

Dowody powyższych własności wynikają wprost z własności wartości oczekiwanej oraz wariancji dla klasycznych zmiennych losowych o wartościach rzeczywistych [24].

Jak zauważyli Liu i Liu, w praktycznych zagadnieniach, częściej wymagane jest pojęcie wartości oczekiwanej jako liczby ostrej, a nie rozmytej, tak aby możliwe było wskazanie wartości, na podstawie której podejmowane będą różnego rodzaju decyzje. Podobnie jak w podejściu zaproponowanym przez Liu i Liu, wprowadzone zostaną definicje ostrej wartości oczekiwanej i wariancji dla skierowanych rozmytych zmiennych losowych z wykorzystaniem operatora wyostrzania, określanego jak operator wartości oczekiwanej.

**Definicja 6.18.** *Operatorem wartości oczekiwanej dla skierowanej liczby rozmytej  $A = (f_A, g_A)$  nazywamy operator wyostrzania  $E: \mathcal{OF} \rightarrow \mathbb{R}$  określony wzorem*

$$E(A) = E(f_A, g_A) = \frac{1}{2} \int_0^1 [f_A(s) + g_A(s)] ds. \quad (6.17)$$

Operator wartości oczekiwanej zdefiniowany powyżej jest równoważny operatorowi wartości oczekiwanej zdefiniowanemu przez Liu i Liu (zobacz def. 6.8). Równoważność tą określa poniższe twierdzenie.

**Twierdzenie 6.3.** *Niech  $A = (f, g)$  będzie skierowaną liczbą rozmytą taką, że istnieją funkcje odwrotne  $f^{-1}$  oraz  $g^{-1}$  oraz  $f(0) \leq f(1) \leq g(1) \leq g(0)$ . Co oznacza, że skierowaną liczbą rozmytą  $A$  można przedstawić jako klasyczną liczbę rozmytą  $A^*$  o funkcji przynależności  $\mu_{A^*}: \mathbb{R} \rightarrow [0, 1]$  postaci*

$$\mu_{A^*}(x) = \begin{cases} f^{-1}(x) & \text{dla } x \in [f(0), f(1)], \\ g^{-1}(x) & \text{dla } x \in [g(1), g(0)], \\ 1 & \text{dla } x \in [f(1), g(1)], \\ 0 & \text{w p.p.} \end{cases} \quad (6.18)$$

Wówczas zachodzi równość

$$E(A^*) = E(A), \quad (6.19)$$

gdzie  $E$  jest operatorem wartości oczekiwanej według definicji 6.8.

*Dowód.* Zgodnie z wzorem (6.10) miara wiarygodności dla liczby  $A^*$  wyraża się wzorem

$$\text{Cr}\{A^* \geq t\} = \begin{cases} 1 & \text{dla } t < f(0), \\ 1 - \frac{1}{2}f^{-1}(t) & \text{dla } t \in [f(0), f(1)], \\ \frac{1}{2} & \text{dla } t \in [f(1), g(1)], \\ \frac{1}{2}g^{-1}(t) & \text{dla } t \in [g(1), g(0)], \\ 0 & \text{dla } t > g(0). \end{cases} \quad (6.20)$$

Z definicji 6.8 operator wartości oczekiwanej liczby rozmytej  $A^*$  wyraża się wzorem

$$E(A^*) = \int_0^{\infty} \text{Cr}\{A^* \geq t\} dt - \int_{-\infty}^0 \text{Cr}\{A^* \leq t\} dt. \quad (6.21)$$

Stosując wzór (6.10) oraz korzystając z własności całki otrzymujemy

$$E(A^*) = \int_{-\infty}^{\infty} \text{Cr}\{A^* \geq t\} dt - \int_{-\infty}^0 1 dt. \quad (6.22)$$

Wstawiając wzór (6.20) do równania (6.22), korzystając z liniowości całki, a następnie grupując odpowiednie całki otrzymujemy

$$E(A^*) = I_1 + I_2, \text{ gdzie}$$

$$I_1 = \int_{-\infty}^{f(0)} 1 \, dt + \int_{f(0)}^{f(1)} 1 \, dt + \frac{1}{2} \int_{f(1)}^{g(1)} 1 \, dt - \int_{-\infty}^0 1 \, dt, \quad (6.23)$$

$$I_2 = -\frac{1}{2} \left[ \int_{f(0)}^{f(1)} f^{-1}(t) \, dt + \int_{g(0)}^{g(1)} g^{-1}(t) \, dt \right]. \quad (6.24)$$

Wyrażenie  $I_1$ , niezależnie od położenia punktu 0 względem punktów  $f(0)$ ,  $f(1)$  oraz  $g(1)$  wynosi

$$I_1 = \frac{1}{2}[f(1) + g(1)]. \quad (6.25)$$

Wyrażenie  $I_2$ , poprzez podstawienia w odpowiednich całkach  $s = f^{-1}(t)$  oraz  $s = g^{-1}(t)$ , a następnie stosując całkowanie przez części do obu całek, sprowadza się do wyrażenia

$$I_2 = -\frac{1}{2}[f(1) + g(1)] + \frac{1}{2} \int_0^1 [f(s) + g(s)] \, ds. \quad (6.26)$$

Ostatecznie sumując wyrażenia  $I_1$  oraz  $I_2$  otrzymujemy prawą stronę równania (6.19).

□

**Definicja 6.19.** *Ostrą wartością oczekiwaną skierowanej rozmytej zmiennej losowej  $\tilde{\xi}: \Omega \rightarrow \mathcal{OF}$ , dla której istnieje rozmyta wartość oczekiwana  $\mathbb{E}[\tilde{\xi}]$ , nazywamy liczbę rzeczywistą  $\mathcal{E}(\tilde{\xi})$  taką, że*

$$\mathcal{E}(\tilde{\xi}) = \mathbb{E}(\mathbb{E}[\tilde{\xi}]) = \frac{1}{2} \int_0^1 [f_{m_{\tilde{\xi}}}(t) + g_{m_{\tilde{\xi}}}(t)] \, dt. \quad (6.27)$$

Ostrą wartość oczekiwaną skierowanej rozmytej zmiennej losowej można obliczyć na dwa sposoby. Wprost z definicji, tzn. obliczyć rozmytą wartość oczekiwaną, a następnie wyostrzyć ją operatorem wartości oczekiwanej. Można, też zrobić odwrotnie, tzn. dla ustalonych  $\omega$  obliczyć wartości wyostrzone skierowanych liczb rozmytych  $\tilde{\xi}(\omega)$ . Wówczas  $\mathbb{E}(\tilde{\xi}(\omega))$  jest zmienną losową o wartościach rzeczywistych, a więc można policzyć jej wartość oczekiwaną w klasycznym sensie. O równości obu metod mówi poniższe twierdzenie.



**Twierdzenie 6.4.** Niech  $\tilde{\xi}$  będzie skierowaną rozmytą zmienną losową określoną na przestrzeni probabilistycznej  $(\Omega, \mathcal{A}, \Pr)$ . Wówczas, jeżeli istnieje  $\mathbb{E}[\tilde{\xi}]$ , to

$$\mathbb{E}(\mathbb{E}[\tilde{\xi}]) = \mathbb{E}[\mathbb{E}(\tilde{\xi})], \quad (6.28)$$

gdzie  $\mathbb{E}$  z prawej strony równania (6.28) oznacza wartość oczekiwaną klasycznej zmiennej losowej.

*Dowód.* Zgodnie z wzorem (6.27), korzystając z wzoru (6.13), z własności całki oraz z tw. Fubiniiego o całkach iterowanych otrzymujemy

$$\begin{aligned} \mathbb{E}(\mathbb{E}[\tilde{\xi}]) &= \frac{1}{2} \int_0^1 [f_{m_{\tilde{\xi}}}(t) + g_{m_{\tilde{\xi}}}(t)] dt = \int_0^1 \frac{1}{2} f_{m_{\tilde{\xi}}}(t) dt + \int_0^1 \frac{1}{2} g_{m_{\tilde{\xi}}}(t) dt \\ &= \int_0^1 \mathbb{E} \left[ \frac{1}{2} f_{\tilde{\xi}}(t) \right] dt + \int_0^1 \mathbb{E} \left[ \frac{1}{2} g_{\tilde{\xi}}(t) \right] dt \\ &= \int_0^1 \left[ \int_{\Omega} \frac{1}{2} f_{\tilde{\xi}}(t, \omega) \Pr(d\omega) \right] dt + \int_0^1 \left[ \int_{\Omega} \frac{1}{2} g_{\tilde{\xi}}(t, \omega) \Pr(d\omega) \right] dt \\ &= \int_{\Omega} \left[ \frac{1}{2} \int_0^1 f_{\tilde{\xi}}(t, \omega) dt + \frac{1}{2} \int_0^1 g_{\tilde{\xi}}(t, \omega) dt \right] \Pr(d\omega) \\ &= \mathbb{E} \left[ \frac{1}{2} \int_0^1 [f_{\tilde{\xi}}(t) + g_{\tilde{\xi}}(t)] dt \right] = \mathbb{E}[\mathbb{E}(\tilde{\xi})]. \end{aligned}$$

□

Ostra wariancja skierowanej rozmytej zmiennej losowej definiowana jest natomiast jako wariancja klasycznej zmiennej losowej powstałej poprzez wyostrzenie skierowanej rozmytej zmiennej losowej. Jest to więc średni kwadrat odchylenia wartości wyostrzonych od ostrej wartości średniej.

**Definicja 6.20.** Ostrą wariancją skierowanej rozmytej zmiennej losowej  $\tilde{\xi}: \Omega \rightarrow \mathcal{OF}$ , dla której istnieje  $\mathcal{E}[\tilde{\xi}]$ , nazywamy liczbę rzeczywistą  $\mathcal{V}(\tilde{\xi})$  taką, że

$$\mathcal{V}(\tilde{\xi}) = \text{Var}[\mathbb{E}(\tilde{\xi})] = \mathbb{E} \left[ \left( \mathbb{E}(\tilde{\xi}) - \mathcal{E}(\tilde{\xi}) \right)^2 \right], \quad (6.29)$$

gdzie  $\mathbb{E}$  i  $\text{Var}$  oznacza wartość oczekiwaną oraz wariancję w klasycznym sensie.

### 6.3. Skierowane rozmyte liczby pseudolosowe

Z punktu widzenia zagadnień symulacji, najważniejszą kwestią jest umiejętność generowania liczb losowych o zadanych rozkładach prawdopodobieństwa. W praktyce, liczby losowe generowane są za pomocą algorytmów deterministycznych, które jedynie wyglądają jak losowe, stąd nazywane są liczbami pseudolosowymi, a algorytmy generatorami liczb pseudolosowych. Tak jak pojęcie rozmytej zmiennej losowej definiowane jest poprzez rozważanie odpowiednio zdefiniowanych klasycznych zmiennych losowych, tak definiowanie rozkładu prawdopodobieństwa rozmytej zmiennej losowej sprowadza się do określenia rozkładów tych właśnie zmiennych losowych [71, 95, 20, 9, 10].

Na tej samej zasadzie, zaproponowana zostanie definicja normalnego rozkładu prawdopodobieństwa dla skierowanej rozmytej zmiennej losowej. Jednakże, w odróżnieniu od klasycznych liczb rozmytych, skierowane liczby rozmyte posiadają dodatkową własność, a mianowicie orientację, którą również należy wziąć pod uwagę.

Niech  $\tilde{\xi}$  będzie skierowaną rozmytą zmienną losową określoną na przestrzeni probabilistycznej  $(\Omega, \mathcal{A}, \Pr)$ .

**Definicja 6.21.** Powiemy, że  $\tilde{\xi}$  ma rozkład normalny o parametrach  $\tilde{\mu}, \tilde{\sigma}^2 \in \mathcal{OF}$ ,  $\sigma^2 \in \mathbb{R}$  oraz  $p \in [0, 1]$  takich, że  $\text{Ord}(\tilde{\mu}) \geq 0$ ,  $\tilde{\sigma}^2 + \sigma^2 > 0$ , co zapisujemy  $\tilde{\xi} \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2, \sigma^2, p)$ , jeżeli dla każdego  $t \in [0, 1]$  zmienne losowe  $f_{\tilde{\xi}}(t)$  oraz  $g_{\tilde{\xi}}(t)$  mają następujące mieszane rozkłady normalne, odpowiednio

$$\begin{aligned} f_{\tilde{\xi}}(t) &\sim p \cdot \mathcal{N}(f_{\tilde{\mu}}(t), f_{\tilde{\sigma}^2}(t) + \sigma^2) + (1 - p) \cdot \mathcal{N}(g_{\tilde{\mu}}(t), g_{\tilde{\sigma}^2}(t) + \sigma^2) \\ g_{\tilde{\xi}}(t) &\sim p \cdot \mathcal{N}(g_{\tilde{\mu}}(t), g_{\tilde{\sigma}^2}(t) + \sigma^2) + (1 - p) \cdot \mathcal{N}(f_{\tilde{\mu}}(t), f_{\tilde{\sigma}^2}(t) + \sigma^2). \end{aligned}$$

Ponadto,  $\Pr(\text{Ord}(\tilde{\xi}) \geq 0) = p$  oraz  $\Pr(\text{Ord}(\tilde{\xi}) < 0) = 1 - p$ , gdzie  $\text{Ord}$  jest operatorem skierowania zdefiniowanym z wykorzystaniem operatora wartości oczekiwanej (zobacz wzór (3.14)).

Warto zaznaczyć, że w odróżnieniu od klasycznego rozkładu normalnego, parametry  $\tilde{\mu}, \tilde{\sigma}^2$  nie są rozmytą wartością oczekiwaną oraz wariancją  $\tilde{\xi}$ . Rozmyta wartość oczekiwana i wariancja  $\tilde{\xi}$  wyrażają się wzorami

$$\mathbb{E}[\tilde{\xi}] = p \cdot (f_{\tilde{\mu}}, g_{\tilde{\mu}}) + (1 - p) \cdot (g_{\tilde{\mu}}, f_{\tilde{\mu}}). \quad (6.30)$$

$$\text{Var}[\tilde{\xi}] = p(1 - p) \cdot ((f_{\tilde{\mu}}, g_{\tilde{\mu}}) - (g_{\tilde{\mu}}, f_{\tilde{\mu}}))^2 + p \cdot (f_{\tilde{\sigma}^2}, g_{\tilde{\sigma}^2}) + (1 - p) \cdot (g_{\tilde{\sigma}^2}, f_{\tilde{\sigma}^2}) + \sigma^2. \quad (6.31)$$

Ponadto, estymatory parametrów dla tak zdefiniowanego rozkładu określone są w następujący sposób. Niech dana będzie próba losowa skierowanej rozmytej zmiennej losowej  $\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_n$  z rozkładu normalnego o parametrach  $\tilde{\mu}, \tilde{\sigma}^2, \sigma^2$  oraz  $p$ . Wówczas:

- Estymatorem rozmytej wartości oczekiwanej jest:

$$\hat{\mathbb{E}}[\tilde{\xi}] = \frac{1}{n} \sum_{i=1}^n \tilde{\xi}_i. \quad (6.32)$$

- Estymatorem rozmytej wariancji jest:

$$\hat{\text{Var}}[\tilde{\xi}] = \frac{1}{n-1} \sum_{i=1}^n (\tilde{\xi}_i - \hat{\mathbb{E}}[\tilde{\xi}])^2. \quad (6.33)$$

- Estymatorem parametru  $\sigma^2$  jest:

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (\mathbb{E}(\tilde{\xi}_i) - \mathbb{E}(\hat{\mathbb{E}}[\tilde{\xi}]))^2. \quad (6.34)$$

- Estymatorem prawdopodobieństwa skierowania  $p$  jest:

$$\hat{p} = \frac{\#\{i \in \{1, 2, \dots, n\} : \text{Ord}(\tilde{\xi}_i) \geq 0\}}{n}. \quad (6.35)$$

- Estymatorami parametrów  $\tilde{\mu}$  oraz  $\tilde{\sigma}^2$  są:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \tilde{\xi}_i^*, \quad (6.36)$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (\tilde{\xi}_i^* - \hat{\mu})^2 - \hat{\sigma}^2, \quad (6.37)$$

gdzie

$$\tilde{\xi}_i^* = \begin{cases} (f_{\tilde{\xi}_i}, g_{\tilde{\xi}_i}), & \text{gdy } \text{Ord}(\tilde{\xi}_i) \geq 0, \\ (g_{\tilde{\xi}_i}, f_{\tilde{\xi}_i}), & \text{gdy } \text{Ord}(\tilde{\xi}_i) < 0. \end{cases} \quad (6.38)$$

Poniżej przedstawione zostaną przykładowe metody generowania skierowanych rozmytych zmiennych losowych o rozkładzie normalnym. Metody te sprowadzają się do zaproponowania odpowiednich procesów stochastycznych dla części  $f$  i  $g$  skierowanej rozmytej zmiennej losowej. Jednym z najprostszych rozwiązań jest zaproponowanie dla skierowanej rozmytej zmiennej losowej  $\tilde{\xi}$  procesów stochastycznych postaci

$$f_{\tilde{\xi}}(t) = g_{\tilde{\xi}}(t) = X, \quad (6.39)$$

dla każdego  $t \in [0, 1]$ , gdzie  $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$  jest klasyczną zmienną losową. Wówczas, tak wygenerowana skierowana rozmyta zmienna losowa ma rozkład normalny o parametrach  $\tilde{\mu} = \mu_X$ ,  $\tilde{\sigma}^2 = 0$ ,  $\sigma^2 = \sigma_X^2$  oraz  $p = 1$ . Może być ona utożsamiona z klasyczną zmienną losową o rozkładzie normalnym.

W celu skonstruowania skierowanej rozmytej zmiennej losowej o rozkładzie normalnym z parametrami rozmytymi, wykorzystane zostaną następujące własności rozkładu normalnego [52].

**Twierdzenie 6.5.** *Niech  $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ ,  $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ ,  $\alpha, \beta \in \mathbb{R}$ ,  $\alpha \neq 0$ , to wówczas*

$$X + Y \sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2), \quad (6.40)$$

$$\alpha X + \beta \sim \mathcal{N}(\alpha\mu_X + \beta, \alpha^2\sigma_X^2). \quad (6.41)$$

Niech ustalone będą parametry  $\tilde{\mu}, \tilde{\sigma}^2 \in \mathcal{OF}$ ,  $\sigma^2 \in \mathbb{R}_+$  takie, że  $\tilde{\mu} \neq 0$  oraz  $\tilde{\sigma}^2 > 0$  (dla każdego  $t \in [0, 1]$ ). Ponadto, niech  $\tilde{\eta}$  oraz  $\tilde{\zeta}$  będą skierowanymi rozmytymi zmiennymi losowymi o procesach stochastycznych  $f_{\tilde{\eta}}, g_{\tilde{\eta}}$  oraz  $f_{\tilde{\zeta}}, g_{\tilde{\zeta}}$  określonych w następujący sposób.

$$f_{\tilde{\eta}}(t) = X_t \quad \wedge \quad g_{\tilde{\eta}}(t) = Y_t, \quad (6.42)$$

$$f_{\tilde{\zeta}}(t) = S \quad \wedge \quad g_{\tilde{\zeta}}(t) = S, \quad (6.43)$$

gdzie  $X_t \sim \mathcal{N}\left(1, \frac{f_{\tilde{\sigma}^2}(t)}{f_{\tilde{\mu}}^2(t)}\right)$ ,  $Y_t \sim \mathcal{N}\left(1, \frac{g_{\tilde{\sigma}^2}(t)}{g_{\tilde{\mu}}^2(t)}\right)$  dla  $t \in [0, 1]$  oraz  $S \sim \mathcal{N}(0, \sigma^2)$ . Wówczas, korzystając z twierdzenia 6.5 skierowana rozmyta zmienna losowa  $\tilde{\xi} = \tilde{\mu} \cdot \tilde{\eta} + \tilde{\zeta}$  posiada rozkład normalny o parametrach  $\tilde{\mu}_{\tilde{\xi}} = \tilde{\mu}$ ,  $\tilde{\sigma}_{\tilde{\xi}}^2 = \tilde{\sigma}^2$  oraz  $\sigma_{\tilde{\xi}}^2 = \sigma^2$ .

Wykorzystując powyższe rozważania, sformułowana zostanie procedura generowania skierowanych liczb pseudolosowych o rozkładzie normalnym z zadanymi parametrami  $\tilde{\mu}, \tilde{\sigma}^2 \in \mathcal{OF}$ ,  $\sigma^2 \in \mathbb{R}_+$  oraz  $p \in [0, 1]$ . Na potrzeby obliczeń numerycznych, założono, że parametry te określone są w sposób dyskretny na przedziale  $[0, 1]$  za pomocą  $T + 1$  punktów,  $t \in \{0, 1, \dots, T\}$  (por. rys. 3.5).

### Algorytm 1:

**Krok 1.** Skonstruuj skierowaną liczbę rozmytą  $\tilde{\eta} := \tilde{\mu} + C$ , gdzie  $C$  jest dodatnią liczbą rzeczywistą taką, aby  $\tilde{\eta} > 0$  dla każdego  $t \in \{0, 1, \dots, T\}$ .

**Krok 2.** Wygeneruj dwa ciągi liczb losowych  $X_t$  oraz  $Y_t$  dla  $t \in \{0, 1, \dots, T\}$  o następujących rozkładach normalnych

$$X_t \sim \mathcal{N}\left(1, \frac{f_{\tilde{\sigma}^2}(t)}{f_{\tilde{\eta}}^2(t)}\right), \quad (6.44)$$

$$Y_t \sim \mathcal{N}\left(1, \frac{g_{\tilde{\sigma}^2}(t)}{g_{\tilde{\eta}}^2(t)}\right), \quad (6.45)$$

oraz zmienną losową  $S \sim \mathcal{N}(0, \sigma^2)$ .

**Krok 3.** Wygeneruj liczbę losową  $r$  z rozkładu jednostajnego na przedziale  $[0, 1]$ . Jeżeli  $r < p$  to wykonujemy krok 4, w przeciwnym kroku 4'.

**Krok 4.** Skonstruuj skierowaną liczbę rozmytą  $\tilde{\xi}$  w taki sposób, że

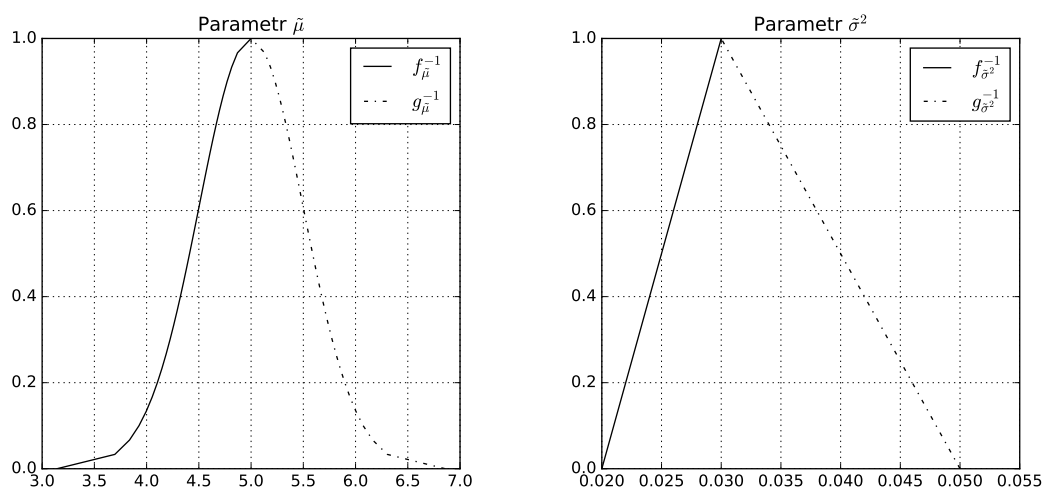
$$(f_{\tilde{\xi}}(t), g_{\tilde{\xi}}(t)) = (X_t, Y_t) \cdot (f_{\tilde{\eta}}(t), g_{\tilde{\eta}}(t)) + S. \quad (6.46)$$

**Krok 4'.** Skonstruuj skierowaną liczbę rozmytą  $\tilde{\xi}$  w taki sposób, że

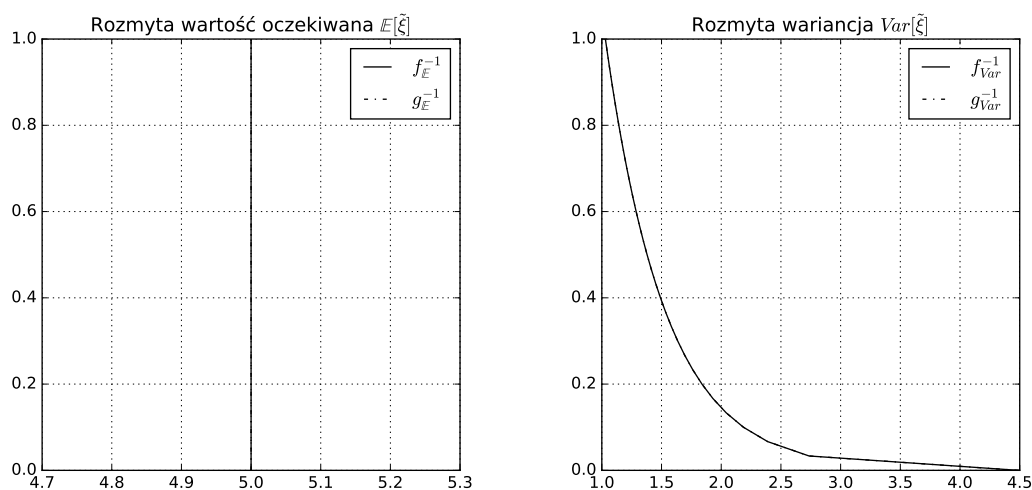
$$(f_{\tilde{\xi}}(t), g_{\tilde{\xi}}(t)) = (X_t, Y_t) \cdot (g_{\tilde{\eta}}(t), f_{\tilde{\eta}}(t)) + S. \quad (6.47)$$

**Krok 5.** Od  $\tilde{\xi}$  z kroku 4 lub 4' odejmij stałą  $C$  określoną w kroku 1. Tak określona skierowana liczba rozmyta jest skierowaną liczbą pseudolosową o rozkładzie normalnym z zadanymi uprzednio parametrami  $\tilde{\mu}, \tilde{\sigma}^2, \sigma^2$  oraz  $p$ .

**Przykład:** Niech  $\tilde{\xi}$  będzie skierowaną rozmytą zmienną losową o rozkładzie normalnym z parametrami  $p = 0.5, \sigma^2 = 1$  oraz  $\tilde{\mu}$  i  $\tilde{\sigma}^2$  określonymi tak jak na rys. 6.2. Parametr  $\tilde{\mu}$  określony jest jako gaussowska skierowana liczba rozmyta z parametrami  $\mu_f = \mu_g = 5$  oraz  $\sigma_f = -\sigma_g = -0.5$ . Parametr  $\tilde{\sigma}^2$  to trójkątna skierowana liczba rozmyta o parametrach  $f(0) = 0.02, f(1) = g(1) = 0.03, g(0) = 0.05$ . Na rys. 6.3 przedstawione zostały skierowane liczby rozmyte będące rozmytą wartością oczekiwaną oraz rozmytą wariancją zmiennej  $\tilde{\xi}$  zgodnie z wzorami (6.30) i (6.31).

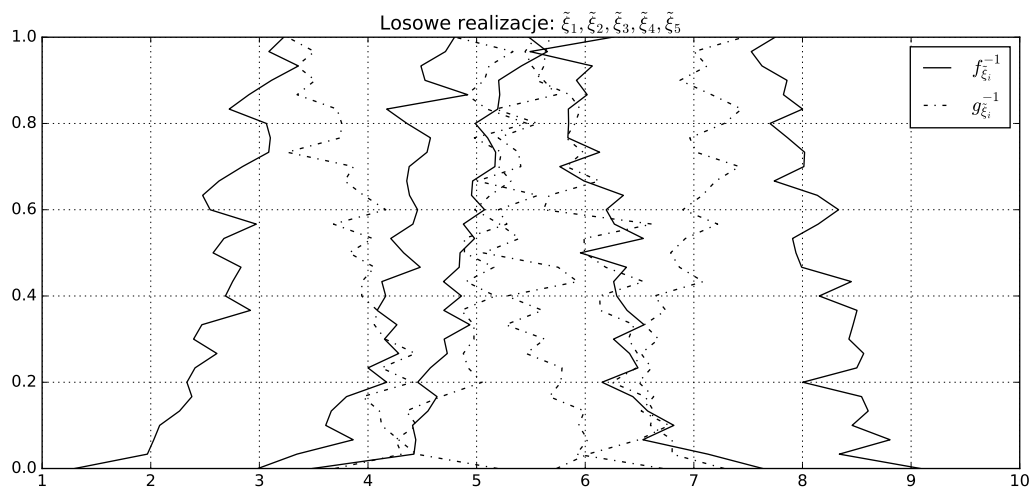


Rysunek 6.2: Rozmyte parametry  $\tilde{\mu}$  oraz  $\tilde{\sigma}^2$  rozkładu normalnego.

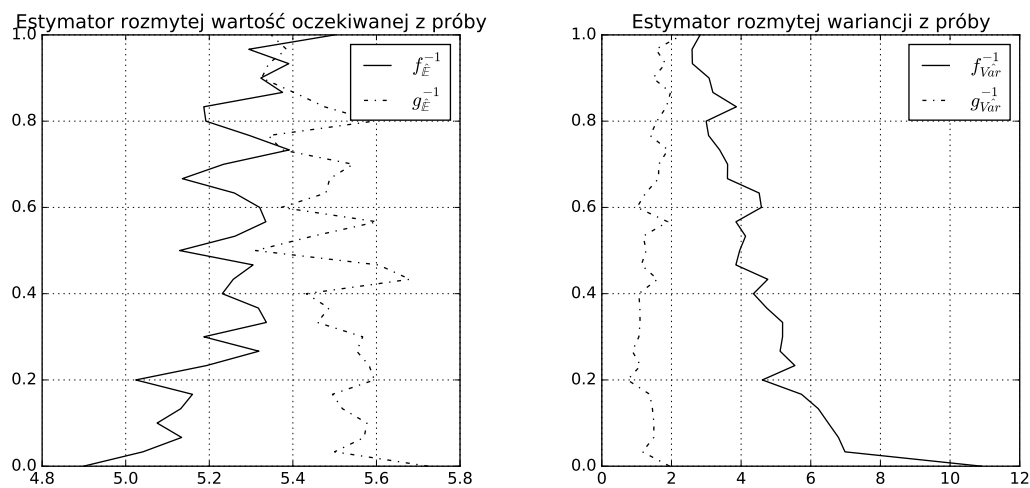


Rysunek 6.3: Rozmyta wartość oczekiwana i rozmyta wariancja zmiennej  $\tilde{\xi}$ .

Rys. 6.4 przedstawia losową próbę pięciu realizacji skierowanej rozmytej zmiennej  $\tilde{\xi}$  wygenerowanych zgodnie z zaproponowanym algorytmem 1. Z kolei, na rys. 6.5 przedstawiona jest rozmyta wartość oczekiwana oraz rozmyta wariancja z próby obliczona według estymatorów (6.33) i (6.34).

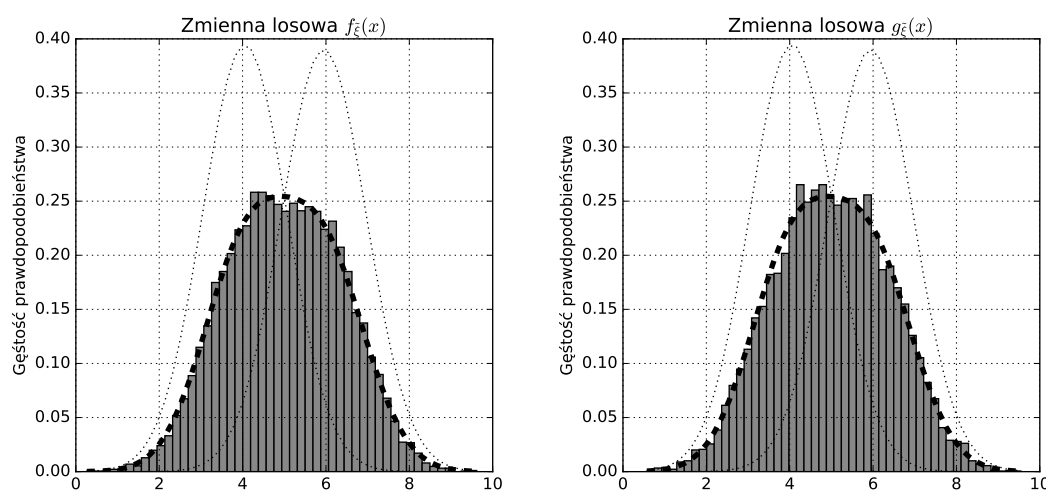


Rysunek 6.4: Pięć losowych realizacji zmiennej  $\tilde{\xi}$ .



Rysunek 6.5: Estymatory rozmytej wartości oczekiwanej i rozmytej wariancji z próby.

Na rys. 6.6 przedstawiono znormalizowane histogramy oraz teoretyczne funkcje gęstości prawdopodobieństwa zmiennych losowych  $f_{\xi_i}(x)$ ,  $g_{\xi_i}(x)$  dla  $x = 0.5$  i  $n = 10000$ . W tabeli 6.1 przedstawiono średnie odległości między teoretycznymi wartościami parametrów  $\mathbb{E}$ ,  $\text{Var}$ ,  $\tilde{\mu}$ ,  $\tilde{\sigma}^2$  oraz  $p$ , a ich estymatorami w zależności od rozmiaru próby (10 symulacji). Odległość między liczbami rzeczywistymi obliczona została jako wartość bezwzględna z różnicy, natomiast między skierowanymi liczbami rozmytymi za pomocą metryki  $d$  danej wzorem (5.16). Otrzymane wyniki pozwalają przypuszczać, że zaproponowany algorytm generowania skierowanych rozmytych liczb pseudolosowych działa poprawnie, a przyjęte estymatory parametrów rozkładu normalnego są estymatorami zgodnymi.



Rysunek 6.6: Znormalizowane histogramy oraz teoretyczne funkcje gęstości prawdopodobieństwa zmiennych losowych  $f_{\xi_i}(x)$ ,  $g_{\xi_i}(x)$  dla  $x = 0.5$ .

$n$	5	10	50	100	1000	10000	100000
$ p - \hat{p} $	0.1800	0.1000	0.0540	0.0430	0.0127	0.0036	0.0014
$d(\mathbb{E}, \hat{\mathbb{E}})$	0.8882	0.7049	0.2684	0.1919	0.0638	0.0201	0.0059
$d(\text{Var}, \hat{\text{Var}})$	0.3905	0.2182	0.1028	0.0695	0.0235	0.0092	0.0078
$ \sigma^2 - \hat{\sigma}^2 $	0.6188	0.3136	0.1841	0.1336	0.0282	0.0134	0.0033
$d(\tilde{\mu}, \hat{\tilde{\mu}})$	1.0572	0.7902	0.2585	0.2352	0.0793	0.0253	0.0069
$d(\tilde{\sigma}^2, \hat{\tilde{\sigma}}^2)$	2.1067	1.1534	0.5697	0.3664	0.1016	0.0376	0.0103

Tabela 6.1: Średnie odległości między teoretycznymi wartościami parametrów  $\mathbb{E}$ ,  $\text{Var}$ ,  $\tilde{\mu}$ ,  $\tilde{\sigma}^2$  oraz  $p$ , a ich estymatorami w zależności od rozmiaru próby.



# Rozdział 7

## Przykłady zastosowań

W niniejszym rozdziale przedstawione zostanie kilka przykładów ilustrujących możliwości wykorzystania zaproponowanych w rozprawie modeli i metod w praktycznych zagadnieniach związanych z rynkiem finansowym.

### 7.1. Wycena opcji europejskiej na indeks WIG20

Pierwszy z przykładów przedstawia wykorzystanie skierowanego rozmytego modelu autoregresyjnego do wyceny europejskiej opcji kupna na indeks WIG20, notowanej na Giełdzie Papierów Wartościowych w Warszawie. Wycena opcji przeprowadzona została zgodnie z podejściem nazywanym *wyceną dla miary lokalnie wolnej od ryzyka* (ang. *Local Risk-Neutral Valuation Relationship - LRNVR*). W podejściu tym wprowadza się miarę prawdopodobieństw  $\mathbf{P}$ , dla procesu nieprzekształconego oraz miarę  $\mathbf{Q}$ , względem której zdyskontowany proces cen instrumentu finansowego jest martyngałem. Podejście to do wyceny instrumentów pochodnych przy użyciu modelu klasy GARCH zaproponował i opisał Duan w [13].

### 7.1.1. Dopasowanie modelu SRAR(p) do notowań

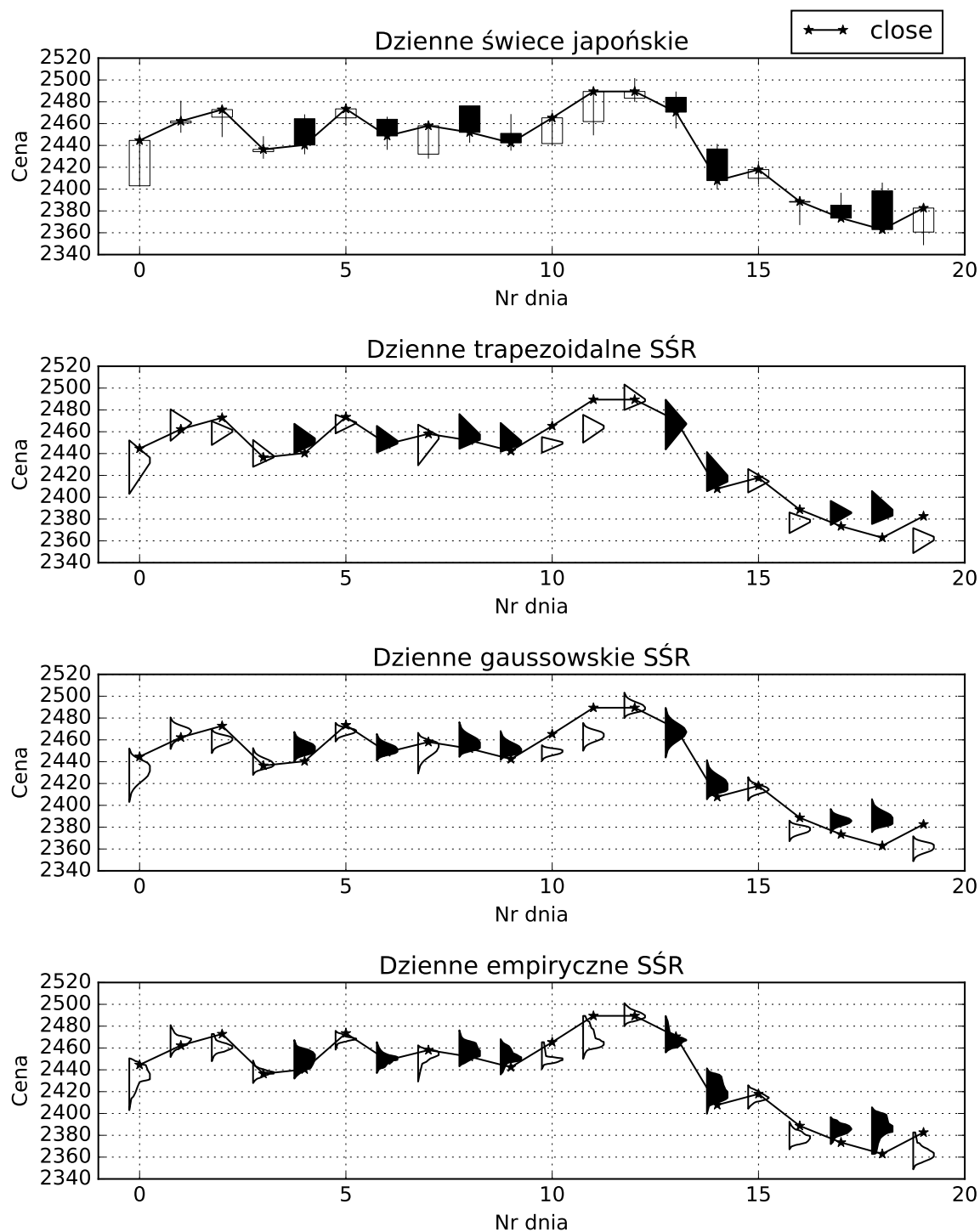
W opisywanym przykładzie wycenione zostaną europejskie opcje kupna na index WIG20 o terminie wygaśnięcia 17.03.2017 dla różnych cen wykonania. Wyceny wszystkich opcji wykonane zostaną na koniec roku 2016 (30.12.2016). Do skonstruowania modeli wykorzystano notowania indeksu WIG20 z okresu od 04.01.2010 do 30.12.2016 o częstotliwości minutowej, co daje 868645 obserwacji w ciągu 1751 dni sesyjnych<sup>1</sup>.

**Etap 1. Rozmywanie notowań:** Każdy dzień notowań reprezentowany jest za pomocą skierowanej świecy rozmytej. Rozważone zostaną trzy modele różniące się typem skierowanej świecy rozmytej (Trapezoidalna, Gaussowska, Empiryczna), dla których parametry  $S_1, S_2$  obliczane są jako minimum i maksimum z trzech różnych średnich oraz  $C_1 = C_2 = 0$ . Wszystkie rodzaje świec reprezentowane są w sposób dyskretny za pomocą 11 równomiernie rozłożonych punktów dla funkcji  $f$  i  $g$ . Na rys. 7.1 przedstawiono fragment notowań indeksu WIG20 (pierwszych 20 sesji) w postaci dziennych świec japońskich, oraz trzech rodzajów świec rozmytych. Świece bez wypełnienia to świece wzrostowe, natomiast z wypełnieniem, świece spadkowe.

**Etap 2. Badanie stacjonarności :** Dla każdego z rozmytych szeregów czasowych skonstruowano 22 ostre szeregi czasowe dla  $x$  od 0.0 do 1.0 z krokiem 0.1 zarówno dla funkcji  $f$  oraz  $g$ . Każdy ze skonstruowanych szeregów czasowych poddano testowi pierwiastka jednostkowego<sup>2</sup>. Otrzymane wartości statystyki testowej  $DF$  mieściły się w przedziale  $(-1.92, -1.61)$  co przy wartości krytycznej na poziomie istotności 0.05 wynoszącej  $-2.86319583097$  oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej mówiącej o niestacjonarności badanych szeregów czasowych. Stąd badane skierowane rozmyte szeregi czasowe uznane zostały za szeregi niestacjonarne.

<sup>1</sup>Wykorzystane dane pochodzą z internetowego serwisu Domu Maklerskiego Banku Ochrony Środowiska ([www.bossa.pl](http://www.bossa.pl)).

<sup>2</sup>Wykorzystano rozszerzony test Dickey-Fullera (Augmented Dickey-Fullera) z modułu Statsmodels ([www.statsmodels.sourceforge.net](http://www.statsmodels.sourceforge.net))

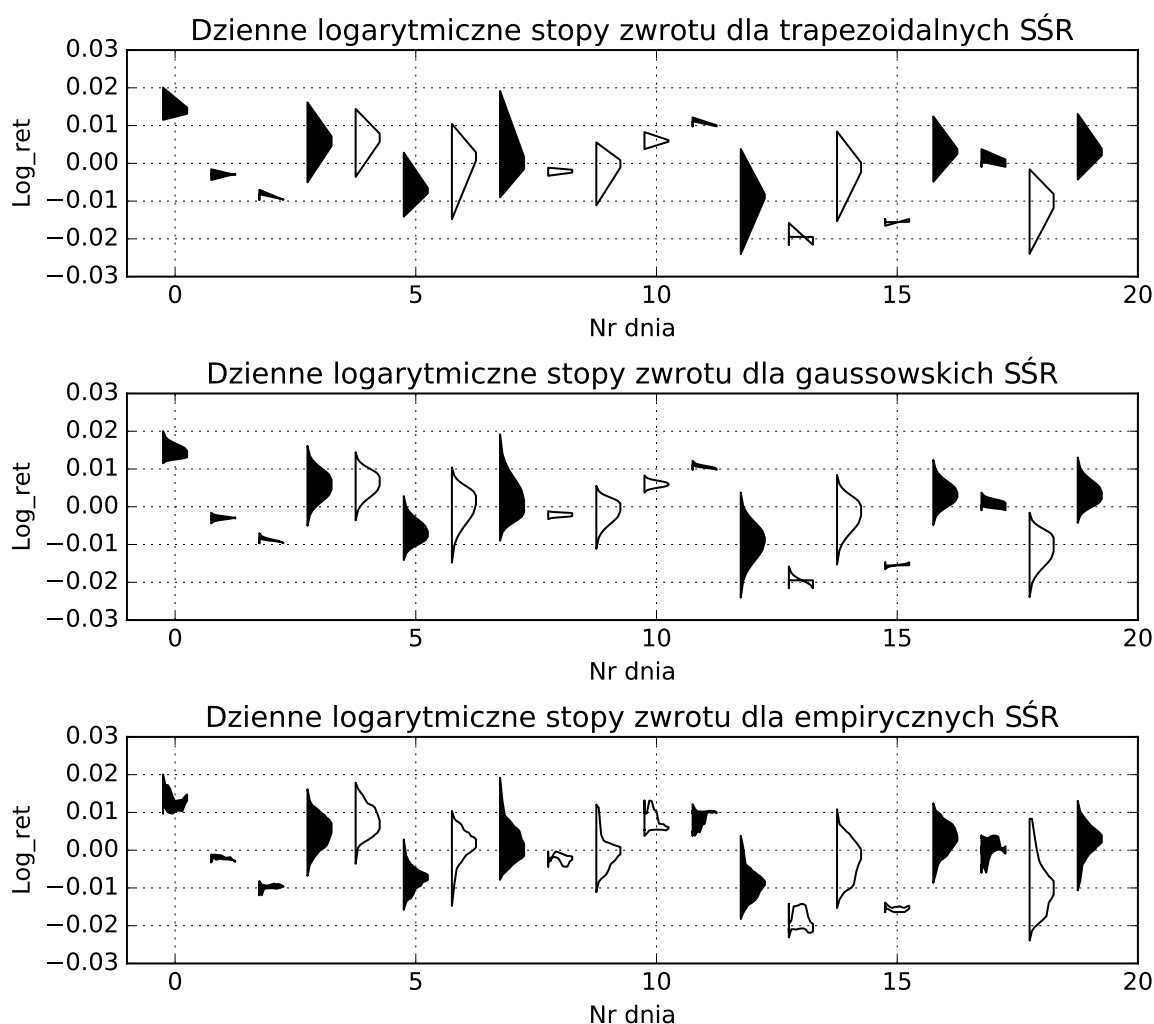


Rysunek 7.1: Dzienne świece notowań indeksu WIG20.

**Etap 3. Logarytmiczne stopy zwrotu:** W celu sprowadzenia rozmytych szeregów czasowych do szeregów stacjonarnych skonstruowane zostały rozmyte szeregi czasowe logarytmicznych stóp zwrotu w następujący sposób

$$\tilde{Y}_t = \ln \left( \frac{\tilde{X}_t}{\tilde{X}_{t-1}} \right), \quad (7.1)$$

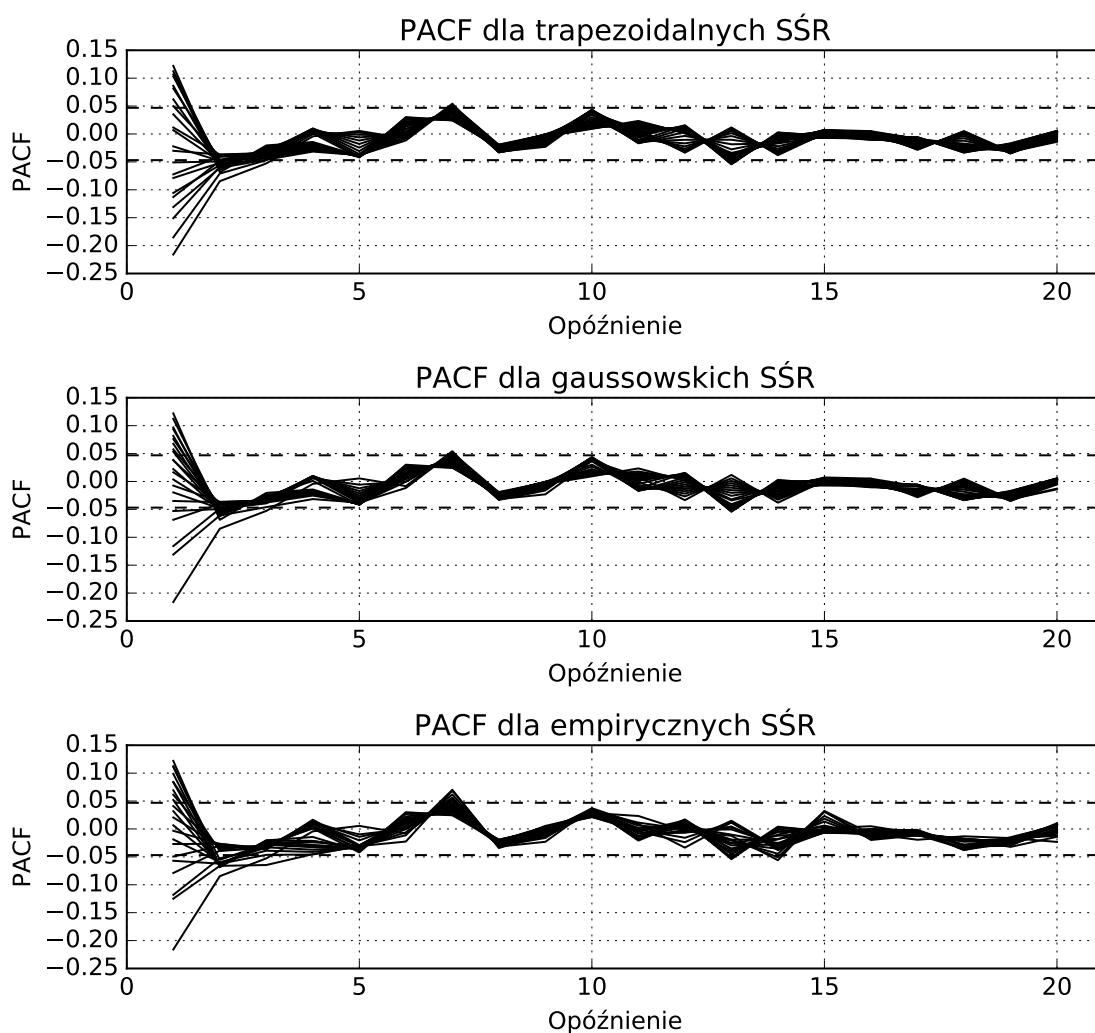
gdzie  $\tilde{X}_t$  to skierowane świece rozmyte skonstruowane w etapie 1. Otrzymane w ten sposób skierowane rozmyte szeregi czasowe przedstawione są na rys. 7.2. Skierowane liczby rozmyte przedstawione bez wypełnienia to liczby rozmyte o skierowaniu dodatnim, natomiast z wypełnieniem, o skierowaniu ujemnym.



Rysunek 7.2: Dzienne rozmyte logarytmiczne stopy zwrotu notowań indeksu WIG20.

Powstałe szeregi  $\tilde{Y}_t$  poddano testowi stacjonarności w ten sam sposób jak w etapie 2. Otrzymane wartości statystyki testowej  $DF$  mieściły się w przedziale  $(-43.93, -15.56)$  co przy wartości krytycznej na poziomie istotności 0.01 wynoszącej  $-3.43409652687$  pozwala na przyjęcie hipotezy alternatywnej mówiącej o stacjonarności badanych rozmytych szeregów czasowych.

**Etap 4. Identyfikacja rzędu i estymacja parametrów:** Jak napisano w rozdziale 5.2.2 do określenia rzędu autoregresji można posłużyć się wykresami funkcji autokorelacji cząstkowej<sup>3</sup> przedstawionymi na rys. 7.3. Z wykresów tych wynika że dla wszystkich



Rysunek 7.3: Wykresy PACF dla rozmytych szeregów logarytmicznych stóp zwrotu.

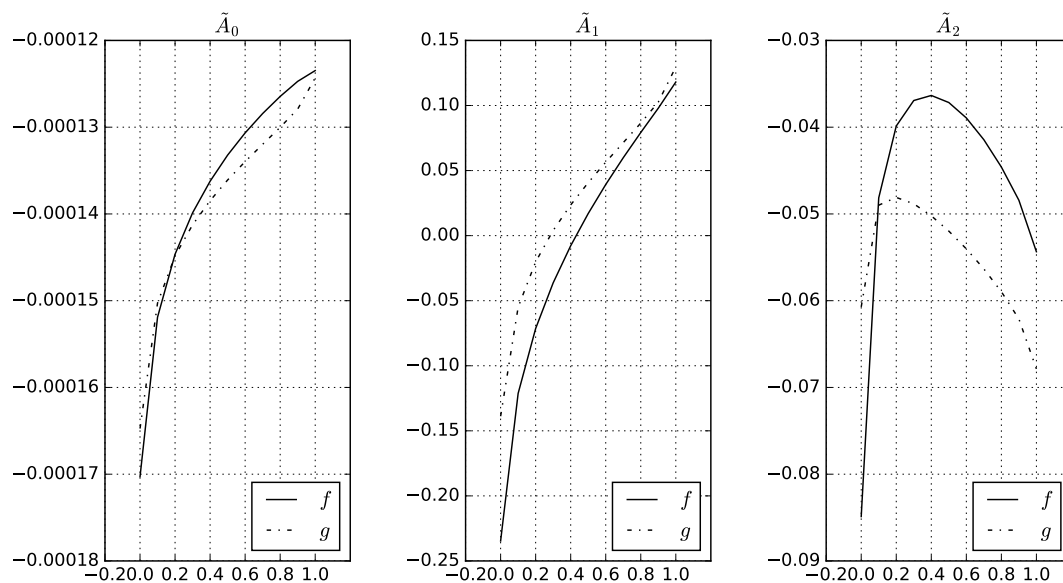
<sup>3</sup>Funkcje PACF obliczono korzystając z modułu Statsmodels ([www.statsmodels.sourceforge.net](http://www.statsmodels.sourceforge.net))

rodzajów świec rozmytych istotne są na pewno opóźnienia 1 i 2. Niemniej, wartości powyżej poziomu istotności osiągają również dużo większe opóźnienia np. 7 i 13. W celu wybrania ostatecznej liczby opóźnień wykorzystano kryterium informacyjne BIC<sup>4</sup>, które oprócz wartości dopasowania modelu uwzględnia również jego złożoność (liczbę parametrów). Dla wszystkich rozmytych szeregów, według kryterium BIC, rząd powinien wynosić 2. Stąd, ostatecznie zdecydowano się na estymację skierowanego rozmytego modelu autoregresyjnego rzędu 2 dla wszystkich rozważanych rozmytych szeregów czasowych. Wobec czego skierowany rozmyty model opisujący notowania indeksu WIG20 względem miary probabilistycznej  $\mathbf{P}$  przyjmuje następującą postać

$$\tilde{X}_t = \tilde{X}_{t-1} \cdot \exp \left( \tilde{A}_0 + \tilde{A}_1 \cdot \tilde{Y}_{t-1} + \tilde{A}_2 \cdot Y_{t-2} + \tilde{\varepsilon}_t \right), \quad (7.2)$$

$$\tilde{Y}_t = \ln \left( \frac{\tilde{X}_t}{\tilde{X}_{t-1}} \right). \quad (7.3)$$

Estymacji modeli dokonano poprzez minimalizację funkcjonału (5.36). Otrzymane współczynniki dla modelu skonstruowanego w oparciu o gaussowskie SŚR przedstawiono na rys. 7.4. Dla pozostałych modeli otrzymane współczynniki wyglądają bardzo podobnie.

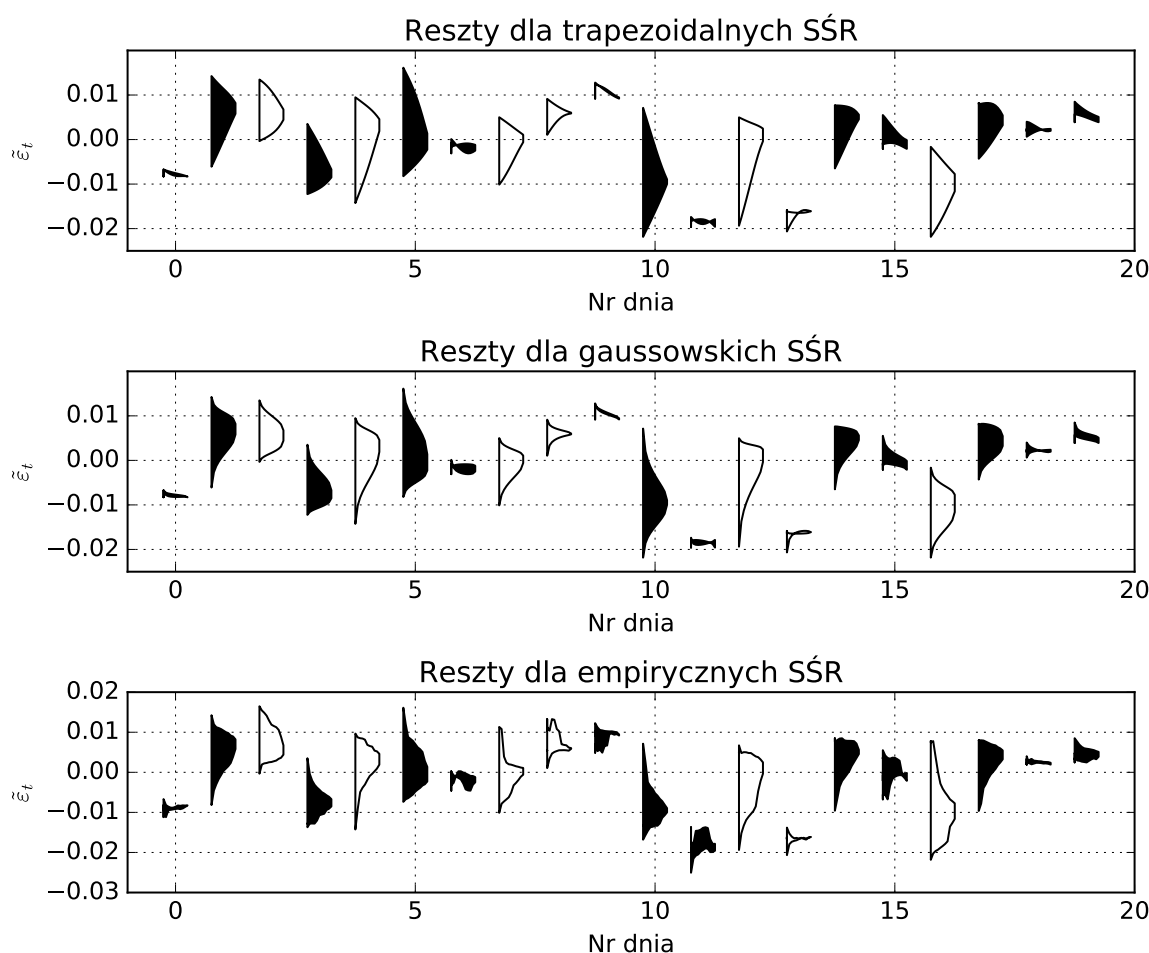


Rysunek 7.4: Współczynniki modelu SRAR(2) opartego na gaussowskich SŚR.

<sup>4</sup>Kryterium BIC obliczono korzystając z modułu Statsmodels ([www.statsmodels.sourceforge.net](http://www.statsmodels.sourceforge.net))

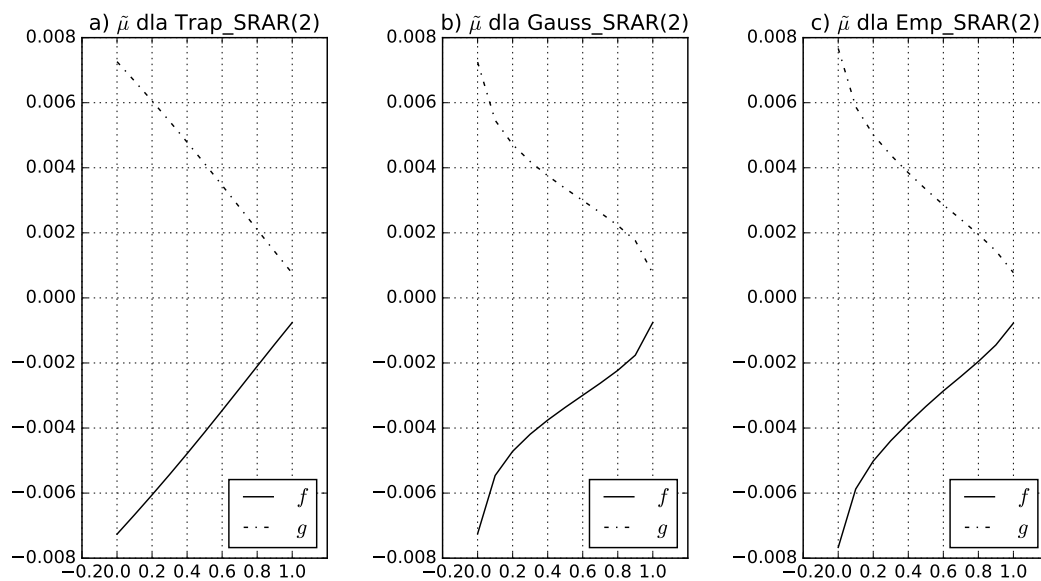
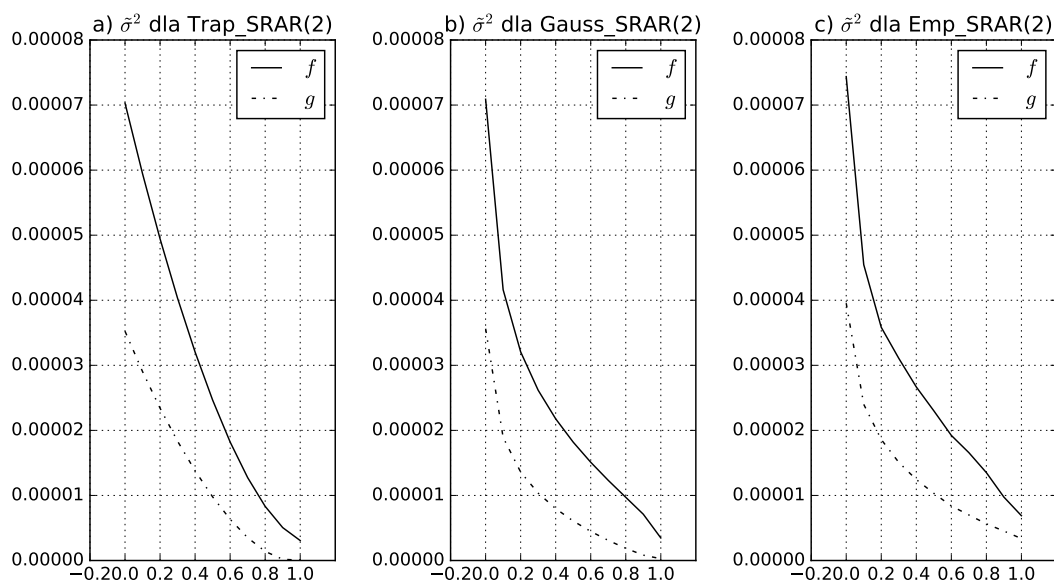
### 7.1.2. Analiza reszt modelu

W zagadnieniach symulacji najistotniejszą rzeczą jest określenie jakimi rozkładami prawdopodobieństwa charakteryzują się reszty (błędy) skonstruowanych modeli. W rozważanym przykładzie, reszty w modelach rozmytych są skierowanymi liczbami rozmytymi. Założono więc, że reszty te charakteryzują się skierowanym rozmytym rozkładem normalnym o parametrach  $\tilde{\mu}$ ,  $\tilde{\sigma}^2$ ,  $\sigma^2$  oraz  $p$ . Dla każdego z rozważanych rozmytych szeregów czasowych odpowiednie parametry oszacowano korzystając z estymatorów zdefiniowanych w rozdziale 6.3. Szeregi czasowe reszt dla rozważanych modeli przedstawione są na rys. 7.5, natomiast wartości parametrów oraz charakterystyki im odpowiadające przedstawione są w tabeli 7.1 oraz na rys. od 7.6 do 7.9.

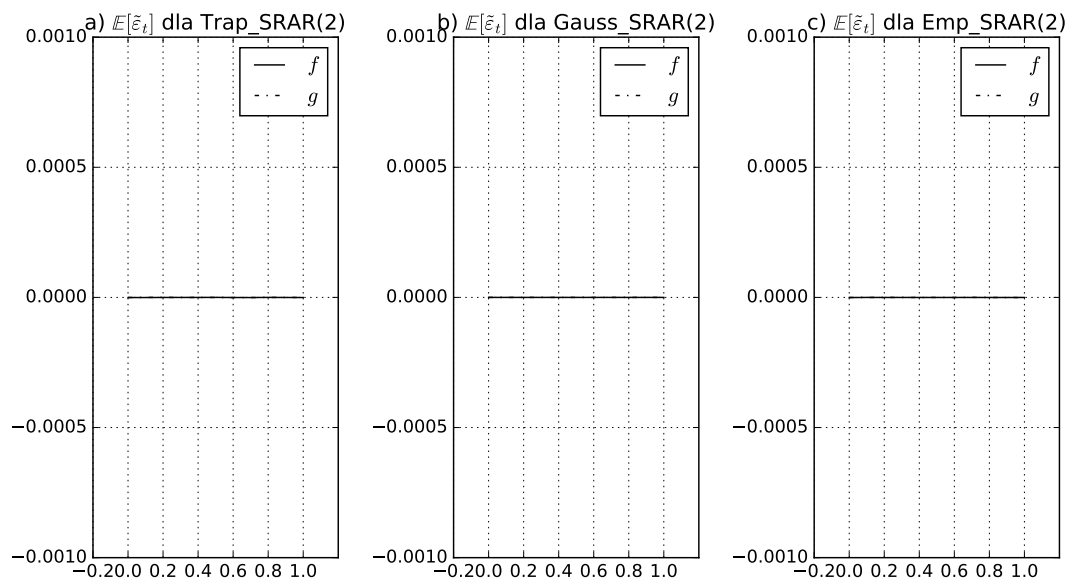
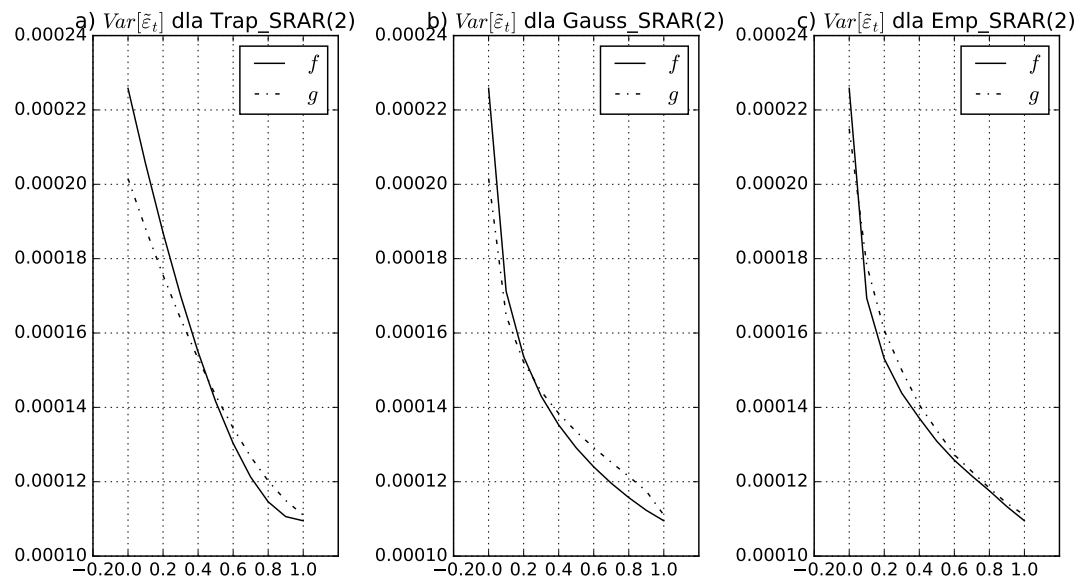


Rysunek 7.5: Rozmyte szeregi czasowe reszt  $\tilde{\epsilon}_t$ .

<i>Model</i>	$\tilde{\mu}$	$\tilde{\sigma}^2$	$\sigma^2$	$p$	$\mathbb{E}$	Var
Trap_SRAR(2)	rys. 7.6a	rys. 7.7a	0.0001081	0.4748	rys. 7.8a	rys. 7.9a
Gauss_SRAR(2)	rys. 7.6b	rys. 7.7b	0.0001078	0.4765	rys. 7.8b	rys. 7.9b
Emp_SRAR(2)	rys. 7.6c	rys. 7.7c	0.0001045	0.4743	rys. 7.8c	rys. 7.9c

Tabela 7.1: Wartości parametrów i charakterystyk dla rozkładów reszt  $\tilde{\varepsilon}_t$ .Rysunek 7.6: Realizacje parametru  $\tilde{\mu}$ .Rysunek 7.7: Realizacje parametru  $\tilde{\sigma}^2$ .



Rysunek 7.8: Rozmyte wartości oczekiwane  $E[\tilde{\epsilon}_t]$ .Rysunek 7.9: Rozmyte wariancje  $Var[\tilde{\epsilon}_t]$ .

### 7.1.3. Wycena opcji europejskiej

Wycena europejskiej opcji kupna na indeks WIG20 w niniejszym przykładzie oparta jest na procedurze Monte Carlo, w której pojedyncza iteracja powtarzana  $N$  razy przebiega według poniższego algorytmu.

**Krok 1.** Wygenerowanie próby  $\{\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_T\}$  skierowanych rozmytych zmiennych losowych o rozkładzie normalnym z parametrami  $\tilde{\mu}, \tilde{\sigma}^2, \sigma^2$  oraz  $p$ , gdzie  $T$  to liczba dni sesyjnych do terminu wygaśnięcia opcji.

**Krok 2.** Wyznaczenie ścieżek skierowanych rozmytych logarytmicznych stóp zwrotu zgodnie ze zmodyfikowanym modelem SRAR(2). Modyfikacja dokonywana jest w taki sposób, aby w każdej chwili  $t$  warunkowa wartość oczekiwana logarytmicznej stopy zwrotu była równa stopie zwrotu wolnej od ryzyka  $r_f$ . Wówczas, względem równoważnej miary martyngałowej  $\mathbf{Q}$  model przyjmuje następującą postać

$$\tilde{Y}_t = r_f + \tilde{\xi}_t - \mathbb{E}[\tilde{\xi}_t]. \quad (7.4)$$

**Krok 3.** Wyznaczenie skierowanych rozmytych cen instrumentu bazowego w terminie wygaśnięcia  $T$  zgodnie z wzorem rekurencyjnym

$$\tilde{X}_T = \tilde{X}_{T-1} \cdot \exp(\tilde{Y}_T), \quad (7.5)$$

gdzie wartość początkowa  $\tilde{X}_0$  to skierowana świeca rozmyta notowań instrumentu bazowego w dniu wyceny.

**Krok 4.** Wyznaczenie wartości wypłaty z opcji kupna zgodnie z wzorem

$$\tilde{H}_n = \max(\tilde{X}_T - K, 0), \text{ dla wyceny rozmytej,} \quad (7.6)$$

$$H_n = \max(\text{Defuzzy}(\tilde{X}_T) - K, 0), \text{ dla wyceny ostrej,} \quad (7.7)$$

gdzie  $n$  to numer iteracji,  $K$  to cena wykonania opcji, *Defuzzy* to pewien ustalony operator wyostrozania, natomiast funkcja  $\max$  dla skierowanych liczb rozmytych zdefiniowana jest w następujący sposób

$$\max(A, 0) = \max((f_A, g_A), 0) = (\max(f_A(x), 0), \max(g(x), 0)) \text{ dla } x \in [0, 1].$$

Po wykonaniu  $N$  iteracji, skierowaną liczbę rozmytą reprezentującą cenę opcji kupna wyznaczana się zgodnie z wzorem (7.8), natomiast cenę w postaci liczby ostrej zgodnie z wzorem (7.9)

$$\tilde{C} = e^{(-r_f T^{(k)})} \cdot \frac{1}{N} \sum_n^N \tilde{H}_n, \quad (7.8)$$

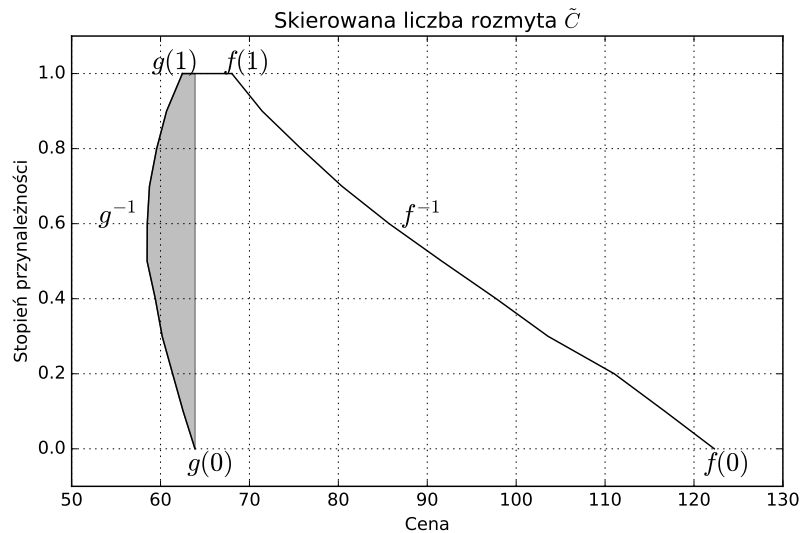
$$C = e^{(-r_f T^{(k)})} \cdot \frac{1}{N} \sum_n^N H_n, \quad (7.9)$$

gdzie  $T^{(k)}$  to liczba dni kalendarzowych do terminu wygaśnięcia podana w latach.

Tabela 7.2 przedstawia wyniki wyceny europejskiej opcji kupna na koniec 30.12.2016 o różnych cenach wykonania oraz z terminem wygaśnięcia 20.01.2017. Wyceny uzyskane z modeli rozmytych  $C_1$  (z wykorzystaniem operatora wartości oczekiwanej),  $C_2$  (z wykorzystaniem operatora środka ciężkości),  $\ker(\tilde{C})$  (jądro zbioru rozmytego) porównano z wyceną z modelu Blacka-Scholesa-Mertona (BSM) oraz z cenami rynkowymi w dniu wyceny (cena zamknięcia z dnia 30.12.2016). W symulacjach generowano po  $N = 10000$  iteracji, a za stopę wolną od ryzyka przyjęto  $r_f = 1.335\%$  (średnia rentowność bonów skarbowych sprzedanych w 2016 roku<sup>5</sup>). Dla modelu BSM ustalono  $\sigma = 18.898\%$  (zmienność historyczna z okresu od początku 2010 do końca 2016 roku). Warto zauważyć, że wyceny z modelu BSM są dla cen wykonania poniżej ceny instrumentu bazowego w dniu wyceny (1942.92) zaniżone, natomiast dla cen powyżej zawyżone. W przypadku modeli rozmytych niezależnie od ceny wykonania są prawie zawsze zaniżone. Ponadto, różnice między wyceną a ceną rynkową są średnio dużo mniejsze dla modeli rozmytych w stosunku do modelu BSM, a w szczególności dla modeli wykorzystujących operator środka ciężkości. Wyceny najbliższe cenie rynkowej zostały w tabeli 7.2 wytłuszczone.

Bardzo interesującym wynikiem jest postać ceny opcji w postaci skierowanej liczby rozmytej. Na rys. 7.10 przedstawiona jest rozmyta cena rozważanej opcji europejskiej z ceną wykonania 1950 uzyskana z modelu Trap\_SRAR(2). Uzyskana cena jest niewłaściwą skierowaną liczbą rozmytą, a powstała z funkcji  $f$  i  $g$  relacja przynależności może być

<sup>5</sup>Źródło: [www.finanse.mf.gov.pl](http://www.finanse.mf.gov.pl)

Rysunek 7.10: Realizacja rozmytej ceny opcji europejskiej  $\tilde{C}$ .

interpretowana w następujący sposób<sup>6</sup>. Podążając przeciwnie do narastania osi OX, od wartości  $f(0)$  do wartości  $g(0)$ , rynek jest zgodny, w jakim stopniu ta wartość przynależy do zbioru  $\tilde{C}$ . Idąc dalej, po przekroczeniu wartości  $g(0)$  na rynku nie ma pewności co do tego, jak powinna wyglądać funkcja przynależności (zamalowany obszar). W tym przedziale wartości funkcji przynależności również są rozmyte (podobnie jak w zbiorach rozmytych 2 typu). Ponadto, w skierowaniu liczby  $\tilde{C}$ , które jest ujemne, zawarta jest informacja mówiąca o spadku w czasie ceny opcji, co może być utożsamione z wartością czasową opcji, cena opcji względem czasu jest funkcją malejącą (przy stałych pozostałych czynnikach).

<sup>6</sup>Wykorzystano interpretację niewłaściwych skierowanych liczb rozmytych zaproponowaną w pracy [94].

cena wykonania	cena rynkowa	BSM	Trap_SRAR(2)			Gauss_SRAR(2)			Emp_SRAR(2)		
			$C_1$	$C_2$	$\ker(\tilde{C})$	$C_1$	$C_2$	$\ker(\tilde{C})$	$C_1$	$C_2$	$\ker(\tilde{C})$
1800	178.4	168.6	171.5	<b>176.9</b>	[165.9, 173.8]	171.8	176.2	[167.1, 174.6]	171.8	176.8	[166.3, 174.6]
1850	135	130.3	130.6	135.7	[125.4, 132.9]	131.1	<b>135.2</b>	[126.9, 133.8]	130.7	135.3	[125.9, 133.7]
1900	101	97.14	95.03	<b>99.53</b>	[90.68, 97.22]	95.01	99.43	[92.11, 98.43]	94.77	98.85	[90.07, 98.06]
1950	71.55	69.68	65.95	69.71	[63.63, 69.01]	66.67	<b>69.72</b>	[63.63, 69.01]	65.31	68.76	[62.38, 68.46]
2000	46.33	48.02	43.75	46.68	[41.01, 45.53]	43.06	<b>46.34</b>	[41.6, 45.92]	42.6	45.26	[40.47, 45.63]
2050	29.50	31.8	27.65	29.83	[25.70, 29.02]	27.45	<b>29.31</b>	[25.77, 28.88]	26.24	28.19	[24.76, 28.8]
2100	18.17	20.2	16.52	<b>18.06</b>	[15.25, 17.55]	16.16	17.48	[15.06, 17.17]	15.18	16.49	[14.33, 17.15]

Tabela 7.2: Wycena europejskiej opcji kupna na indeks WIG20 z terminem wygaśnięcia 17.03.2017.

## 7.2. Dywersyfikacja portfela akcji

Zagadnienie dywersyfikacji portfela inwestycyjnego polega na rozdzieleniu dostępnego kapitału pośród dostępne aktywa w taki sposób, aby osiągnąć założony oczekiwany zwrot z inwestycji przy jak najmniejszym ryzyku, lub odwrotnie, zmaksymalizować oczekiwany zysk przy ustalonym poziomie ryzyka. W klasycznej teorii portfela zapoczątkowanej przez Markowitza [54, 55, 56], ryzyko portfela mierzone jest za pomocą wariancji stóp zwrotu z portfela, natomiast oczekiwany zysk za pomocą wartości oczekiwanej stóp zwrotu. Stąd problem ten nazywany jest problemem *Mean-Variance* (MV).

Niech wektor  $\mathbf{x} = [x_1, x_2, \dots, x_n]$  oznacza procent kapitału alokowanego w aktywa  $i$ . Wówczas problem MV można zapisać w następujący sposób

$$\begin{aligned} & \text{Minimalizuj } \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \\ & \text{Przy ograniczeniach} \\ & \sum_{i=1}^n \mu_i x_i = \rho \quad (\text{założona oczekiwana stopa zwrotu}) \\ & \sum_{i=1}^n x_i = 1 \quad (\text{ograniczenie na budżet}) \end{aligned} \tag{7.10}$$

gdzie  $n$  to liczba dostępnych aktywów,  $\mu_i$  to oczekiwana stopa zwrotu z aktywa  $i$  oraz  $\sigma_{ij}$  to kowariancja stóp zwrotu aktywów  $i$  i  $j$ .

Problem (7.10) jest wypukłym problemem programowania kwadratowego z ograniczeniami, które może być rozwiązane przez szeroką gamę algorytmów do tego przeznaczonych, nawet dla dużej liczby aktywów. W niniejszym przykładzie rozważony zostanie bardziej realistyczny problem z dodatkowymi ograniczeniami, takimi jak ograniczenie na maksymalną liczbę aktywów  $K$  w portfelu oraz ograniczenia dolne i górne  $[l_i, u_i]$  na procentowy udział  $i$ -tego aktywa w portfelu. Ponadto, uwzględniane są również koszty transakcyjne związane ze zmianą struktury portfela. Tak zmodyfikowany problem nie jest już wypukły,

dlatego rozważony zostanie w postaci zagadnienia optymalizacji wielokryterialnej

$$\begin{aligned}
 & \text{Minimalizuj } \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \\
 & \text{Maksymalizuj } \sum_{i=1}^n \mu_i x_i - \sum_{i=1}^n c_i |x_i - x_i^0| \\
 & \text{Przy ograniczeniach} \\
 & \sum_{i=1}^n x_i = 1 \quad (\text{ograniczenie na budżet}) \\
 & x_i = 0 \text{ lub } l_i \leq x_i \leq u_i \quad (\text{ograniczenie na udziały}) \\
 & \#\{i: x_i > 0\} \leq K \quad (\text{ograniczenie na liczbę aktywów}) \\
 & i = 1, \dots, n
 \end{aligned} \tag{7.11}$$

gdzie wektor  $\mathbf{x}^0 = [x_1^0, x_2^0, \dots, x_n^0]$  to obecna struktura portfela, natomiast  $c_i$  to koszt transakcyjny związany z  $i$ -tym aktywem.

### 7.2.1. Rozmyte stopy zwrotu

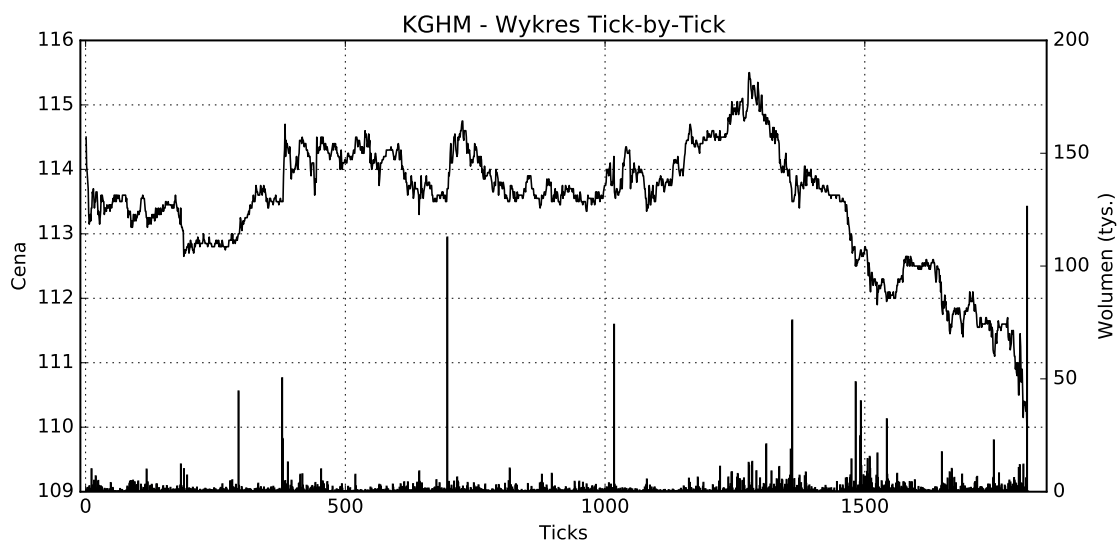
Klasycznie, stopy zwrotu z aktywów finansowych opisywane są za pomocą zmiennych losowych z pewnym rozkładem prawdopodobieństwa. Niemniej, jak wspomniano już wcześniej w rozprawie, istnieje wiele różnych czynników wpływających na rynki finansowe. Stąd, uzasadnione jest modelowanie stóp zwrotu z uwzględnieniem niepewności. W niniejszym przykładzie stopy zwrotu modelowane są za pomocą koncepcji skierowanych świec rozmytych z wykorzystaniem danych wysokiej częstotliwości.

Niech  $X_{pt}^{(i)}$  dla  $t = 1, \dots, T$  będzie szeregiem czasowym notowań  $i$ -tego aktywa dla  $p$ -tego przedziału czasu (np. dnia, tygodnia, miesiąca itp.),  $p = 1, \dots, P$ . Dla każdego przedziału czasu  $p$  konstruowany jest proces tworzenia się stóp zwrotu  $R_{pt}^{(i)}$  w następujący sposób

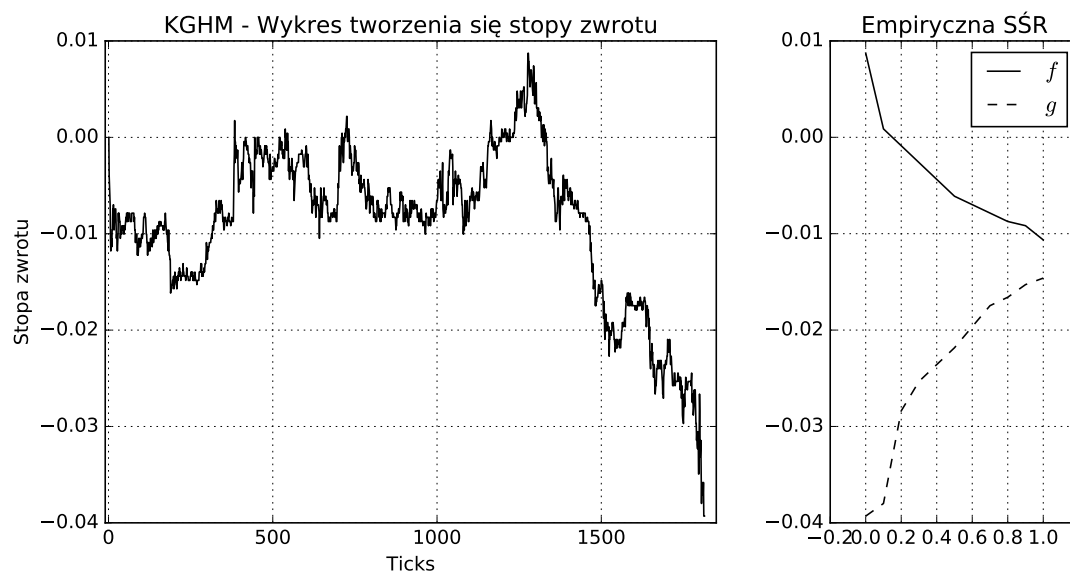
$$R_{pt}^{(i)} = \frac{X_{pt}^{(i)} - X_{p1}^{(i)}}{X_{p1}^{(i)}} \quad t = 1, \dots, T. \tag{7.12}$$

Następnie, każdy z procesów  $R_{pt}^{(i)}$  utożsamiany jest ze skierowaną świecą rozmytą  $\tilde{R}_p^{(i)}$ , zgodnie z metodami przedstawionymi w rozdziale 4.2. Przykładowo na rys. 7.11 przedstawiony jest wykres notowań tick-by-tick spółki KGHM wraz z wolumenem w okresie

jednego tygodnia od 20.01.2014 do 24.01.2014. Natomiast, rys. 7.12 przedstawia szereg tworzenia się tygodniowej stopy zwrotu dla spółki KGHM w tym samym okresie, wraz z odpowiadającą mu empiryczną skierowaną świecą rozmytą.



Rysunek 7.11: Notowania spółki KGHM w okresie 20-24.01.2017.



Rysunek 7.12: Empiryczna ŚSR dla szeregu tworzenia się stóp zwrotu spółki KGHM w okresie 20-24.01.2017.



### 7.2.2. Rozmyta dywersyfikacja portfela akcji

Zagadnienie dywersyfikacji portfela z uwzględnieniem niepewności definiowane jest w analogiczny sposób jak zagadnienie (7.11), z tą różnicą, że minimalizowana jest ostra wariancja rozmytej stopy zwrotu z portfela, przy maksymalizacji ostrej wartości oczekiwanej rozmytej stopy zwrotu z portfela. Z definicji 6.19 oraz 6.20 wynika, że ostra wartość oczekiwana i ostra wariancja skierowanych rozmytych zmiennych losowych jest równoważna klasycznej wartości oczekiwanej i wariancji zmiennych losowych powstałych przez wyostrenie operatorem wartości oczekiwanej skierowanych rozmytych zmiennych losowych. Wobec czego rozmyte zagadnienie dywersyfikacji portfela można zapisać w postaci poniższego problemu optymalizacji wielokryterialnej

$$\begin{aligned}
 & \text{Minimalizuj } \sum_{i=1}^n \sum_{j=1}^n \sigma_{\tilde{r}_i \tilde{r}_j} x_i x_j \\
 & \text{Maksymalizuj } \sum_{i=1}^n \mu_{\tilde{r}_i} x_i - \sum_{i=1}^n c_i |x_i - x_i^0| \\
 & \text{Przy ograniczeniach} \\
 & \sum_{i=1}^n x_i = 1 \quad (\text{ograniczenie na budżet}) \\
 & x_i = 0 \text{ lub } l_i \leq x_i \leq u_i \quad (\text{ograniczenie na udziały}) \\
 & \#\{i : x_i > 0\} \leq K \quad (\text{ograniczenie na liczbę aktywów}) \\
 & i = 1, \dots, n
 \end{aligned} \tag{7.13}$$

gdzie  $\tilde{r}_i$  to ostre stopy zwrotu powstałe przez wyostrenie operatorem wartości oczekiwanej rozmytych stóp zwrotu  $\tilde{R}^{(i)}$ ,  $\mu_{\tilde{r}_i}$  to wartość oczekiwana stóp zwrotu  $\tilde{r}_i$ ,  $\sigma_{\tilde{r}_i \tilde{r}_j}$  to kowariancja stóp zwrotu  $\tilde{r}_i$  oraz  $\tilde{r}_j$ . Pozostałe oznaczenia są takie same jak w zagadnieniach (7.10) i (7.11).

### 7.2.3. Wyniki empiryczne

W przykładzie empirycznym rozważony zostanie klasyczny oraz rozmyty problem dynamicznej (cotygodniowej) dywersyfikacji portfela konstruowanego z akcji spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie wchodzących w skład indeksu

WIG30<sup>7</sup>. W tym celu, dla każdej z 30 spółek w okresie od 05.01.2015 do 24.02.2017 konstruowane zostały tygodniowe klasyczne stopy zwrotu, obliczone na podstawie cen otwarcia z początku tygodnia (poniedziałek) i cen zamknięcia na koniec tygodnia (piątek)<sup>8</sup>. Oraz stopy zwrotu będące wyostrzonymi wartościami rozmytych stóp zwrotu konstruowanych zgodnie z podrozdziałem 7.2.1. Rozważone zostaną trzy rodzaje rozmytych stóp zwrotu różniące się typem skierowanej świecy rozmytej (Trapezoidalna, Gaussowska, Empiryczna), dla których parametry  $S_1, S_2$  obliczane są jako minimum i maksimum z trzech różnych średnich, parametry  $C_1, C_2$  jako odchylenie standardowe oraz uwzględniana jest informacja o wolumenie transakcji.

W celu przeprowadzenia dywersyfikacji założono, że inwestor zamierza ulokować w akcjach kapitał wynoszący 100000 PLN, maksymalna ilość różnych aktywów w portfelu wynosi  $K = 20$  oraz ograniczenia na udział poszczególnych aktywów ustalone są na poziomach  $l_i = 0.014$  i  $u_i = 0.2$ . Koszt transakcyjny dla każdego aktywa jest taki sam i wynosi 0.38% wartości transakcji lecz nie mniej niż 5 PLN i jest on dodatkowym kosztem inwestora. Proces dywersyfikacji wykonywany jest na początku każdego tygodnia przez okres ośmiu tygodni od 02.01.2017 do 24.02.2017. W taki sposób, że na koniec każdego tygodnia określana jest aktualna struktura portfela według piątkowych cen zamknięcia, następnie w procesie dywersyfikacji ustalana jest nowa struktura portfela na bazie stóp zwrotu z okresu dwóch lat do tyłu (104 tygodnie). Zmiana struktury portfela tzn. sprzedaż lub dokupienie akcji odbywa się po poniedziałkowych cenach otwarcia. Ponieważ, możliwe jest kupno lub sprzedaż jedynie całkowitej liczby akcji, niewykorzystana część kapitału lokowana jest na lokacie o oprocentowaniu stałym równym stopie wolnej od ryzyka  $r_f = 1.335\%$  (średnia rentowność bonów skarbowych sprzedanych w 2016 roku<sup>9</sup>)

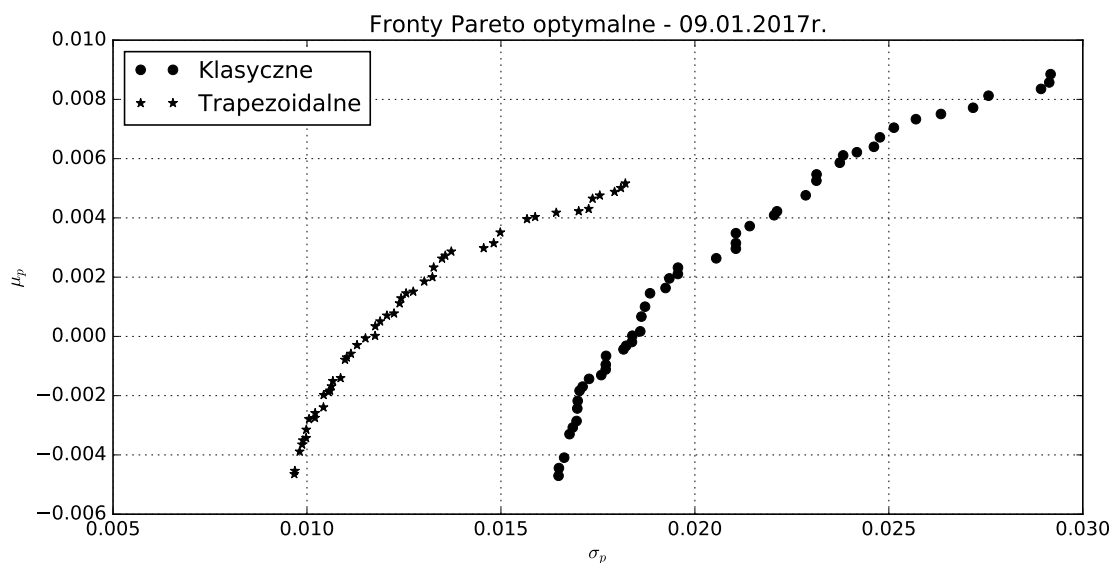
Rozwiązaniem zagadnienia dywersyfikacji portfela (7.11) lub (7.13) jest zbiór portfeli efektywnych (Pareto optymalnych). Tak, więc dodatkowym problemem jest wybór

<sup>7</sup>Skład portfela indeksu WIG30 na dzień 02.01.2017r. Wykorzystane notowania spółek WIG30 poprano z serwisu [www.bossa.pl](http://www.bossa.pl)

<sup>8</sup>Jeżeli poniedziałek lub piątek nie jest dniem handlowym to brane są ceny z odpowiednio następnego lub poprzedniego dnia notowań.

<sup>9</sup>Oprocentowanie w skali roku.

jednego z pośród nich. Wyboru tego dokonuje się w praktyce zgodnie z preferencjami inwestora, np. dla ustalonego poziomu ryzyka wybiera się portfel efektywny o największej oczekiwanej stopie zwrotu lub odwrotnie, dla ustalonej stopy zwrotu wybiera się portfel efektywny o najmniejszym ryzyku. Niemniej, w celu porównania podejścia klasycznego z rozmytym taki wybór może prowadzić do błędnych wniosków. Wynika to z faktu, że oczekiwana stopa zwrotu na poziomie np. 5% oznacza zupełnie co innego dla klasycznych stóp zwrotu, a co innego dla stóp zwrotu w podejściu rozmytym, podobnie jest z poziomem ryzyka. W tabeli 7.4 podane zostały wartości oczekiwane oraz odchylenia standardowe dla wszystkich skonstruowanych tygodniowych stóp zwrotu w rozważanym okresie. Jak można zauważyć, odchylenia standardowe dla stóp zwrotu w podejściu rozmytym są dużo mniejsze niż w klasycznym. Co wcale nie oznacza, że ryzyko jest mniejsze, a jedynie to, że jest to inna miara ryzyka. Widać to również na wyznaczonych frontach (por. rys. 7.13)<sup>10</sup>.



Rysunek 7.13: Fronty Pareto optymalne dla podejścia klasycznego i rozmytego z wykorzystaniem trapezoidalnych SŚR otrzymane w dniu 09.01.2017.

W niniejszym przykładzie w procesie dywersyfikacji wybierany będzie portfel efektywny o największym wskaźniku Sharpe'a (ang. *tangency portfolio*). Wskaźnik Sharpe'a, wprowadzony przez W. F. Sharpe'a [74], to stosunek stopy zwrotu z portfela pomniejszo-

<sup>10</sup>Fronty z rysunku 7.13 jak i wszystkie późniejsze portfele optymalne wyznaczone zostały przy wykorzystaniu algorytmu optymalizacji wielokryterialnej IMGAMO (IMMune GAME theory MultiObjective) zaproponowanego przez P. Jarosza et al. w [26, 27, 28].

nej o stopę wolną od ryzyka do odchylenia standardowego z portfela. Wskaźnik Sharpe'a pozwala w łatwy sposób ocenić wysokość premii uzyskiwanej z danego portfela w stosunku do poniesionego ryzyka. Wyliczany jest on za pomocą następującego wzoru

$$SR = \frac{\mu_p - r_f}{\sigma_p}, \quad (7.14)$$

gdzie  $\mu_p, \sigma_p$  to odpowiednio oczekiwana stopa zwrotu i odchylenie standardowe z portfela, natomiast  $r_f$  to stopa wolna od ryzyka.

Wyniki portfeli zarządzanych według rozważanych podejść ocenione zostały za pomocą następujących 5 wskaźników: wartość średnia ( $\mu$ ), odchylenie standardowe ( $\sigma$ ), wskaźnik Sharpe'a (SR), wskaźnik informacyjny (IR) oraz współczynnik beta (Beta), wyliczonych na podstawie zrealizowanych w okresie dywersyfikacji tygodniowych stóp zwrotu. Dla wskaźników IR i Beta jako benchmark przyjęto tygodniowe stopy zwrotu z indeksu WIG30. Wartości wskaźników przedstawione są w tabeli 7.3, czcionką wytłuszczoną zaznaczone zostały najlepsze z nich.

Jak można zauważyć, wartości wskaźników dla portfeli zarządzanych w podejściu rozmytym są do siebie zbliżone i lepsze od wartości dla podejścia klasycznego (oprócz odchylenia standardowego). Wskaźniki SR i IR mierzące stosunek średnich dodatkowych stóp zwrotu (ponad stopę wolną od ryzyka dla SR oraz ponad stopy zwrotu benchmarku dla IR) do poniesionego ryzyka (odchylenie standardowe dodatkowych stóp zwrotu) są najlepsze dla podejścia rozmytego budowanego w oparciu o gaussowskie SŚR. Wartości współczynnika beta, który mierzy wrażliwość zmiany wartości portfela na zmianę wartości przyjętego benchmarku, są większe od jedności. Oznacza to, że wartości portfeli zmieniają

Podejście	$\mu$	$\sigma$	SR	IR	Beta
Klasyczne	0.01972	<b>0.02500</b>	0.77843	0.17778	1.14987
Trapezoidalne	0.02143	0.02552	0.82976	0.25116	<b>1.07528</b>
Gaussowskie	<b>0.02182</b>	0.02584	<b>0.83439</b>	<b>0.26823</b>	1.09107
Empiryczne	0.02129	0.02566	0.81970	0.24342	1.08751

Tabela 7.3: Wskaźniki oceny jakości skonstruowanych portfeli na bazie zrealizowanych tygodniowych stóp zwrotu.

się zgodnie ze zmianą benchmarku ale bardziej dynamicznie, co w przypadku wzrostów pozwala osiągnąć lepsze stopy zwrotu, natomiast w przypadku spadków powoduje dużo większe straty w stosunku do benchmarku. Pożądanymi portfelami są więc portfele o dodatniej stopie zwrotu, wysokich wskaźnikach SR i IR oraz współczynnika beta bliskim zeru. Warto zaznaczyć, że w analizowanym okresie panowała dobra koniunktura na rynku, a indeks WIG30 osiągnął w tym okresie stopę zwrotu na poziomie 14.42% (średnia tygodniowa stopa zwrotu 1.71%).

W tabelach 7.5 oraz 7.6 przedstawiono otrzymane w wyniku dywersyfikacji struktury portfeli dla każdego z tygodni i dla każdego z rozważanych podejść. Jak można zauważyć, niezależnie od podejścia, średnio ponad 80% każdego z portfeli stanowią akcje 6 spółek CDR, JSW, KER, LTS, PGN oraz PKN, a różnice między ich procentowymi udziałami w portfelach względem podejść są niewielkie. Dużo większe różnice między podejściem klasycznym a rozmytym można zauważyć w wyborze spółek i ich udziałów do pozostałej części portfela. Z kolei, tabele 7.7 oraz 7.8 przedstawiają szczegółowe zrealizowane wyniki portfeli na koniec każdego z tygodni.

Spółka	Klasyczne		Trapezoidalne		Gaussowskie		Empiryczne	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
ALR	0.00032	0.04428	0.00049	0.02511	0.00047	0.02592	0.00011	0.02762
ACP	0.00214	0.02618	0.00258	0.01533	0.00266	0.01569	0.00261	0.01690
LWB	0.00021	0.06303	-0.00055	0.03656	-0.00051	0.03783	-0.00099	0.04106
BZW	0.00075	0.03710	0.00029	0.02231	0.00035	0.02308	0.00020	0.02428
CCC	0.00648	0.04322	0.00266	0.02649	0.00285	0.02737	0.00338	0.02907
CDR	0.00891	0.05331	0.00534	0.03289	0.00561	0.03400	0.00603	0.03667
CPS	0.00183	0.03268	0.00196	0.02236	0.00204	0.02285	0.00199	0.02400
ENA	-0.00009	0.03930	0.00002	0.02387	0.00003	0.02469	0.00023	0.02650
ENG	-0.00181	0.04549	-0.00134	0.02696	-0.00135	0.02797	-0.00121	0.02979
EUR	0.00015	0.04844	-0.00121	0.02698	-0.00122	0.02800	-0.00070	0.03070
KGH	0.00196	0.05138	0.00135	0.03129	0.00142	0.03239	0.00145	0.03435
ATT	0.00278	0.05003	0.00332	0.03232	0.00337	0.03329	0.00324	0.03470
GTC	0.00240	0.03764	-0.00123	0.02474	0.00113	0.02546	-0.00106	0.02716
ING	0.00313	0.03118	0.00212	0.02065	0.00221	0.02126	0.00258	0.02263
JSW	0.00675	0.08182	0.00794	0.05515	0.00811	0.05711	0.00776	0.05941
KER	0.00844	0.05443	0.00332	0.03323	0.00350	0.03426	0.00331	0.03656
LTS	0.00360	0.03626	0.00236	0.02221	0.00243	0.02300	0.00270	0.02415
LPP	-0.00216	0.05373	0.00010	0.03214	0.00013	0.03344	-0.00035	0.03514
MBK	-0.00101	0.04126	-0.00036	0.02472	-0.00032	0.02558	-0.00045	0.02678
MIL	0.00031	0.04853	-0.00140	0.02979	-0.00134	0.03088	-0.00099	0.03222
OPL	-0.00446	0.03657	-0.00251	0.02178	-0.00263	0.02256	-0.00274	0.02427
PEO	-0.00093	0.02942	-0.00041	0.01823	-0.00040	0.01881	-0.00051	0.02010
PGE	-0.00153	0.03826	-0.00152	0.02324	-0.00157	0.02404	-0.00167	0.02508
PGN	0.00255	0.03675	0.00352	0.02121	0.00356	0.02186	0.00370	0.02352
PKN	0.00581	0.03715	0.00489	0.02174	0.00500	0.02246	0.00500	0.02433
PKO	0.00008	0.03433	0.00039	0.01965	0.00045	0.02041	0.00046	0.02175
PKP	-0.00091	0.04446	-0.00078	0.02767	-0.00068	0.02871	-0.00093	0.02999
PZU	-0.00030	0.03214	-0.00030	0.01901	0.00032	0.01969	-0.00025	0.02078
SNS	0.00095	0.03654	0.00096	0.02195	0.00102	0.02276	0.00085	0.02430
TPE	-0.00139	0.03953	-0.00002	0.02395	-0.00007	0.02477	-0.00030	0.02639

Tabela 7.4: Wartości oczekiwane i odchylenia standardowe tygodniowych stóp zwrotu dla spółek z indeksu WIG30.

Tydzień	Klasyczne															
	ACP	BZW	CCC	CDR	CPS	EUR	GTC	ATT	ING	JSW	KER	LTS	LPP	PGN	PKN	SNS
02.01.2017	0.014	0	0.077	0.2	0	0.014	0.014	0	0.014	0.2	0.2	0.039	0	0.014	0.2	0.014
09.01.2017	0	0	0.088	0.2	0	0.023	0.027	0	0	0.174	0.2	0.053	0	0.02	0.2	0.014
16.01.2017	0	0	0.073	0.2	0	0.021	0.025	0	0	0.175	0.2	0.062	0	0.02	0.197	0.027
23.01.2017	0	0	0.062	0.2	0	0.015	0.039	0	0	0.15	0.2	0.074	0	0.029	0.19	0.041
30.01.2017	0	0	0.045	0.2	0	0.014	0.051	0	0	0.141	0.2	0.083	0	0.036	0.175	0.055
06.02.2017	0	0	0.032	0.2	0	0.014	0.063	0	0	0.129	0.2	0.1	0	0.037	0.165	0.061
13.02.2017	0	0	0.042	0.2	0	0.014	0.073	0	0	0.11	0.2	0.117	0	0.029	0.149	0.066
20.02.2017	0	0	0.029	0.2	0	0	0.083	0	0	0.119	0.2	0.132	0	0.018	0.169	0.051

Tydzień	Trapezoidalne															
	ACP	BZW	CCC	CDR	CPS	EUR	GTC	ATT	ING	JSW	KER	LTS	LPP	PGN	PKN	SNS
02.01.2017	0.016	0.014	0.015	0.2	0.015	0	0	0.014	0.014	0.2	0.2	0.036	0.014	0.044	0.2	0.015
09.01.2017	0.03	0	0.019	0.2	0.015	0	0	0	0	0.2	0.2	0.054	0	0.058	0.2	0.024
16.01.2017	0.041	0	0.014	0.2	0	0	0	0	0	0.2	0.2	0.064	0	0.067	0.197	0.017
23.01.2017	0.053	0	0	0.2	0	0	0	0	0	0.184	0.191	0.075	0	0.078	0.2	0.017
30.01.2017	0.065	0	0	0.2	0	0	0	0	0	0.175	0.19	0.084	0	0.073	0.2	0.014
06.02.2017	0.078	0	0	0.2	0	0	0	0	0	0.163	0.174	0.1	0	0.086	0.2	0
13.02.2017	0.085	0	0	0.2	0	0	0	0	0	0.141	0.158	0.118	0	0.097	0.2	0
20.02.2017	0.091	0	0	0.2	0	0	0	0	0	0.128	0.141	0.132	0	0.107	0.2	0

Tabela 7.5: Struktury otrzymanych portfeli dla podejścia klasycznego i rozmytego z wykorzystaniem trapezoidalnych ŚSR.

Tydzień	Gaussowskie															
	ACP	BZW	CCC	CDR	CPS	EUR	GTC	ATT	ING	JSW	KER	LTS	LPP	PGN	PKN	SNS
02.01.2017	0.015	0.014	0.014	0.2	0.014	0	0.014	0.014	0.014	0.2	0.2	0.042	0.014	0.028	0.2	0.016
09.01.2017	0.029	0	0.026	0.2	0	0	0	0	0.015	0.2	0.2	0.057	0	0.042	0.2	0.03
16.01.2017	0.041	0	0.014	0.2	0	0	0	0	0	0.2	0.2	0.066	0	0.052	0.197	0.031
23.01.2017	0.053	0	0	0.2	0	0	0	0	0	0.18	0.193	0.077	0	0.063	0.2	0.034
30.01.2017	0.064	0	0	0.2	0	0	0	0	0	0.171	0.191	0.086	0	0.072	0.194	0.023
06.02.2017	0.074	0	0	0.2	0	0	0	0	0	0.159	0.172	0.102	0	0.085	0.195	0.014
13.02.2017	0.081	0	0	0.2	0	0	0	0	0	0.137	0.157	0.12	0	0.095	0.196	0.014
20.02.2017	0.088	0	0	0.2	0	0	0	0	0	0.133	0.14	0.134	0	0.105	0.2	0

Tydzień	Empiryczne															
	ACP	BZW	CCC	CDR	CPS	EUR	GTC	ATT	ING	JSW	KER	LTS	LPP	PGN	PKN	SNS
02.01.2017	0.014	0	0.015	0.2	0.014	0	0.015	0.014	0.015	0.2	0.2	0.022	0	0.077	0.2	0.014
09.01.2017	0.028	0	0.028	0.2	0	0	0	0	0	0.197	0.2	0.035	0	0.09	0.2	0.022
16.01.2017	0.034	0	0.014	0.2	0	0	0	0	0	0.199	0.2	0.045	0	0.097	0.197	0.014
23.01.2017	0.046	0	0	0.2	0	0	0	0	0	0.176	0.19	0.057	0	0.108	0.2	0.022
30.01.2017	0.044	0	0	0.2	0	0	0	0	0	0.166	0.189	0.067	0	0.12	0.2	0.014
06.02.2017	0.056	0	0	0.2	0	0	0	0	0	0.157	0.17	0.083	0	0.134	0.2	0
13.02.2017	0.065	0	0	0.2	0	0	0	0	0	0.136	0.155	0.099	0	0.145	0.199	0
20.02.2017	0.067	0	0	0.2	0	0	0	0	0	0.128	0.138	0.113	0	0.154	0.2	0

Tabela 7.6: Struktury otrzymanych portfeli dla podejścia rozmytego z wykorzystaniem gaussowskich i empirycznych SŚR.



Klasyczne						
Tydzień	Wartość portfela akcji	Wartość lokaty	Całkowity koszt	Tygodniowa stopa zwrotu	Całkowita stopa zwrotu	
06.01.2017	99717.66	245.9	469.08	-0.0051	-0.0051	
13.01.2017	105263.52	431.75	619.33	0.0561	0.0508	
20.01.2017	104702.04	215.53	770.6	-0.0088	0.0415	
27.01.2017	107850.69	343.61	923.08	0.03	0.0727	
03.02.2017	106662.96	326.99	1074.06	-0.0126	0.0592	
10.02.2017	112253.12	329.41	1224.54	0.0514	0.1136	
17.02.2017	114437.5	262.89	1377.16	0.0176	0.1332	
24.02.2017	117736.82	420.7	1528.88	0.0292	0.1663	

Trapezoidalne						
Tydzień	Wartość portfela akcji	Wartość lokaty	Całkowity koszt	Tygodniowa stopa zwrotu	Całkowita stopa zwrotu	
06.01.2017	97812.43	1886.89	458.06	-0.0076	-0.0076	
13.01.2017	105538.27	161.88	609.25	0.0589	0.0509	
20.01.2017	104224.84	486.71	761.34	-0.0109	0.0395	
27.01.2017	108248.81	41.11	914.06	0.033	0.0738	
03.02.2017	107279.02	145.92	1064.51	-0.0095	0.0636	
10.02.2017	112294.57	166.58	1215.15	0.0459	0.1125	
17.02.2017	115225.44	216.89	1368.15	0.0254	0.1407	
24.02.2017	119593.04	116.75	1518.19	0.0361	0.1819	

Tabela 7.7: Zrealizowane wyniki portfeli dla podejścia klasycznego i rozmytego z wykorzystaniem trapezoidalnych SSR.

Gaussowskie					
Tydzień	Wartość portfela akcji	Wartość lokaty	Całkowity koszt	Tygodniowa stopa zwrotu	Całkowita stopa zwrotu
06.01.2017	997702.31	2008.07	453.01	-0.0074	-0.0074
13.01.2017	105570.01	300.69	605.17	0.0605	0.0527
20.01.2017	104637.31	331.41	757.53	-0.01	0.0421
27.01.2017	108342.4	267.44	910.2	0.0335	0.077
03.02.2017	107471.96	238.14	1060.65	-0.0098	0.0665
10.02.2017	112697.99	158.2	1210.83	0.0468	0.1165
17.02.2017	115536.23	220.1	1363.69	0.0246	0.1439
24.02.2017	119939.84	118.86	1514.56	0.0363	0.1854

Empiryczne					
Tydzień	Wartość portfela akcji	Wartość lokaty	Całkowity koszt	Tygodniowa stopa zwrotu	Całkowita stopa zwrotu
06.01.2017	99180.44	427.82	463.39	-0.0086	-0.0086
13.01.2017	105226.63	273.95	614.99	0.0579	0.0486
20.01.2017	104280.99	201.3	766.27	-0.0112	0.0372
27.01.2017	107958.59	205.51	919.41	0.034	0.0724
03.02.2017	107000.15	249.83	1070.27	-0.0099	0.0618
10.02.2017	112004.97	234.37	1221.07	0.0456	0.1102
17.02.2017	115093.52	156.25	1373.69	0.0257	0.1388
24.02.2017	119456.69	124.9	1523.7	0.0367	0.1806

Tabela 7.8: Zrealizowane wyniki portfeli dla podejścia rozmytego z wykorzystaniem gaussowskich i empirycznych ŚŚR.

## 7.3. Analiza techniczna rynków finansowych

„Analiza techniczna to badanie zachowań rynku, przede wszystkim przy użyciu wykresów, którego celem jest przewidywanie przyszłych trendów cenowych”<sup>11</sup>. Jak wynika z zacytowanej definicji analizy technicznej, podstawę badań analizy technicznej stanowią wykresy. Są to zarówno wykresy przedstawiające bezpośrednio ceny jak i również wykresy przedstawiające różnego rodzaju wskaźniki analizy technicznej. Wykresy te tworzone są najczęściej z ustalonym interwałem czasu (np. godzina, dzień itp.), co wymusza ograniczenie się do wykorzystania tylko kilku wartości cen z danego okresu. Na przykład, tworząc wykres liniowy wykorzystuje się tylko jedną cenę, najczęściej jest to cena zamknięcia, czyli ostatnia w danym okresie. Do tworzenia wykresów słupkowych lub świecowych wykorzystywane są cztery ceny z danego interwału (otwarcia, najmniejsza, największa i zamknięcia).

### 7.3.1. Rozmyte wskaźniki analizy technicznej

Jak pokazano w pierwszym z przykładów, do tworzenia wykresów cenowych również dobrze mogą zostać wykorzystane skierowane świece rozmyte (zobacz rys. 7.1). Wówczas na kształt świecy mają wpływ wszystkie ceny z danego okresu, a nie tylko wybrane. Co więcej, tak skonstruowane świece są skierowanymi liczbami rozmytymi. Korzystając więc z dobrze zdefiniowanej arytmetyki skierowanych liczb rozmytych możliwe staje się tworzenie wskaźników analizy technicznej, które w pewnym sensie zawierać będą informacje o całym przebiegu notowań w danym okresie.

Jednym z najprostszych i zarazem najpopularniejszym wskaźnikiem analizy technicznej jest *średnia krocząca* (ang. *moving average*). Prosta średnia krocząca rzędu  $p$  to zwykła średnia arytmetyczna cen (najczęściej cen zamknięcia) z ostatnich  $p$  okresów. Jej skierowanym rozmytym odpowiednikiem jest również zwykła średnia arytmetyczna, jednak tym razem jest to średnia ze skierowanych świec rozmytych zbudowanych na bazie ostatnich  $p$  okresów. Wartość skierowanej rozmytej prostej średniej ruchomej rzędu  $p$  w chwili  $t$

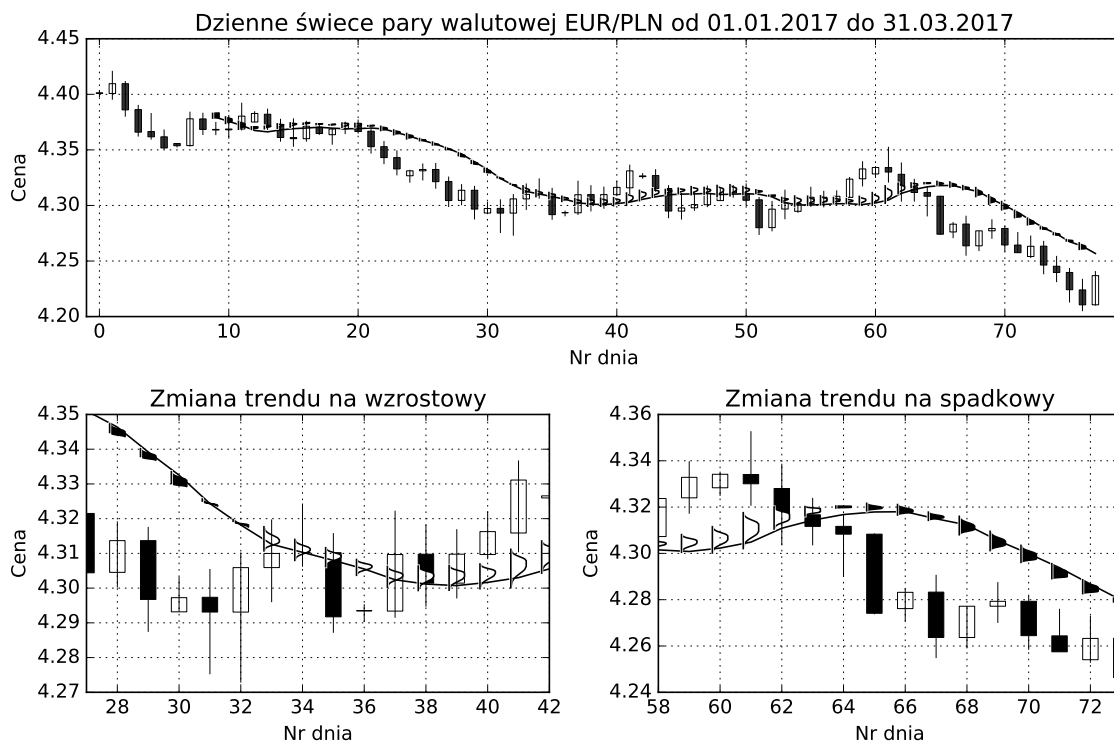
<sup>11</sup>John J. Murphy: *Analiza techniczna rynków finansowych*, przekład Wojciech Madej. Wydanie oryginalne [61].

wyliczana jest zgodnie z wzorem

$$SRSMA(p) = \frac{1}{p} (\tilde{X}_t + \tilde{X}_{t-1} + \dots + \tilde{X}_{t-p+1}), \quad (7.15)$$

gdzie  $\tilde{X}_t$  to skierowana świeca rozmyta skonstruowana dla okresu  $t$ .

Na rys. 7.14 na górnym wykresie przedstawiono przykładową realizację klasycznej (po cenach zamknięcia, linia ciągła) oraz rozmytej (gaussowskie SSR) 9-dniowej średniej ruchomej dla pary walutowej EUR/PLN w okresie od 01.01.2017 do 31.03.2017 na tle notowań w postaci dziennych świec japońskich. W analizie technicznej średnie ruchome najczęściej wykorzystywane są do określania aktualnie panującego trendu na rynku. Jeśli ceny znajdują się nad linią średniej, to na rynku panuje trend wzrostowy, jeżeli poniżej, to trend spadkowy. W przypadku skierowanej rozmytej średniej ruchomej zastosowanie jest takie samo, z tą różnicą, że trend może być określony na dwa sposoby. W pierwszym, tworzy się ostrą linię średniej ruchomej poprzez wyostrenie skierowanych liczb rozmytych, a interpretacja trendu odbywa się w taki sam sposób jak w przypadku kla-

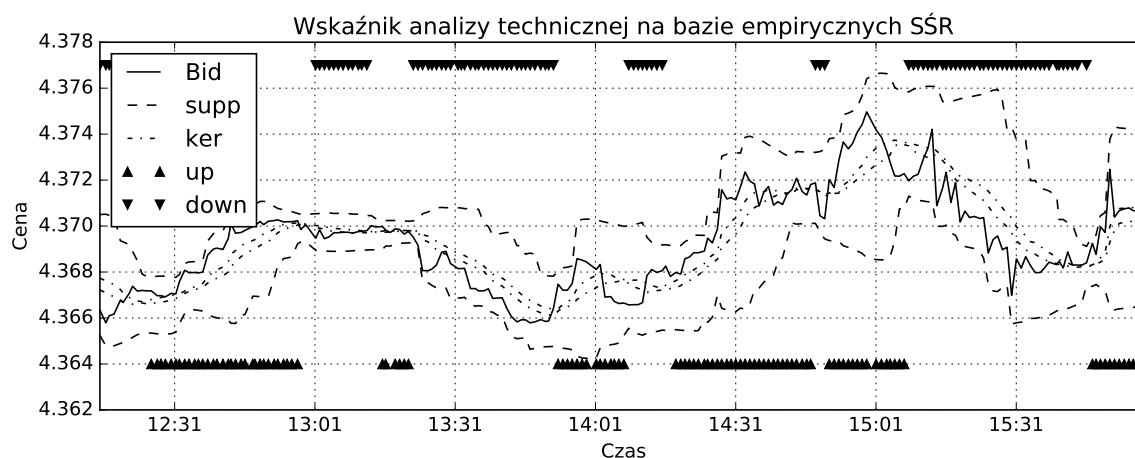


Rysunek 7.14: Realizacja skierowanej rozmytej średniej ruchomej.

szym. W drugim sposobie, trend wyznaczony jest za pomocą orientacji skierowanych liczb rozmytych. Jeżeli orientacja jest dodatnia to panuje trend wzrostowy, jeśli ujemna to spadkowy. Na rys. 7.14 na dolnych wykresach, przybliżone zostały miejsca, w który następuje zmiana trendu. Warto zwrócić uwagę na fakt, że sygnalizacja zmiany trendu poprzez liczby rozmyte następuje w sposób płynny w tym sensie, że szerokość liczby rozmytej przed zmianą trendu ulega zwężeniu, a po zmianie rozszerzeniu. W analogiczny sposób jak średnia ruchoma mogą zostać rozmyte inne wskaźniki analizy technicznej.

### 7.3.2. Skierowana świeca rozmyta jako wskaźnik

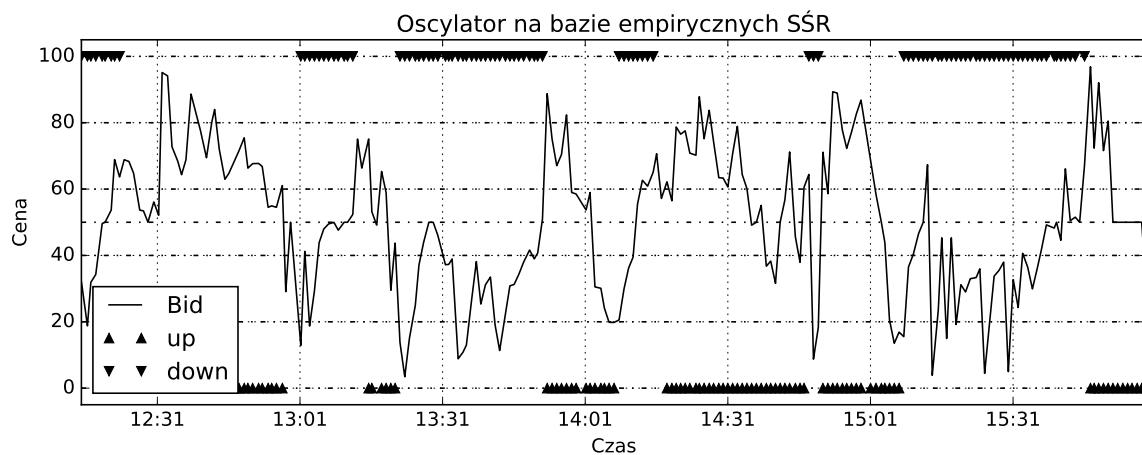
Innym sposobem wykorzystania skierowanych świec rozmytych w analizie technicznej jest użycie metod konstrukcji skierowanych świec rozmytych bezpośrednio jako wskaźnika analizy technicznej. W tym celu obliczanie skierowanej świecy rozmytej dokonywane jest na ruchomym oknie obserwacji o określonym rozmiarze. Rozmiar okna może być określony poprzez liczbę obserwacji (np. ostatnie 100 notowań) jak również poprzez przedział czasu (np. ostatnie 10 minut). Na rys. 7.15 przedstawiono wskaźnik analizy technicznej skonstruowany na bazie empirycznych skierowanych świec rozmytych przy użyciu okna długości 15 minut, rysowany na wykresie co 1 minutę dla notowań pary walutowej EUR/PLN z dnia 11.01.2017 w godzinach od 12:00 do 16:00. Na wykresie umieszczono linie wyznaczające nośnik (supp) oraz jądro (ker) skierowanej liczby rozmytej, linię odpowiadającą cenie



Rysunek 7.15: Wskaźnik analizy technicznej na bazie empirycznych ŚSR.

kupna (Bid) oraz za pomocą odpowiednio zwróconych trójkątów informację o skierowaniu liczby rozmytej w danym momencie. Jak można zauważyć, cena kupna porusza się zawsze w obszarze ograniczonym przez nośnik (wynika to wprost z definicji empirycznej SŚR).

Informacje jakie można odczytać z tak skonstruowanego wskaźnika, to na przykład informacja, w jakim stopniu aktualna cena kupna należy do skonstruowanego zbioru rozmytego lub jakie poziomy przynależności przebija i z której strony. W celu zwiększenia czytelności poziomów na jakich znajduje się cena, wskaźnik ten można przedstawić w postaci oscylatora. W tym celu należy przeskalować otrzymane wartości do ustalonego zakresu np. od 0 do 100, zachowując przy tym kształty funkcji  $f$  i  $g$ . Na rys. 7.16 przedstawiony jest wskaźnik z rys. 7.15 w postaci oscylatora o wartościach w przedziale  $[0, 100]$  z wykorzystaniem skalowania liniowego między poziomami przynależności z krokiem 0.1.



Rysunek 7.16: Wskaźnik na bazie empirycznych SŚR w postaci oscylatora.

# Rozdział 8

## Podsumowanie i wnioski

Problematyka badawcza niniejszej rozprawy związana jest z modelowaniem finansowych szeregów czasowych wysokiej częstotliwości z uwzględnieniem niepewności związanej zarówno ze specyfiką danych wysokiej częstotliwości jak również z otoczeniem, z którego dane pochodzą. Głównym celem pracy było opracowanie metod, modeli i algorytmów umożliwiających modelowanie i symulację finansowych szeregów czasowych z wykorzystaniem modelu skierowanych liczb rozmytych jako modelu niepewności.

W ramach rozprawy opracowano model skierowanych świec rozmytych, pozwalający na reprezentowanie finansowych szeregów czasowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych. Na podstawie przedstawionych przykładów skierowanych świec rozmytych w rozdziałach 4.3 i 4.4 oraz uzyskanych wynikach w praktycznych zagadnieniach z rozdziału 7 można stwierdzić, że zaproponowane metody konstrukcji skierowanych świec rozmytych pozwalają na zmniejszenie rozmiaru danych poprzez zmniejszenie ich częstotliwości, przy jednoczesnym ograniczeniu utraty informacji o charakterze oryginalnego szeregu. Wobec czego można uznać, że pierwsza z tez niniejszej rozprawy mówiąca, iż *możliwe jest przedstawienie danych finansowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych zachowując przy tym więcej informacji o danym szeregu czasowym (więcej niż w przypadku klasycznych reprezentacji notowań giełdowych takich jak: ceny zamknięcia, słupki OHLC czy świece japońskie)*, została wykazana.

Wykorzystanie modelu skierowanych liczb rozmytych wraz z rozwinięciem go o pojęcie skierowanej rozmytej zmiennej losowej pozwoliło autorowi na opracowanie w ramach rozprawy metodologii w pełni rozmytych stochastycznych modeli szeregów czasowych bazujących na skierowanych liczbach rozmytych, w postaci klasycznych równań. W rozdziale 5.2 omówiono aspekty estymacji współczynników oraz prognozowania w takich modelach na przykładzie skierowanego rozmytego modelu regresji liniowej oraz modelu autoregresyjnego. Opracowaną metodę estymacji współczynników oparto na metodzie najmniejszych kwadratów. Bazując na przedstawionych przykładach, łatwo jest wyobrazić sobie w pełni rozmyte wersje innych dobrze znanych modeli szeregów czasowych, jak na przykład modele klasy GARCH. Ponadto, wprowadzone przez autora w rozdziale 6.2 pojęcie skierowanej rozmytej zmiennej losowej wraz z opracowanym algorytmem generowania skierowanych rozmytych liczb pseudolosowych o rozkładzie normalnym, pozwala na stochastyczną analizę reszt konstruowanych modeli rozmytych oraz wykorzystanie ich w zagadnieniach symulacji. Zakładając skierowany rozmyty rozkład błędów modelu możliwe staje się też zaadaptowanie metody największej wiarygodności do celów estymacji współczynników w modelach rozmytych. Powołując się na przykłady z rozdziału 7 oraz otrzymane w nich wyniki można stwierdzić, że druga z tez niniejszej rozprawy mówiąca, iż *skonstruowane na bazie skierowanych liczb rozmytych w pełni rozmyte modele szeregów czasowych pozwalają na efektywne modelowanie i symulację finansowych szeregów czasowych oraz możliwe jest wykorzystanie ich w praktycznych zagadnieniach inżynierii finansowej, takich jak np. wycena instrumentów finansowych*, została potwierdzona.

Warto zaznaczyć, że opracowane w ramach rozprawy modele, metody i algorytmy zostały również w pełni zaimplementowane w języku programowania Python. Przy ich implementacji wykorzystano pakiety do obliczeń naukowych w Pythonie m.in. takie jak NumPy, SciPy, Pandas, Statsmodels oraz Matplotlib. Powstały kod źródłowy stanowiący swego rodzaju bibliotekę do obliczeń z wykorzystaniem skierowanych liczb rozmytych dołączony został w osobnych plikach do rozprawy w wersji elektronicznej.



# Bibliografia

- [1] Atsalakis G.: *Exchange rate forecasting by neuro-fuzzy techniques*. Journal of Financial Decision Making, 1(2), pp. 15–26, 2006.
- [2] Aumann R.J.: *Integrals of set-valued functions*. J. Math. Anal. Appl. 12, pp. 1–12, 1965.
- [3] Box G. E. P., Jenkins G.: *Time Series Analysis. Forecasting and Control*. HoldenDay, San Francisco, 1970.
- [4] Brooks C.: *Introductory Econometric for Finance*. Second edition, Cambridge, New York, 2008.
- [5] Brzeszczyński J., Kelm R.: *Ekonometryczne modele rynków finansowych. Modele kursów giełdowych i kursów walutowych*. WIG-Press, Warszawa, 2002.
- [6] Chen S-M.: *Forecasting enrollments based on fuzzy time series*. Fuzzy Sets and Systems, vol. 81, issue 3, pp. 311–319, 1996.
- [7] Chen S-M.: *Forecasting enrollments based on high-order fuzzy time series*. Cybernetics and Systems, vol. 33, issue 1, pp. 1–16, 2010.
- [8] Chen S-M.: *Handling forecasting problems based on high-order fuzzy logical relationships*. Expert Systems with Applications, vol. 38, issue 4, pp. 3857–3864, 2011.
- [9] Colubi A., Fernandez-Garcia C., Gil M.A.: *Simulation of Fuzzy Random variables: An Empirical Approach to Statistical/Probabilistic Studies with Fuzzy Experimental Data* IEEE Transactions On Fuzzy Systems, 10, pp. 384–390, 2002.
- [10] Colubi A., Coppi R., D'urso P., Gil M.A.: *Statistics with fuzzy random variables*. METRON - International Journal of Statistics, LXV(3), pp. 277–303, 2007.
- [11] Czogała E., Pedrycz W.: *Elementy i metody teorii zbiorów rozmytych*. PWN, 1985.
- [12] Diamond P.: *Fuzzy least squares*. Information Sciences 46(3), pp. 141-157, 1988.
- [13] Duan J.-C.: *The GARCH Option Pricing Model*. Mathematical Finance, 5, pp. 13–32, 1995.

- [14] Dubois D., Prade H.: *Operations on fuzzy numbers*. Int. J. System Science, 9, pp. 576-578, 1978.
- [15] Enders W.: *Applied Econometric Time Series*. Second edition, Wiley, New York, 2004.
- [16] Engle R. F., Russell J. R.: *Analysis of high frequency financial data*. In Handbook of Financial Econometrics (L. P. Hansen, Y. Ait-Sahalia, eds.), North Holland, Amsterdam, 2005.
- [17] Fama E.F.: *Efficient Capital Markets: A Review of Theory and Empirical Work*. Journal of Finance, 25 (2), pp. 383-417, 1970.
- [18] Feng Y, Hu L., Shu H.: *The Variance and Covariance of Fuzzy Random Variables and Their Applications*. Fuzzy Sets and Systems 120, pp. 487-497, 2001.
- [19] Gil M.A.: *Fuzzy random variables: development and state of the art*. Mathematics of Fuzzy Systems, Linz Seminar on Fuzzy Set Theory, Linz Austria, 2004.
- [20] Gil M.A., López-Díaz M., Ralescu D.A.: *Overview on the development of fuzzy random variables*. Fuzzy Sets and Systems, 157, pp.2546-2557, 2006.
- [21] Goetschel R., Voxman W.: *Elementary fuzzy calculus*. Fuzzy Sets and Systems 18, pp. 31-43, 1986.
- [22] Ishibuchi H.: *Fuzzy regression analysis*. Fuzzy Theory and Systems, 4, pp. 137-148, 1992.
- [23] Jajuga K. Jajuga T.: *Inwestycje. Instrumenty finansowe, ryzyko finansowe, inżynieria finansowa*. PWN, Warszawa, 1996.
- [24] Jakubowski J., Sztencel R.: *Wstęp do teorii prawdopodobieństwa*. Wydanie II, Script, Warszawa, 2001.
- [25] Jang J-S. R.: *ANFIS: Adaptive-network-based fuzzy inference systems*. IEEE Transactions on Systems, Man and Cybernetics, 23(3), pp. 665-685, 1993.
- [26] Jarosz P., Burczyński T.: *Coupling of Immune Algorithms and Game Theory in Multiobjective Optimization*. In: L. Rutkowski et al. (Eds.), ICAISC 2010, LNAI 6114, Part II, pp. 500-507, Springer-Verlag, Berlin Heidelberg, 2010.
- [27] Jarosz P., Burczyński T.: *Biologically-inspired Methods and Game Theory in Multi-criterion Decision Processes*. In: P. Bouvry et al. (Eds.), INTELLIGENT DECISION SYSTEMS IN LARGE-SCALE DISTRIBUTED ENVIRONMENTS, Studies

- in Computational Intelligence 362, pp. 101–124, Springer-Verlag, Berlin Heidelberg, 2011.
- [28] Jarosz P.: *Sztuczne systemy immunologiczne i teoria gier w optymalizacji wielokryterialnej*. Rozprawa doktorska, IPI PAN, Warszawa, 2011.
- [29] Kacprzyk J.: *Zbiory rozmyte w analizie systemowej*. PWN, 1986.
- [30] Kahraman C., Beskese A., Tunc Bozbura F.: *Fuzzy Regression Approaches and Applications*. StudFuzz 201, pp. 589–615, Springer-Verlag, Berlin Heidelberg, 2006.
- [31] Kay S.M.: *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [32] Klir G. J.: *Fuzzy arithmetic with requisite constraints*. Fuzzy Sets and Systems 91, pp. 165–175, 1997.
- [33] Kosiński W., Piechór K., Prokopowicz P., Tyburek K.: *On algorithmic approach to operations on fuzzy numbers*. In: T. Burczyński, W. Cholewa, (Eds.), *Methods of Artificial Intelligence in Mechanics and Mechanical Engineering*, PACM, pp. 95-98, Gliwice, 2001.
- [34] Kosiński W., Prokopowicz P., Ślęzak D.: *Drawback of fuzzy arithmetic - New intuitions and propositions*. In: T. Burczyński, W. Cholewa, W. Moczulski (Eds.), *Proc. Methods of Artificial Intelligence*, PACM, pp. 231-237, Gliwice, 2002.
- [35] Kosiński W., Prokopowicz P., Ślęzak D.: *On algebraic operations on fuzzy numbers*. In: M. Kłopotek, S.T. Wierzchoń, K. Trojanowski, (Eds.), *Intelligent Information Processing and Web Mining*, Proc. Int. Symp. IIS: IIPWM'03, Zakopane, Poland, Heidelberg: Physica Verlag, pp. 353-362, 2003.
- [36] Kosiński W., Prokopowicz P., Ślęzak D.: *Ordered fuzzy numbers*. Bull. Polish Acad. Sci., Ser. Sci. Math., Vol. 51, No. 3, pp. 327-338, 2003.
- [37] Kosiński W., Prokopowicz P.: *Algebra liczb rozmytych*. Matematyka Stosowana. Matematyka dla Społeczeństwa, Vol. 5, No. 46, s. 37-63, 2004.
- [38] Kosiński W.: *On soft computing and modelling*. Image Processing Communications 11(1), pp. 71-82, 2006.
- [39] Kosiński W., Prokopowicz P., Kacprzak D.: *Fuzziness representation of dynamic changes by ordered fuzzy numbers*. In: Seising, R. (ed.) *Views on Fuzzy Sets and Systems*. STUDFUZZ, vol. 243, pp. 485–508, Springer Heidelberg, 2009.

- [40] Kosiński W., Frischmuth K., Wilczyńska-Sztyma D.: *A new Fuzzy Approach to Ordinary Differential Equations*. In: L. Rutkowski et al. (Eds.), ICAISC 2010, Part I, LNAI 6113, Springer-Verlag Berlin Heidelberg, pp. 120-127, 2010.
- [41] Körner R.: *Linear Models with Random Fuzzy Variables*. Technischen Universität Bergakademie Freiberg, 1997.
- [42] Kruse R., Meyer K.D.: *Statistics with Vague Data*. D. Reidel Pub. Co., 1987.
- [43] Kwakernaak H.: *Fuzzy Random Variables - I. Definitions and Theorems*. Information Sciences 15(1), pp. 1-29, 1978.
- [44] Kwakernaak H.: *Fuzzy Random Variables - II. Algorithms and Examples for the Discrete Case*. Information Sciences 17(3), pp. 253-278, 1979.
- [45] Lin Ch.-T., Lee C.S.G.: *Neural fuzzy systems. A neuro-fuzzy synergism to intelligent systems*. Prentice Hall PTR, 1996.
- [46] Liu B., Liu Y.-K.: *Expected value of fuzzy variable and fuzzy expected value models*. IEEE Transaction on Fuzzy Systems 10, pp. 445-450, 2002.
- [47] Liu Y.-K., Liu B.: *Fuzzy Random Variables: A Scalar Expected Value Operator*. Fuzzy Optimization and Decision Making 2(2), pp. 143-160, 2003.
- [48] Łachwa, A.: *Rozmyty świat zbiorów, liczb, relacji, faktów, reguł i decyzji*. AOW EXIT, Warszawa, 2001.
- [49] Łęski, J.: *Systemy neuronowo-rozmyte*. WNT, Warszawa, 2008.
- [50] Lo A.W., MacKinlay A.C.: *Stock market prices do not follow random walks: Evidence from a simple specification test*. The Review of Financial Studies, Volume 1, Issue 1, pp. 41-66, 1988.
- [51] Lubiano M.A, Gil M.A., López-Díaz M., López M.T.: *The  $\vec{\lambda}$ -mean squared dispersion associated with a fuzzy random variable*. Fuzzy Sets and Systems 111, pp. 307-317, 2000.
- [52] Magiera R.: *Modele i metody statystyki matematycznej. Część I: Rozkłady i symulacja stochastyczna*. Wydanie drugie rozszerzone, Oficyna Wydawnicza GiS, Wrocław, 2005.
- [53] Mamdani E. H., Assilian S.: *An experiment in linguistic synthesis with a fuzzy logic controller*. Int. J. Man-machine Studies, Vol. 7, pp. 1-13, 1975.
- [54] Markowitz H.M.: *Portfolio selection*. J. Financ. 7, no. 1, pp. 77-91, 1952.

- [55] Markowitz H.M.: *Portfolio selection: Efficient diversification of investments*. Cowles Foundation for Research in Economics at Yal University. Monograph 16, John Wiley & Sons Inc., New York, 1959.
- [56] Markowitz H.M.: *Mean-variance analysis in portfolio choice and capital markets*. Basil Blackwell, Oxford, 1987.
- [57] Matheron G.: *Random Sets and Integral Geometry*. Wiley, New York, 1975.
- [58] Marszałek A., Burczyński T.: *Financial fuzzy time series models based on ordered fuzzy numbers*. In: W. Pedrycz, S.-M. Chen (Eds.), *Time Series Analysis, Model & Applications*, ISRL 47, pp. 77–95, Springer-Verlag, Berlin Heidelberg, 2013.
- [59] Marszałek A., Burczyński T.: *Modelling financial high frequency data using ordered fuzzy numbers*. W: L. Rutkowski et al. (Eds.), *ICAISC 2013*, LNAI 7894, Part I, pp. 345–352, Springer-Verlag, Berlin Heidelberg, 2013.
- [60] Marszałek A., Burczyński T.: *Modeling and forecasting financial time series with ordered fuzzy candlesticks*. *Information Science*, Vol. 273, pp. 144–155, 2014.
- [61] Murphy J.J: *Technical Analysis of the Financial Markets*. New York Institute of Finance, New York, 1999.
- [62] Nahmias S.: *Fuzzy variables*. *Fuzzy Sets and Systems*, Vol. 1, pp. 97–110, 1978.
- [63] Nison S.: *Japanese Candlestick Charting Techniques*. New York Institute of Finance, New York, 1991.
- [64] Ozawa K., Watanabe T., Kanke M.: *Forecasting Fuzzy Time Series with Fuzzy AR Model*. In: *Proceedings, the 20th International Conference on Computers & Industrial Engineering*, pp. 105–108, 1996.
- [65] Papoulis A.: *Prawdopodobieństwo, zmienne losowe i procesy stochastyczne*. WNT, Warszawa, 1972.
- [66] Piegat A.: *Modelowanie i sterowanie rozmyte*. AOW EXIT, Warszawa, 1999.
- [67] Prokopowicz P.: *Flexible and Simple Methods of Calculations on Fuzzy Numbers with the Ordered Fuzzy Numbers Model*. In: L. Rutkowski et al. (Eds.), *ICAISC 2013*, LNAI 7894, Part I, pp. 365–375, Springer-Verlag, Berlin Heidelberg, 2013.
- [68] Prokopowicz P., Pedrycz W.: *The Directed Compatibility Between Ordered Fuzzy Numbers - A Base Tool for a Direction Sensitive Fuzzy Information Processing*. *Proc. of ICAISC 2015*, Part I, LNAI, vol. 9119 pp. 249–259. Springer International Publishing Switzerland, 2015

- [69] Prokopowicz P.: *The Directed Inference for the Kosinski's Fuzzy Number Model*. Proc. of AECIA 2015, Part III, AISC, vol. 427 pp. 493–503. Springer International Publishing Switzerland, 2016
- [70] Puri L., Ralescu D.A.: *Fuzzy Random Variables*. Journal of Mathematical Analysis and Applications 114, pp. 409–422, 1986.
- [71] Puri M.L., Ralescu D.A.: *The concept of normality for fuzzy random variables*. Annals of Probability, 13, pp. 1372–1379, 1985.
- [72] Rutkowska D., Piliński M., Rutkowski L.: *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*. PWN, Warszawa, 1997.
- [73] Sanchez E.: *Solutions of fuzzy equations with extended operations*. Fuzzy Sets and Systems 12, pp. 237–248, 1984.
- [74] Sharpe W.F.: *The Sharpe Ratio*. The Journal of Portfolio Management. 21 (1) pp. 49–58, 1994.
- [75] Song Q., Chissom B.S.: *Fuzzy time series and its models*. Fuzzy Sets and Systems, vol. 54, issue 3, pp. 269–277, 1993.
- [76] Song Q., Chissom B.S.: *Forecasting enrollments with fuzzy time series — Part I*. Fuzzy Sets and Systems, vol. 54, issue 1, pp. 1–9, 1993.
- [77] Song Q., Chissom B.S.: *Forecasting enrollments with fuzzy time series — Part II*. Fuzzy Sets and Systems, vol. 62, issue 1, pp. 1–8, 1994.
- [78] Sharma S., Chouhan M.: *A Review: Fuzzy Time Series Model for Forecasting*. IJAST, vol. 2, issue 3, pp. 32–35, 2014.
- [79] Singh P.: *A brief review of modeling approaches based on fuzzy time series*. Int. J. Mach. Learn. & Cyber., pp. 1–24, 2015.
- [80] Takagi T., Sugeno M.: *Fuzzy identification of systems and its applications to modeling and control*. IEEE Trans. Systems, Man, and Cybernetics, Vol. 15, pp. 116–132, 1985.
- [81] Tanaka H., Uejima S., Asai K.: *Linear regression analysis with fuzzy model*. IEEE Transactions on Systems, Man and Cybernetics 12(6), pp. 903–907, 1982.
- [82] Tanaka H.: *Fuzzy data analysis by possibilistic linear models*. Fuzzy Sets and Systems, vol. 24, pp. 363–375, 1987.
- [83] Tanaka H., Hayashi I., Watada J.: *Possibilistic Linear Regression Analysis for Fuzzy Data*. Eur. J. Oper. Res., vol. 40, pp. 389–396, 1989.

- [84] Tanaka K.: *An introduction to fuzzy logic for practical applications*. Springer, 1997.
- [85] Tsay R.S.: *Analysis of Financial Time Series*. Second edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.
- [86] Tseng F.M., et al.: *Fuzzy ARIMA model for forecasting the foreign exchange market*. Fuzzy Sets and Systems 118(1), pp. 9–19, 2001.
- [87] Tseng F.M., Tzeng G.H.: *A fuzzy seasonal ARIMA model for forecasting*. Fuzzy Sets and Systems 126(3), pp. 367–376, 2002.
- [88] Ubale A.B., Sananse S.L.: *Fuzzy Regression Model and Its Application: A Review*. IJRSET, vol. 4, issue 11, 2015.
- [89] Wagenknecht M.: *On the approximate treatment of fuzzy arithmetics by inclusion, linear regression and information content estimation*. W: Zbiory rozmyte i ich zastosowania, J. Chojcan, J. Łęski (red.), Wyd. Politechniki Śląskiej, pp. 291–310, Gliwice, 2001.
- [90] Wagenknecht M., Hampel R., Schneider V.: *Computational aspects of fuzzy arithmetic based on Archimedean  $t$ -norms*. Fuzzy Sets and Systems 123, pp. 49–62, 2001.
- [91] Watada J.: *Fuzzy time series analysis and forecasting of sales volume*. In: Kacprzyk J., Fedrizzi M. (eds.) *Fuzzy Regression Analysis*, pp. 211–227. Omnitech Press, Physica-Verlag, Warsaw, Heidelberg, 1992.
- [92] Watada J.: *Possibilistic Time-series Analysis and Its Analysis of Consumption*. In: Dubois D., Prade H. Yager R.R. (eds.), pp. 187–217, John Wiley & Sons, INC, 1996.
- [93] Wesołowska M.Z.: *O szeregach czasowych i ich prognozowaniu*. Wydawnictwo SGGW, Warszawa 2008.
- [94] Wilczyńska-Sztyrna D.: *Idea skierowanych liczb rozmytych, przykłady interpretacji*. XI International PhD Workshop OWD 2009, s. 17–20 October 2009.
- [95] Wu H.C.: *Statistical confidence intervals for fuzzy data*. Expert Systems with Applications, 36, pp. 2670–2676, 2009.
- [96] Yager R.R.: *A characterisation of the extension principle*. Fuzzy Sets and Systems, 18, pp. 205–217, 1986.
- [97] Yager R.R., Filev D.P.: *Podstawy modelowania i sterowania rozmytego*. WNT, Warszawa, 1995.
- [98] Yan B., Zivot E.: *Analysis of High-Frequency Financial Data with S-PLUS*. Working Papers UWEC-2005-03, University of Washington, Department of Economics, 2003.

- 
- [99] Zadeh L.A.: *Fuzzy sets*. Information and Control, 8, pp. 338–353, 1965.
- [100] Zadeh L.A.: *The concept of a linguistic variable and its application to approximate reasoning*. Part I, Inform. Sci., 8, pp. 199–249, 1975.
- [101] Zadeh L.A.: *Fuzzy set as the basis for the theory of possibility*. Fuzzy Sets and Systems, 1, pp. 3–28, 1978.
- [102] Zadeh L.A.: *The role of fuzzy logic in the management of uncertainty in expert systems*. Fuzzy Sets and Systems, 11, pp. 199–227, 1983.



# Spis rysunków

2.1. Nośnik, jądro i $\alpha$ -przekrój wypukłego zbioru rozmytego $A$ . . . . .	10
2.2. Trójkątna funkcja przynależności . . . . .	11
2.3. Trapezowa funkcja przynależności . . . . .	11
2.4. Gaussowska i uogólniona dzwonowa funkcja przynależności . . . . .	12
2.5. Sigmoidalna funkcja przynależności . . . . .	12
2.6. Rozmyty singleton . . . . .	13
2.7. Funkcja przynależności typu S . . . . .	13
2.8. Suma liczby rozmytej i liczby przeciwnej do niej . . . . .	17
2.9. Iloczyn liczby rozmytej i liczby odwrotnej do niej . . . . .	18
2.10. Iloczyn liczb rozmytych typu $L-R$ . . . . .	20
3.1. Graficzna interpretacja skierowanej liczby rozmytej . . . . .	24
3.2. Pozytywnie i negatywnie zorientowane skierowane liczby rozmyte . . . . .	25
3.3. Suma skierowanej liczby rozmytej i liczby przeciwnej do niej . . . . .	26
3.4. Iloczyn skierowanej liczby rozmytej i liczby odwrotnej do niej . . . . .	26
3.5. Dyskretyzacja skierowanej liczby rozmytej . . . . .	28
3.6. Niewłaściwe skierowane liczby rozmyte . . . . .	31
4.1. Szkic koncepcji skierowanej świecy rozmytej . . . . .	35
4.2. Przykłady trapezoidalnej skierowanej świecy rozmytej . . . . .	39
4.3. Przykłady gaussowskiej skierowanej świecy rozmytej . . . . .	41
4.4. Przykłady empirycznej skierowanej świecy rozmytej . . . . .	41
4.5. Dwa szeregi czasowe o identycznych świecach japońskich . . . . .	42
4.6. Realizacje SŚR dla szeregów czasowych o identycznych świecach japońskich	43

4.7. Realizacje SŚR bez oraz z informacją o wolumenie . . . . .	44
4.8. Realizacje SŚR bez oraz z informacją o spreadzie . . . . .	45
5.1. Skierowany rozmyty model regresji dla danych Ishibuchi - brak rozmycia . . . . .	56
5.2. Skierowany rozmyty model regresji dla danych Ishibuchi - rozmycie $y_t$ . . . . .	57
5.3. Skierowany rozmyty model regresji dla danych Ishibuchi - rozmycie $y_t$ plus skierowanie . . . . .	57
5.4. Skierowany rozmyty model regresji dla danych Ishibuchi - rozmycie $y_t$ plus skierowanie . . . . .	58
6.1. $\alpha$ -przekrój rozmytej zmiennej losowej: a) Kwakernaak, b) Puri i Ralescu . . . . .	67
6.2. Rozmyte parametry $\tilde{\mu}$ oraz $\tilde{\sigma}^2$ rozkładu normalnego . . . . .	81
6.3. Rozmyta wartość oczekiwana i rozmyta wariancja zmiennej $\tilde{\xi}$ . . . . .	81
6.4. Pięć losowych realizacji zmiennej $\tilde{\xi}$ . . . . .	82
6.5. Estymatory rozmytej wartości oczekiwanej i rozmytej wariancji z próby . . . . .	82
6.6. Znormalizowane histogramy oraz teoretyczne funkcje gęstości prawdopodobieństwa zmiennych losowych $f_{\tilde{\xi}_i}(x)$ , $g_{\tilde{\xi}_i}(x)$ dla $x = 0.5$ . . . . .	83
7.1. Dienne świece notowań indeksu WIG20 . . . . .	86
7.2. Dienne rozmyte logarytmiczne stopy zwrotu notowań indeksu WIG20 . . . . .	87
7.3. Wykresy PACF dla rozmytych szeregów logarytmicznych stóp zwrotu . . . . .	88
7.4. Współczynniki modelu SRAR(2) opartego na gaussowskich SŚR . . . . .	89
7.5. Rozmyte szeregi czasowe reszt $\tilde{\varepsilon}_t$ . . . . .	90
7.6. Realizacje parametrów $\tilde{\mu}$ . . . . .	91
7.7. Realizacje parametrów $\tilde{\sigma}^2$ . . . . .	91
7.8. Rozmyte wartości oczekiwane $\mathbb{E}[\tilde{\varepsilon}_t]$ . . . . .	92
7.9. Rozmyte wariancje $\text{Var}[\tilde{\varepsilon}_t]$ . . . . .	92
7.10. Realizacja rozmytej ceny opcji europejskiej $\tilde{C}$ . . . . .	95
7.11. Notowania spółki KGHM w okresie 20-24.01.2017 . . . . .	99
7.12. Empiryczna SŚR dla szeregu tworzenia się stóp zwrotu spółki KGHM w okresie 20-24.01.2017 . . . . .	99
7.13. Fronty Pareto optymalne dla podejścia klasycznego i rozmytego z wykorzystaniem trapezoidalnych SŚR otrzymane w dniu 09.01.2017 . . . . .	102

7.14. Realizacja skierowanej rozmytej średniej ruchomej . . . . .	111
7.15. Wskaźnik analizy technicznej na bazie empirycznych SŚR . . . . .	112
7.16. Wskaźnik na bazie empirycznych SŚR w postaci oscylatora . . . . .	113

# Spis tabel

5.1. Zbiór danych Ishibuchi z losowym błędem pomiarowym. . . . .	55
6.1. Średnie odległości między teoretycznymi wartościami parametrów $\mathbb{E}$ , $\text{Var}$ , $\tilde{\mu}$ , $\tilde{\sigma}^2$ oraz $p$ , a ich estymatorami w zależności od rozmiaru próby. . . . .	83
7.1. Wartości parametrów i charakterystyk dla rozkładów reszt $\tilde{\varepsilon}_t$ . . . . .	91
7.2. Wycena europejskiej opcji kupna na indeks WIG20 z terminem wygaśnięcia 17.03.2017. . . . .	96
7.3. Wskaźniki oceny jakości skonstruowanych portfeli na bazie zrealizowanych tygodniowych stóp zwrotu. . . . .	103
7.4. Wartości oczekiwane i odchylenia standardowe tygodniowych stóp zwrotu dla spółek z indeksu WIG30. . . . .	105
7.5. Struktury otrzymanych portfeli dla podejścia klasycznego i rozmytego z wykorzystaniem trapezoidalnych SŚR. . . . .	106
7.6. Struktury otrzymanych portfeli dla podejścia rozmytego z wykorzystaniem gaussowskich i empirycznych SŚR. . . . .	107
7.7. Zrealizowane wyniki portfeli dla podejścia klasycznego i rozmytego z wy- korzystaniem trapezoidalnych SŚR. . . . .	108
7.8. Zrealizowane wyniki portfeli dla podejścia rozmytego z wykorzystaniem gaussowskich i empirycznych SŚR. . . . .	109

# Skierowane liczby rozmyte w modelowaniu i symulacji finansowych szeregów czasowych

## Streszczenie

Obszar badawczy niniejszej rozprawy koncentruje się na zagadnieniach związanych z modelowaniem niepewności oraz modelowaniem finansowych szeregów czasowych. Dane finansowe wysokiej częstotliwości to obserwacje zmiennych finansowych, takich jak notowania akcji, kontraktów futures lub par walutowych, notowanych z częstotliwością wyższą niż dzienna. Dane finansowe wysokiej częstotliwości posiadają unikalne cechy nieobecne w danych mierzonych w niższych częstotliwościach. Ponadto, ogromna ilość informacji znajduje odzwierciedlenie na rynku finansowym. Można zatem stwierdzić, że dane giełdowe powinny być rozpatrywane z uwzględnieniem niepewności.

Głównym celem pracy było opracowanie modeli, metod i algorytmów umożliwiających modelowanie i symulację finansowych szeregów czasowych z wykorzystaniem koncepcji skierowanej liczby rozmytej jako modelu niepewności. Wykazano, że możliwa jest reprezentacja finansowych szeregów czasowych przy użyciu skierowanych liczb rozmytych przy jednoczesnym utrzymaniu większej ilości informacji o zmianach cen w założonym przedziale czasowym w porównaniu do powszechnie stosowanych wykresów cenowych. Co więcej, dzięki dobrze zdefiniowanej arytmetyce skierowanych liczb rozmytych oraz opracowanej przez autora koncepcji skierowanej rozmytej zmiennej losowej, możliwe jest skonstruowanie w pełni rozmytych modeli szeregów czasowych i wykorzystanie ich do zagadnień symulacji.

Rozprawa składa się z ośmiu rozdziałów. W pierwszym z nich przedstawione są motywacje oraz sformułowane zostały cele i tezy pracy. W drugim i trzecim rozdziale omówiono podstawowe pojęcia dotyczące zbiorów rozmytych oraz skierowanych liczb rozmytych. Rozdział czwarty zawiera opis zaproponowanych przez autora metod reprezentacji finansowych szeregów czasowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych. W rozdziale piątym przedstawiono metodologię w pełni rozmytych modeli szeregów czasowych. Nowe pojęcie rozmytej zmiennej losowej dla skierowanych liczb rozmytych przedstawiono w rozdziale szóstym. Rozdział siódmy poświęcony jest praktycznym przykładom zastosowań proponowanych modeli. W ostatnim rozdziale przedstawiono podsumowanie pracy i przedstawiono wnioski.

# Ordered fuzzy numbers in modelling and simulation of financial time series

## Abstract

The research area of this dissertation is centered on the issues in the field of uncertainty modelling and analysis of financial time series. High-frequency financial data are the observations on financial variables such as quotations of shares, futures or currency pairs, quoted at higher than the daily time scale. High-frequency financial data possess unique features absent in data measured at lower frequencies. Moreover, an enormous amount of information is reflected in the financial market. It can be concluded that stock market data should be considered in the framework of uncertainties.

The main goal of the work was to develop models, methods and algorithms to modeling and simulation of financial time series using the concept of ordered fuzzy number as an uncertainty model. It was proved that it is possible to represent the financial time series by using ordered fuzzy numbers and maintain more information about price movements at assumed time interval compared to commonly used price charts. Moreover, thanks to well-defined arithmetic of ordered fuzzy numbers and the concept of ordered fuzzy random variable developed by the author, it is possible to construct full, fuzzy, stochastic time series models and use them for simulation purposes.

The dissertation consists of eight chapters. In the first one, motivations and goals are formulated. In the second and third chapters, the fundamental concepts of fuzzy numbers and ordered fuzzy numbers are reviewed. The fourth chapter includes a description of new ideas and methods for representation of financial high-frequency time series by using ordered fuzzy numbers. In the fifth chapter, the methodology of full, fuzzy models of time series is presented. The new concept of the fuzzy random variables for ordered fuzzy numbers is provided in the sixth chapter. The seventh chapter is dedicated to presenting several practical applications of the proposed models. In the last chapter, the summary of this work is presented and conclusions are drawn.