

ANALIZA PORTFELOWA

Andrzej Palczewski
Jan Palczewski

Notatki spisane przez:
Katarzynę Kańską, Piotra Obarskiego, Sonię Strokę i Kingę Trychtę

Spis treści

Wstęp	5
1 Teoria użyteczności	9
1.1 Funkcja użyteczności	9
1.2 Awersja do ryzyka	21
1.3 Standardowe funkcje użyteczności	23
2 Portfele optymalne bez ograniczenia na krótką sprzedaż	25
2.1 Optymalizacja portfela instrumentów ryzykownych	25
2.2 Funkcje preferencji	27
2.3 Granica portfelowa	31
2.4 Optymalizacja z instrumentem wolnym od ryzyka	35
3 Model wyceny dóbr kapitałowych CAPM	43
3.1 Portfele ortogonalne	43
3.2 Model CAPM	48
4 Portfele optymalne z ograniczeniem na krótką sprzedaż	53
4.1 Optymalizacja z ograniczeniem krótkiej sprzedaży	53
4.2 Algorytm prostej krytycznej	60
4.3 Optymalizacja z instrumentem wolnym od ryzyka i ograniczeniem krótkiej sprzedaży	67
5 Estymatory portfeli optymalnych Markowitza	71
5.1 Rynek finansowy	71
5.2 Portfel optymalny dla jednego instrumentu ryzykownego	74
5.3 Portfele optymalne dla wielu instrumentów ryzykownych	78
6 Model Blacka–Littermana	83
6.1 Założenia modelu	83
Rozkład równowagowy	84
Rozkład prognoz	86
6.2 Wzór Blacka-Littermana	87
7 Portfele minimalizujące ryzyko	93
7.1 Kryteria bezpieczeństwa	93
Kryterium Roya	93
Kryterium Telsera	95
Kryterium Kataoki	97

	Kryteria bezpieczeństwa w obecności instrumentu wolnego od ryzyka . . .	98
7.2	Miary ryzyka	99
	Wartość narażona na ryzyko (Value at Risk)	103
	Warunkowa wartość narażona na ryzyko (Conditional VaR)	105
7.3	Portfele minimalizujące VaR lub CVaR	107
8	Wiadomości uzupełniające ze statystyki	113
8.1	Własności rozkładów empirycznych	113
8.2	Własności próbek wielowymiarowych	117
9	Podstawowe fakty z optymalizacji nieliniowej	125
9.1	Optymalizacja z ograniczeniami równościowymi	125
9.2	Optymalizacja wypukła	127
9.3	Optymalizacja z ograniczeniami mieszanymi	128
9.4	Punkty siodłowe funkcji Lagrange'a	130
9.5	Funkcja wrażliwości	131
	Literatura	133

Wstęp

Analiza portfelowa jest jednym z filarów metod ilościowych w finansach. Stanowić więc powinna także istotny element kształcenia w zakresie matematyki finansowej. Choć problemem skutecznego inwestowania posiadanego kapitału zajmowano się od zawsze, to systemowe podejście do problemu pojawiło się dopiero w połowie XX wieku. Za prekursora analizy portfelowej uważa się powszechnie Harry'ego Markowitza, który w swojej pracy doktorskiej sprowadził problem optymalnych inwestycji do rozwiązania pewnego zadania optymalizacji kwadratowej. Tak rozumiana analiza portfelowa jest opisana w wielu podręcznikach akademickich. Przykładem mogą być tu wymienione w spisie literatury książki Prigenta lub Szegö'ego. Także w tym tekście klasycznej teorii Markowitza poświęciliśmy wiele miejsca opisując różne jej warianty: bez obecności instrumentu wolnego od ryzyka, z obecnością tego instrumentu, bez ograniczenia na "krótką sprzedaż" oraz z tym ograniczeniem. Ta tematyka obejmuje rozdziały 2 i 4.

W rozdziale 4 opisaliśmy szczegółowo pochodzący od Markowitza algorytm znajdowania portfeli optymalnych z ograniczeniem na "krótką sprzedaż" (*Critical Line Algorithm*), który współcześnie nie ma wielkiego znaczenia numerycznego (w wielu pakietach dostępne są bardzo skuteczne numeryczne procedury optymalizacji kwadratowej), ale stanowi narzędzie przy dowodzeniu pewnych własności rozwiązań. Ze względu na głównie historyczne znaczenie CLA, jego opis trudno znaleźć we współczesnych podręcznikach poświęconych analizie portfelowej. Ten tekst stanowi tu pewien wyjątek, a ponieważ CLA, mimo że zawiera skomplikowane rozumowania matematyczne, jest wygodną drogą do udowodnienia pewnych własności rozwiązań, włączenie opisu algorytmu uznaliśmy za celowe w tekście adresowanym do matematyków.

Metoda Markowitza znajdowania optymalnych portfeli inwestycyjnych dość szybko spotkała się z krytyką, która najpierw wyszła ze strony praktyków rynku finansowego a potem znalazła potwierdzenie w wielu tekstach teoretycznych. Istota krytyki sprowadza się do konstatacji, że teoria Markowitza wymaga znajomości momentów rozkładu przyszłych stóp zwrotu instrumentów inwestycyjnych. Tymczasem jedyne czym praktycznie możemy dysponować, to estymatory tych momentów. Jak się okazuje, wynik optymalizacji jest bardzo wrażliwy na (nieuniknione) błędy estymacji, co stawia praktyczną użyteczność teorii Markowitza pod znakiem zapytania. Szczegółowemu opisowi tego zjawiska poświęcony jest rozdział 5, w którym wykorzystaliśmy także nasze własne wyniki badań.

Teoria Markowitza, jako metoda znajdowania inwestycji optymalnych, ma głównie historyczne znaczenie. Jednak pewne pojęcia powstałe w tej teorii, takie jak "granica portfelowa", "granica efektywna", "portfel efektywny", weszły na trwałe do języka finansów. Jednocześnie to na gruncie teorii Markowitza powstał model wyceny dóbr kapitałowych (*Capital Asset Pricing Model*), który jest powszechnie wykorzystywany przez praktyków i badany teoretycznie. Pojęcia stworzone w tym modelu, takie jak "two fund theorem", "capital market line", "portfel rynkowy", " β portfela", "aktywna α ", są powszechnie używane przez praktyków rynku finansowego i wykorzystywane w ich codziennej działalności inwestycyjnej. Uznając to za ważny element wykształcenia osoby zajmującej się matematyką finansową poświęciliśmy tej tematyce rozdział 3.

Ze względu na małą praktyczną przydatność teorii Markowitza od dawna pracowano nad znalezieniem lepszych metod poszukiwania portfeli optymalnych. Postanowiliśmy opisać jedną z takich metod – model Blacka-Littermana, który od dłuższego czasu cieszy się dużym powodzeniem wśród praktyków rynku finansowego. W ostatnich latach powstało na temat tego modelu także szereg nowych prac teoretycznych. W opisie ograniczyliśmy się do przedstawienia klasycznej wersji modelu, ale czytelnik powinien mieć świadomość, że jest to jedynie wstęp do obecnie intensywnie rozwijającej się dziedziny badań teoretycznych.

Jak napisaliśmy wyżej, teoria Markowitza utraciła status skutecznej metody znajdowania optymalnych inwestycji. Znaczne zainteresowanie budzi obecnie podejście bazujące na teorii użyteczności. Jest to obszar zarówno ciekawych prac teoretycznych jak też intensywnych badań aplikacyjnych. Wstęp do takiego opisu inwestycji optymalnych zawarliśmy w rozdziale 1. Konsekwentne przedstawienie całej teorii może znaleźć czytelnik w książce Föllmera i Schieda.

Rozwój współczesnego rynku finansowego doprowadził do powstania nowego typu instytucji – banku inwestycyjnego. Co więcej, większość tradycyjnych banków rozwinęło własną działalność inwestycyjną. To nowe zjawisko doprowadziło do pojawienia się nowego rodzaju ryzyka związanego z portfelem inwestycyjnym banku. W efekcie obok problemu inwestycji optymalnych pojawił się problem inwestycji bezpiecznych, tj. inwestycji nie obciążonych nadmiernym ryzykiem. Problem bezpieczeństwa inwestycji był oczywiście znany od dawna a odpowiednie kryteria bezpieczeństwa powstały w latach 50-tych XX wieku. Jednak problem ten nabrał nowego znaczenia po pojawieniu się banków inwestycyjnych. W efekcie pod koniec lat 90-tych XX wieku powstało pojęcie koherentnej miary ryzyka oraz konkretne miary, które zostały narzucane bankom przez instytucje regulujące rynek finansowy jako właściwe narzędzia oceny ryzyka portfeli inwestycyjnych. Problematyce tej oraz omówieniu najpopularniejszych miar – VaR i CVaR – poświęciliśmy rozdział 7.

Zasadniczy wykład został uzupełniony dwoma rozdziałami zawierającymi informacje z innych działów matematyki. Rozdział 8 zawiera uzupełniające informacje ze statystyki niezbędne dla dowodów twierdzeń w rozdziale 5. W rozdziale 9 zgromadziliśmy najważniejsze fakty z optymalizacji nieliniowej, które wykorzystujemy w wielu miejscach w dowodach. Rozdział ten powstał z fragmentów

skryptu "Optymalizacja II", którego jeden z nas (JP) jest autorem.

Historia powstania notatek Pierwsza wersja notatek powstała w czasie naszego wspólnego wykładu monograficznego "Współczesne metody analizy portfelowej" jaki wygłosiliśmy na Wydziale Matematyki, Informatyki i Mechaniki UW w roku akademickim 2011/12. W następnym roku akademickim jeden z nas (AP) zaczął prowadzić wykład fakultatywny "Analiza portfelowa I". Wykład ten został oczywiście oparty na notatkach, jakie wspólnie zrobiliśmy do wykładu monograficznego z poprzedniego roku. W kolejnych latach notatki te były uzupełniane a dowody upraszczane. Znaczny wkład w wygładzenie notatek miało kilka szkół (letnich lub zimowych), na których każdy z nas (osobno) wykładał jakiś fragment analizy portfelowej. W dalszym ciągu były to jednak odręczne notatki pisane przez dwie osoby, z których każda starała się poprawić błędy drugiej.

Zasadnicza zmiana nastąpiła w roku akademickim 2014/15 kiedy Pani Katarzyna Kańska zaoferowała gotowość zajęcia się przepisaniem notatek odręcznych w LaTeXu. W efekcie zorganizowała ona zespół "przepisywaczy", który sukcesywnie dokonał przepisania w LaTeXu prawie całych notatek. Oczywiście odręczna wersja notatek była przeznaczona jako materiał pomocniczy dla wykładowcy, który wie o czym chce mówić. Notatki zawierały więc sformułowania twierdzeń oraz zasadnicze rachunki potrzebne w dowodach. Jako wykładowca "Analizy Portfelowej I" uznałem (AP), że takie notatki są mało przydatne jako podręcznik dla studentów. W efekcie przepisany tekst został uzupełniony licznymi komentarzami, które mają ułatwić czytelnikowi zrozumienie materiału. Przy uzupełnianiu tekstu okazało się, że w trakcie wykładu często zmienialiśmy oznaczenia. Problem ten, mało istotny w wykładzie wygłaszanym na żywo, gdzie można zawsze podać dodatkowe wyjaśnienia, okazał się trudny do rozwiązania w wersji drukowanej. W efekcie część oznaczeń została zmieniona, tak aby uniknąć kolizji. Jednak niektóre oznaczenia nie są jednoznaczne: litery α , β , γ i δ zostały użyte do oznaczenia pewnych stałych występujących w rozwiązywaniu zadania optymalizacyjnego Markowitza (bo takie oznaczenia są używane w wielu podręcznikach), ale jednocześnie w rozdziale 7, α oznacza poziom ufności miar VaR lub CVaR, a stała β jest używane w rozdziale 3 dla oznaczenia "bety" instrumentu względem portfela rynkowego. Mamy nadzieję, że te kolizje nie będą stanowiły problemu przy czytaniu tekstu.

W portalu "Matematyka stosowana" (mst.mimuw.edu.pl) dostępny jest skrypt "Analiza Portfelowa i Rynki Kapitałowe 1" autorstwa Piotra Mormula i Mariusza Baryło. Zasadniczo między naszym tekstem a tekstem skryptu nie ma żadnej zależności. Zbiór faktów i twierdzeń zawartych w naszych rozdziałach 2 i 4 jest częściowo podobny do twierdzeń ze skryptu (choć dowody są najczęściej inne). Ale reszty naszego materiału nie ma we wspomnianym skrypcie. Daleko idące podobieństwo dotyczy prezentacji algorytmu CLA Markowitza. Dowód tego algorytmu wygodnie jest zaczynać od jednego z dwóch punktów granicy portfelowej: od punktu, w którym portfel ma najmniejszą stopę zwrotu, albo od punktu, gdzie portfel ma największą stopę zwrotu. Konsekwencją wyboru punktu początkowego

jest przyjęcie odpowiedniej postaci dodatkowych założeń gwarantujących brak degeneracji (innego dla każdego punktu startu). W naszej prezentacji, podobnie jak w skrypcie Mormula i Baryły zaczynamy od portfela największej stopy zwrotu, co oznacza przyjęcie takich samych założeń o braku degeneracji. Jednak w przeciwieństwie do wspomnianego skryptu, wszystkie dowody doprowadzamy do końca, a przypadkach braku rozstrzygających argumentów pokazujemy kontrprzykłady.

Podziękowania Jak już wspomniano wcześniej, drukowana wersja tych notatek nie powstałaby bez inicjatywy oraz zaangażowania Katarzyny Kańskiej. Jej zasługą są także wszystkie rysunki w rozdziałach 2–4, co wymagało dużej dozy cierpliwości przy robieniu licznych poprawek (bo odręczne rysunki, które miały być pierwowzorem, nie były zbyt precyzyjne, co ujawniało się dopiero w trakcie pracy). W przepisywaniu tekstu Panią Katarzynę Kańską wspierało kilku słuchaczy wykładu w roku 2014/15 i 2015/16. Są to: Piotr Obarski, Sonia Stroka i Kinga Trychta. Wszystkim im składamy gorące podziękowania.

Andrzej Palczewski i Jan Palczewski, Warszawa 2016

Rozdział 1

Teoria użyteczności

Cel analizy portfelowej formułuje się jako znalezienie "optymalnej inwestycji finansowej w warunkach niepewności". Samo to sformułowanie zawiera pojęcia, które wymagają poprawnego zdefiniowania. W dalszym ciągu będziemy ustalali kryteria optymalności inwestycji a także precyzowali jak rozumiemy "niepewność". To precyzowanie pojęć rozpoczniemy od ustalenia modelu rynku finansowego. Na razie będzie to raczej definicja ekonomiczna niż matematyczna, ale pozwoli nam ona na wyrobienie sobie ważnych intuicji.

DEFINICJA. 1.1 (Model rynku finansowego) *Rynek finansowy to:*

- o *d instrumentów „ryzykownych” (akcje, instrumenty pochodne).*
- o *1 instrument „wolny od ryzyka” (rachunek bankowy).*

Co oznacza „ryzykowny”, „wolny od ryzyka” zostanie doprecyzowane później. Na użytek tej definicji możemy przyjąć, że instrument wolny od ryzyka to taki, którego ceny są znaną funkcją deterministyczną czasu. Ceny instrumentów ryzykownych są natomiast realizacjami pewnych procesów stochastycznych.

Podjęcie decyzji inwestycyjnej oznacza ustalenie jaki kapitał zainwestować w jakie instrumenty. Bardziej formalnie będziemy to rozumieli jako wyznaczenie strategii inwestycyjnej (będziemy ją też nazywali portfelem inwestycyjnym), która będzie optymalną inwestycją na określony okres czasu.

W naszych rozważaniach ograniczymy się do problemu statycznego – znamy stan rynku dziś i chcemy wybrać portfel, który będzie optymalny po jednostce czasu (dzień, miesiąc, rok). Ważne, że niezależnie od długości tego czasu jest to jedna decyzja: portfel ustalony dziś jest niezmienny do czasu oceny inwestycji (będziemy mówili o czasie T_0 i T_1 lub po prostu 0 i 1).

1.1 Funkcja użyteczności

Obecnie zajmiemy się bardziej precyzyjną odpowiedzią na pytanie "Co oznacza optymalny portfel?" Aby odpowiedzieć na to pytanie potrzebujemy relacji po-

rzędu w zbiorze możliwych portfeli.

Aby porównywać portfele musimy zrobić pewne dodatkowe założenia (intuicyjne oczywiste):

1. Porównujemy jedynie portfele, które startują z tego samego kapitału początkowego V_0 .
2. Musimy założyć pewną znajomość rynku – znamy przestrzeń probabilistyczną $(\Omega, \mathcal{F}, \mathbb{P})$ opisującą stan rynku, tzn. przestrzeń probabilistyczną, na której są określone zmienne losowe opisujące ceny instrumentów ryzykownych.

Wprowadźmy następujące oznaczenia – niech S_0^i będą obecnymi cenami instrumentów, a S^i , $i = 1, 2, \dots, d$, przyszłymi cenami na końcu okresu inwestycyjnego. S^i są zmiennymi losowymi na przestrzeni $(\Omega, \mathcal{F}, \mathbb{P})$ przyszłego stanu rynku. Jeśli dysponujemy kapitałem początkowym V_0 , to kupujemy n^i jednostek i -tego instrumentu tak, aby $V_0 = \sum_{i=1}^d n^i S_0^i$. Na końcu procesu inwestycyjnego dysponujemy tą samą liczbą jednostek i -tego instrumentu, ale już z nowymi cenami S^i , czyli kapitałem $V = \sum_{i=1}^d n^i S^i$. Oczywiście V też jest zmienną losową.

Opis inwestycji przez podawanie ilości jednostek poszczególnych instrumentów w portfelu nie jest jedynym możliwym. Inwestycje można także opisywać przez podawanie ich stóp zwrotu.

DEFINICJA. 1.2 *Stopą zwrotu z inwestycji w i -ty instrument nazywamy wielkość*

$$R^i = \frac{S^i - S_0^i}{S_0^i} = \frac{S^i}{S_0^i} - 1.$$

Tak zdefiniowana stopa zwrotu nazywana jest arytmetyczną stopą zwrotu. Często wprowadzana jest także logarytmiczna stopa zwrotu $R^i = \ln \frac{S^i}{S_0^i}$.

Korzystając z rozwinięcia $\ln(1+x) = x - \frac{x^2}{2} + \dots$ widzimy, że w pierwszym przybliżeniu (małe x) obie definicje pokrywają się.

Jeżeli policzymy stopę zwrotu z całej inwestycji, to otrzymamy

$$\bar{R} = \frac{V - V_0}{V_0} = \frac{V}{V_0} - 1 = \sum_{i=1}^d \frac{n^i S_0^i}{V_0} \frac{S^i}{S_0^i} - 1.$$

Niech $x_i = \frac{n^i S_0^i}{V_0}$. Wtedy x_i oznacza część kapitału początkowego zainwestowaną w i -ty instrument. Oczywiście $\sum_{i=1}^d x_i = 1$. Stąd

$$\bar{R} = \sum_{i=1}^d x_i \frac{S^i}{S_0^i} - 1 = \sum_{i=1}^d x_i \frac{S^i}{S_0^i} - \sum_{i=1}^d x_i = \sum_{i=1}^d x_i \left(\frac{S^i}{S_0^i} - 1 \right) = \sum_{i=1}^d x_i R^i.$$

DEFINICJA. 1.3 *Strategią inwestycyjną (portfelem inwestycyjnym) nazywać będziemy wektor $x = (x_1, \dots, x_d)$, taki że $\sum_{i=1}^d x_i = 1$. Zmienną losową $\bar{R} = \sum_{i=1}^d x_i R^i$ nazywać będziemy stopą zwrotu strategii x (jeśli będziemy chcieli wyraźnie zaznaczyć dla jakiej strategii jest to stopa zwrotu, to będziemy pisali \bar{R}_x).*

Jeśli założyć dodatkowo, że wartość inwestycji jest proporcjonalna do zainwestowanego kapitału początkowego, tzn. $V \equiv V(\lambda V_0) = \lambda V(V_0)$, to można zapisać V jako funkcję stopy zwrotu portfela.

Rzeczywiście, jeśli $V(\lambda V_0) = \lambda V(V_0)$, to $\bar{R} = \frac{V(V_0) - V_0}{V_0} = \frac{V(\lambda V_0) - \lambda V_0}{\lambda V_0}$.

Czyli stopa zwrotu portfela nie zależy od zainwestowanego kapitału początkowego oraz $V = (1 + \bar{R})V_0$.

Uwaga. 1 *W obecnych rozważaniach pomijamy instrument pozbawiony ryzyka, ponieważ inwestowanie w ten instrument nie jest obciążone niepewnością.*

Niech \mathcal{V} oznacza zbiór wartości inwestycji (wartość inwestycji mierzymy jej kapitałem końcowym V bo założyliśmy, że wszystkie inwestycje startują z tej samej wartości kapitału początkowego V_0). Pytanie o optymalność strategii inwestycyjnej można sformułować jak pytanie o istnienie relacji porządku w zbiorze \mathcal{M} rozkładów prawdopodobieństwa na $(\mathcal{V}, \mathcal{W})$, gdzie \mathcal{W} jest σ -ciałem zbiorów borelowskich na \mathcal{V} . Te rozkłady prawdopodobieństwa traktujemy jako miary na $(\mathcal{V}, \mathcal{W})$, czyli rozpatrujemy odpowiednią przestrzeń miar $\mathcal{M}(\mathcal{V}, \mathcal{W}) = \mathcal{M}(\mathcal{V})$.

DEFINICJA. 1.4 *Relacją preferencji na \mathcal{M} nazywamy relację dwuargumentową \succ w zbiorze \mathcal{M} , która spełnia warunki:*

1. *Asymetrii: jeśli zachodzi relacja $X \succ Y$, to nie zachodzi relacja $Y \succ X$;*
2. *Negatywnej przechodniości: jeśli $X \succ Y$ i $Z \in \mathcal{M}$, to zachodzi przynajmniej jedna z relacji $X \succ Z$, $Z \succ Y$.*

DEFINICJA. 1.5 *Relacją słabej preferencji na \mathcal{M} nazywamy relację dwuargumentową \succeq w zbiorze \mathcal{M} , która spełnia warunki:*

1. *Zupełności: $\forall X, Y \in \mathcal{M}$ zachodzi relacja $X \succeq Y$ lub relacja $Y \succeq X$;*
2. *Przechodniości: jeśli $X \succeq Y$ i $Y \succeq Z$, to zachodzi relacja $X \succeq Z$.*

Relacja słabej preferencji pozwala zdefiniować także relację nieodróżnialności: mówimy, że $X \sim Y$, jeśli $X \succeq Y$ oraz $Y \succeq X$.

Uwaga. 2 *Relacje preferencji oraz słabej preferencji są wzajemnie powiązane następującymi zależnościami: jeśli \succ jest relacją preferencji na \mathcal{M} , to $X \succeq Y$ zdefiniowana jako $Y \not\succeq X$ jest relacją słabej preferencji; podobnie jeśli \succeq jest relacją słabej preferencji na \mathcal{M} , to $X \succ Y$ zdefiniowana jako $Y \not\succeq X$ jest relacją preferencji.*

DEFINICJA. 1.6 Niech dana będzie relacja preferencji \succ na wypukłym zbiorze \mathcal{M} . Mówimy, że relacja ta spełnia:

1. Aksjomat niezależności: jeśli $\forall X, Y \in \mathcal{M}$, takich że $X \succ Y$ zachodzi relacja

$$tX + (1-t)Z \succ tY + (1-t)Z,$$

dla każdego $Z \in \mathcal{M}$ oraz $t \in (0, 1]$.

2. Aksjomat Archimedesesa: jeśli $\forall X, Y, Z \in \mathcal{M}$, takich że $X \succ Z \succ Y$, to istnieją $t_1, t_2 \in (0, 1)$, takie że zachodzą relacje

$$t_1X + (1-t_1)Y \succ Z \succ t_2X + (1-t_2)Y.$$

Ponieważ rozkłady z \mathcal{M} zdefiniowane są na zbiorze \mathcal{V} , który jest podzbiorem liczb rzeczywistych, to relację preferencji można powiązać z relacją liniowego porządku liczb rzeczywistych.

DEFINICJA. 1.7 Funkcję $U : \mathcal{M} \rightarrow \mathbb{R}$ nazywa się numeryczną reprezentacją relacji preferencji \succ , jeśli

$$X \succ Y \iff U(X) > U(Y).$$

TWIERDZENIE. 1.8 Niech zbiór \mathcal{M} będzie wypukły a relacja preferencji \succ na \mathcal{M} spełnia aksjomat niezależności i aksjomat Archimedesesa. Wtedy istnieje na \mathcal{M} numeryczna reprezentacja U relacji \succ spełniająca następujący warunek afiniczności

$$\forall X, Y \in \mathcal{M} \quad U(tX + (1-t)Y) = tU(X) + (1-t)U(Y), \quad \text{dla } t \in [0, 1].$$

Funkcja U jest wyznaczona jednoznacznie z dokładnością do dodatniego przeskalowania afinicznego, tj. jeśli \hat{U} jest inną taką funkcją, to $\hat{U} = aU + b$, gdzie $a > 0, b \in \mathbb{R}$.

Dowód tego twierdzenia poprzedzimy następującym lematem.

LEMAT. 1.9 Niech zbiór \mathcal{M} będzie wypukły a relacja preferencji \succ na \mathcal{M} spełnia aksjomat niezależności i aksjomat Archimedesesa. Wtedy prawdziwe są następujące relacje:

1. Jeśli $X \succ Y$, to dla wszystkich $0 \leq t < s \leq 1$ zachodzi relacja $sX + (1-s)Y \succ tX + (1-t)Y$.
2. Jeśli $X \succ Y$ i $X \succeq Z \succeq Y$, to istnieje dokładnie jedna stała $t \in [0, 1]$, taka że $Z \sim tX + (1-t)Y$.
3. Jeśli $X \sim Y$, to $tX + (1-t)Z \sim tY + (1-t)Z$ dla wszystkich $t \in [0, 1]$ i wszystkich $Z \in \mathcal{M}$.

Dowód. 1. Wykorzystując aksjomat niezależności, w którym s zastąpi t a $Z = Y$ dostajemy $S := sX + (1-s)Y \succ Y$. Wykorzystując ponownie ten aksjomat z $S = X = Z$ oraz zastępując t przez $1-u$ dostajemy $(1-u)S + uS \succ (1-u)Y + uS$. Ponieważ w ostatniej relacji lewa strona równa jest S a prawa $suX + (1-su)Y$, to biorąc $u = t/s \in [0, 1)$ dostajemy poszukiwaną relację.

2. Punkt 1 gwarantuje, że jeśli stała t istnieje, to jest ona wyznaczona jednoznacznie. Pokażemy więc wyłącznie istnieje tej stałej. Dowodzić będziemy jedynie przypadku, gdy $X \succ Z \succ Y$, bo w pozostałych przypadkach wystarczy wziąć $t = 0$ lub $t = 1$, aby otrzymać poszukiwaną relację. Naturalnym kandydatem na stałą t jest

$$t = \sup\{u \in [0, 1]: Z \succeq uX + (1-u)Y\}.$$

Przypuśćmy, że jest to niewłaściwa stała. Oznacza to, że musi zachodzić jedna z relacji

$$Z \succ tX + (1-t)Y \quad \text{lub} \quad tX + (1-t)Y \succ Z. \quad (1.1)$$

Rozważmy pierwszą z relacji w (1.1). Wyklucza ona możliwość $t = 1$, czyli $t < 1$. Wykorzystując aksjomat Archimedesesa dla $t_2 = 1-s$, gdzie $s \in (0, 1)$, dostajemy

$$Z \succ (1-s)X + s(tX + (1-t)Y) = (1-s+ts)X + (s-ts)Y.$$

Z definicji t wynika, że $t \geq 1-s+ts$, co w połączeniu z warunkiem $s < 1$ daje $t \geq 1$, czyli sprzeczność.

Rozważmy teraz drugą relację w (1.1). Wykorzystując aksjomat Archimedesesa dla $t_1 = s \in (0, 1)$ dostajemy

$$s(tX + (1-t)Y) + (1-s)Y = stX + (1-st)Y \succ Z. \quad (1.2)$$

Ponieważ $st < t$, więc istnieje $u \in (st, t]$ takie że $Z \succeq uX + (1-u)Y$. Z faktu $st < u$ oraz punktu 1 wynika

$$Z \succeq uX + (1-u)Y \succ stX + (1-st)Y,$$

co prowadzi do sprzeczności z relacją (1.2).

3. Załóżmy, że istnieje $S \in \mathcal{M}$ takie że $S \approx X$, bo w przeciwnym przypadku rezultat jest oczywisty. Jeśli odpowiednie S istnieje, to zakładamy $S \succ X$ (przeciwny przypadek wymaga analogicznych rozważań). Wtedy także $S \succ Y$ i stosując aksjomat niezależności z $Z = Y$ dostajemy

$$sS + (1-s)Y \succ sY + (1-s)Y = Y,$$

dla każdego $s \in (0, 1]$. Wykorzystując ten aksjomat ponownie z $t \in (0, 1]$ dostajemy (wykorzystujemy relację $X \sim Y$)

$$t(sS + (1-s)Y) + (1-t)Z \succ tX + (1-t)Z. \quad (1.3)$$

Gdyby teza była nieprawdziwa, to musiałaby zachodzić jedna z relacji $tX + (1-t)Z \succ tY + (1-t)Z$ lub $tY + (1-t)Z \succ tX + (1-t)Z$. Rozważmy relację

$tX + (1 - t)Z \succ tY + (1 - t)Z$ (druga analizuje się w analogiczny sposób). Wykorzystując punkt 2 wiemy, że istnieje jednoznacznie wyznaczona stała $u \in (0, 1)$, taka że

$$\begin{aligned} tX + (1 - t)Z &\sim u(t(sS + (1 - s)Y) + (1 - t)Z) + (1 - u)(tY + (1 - t)Z) = \\ &= t(suS + (1 - su)Y) + (1 - t)Z. \end{aligned}$$

Ponieważ relacja (1.3) jest prawdziwa dla każdego $s \in (0, 1]$, więc w szczególności jest prawdziwa, gdy s zastąpić przez su . Wtedy jednak dostajemy

$$tX + (1 - t)Z \succ tX + (1 - t)Z,$$

co prowadzi do sprzeczności. ■

Dowód. (twierdzenia 1.8) Dla $S, Z \in \mathcal{M}$, $Z \succ S$, zdefiniujmy

$$\mathcal{M}(Z, S) = \{X \in \mathcal{M}: Z \succeq X \succeq S\}.$$

Z punktu 2 lematu 1.9 wynika istnienie jednoznacznie wyznaczonej stałej $t = t(X)$ zależnej od X , takiej że $X \sim tZ + (1 - t)S$. Stałą tę wykorzystujemy do zdefiniowania numerycznej reprezentacji U kładąc $U(X) = t(X)$. Zauważmy, że $U(Z) = 1$ a $U(S) = 0$. Musimy teraz wykazać, że ta konstrukcja jest rzeczywiście numeryczną reprezentacją relacji \succ na $\mathcal{M}(Z, S)$. W tym celu musimy udowodnić, że dla dowolnych $X, Y \in \mathcal{M}(Z, S)$ nierówność $U(X) > U(Y)$ jest równoważna relacji $X \succ Y$. Jeśli $U(X) > U(Y)$, to z punktu 1 lematu 1.9 mamy

$$X \sim U(X)Z + (1 - U(X))S \succ U(Y)Z + (1 - U(Y))S \sim Y,$$

czyli $X \succ Y$. Jeśli $X \succ Y$, to wykazana relacja wyklucza nierówność $U(Y) > U(X)$. Wystarczy więc wykazać, że niemożliwa jest równość $U(Y) = U(X)$. Jeśli jednak zachodzi równość $U(Y) = U(X)$, to z definicji U mamy $X \sim Y$, co przeczy relacji $X \succ Y$.

Pokażemy teraz, że $\mathcal{M}(Z, S)$ jest zbiorem wypukłym. Weźmy $X, Y \in \mathcal{M}(Z, S)$ oraz $t \in (0, 1)$. Ponieważ $Z \succ Y$, z aksjomatu niezależności wynika relacja $Z = (1 - t)Z + tZ \succ (1 - t)Y + tZ$. A ponieważ mamy też $Z \succ X$, więc zastosowanie tego samego aksjomatu daje $tZ + (1 - t)Y \succ tX + (1 - t)Y$. Łącząc te relacje dostajemy $Z \succ tX + (1 - t)Y$. Analogicznie dowodzimy, że $tX + (1 - t)Y \succ S$. Do dowodu wypukłości zbioru $\mathcal{M}(Z, S)$ potrzebujemy jeszcze rozpatrzyć przypadki $Z \sim Y$ lub/i $Z \sim X$. Wypukłość w tych przypadkach wynika z punktu 3 lematu 1.9. Wykazaliśmy więc wypukłość zbioru $\mathcal{M}(Z, S)$.

Wykażemy teraz, że funkcja U jest afiniczna. Z wypukłości zbioru $\mathcal{M}(Z, S)$ wynika, że wyrażenie $U(tX + (1 - t)Y)$ jest dobrze zdefiniowane. Ponieważ z definicji $U(X)$ wynika relacja $U(X)Z + (1 - U(X))S \sim X$ a z definicji $U(Y)$ relacja $U(Y)Z + (1 - U(Y))S \sim Y$, więc dwukrotne zastosowanie punktu 3 lematu 1.9 daje

$$tX + (1 - t)Y \sim t(U(X)Z + (1 - U(X))S) + (1 - t)(U(Y)Z + (1 - U(Y))S).$$

Porządkując wyrazy po prawej stronie dostajemy

$$(tU(X) + (1-t)U(Y))Z + (1-tU(X) - (1-t)U(Y))S.$$

Wtedy z definicji U oraz jednoznaczności w punkcie 2 lematu 1.9 mamy poszukiwaną równość

$$U(tX + (1-t)Y) = tU(X) + (1-t)U(Y).$$

Udowodnimy teraz, że otrzymana funkcja U jest wyznaczona jednoznacznie z dokładnością do przekształcenia afinicznego. Załóżmy więc, że istnieje inna funkcja \tilde{U} będąca numeryczną reprezentacją relacji \succ . Niech $X \in \mathcal{M}(Z, S)$. Zdefiniujmy

$$\hat{U}(X) = \frac{\tilde{U}(X) - \tilde{U}(S)}{\tilde{U}(Z) - \tilde{U}(S)}.$$

Jest to afiniczne przekształcenie \tilde{U} przyjmujące wartości $\hat{U}(Z) = 1$ oraz $\hat{U}(S) = 0$, czyli takie same jak funkcja U . Dostajemy więc

$$\hat{U}(X) = \hat{U}(U(X)Z + (1-U(X))S) = U(X)\hat{U}(Z) + (1-U(X))\hat{U}(S) = U(X).$$

Dowodzi to, że funkcja U jest afiniczną transformacją funkcji \tilde{U} .

Dotychczas wykazaliśmy istnienie numerycznej reprezentacji relacji \succ jedynie na zbiorze $\mathcal{M}(Z, S)$. Teraz chcemy pokazać, jak rozszerzyć ją na całe \mathcal{M} . W tym celu weźmy $Z_1, S_1 \in \mathcal{M}$, takie że $\mathcal{M}(Z_1, S_1) \supset \mathcal{M}(Z, S)$. Z dotychczasowego dowodu wiemy, że istnieje numeryczna reprezentacja U_1 na $\mathcal{M}(Z_1, S_1)$. Przez przekształcenie afiniczne możemy ją wybrać tak, aby $U_1(Z) = 1$ a $U_1(S) = 0$. Oznacza to, że U_1 na zbiorze $\mathcal{M}(Z, S)$ pokrywa się z U , czyli U_1 jest rozszerzeniem U . Ponieważ zbiór wszystkich podzbiorów \mathcal{M} postaci $\mathcal{M}(Z, S)$ jest częściowo uporządkowany przez relację inkluzji zbiorów a opisane rozszerzenie jest liniowym porządkiem w zbiorach $\mathcal{M}(Z, S) \subset \mathcal{M}(Z_1, S_1) \subset \dots \subset \mathcal{M}(Z_k, S_k) \subset \dots$, to wystarczy zauważyć, że każdy taki liniowo uporządkowany zbiór zawiera element maksymalny. Rzeczywiście, jeśli istnieją \hat{Z}, \hat{S} , takie że $\hat{Z} \succeq Z_k$ i $S_k \succeq \hat{S}$, dla każdego k , a przynajmniej jedna z tych relacji jest ostra, to $\mathcal{M}(\hat{Z}, \hat{S})$ jest elementem maksymalnym. Jeśli ani \hat{Z} ani \hat{S} nie istnieją, to istnieją Z_{k_0} i S_{k_0} , takie że $\mathcal{M}(Z_{k_0}, S_{k_0}) \supset \mathcal{M}(Z_k, S_k)$ dla wszystkich $k \neq k_0$. Wtedy $\mathcal{M}(Z_{k_0}, S_{k_0})$ jest elementem maksymalnym. Stąd U można rozszerzyć na całe \mathcal{M} z lematu Kuratowskiego-Zorna. ■

DEFINICJA. 1.10 *Reprezentacją von Neumanna-Morgensterna funkcji U nazywamy funkcję rzeczywistą u zdefiniowaną na \mathcal{V} , taką że*

$$U(X) = \int_{\mathcal{V}} u(s)X(ds).$$

Funkcję u nazywa się funkcją użyteczności von Neumanna-Morgensterna, lub krócej funkcją użyteczności.

Aby wykazać istnienie reprezentacji von Neumanna-Morgensterna musimy wprowadzić dodatkową strukturę topologiczną w przestrzeni \mathcal{M} . Pozwoli nam to wykazać także, że funkcja użyteczności von Neumanna-Morgensterna jest ciągła.

DEFINICJA. 1.11 Załóżmy, że przestrzeń \mathcal{V} wartości inwestycji jest ośrodkowa. Słabą topologię w przestrzeni $\mathcal{M}(\mathcal{V})$ definiujemy jako najstabszą topologię, przy której odwzorowania

$$\mathcal{M}(\mathcal{V}) \ni X \mapsto \int_{\mathcal{V}} f(s)X(ds)$$

są ciągłe dla każdego $f \in C_b(\mathcal{V})$, gdzie $C_b(\mathcal{V})$ oznacza przestrzeń funkcji ciągłych i ograniczonych na \mathcal{V} .

Jeśli przestrzeń \mathcal{V} jest ośrodkowa i metryzowalna, to przestrzeń $\mathcal{M}(\mathcal{V})$ jest też ośrodkowa a zdefiniowana słaba topologia jest metryzowalna. Co więcej, jeśli \mathcal{V}_0 jest gęstym podzbiorem \mathcal{V} , to zbiór miar prostych na \mathcal{V}_0 z wymiernymi wagami a_i

$$\left\{ \sum_{i=1}^n a_i \delta_{x_i} : a_i \in \mathbb{Q}^+, x_i \in \mathcal{V}_0, n \in \mathbb{N} \right\}$$

jest gęsty w $\mathcal{M}(\mathcal{V})$ w słabej topologii, gdzie δ_x oznacza miarę Diraca, tj. jeśli A jest zbiorem mierzalnym, to $\delta_x(A) = \mathbf{1}_A(x)$.

DEFINICJA. 1.12 Niech \mathcal{M} będzie przestrzenią topologiczną. Relacja preferencji \succ na \mathcal{M} jest ciągła, jeśli $\forall X \in \mathcal{M}$ zbiory

$$\underline{\mathcal{B}}(X) = \{Y \in \mathcal{M} : X \succ Y\} \quad \text{oraz} \quad \overline{\mathcal{B}}(X) = \{Y \in \mathcal{M} : Y \succ X\}$$

są otwarte.

TWIERDZENIE. 1.13 Niech \mathcal{M} oznacza przestrzeń wszystkich miar probabilistycznych na $(\mathcal{V}, \mathcal{W})$ ze słabą topologią zdefiniowaną wcześniej. Zakładamy, że \mathcal{V} jest przestrzenią ośrodkową i metryzowalną. Niech \succ będzie ciągłą relacją preferencji na \mathcal{M} spełniającą aksjomaty niezależności i Archimedesesa. Wtedy relacja \succ posiada reprezentację von Neumanna-Morgensterna

$$U(X) = \int_{\mathcal{V}} u(s)X(ds),$$

gdzie funkcja $u: \mathcal{V} \rightarrow \mathbb{R}$ jest ciągła i ograniczona. Reprezentacja U oraz funkcja u są wyznaczone jednoznacznie z dokładnością do przekształcenia afinicznego.

Dowód. Rozważmy podprzestrzeń \mathcal{M}_S prostych miar probabilistycznych

$$\mathcal{M}_S = \left\{ X \in \mathcal{M} : X = \sum_{i=1}^n t_i \delta_{x_i}, t_i \in \mathbb{R}^+, \sum_{i=1}^n t_i = 1, x_i \in \mathcal{V}, n \in \mathbb{N} \right\}.$$

Z twierdzenia 1.8 wynika, że relacja preferencji \succ posiada numeryczną reprezentację U . Zdefiniujmy $u(x) = U(\delta_x)$. Z afiniczności U dostajemy

$$U(X) = \sum_{i=1}^n t_i u(x_i) = \int_{\mathcal{V}} u(s)X(ds).$$

Skonstruowaliśmy więc na podprzestrzeni \mathcal{M}_S reprezentację von Neumanna-Morgensterna. Pokażemy teraz, że funkcja u otrzymana w tej konstrukcji jest ciągła i ograniczona.

Założmy, że u jest nieograniczona. Wtedy istnieje ciąg $x_k \in \mathcal{V}$, taki że $u(x_k)$ jest ściśle rosnący oraz $u(x_k) > k$. Niech

$$X_k = \left(1 - \frac{1}{\sqrt{k}}\right)\delta_{x_1} + \frac{1}{\sqrt{k}}\delta_{x_k}.$$

Łatwo można zauważyć, że X_k zbiega słabo do δ_{x_1} . Z monotonicznego wzrostu $u(x_k)$ i ciągłości \succ wynika $\delta_{x_2} \succ \delta_{x_1}$ oraz $\delta_{x_2} \succ X_k$ dla dostatecznie dużych k , co oznacza nierówność $U(X_k) < u(x_2)$. Ponieważ $U(X_k) > \left(1 - \frac{1}{\sqrt{k}}\right)u(x_1) + \frac{1}{\sqrt{k}}u(x_k)$ dla wszystkich k , dostajemy sprzeczność dla dostatecznie dużych k .

Założmy teraz, że funkcja u nie jest ciągła. Istnieje wtedy ciąg $x_k \in \mathcal{V}$ zbieżny do $x \in \mathcal{V}$, taki że $u(x_k)$ nie jest zbieżny do $u(x)$. Niech $\limsup u(x_k) < u(x)$, przechodząc do podciągu (który znowu oznaczymy x_k) mamy $\lim u(x_k) = a < u(x)$. Niech $u(x) - a = \varepsilon$. Wtedy istnieje m , takie że $|u(x_k) - a| < \varepsilon/3$ dla $k > m$. Niech $X = \frac{1}{2}(\delta_x + \delta_{x_m})$. Wtedy dla $k > m$

$$U(\delta_x) = a + \varepsilon > a + \frac{2}{3}\varepsilon > \frac{1}{2}(u(x) + u(x_m)) = U(X) > a + \varepsilon/3 > U(\delta_{x_k}).$$

Stąd $\delta_x \succ X \succ \delta_{x_k}$ mimo założenia, że δ_{x_k} zbiega słabo do δ_x . Otrzymujemy więc sprzeczność z ciągłością relacji \succ .

Na zakończenie pokażemy, że

$$U(X) = \int_{\mathcal{V}} u(s)X(ds), \quad \forall X \in \mathcal{M},$$

definiuje numeryczną reprezentację relacji \succ na całym \mathcal{M} . Ponieważ funkcja u jest ciągła i ograniczona, odwzorowanie U jest ciągle w słabej topologii \mathcal{M} . Co więcej, z definicji 1.11 wiemy, że podprzestrzeń \mathcal{M}_S jest gęsta w \mathcal{M} . Należy jedynie wykazać, że jeśli U jest numeryczną reprezentacją relacji \succ na \mathcal{M}_S , to jest ona także numeryczną reprezentacją na \mathcal{M} . Ten fakt wykażemy w lemacie 1.14. ■

Uwaga. 3 Założenie w powyższym twierdzeniu, że relacja preferencji \succ spełnia aksjomat Archimedesa, jest zbędne i zostało dodane tylko dla skrócenia dowodu. Można wykazać, że w przestrzeni miar probabilistycznych \mathcal{M} ze słabą topologią każda ciągła relacja preferencji spełnia aksjomat Archimedesa.

LEMAT. 1.14 *Niech \mathcal{M} będzie spójną przestrzenią metryczną z ciągłą relacją preferencji \succ . Jeśli $U: \mathcal{M} \rightarrow \mathbb{R}$ jest ciągłą funkcją, której obcięcie do pewnego gęstego podzbioru \mathcal{Z} jest numeryczną reprezentacją relacji \succ obciętej do \mathcal{Z} , to U jest także numeryczną reprezentacją relacji \succ na całym \mathcal{M} .*

Dowód. Dowód polega na wykazaniu, że dla $X, Y \in \mathcal{M}$ relacja $Y \succ X$ jest równoważna nierówności $U(Y) > U(X)$. Wykażemy w pierwszej kolejności, że jeśli $Y \succ X$, to $U(Y) > U(X)$. Na wstępie pokażemy, że istnieje $Z_0 \in \mathcal{Z}$, takie że $Y \succ Z_0 \succ X$. W tym celu rozważmy zbiory $\overline{\mathcal{B}}(X)$ oraz $\underline{\mathcal{B}}(Y)$ z definicji 1.12. Ponieważ $Y \in \overline{\mathcal{B}}(X)$ a $X \in \underline{\mathcal{B}}(Y)$, to oba zbiory są niepuste. Z negatywnej przechodniości relacji \succ wynika, że $\mathcal{M} = \overline{\mathcal{B}}(X) \cup \underline{\mathcal{B}}(Y)$. Ponieważ \mathcal{M} jest zbiorem spójnym, to zbiory $\overline{\mathcal{B}}(X)$, $\underline{\mathcal{B}}(Y)$ nie mogą być rozłączne. Przekrój tych zbiorów $\overline{\mathcal{B}}(X) \cap \underline{\mathcal{B}}(Y)$ zawiera więc element Z_0 należący do gęstego podzbioru \mathcal{Z} , taki że $Y \succ Z_0 \succ X$. Powtarzając powyższe rozumowanie znajdujemy $Z'_0 \in \mathcal{Z}$, taki że $Z_0 \succ Z'_0 \succ X$. W ten sposób tworzymy dwa ciągi Z_k oraz Z'_k należące do \mathcal{Z} oraz zbieżne odpowiednio $Z_k \rightarrow Y$ a $Z'_k \rightarrow X$. Z ciągłości relacji \succ wynika ciąg relacji $Z_k \succ Z_0 \succ Z'_0 \succ Z'_k$ oraz ciąg nierówności $U(Z_k) > U(Z_0) > U(Z'_0) > U(Z'_k)$. Z ciągłości U wynikają zbieżności $U(Z_k) \rightarrow U(Y)$, $U(Z'_k) \rightarrow U(X)$. Mamy więc

$$U(Y) \geq U(Z_0) > U(Z'_0) \geq U(X),$$

co dowodzi rozpatrywanej implikacji.

Aby udowodnić implikację przeciwną przyjmijmy, że dane są $X, Y \in \mathcal{M}$, takie że $U(Y) > U(X)$. Z ciągłości U zbiory

$$\overline{\mathcal{U}}(X) = \{Z \in \mathcal{M}: U(Z) > U(X)\}, \quad \underline{\mathcal{U}}(Y) = \{Z \in \mathcal{M}: U(Y) > U(Z)\}$$

są otwartymi, niepustymi podzbiórmi \mathcal{M} . Podobnie jak poprzednio dostajemy $\mathcal{M} = \overline{\mathcal{U}}(X) \cup \underline{\mathcal{U}}(Y)$, a ze spójności \mathcal{M} mamy $\overline{\mathcal{U}}(X) \cap \underline{\mathcal{U}}(Y) \neq \emptyset$. Implikuje to istnienie $Z_0, Z'_0 \in \mathcal{Z}$, takich że $U(Y) > U(Z_0) > U(Z'_0) > U(X)$. Ponieważ zbiór \mathcal{Z} jest gęsty w \mathcal{M} , to istnieją ciągi $Z_k, Z'_k \in \mathcal{Z}$ zbieżne odpowiednio $Z_k \rightarrow Y$ oraz $Z'_k \rightarrow X$ a także spełniające warunki $U(Z_k) > U(Z_0)$, $U(Z'_0) > U(Z'_k)$. Ponieważ U jest numeryczną reprezentacją relacji \succ na \mathcal{Z} , mamy stąd $Z_k \succ Z_0 \succ Z'_0 \succ Z'_k$. Z ciągłości relacji \succ oraz zbieżności $Z_k \rightarrow Y$ i $Z'_k \rightarrow X$ żadna z dwóch relacji $Z_0 \succ Y$ oraz $X \succ Z'_0$ nie może być prawdziwa. Z negatywnej przechodniości wynika więc $Y \succ X$. ■

Funkcja użyteczności, której istnienie wykazaliśmy w twierdzeniu 1.13 jest ograniczona. Tymczasem bardzo często do opisu preferencji potrzebne są nieograniczone funkcje użyteczności. Aby wykazać istnienie reprezentacji von Neumanna-Morgensterna z nieograniczoną funkcją użyteczności potrzebna jest modyfikacja twierdzenia 1.13. W tym celu wybieramy pewną ciągłą funkcję ψ na ośrodkowej metrycznej przestrzeni \mathcal{V} o wartościach w $[1, \infty)$ i definiujemy podprzestrzeń przestrzeni \mathcal{M}

$$\mathcal{M}^\psi(\mathcal{V}) = \left\{ X \in \mathcal{M}(\mathcal{V}): \int_{\mathcal{V}} \psi(s)X(ds) < \infty \right\}$$

oraz przestrzeń funkcji próbnych dla miar z $\mathcal{M}^\psi(\mathcal{V})$

$$C^\psi(\mathcal{V}) = \{f \in C(\mathcal{V}): \exists c \in \mathbb{R}^+ |f(s)| \leq c\psi(s) \forall s \in \mathcal{V}\}.$$

Z pomocą funkcji próbnych z $C^\psi(\mathcal{V})$ definiujemy słabą topologię w $\mathcal{M}^\psi(\mathcal{V})$, którą nazwiemy ψ -słabą topologią: ciąg miar $X_k \in \mathcal{M}^\psi$ jest zbieżny do $X \in \mathcal{M}^\psi$ wtedy i tylko wtedy, gdy

$$\int_{\mathcal{V}} f(s)X_k(ds) \rightarrow \int_{\mathcal{V}} f(s)X(ds) \quad \forall f \in C^\psi(\mathcal{V}).$$

Wykorzystując tę konstrukcję otrzymujemy następujący wniosek z twierdzenia 1.13.

WNIOSEK. 1.15 *Niech \succ będzie relacją preferencji na przestrzeni $\mathcal{M}^\psi(\mathcal{V})$, która jest ciągła względem opisanej wyżej ψ -słabej topologii oraz spełnia aksjomaty niezależności i Archimedesza. Wtedy relacja \succ posiada reprezentację von Neumanna-Morgensterna*

$$U(X) = \int_{\mathcal{V}} u(s)X(ds),$$

gdzie funkcja $u \in C^\psi(\mathcal{V})$. Reprezentacja U oraz funkcja u są wyznaczone jednoznacznie z dokładnością do przekształcenia afinicznego.

Ekonomiści zwykle oczekują, że funkcja użyteczności jest ściśle rosnąca. Aby skonstruowana w twierdzeniu 1.13 funkcja użyteczności była ściśle rosnąca potrzebne są dodatkowe założenia.

DEFINICJA. 1.16 *Relację preferencji \succ nazywamy monotoniczną, jeśli dla każdego $x, y \in \mathcal{V}$ zachodzi implikacja: $x \succ y \Rightarrow \delta_x \succ \delta_y$.*

TWIERDZENIE. 1.17 *Założmy, że relacja preferencji \succ na \mathcal{M} posiada reprezentację von Neumanna-Morgensterna*

$$U(X) = \int_{\mathcal{V}} u(s)X(ds)$$

dla pewnej funkcji borelowskiej u . Relacja \succ jest monotoniczna wtedy i tylko wtedy, gdy funkcja u jest ściśle rosnąca.

Dowód. Ponieważ $U(\delta_x) = u(x)$ a funkcja u jest ściśle rosnąca, to nierówność $x \succ y$ jest równoważna nierówności $u(x) > u(y)$, która jest równoważna nierówności $U(\delta_x) > U(\delta_y)$ a ta z kolei jest równoważna relacji $\delta_x \succ \delta_y$. ■

W ciągu dalszego wykładu będziemy rozpatrywali wyłącznie relacje preferencji na \mathcal{M} zadane przez reprezentacje von Neumanna-Morgensterna z funkcjami użyteczności, które są ciągłe i ściśle rosnące.

Uwaga. 4 *Relacje preferencji określiliśmy na rozkładach a nie na zmiennych losowych, mimo że same inwestycje a także ich stopy zwrotu traktujemy jako zmienne losowe. Taki sposób postępowania wynika z faktu, że przestrzeń \mathcal{V} , na której są zdefiniowane rozkłady, ma naturalną interpretację ekonomiczną (zbiór wartości inwestycji), gdy tymczasem przestrzeń probabilistyczna $(\Omega, \mathcal{F}, \mathbb{P})$, na której inwestycje oraz ich stopy zwrotu są zdefiniowane, nie ma oczywistej interpretacji ekonomicznej (zwykle mówi się, że zmienne te są zdefiniowane na "stanach rynku", ale nie jest oczywiste jaki jest zbiór tych stanów). W dalszym ciągu będziemy jednak utożsamiali rozkłady X ze zmiennymi losowymi (które będziemy oznaczali tym samym symbolem X). Takie postępowanie nie będzie prowadziło do nieporozumień, ponieważ w przypadku zmiennych losowych będą nas interesowały jedynie ich momenty, które są jednoznacznie wyznaczone przez rozkłady zmiennych losowych.*

W ramach tej umowy reprezentację von Neumanna-Morgensterna z definicji 1.10 można zapisać w postaci

$$U(X) = \mathbb{E}[u(X)].$$

Uwaga. 5 *Przy założeniu, że wartość inwestycji jest proporcjonalna do zainwestowanego kapitału początkowego, funkcję użyteczności można zdefiniować na zbiorze rozkładów stóp zwrotu portfeli inwestycyjnych \bar{R}_x . Definiuje ona relację preferencji wśród strategii/portfeli inwestycyjnych, którą też oznaczymy symbolem \succ . Wtedy $x \succ y \iff \mathbb{E}(u(\bar{R}_x)) > \mathbb{E}(u(\bar{R}_y))$.*

Uwaga. 6 *Zbiór wszystkich portfeli inwestycyjnych można utożsamić ze zbiorem wszystkich wektorów $(n^1, \dots, n^d) \in \mathbb{R}^d$ (lub wektorów $(x_1, \dots, x_d) \in \mathbb{R}^d$, takich że $\sum_{i=1}^d x_i = 1$). Ograniczenia natury ekonomicznej (lub matematycznej) mogą jednak wymusić ograniczenia zbioru inwestycji. Na przykład ograniczenie $V \geq 0$ oznacza ekonomicznie, że wykluczamy bankructwo. Ten warunek wprowadza ograniczenie na wektor (n^1, \dots, n^d) , chociaż trudno to ograniczenie efektywnie wypisać, bo ma ono postać $n^\top S \geq 0$, czyli zależy od wektora cen instrumentów, które są zmiennymi losowymi (analogiczny warunek dla portfela x ma postać $\bar{R} \geq -1$). Dlatego w dalszym ciągu będziemy rozwiązywać zadanie optymalizacji w zbiorze $\mathcal{Z} \subset \mathbb{R}^d$ (dla wektorów (n^1, \dots, n^d)) lub w zbiorze $\mathbb{X} \subset \{x \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1\}$ (dla wektorów (x_1, \dots, x_d)), które nazwiemy zbiorami dopuszczalnych strategii/portfeli inwestycyjnych a odpowiadające im inwestycje – inwestycjami dopuszczalnymi. Zbiór inwestycji dopuszczalnych stanowiących podzbiór \mathcal{M} , oznaczać będziemy $\mathcal{M}_{\mathcal{Z}}$. Szczegółową postać strategii dopuszczalnych (lub inwestycji dopuszczalnych) będziemy definiować w trakcie wykładu.*

DEFINICJA. 1.18 *Inwestycja optymalna to rozwiązanie następującego zadania*

$$\sup_{(n^1, \dots, n^d) \in \mathcal{Z}} \mathbb{E}(u(V)) \tag{1.4}$$

przy założeniu, że kapitał początkowy wynosi V_0 .

To zadanie optymalizacyjne nie zawsze posiada rozwiązanie, a często jeśli posiada, to niejednoznaczne.

Uwaga. 7 *Jeśli wartość inwestycji jest proporcjonalna do kapitału początkowego, to zadanie (1.4) sprowadza się do*

$$\sup_{x \in \mathbb{X}} \mathbb{E}(u((1 + \sum_{i=1}^d x_i R^i) V_0)).$$

1.2 Awersja do ryzyka

W klasycznej ekonomii z każdą inwestycją wiąże się dwie wielkości:

1. korzyść z inwestycji,
2. ryzyko inwestycji.

Pojęcia te pozostawimy na poziomie intuicji, precyzując dalej konkretne miary korzyści i ryzyka. Podkreślmy tylko, że w naszych rozważaniach ograniczymy się tylko do tzw. *ryzyka rynkowego inwestycji*.

Para-definicja: Ryzykiem rynkowym nazywamy możliwość wyniku inwestycji różnego od oczekiwanych korzyści na skutek zmian stanu rynku.

DEFINICJA. 1.19 *Niech dana będzie inwestycja dopuszczalna X , taka że $\mathbb{E}(X) < +\infty$, $\mathbb{E}(u(X)) < +\infty$.*

Pewnym ekwiwalentem (ang. certainty equivalent) X nazywamy stałą $C[X]$, taką że $u(C[X]) = \mathbb{E}(u(X))$.

Różnicę $\mathbb{E}(X) - C[X] = \pi(X)$ nazywamy premią za ryzyko (ang. Pratt risk premium).

DEFINICJA. 1.20 *Jeśli jako inwestora rozumiemy osobę mającą ustaloną funkcję użyteczności, to:*

1. *Inwestor ma awersję do ryzyka, jeśli $C[X] \leq \mathbb{E}(X)$.*
2. *Inwestor jest neutralny względem ryzyka, jeśli $C[X] = \mathbb{E}(X)$.*
3. *Inwestor poszukuje ryzyka, jeśli $C[X] \geq \mathbb{E}(X)$.*

Dla wszystkich dopuszczalnych inwestycji $X \in \mathcal{M}_{\mathbb{Z}}$.

TWIERDZENIE. 1.21 *Niech u będzie funkcją użyteczności inwestora. Wtedy*

1. *u jest wklęsła \Leftrightarrow inwestor ma awersję do ryzyka,*
2. *u jest liniowa \Leftrightarrow inwestor jest neutralny względem ryzyka,*
3. *u jest wypukła \Leftrightarrow inwestor poszukuje ryzyka.*

Dowód. Udowodnimy tylko przypadek 1.

„ \Rightarrow ” Z definicji $u(C[X]) = \mathbb{E}(u(X))$, zaś z wklęsłości u (nierówność Jensena) $u(C[X]) = \mathbb{E}(u(X)) \leq u(\mathbb{E}(X))$. Ponieważ u jest rosnąca, to $C[X] \leq \mathbb{E}(X)$.

„ \Leftarrow ” Ponieważ $C[X] \leq \mathbb{E}(X)$ dla każdego $X \in \mathcal{M}_{\mathcal{Z}}$, to dla u rosnącej $u(C[X]) = \mathbb{E}(u(X)) \leq u(\mathbb{E}(X))$. Z czego wynika, że u jest wklęsła. ■

DEFINICJA. 1.22 Niech u i v będą dwiema funkcjami użyteczności. Mówimy, że funkcji u odpowiada większa awersja do ryzyka niż funkcji v , jeśli $\pi_u(X) \geq \pi_v(X)$ dla wszystkich dopuszczalnych inwestycji $X \in \mathcal{M}_{\mathcal{Z}}$.

TWIERDZENIE. 1.23 Niech u i v będą dwiema funkcjami użyteczności o tej samej dziedzinie, które są klasy C^2 i mają dodatnie pierwsze pochodne. Wtedy następujące warunki są równoważne:

1. $-\frac{u''}{u'}(x) \geq -\frac{v''}{v'}(x)$ dla każdego $x \in \mathbb{R}$,
2. istnieje funkcja wklęsła Φ , taka że $u(x) = \Phi(v(x))$,
3. $\pi_u(X) \geq \pi_v(X)$ dla wszystkich dopuszczalnych inwestycji $X \in \mathcal{M}_{\mathcal{Z}}$.

Dowód.

„1 \Rightarrow 2”

Z dodatniości pochodnej wynika ścisła monotoniczność v , a ze ścisłej monotoniczności istnienie v^{-1} . Niech $\Phi(y) = u \circ v^{-1}(y)$. Wtedy $u(x) = \Phi(v(x))$. Z różniczkowalności u i v wynika różniczkowalność Φ . Stąd $u'(x) = \Phi'(v(x))v'(x)$. Z dodatniości u' i v' wynika, że $\Phi' > 0$.

$$u''(x) = \Phi''(v(x))(v'(x))^2 + \Phi'(v(x))v''(x) = \Phi''(v')^2 + \frac{u'v''}{v'}.$$

Stąd $\frac{v''}{v'} - \frac{u''}{u'} = -\Phi''\frac{(v')^2}{u'}$. Ponieważ $\frac{(v')^2}{u'} > 0$ oraz $\frac{v''}{v'} - \frac{u''}{u'} \geq 0$, to $\Phi'' \leq 0$. Czyli Φ jest wklęsła.

„2 \Rightarrow 3”

Ponieważ $u(x) = \Phi(v(x))$, to $u(C_u[X]) = \mathbb{E}(\Phi(v(X)))$. Z wklęsłości Φ dostajemy $\mathbb{E}(\Phi(v(X))) \leq \Phi(\mathbb{E}(v(X)))$. Stąd $u(C_u[X]) \leq \Phi(v(C_v[X]))$. Ponieważ $u(\cdot) = \Phi(v(\cdot))$ jest rosnąca, to $C_u[X] \leq C_v[X]$, czyli $\pi_u(X) \geq \pi_v(X)$.

„3 \Rightarrow 1”

Przez zaprzeczenie „ $\neg 1 \Rightarrow \neg 3$ ”. Skoro „ $\neg 1$ ”, to istnieje a , takie że $-\frac{u''}{u'}(a) < -\frac{v''}{v'}(a)$. Z ciągłości istnieje otoczenie $N(a)$, w którym ta nierówność zachodzi. Niech X będzie zmienną losową, która w $N(a)$ przyjmuje wartości niezerowe i poza $N(a)$ jest zero. Pokażemy, że Φ z pierwszej części dowodu jest ściśle wypukła na $N(a)$. Z pierwszej części mamy $\frac{v''}{v'}(x) - \frac{u''}{u'}(x) = -k\Phi''(v(x))$, $k > 0$. Na $N(a)$ mamy $\frac{v''}{v'}(x) - \frac{u''}{u'}(x) < 0$, więc $\Phi'' > 0$. Czyli Φ jest ściśle wypukła na $N(a)$. Z drugiej części dowodu wynika, że jeśli Φ jest ściśle wypukła na $N(a)$, to dla X na $N(a)$ zachodzi $\pi_u(X) < \pi_v(X)$, czyli $\neg 3$. ■

DEFINICJA. 1.24 Wyrażenie $-\frac{u''}{u'}(x)$ nazywane jest współczynnikiem absolutnej awersji do ryzyka Arrowa–Pratta (ARA).

Wyrażenie $-x\frac{u''}{u'}(x)$ nazywane jest współczynnikiem względnej awersji do ryzyka Arrowa–Pratta (RRA).

Uwaga. 8 Premię za ryzyko inwestycji X można przy pewnych dodatkowych założeniach przedstawić jako liniową funkcję współczynnika awersji do ryzyka. Ponieważ definicja współczynnika awersji do ryzyka wymaga, aby funkcja użyteczności była dwukrotnie różniczkowalna, wykorzystamy tę własność rozwijając tę funkcję w szereg Taylora wokół punktu $\mathbb{E}(X)$

$$u(X) \approx u(\mathbb{E}(X)) + (X - \mathbb{E}(X))u'(\mathbb{E}(X)) + \frac{1}{2}(X - \mathbb{E}(X))^2 u''(\mathbb{E}(X)).$$

Biorąc wartość oczekiwaną powyższego wyrażenia otrzymujemy

$$\mathbb{E}(u(X)) \approx u(\mathbb{E}(X)) + \frac{1}{2}\text{Var}(X)u''(\mathbb{E}(X)).$$

Z drugiej strony, wykorzystując definicję pewnego ekwiwalentu X i dokonując ponownie rozwinięcia w szereg Taylora wokół punktu $\mathbb{E}(X)$ mamy

$$\mathbb{E}(u(X)) = u(C[X]) \approx u(\mathbb{E}(X)) + (C[X] - \mathbb{E}(X))u'(\mathbb{E}(X)).$$

Przyrównując dwa ostatnie wyrażenia dostajemy przybliżoną równość, która wiąże premię za ryzyko ze współczynnikiem absolutnej awersji do ryzyka

$$\pi(X) = \mathbb{E}(X) - C[X] \approx -\frac{1}{2} \frac{u''(\mathbb{E}(X))}{u'(\mathbb{E}(X))} \text{Var}(X). \quad (1.5)$$

1.3 Standardowe funkcje użyteczności

DEFINICJA. 1.25 Funkcja użyteczności u definiuje harmoniczną absolutną awersję do ryzyka HARA (ang. *harmonic absolute risk aversion*), jeśli odwrotność jej absolutnej awersji do ryzyka jest funkcją liniową: $(-\frac{u''}{u'})^{-1}(x) = \lambda_1 x + \lambda_2$.

TWIERDZENIE. 1.26 Jeśli u jest funkcją użyteczności typu HARA, to

$$u(x) = a\left(b + \frac{x}{c}\right)^{1-c}.$$

Funkcja ta jest zdefiniowana na zbiorze $b + \frac{x}{c} > 0$, czyli $x > -bc$.

Współczynniki a i c muszą spełniać przy tym warunek $a\frac{1-c}{c} > 0$, który gwarantuje, że $u' > 0$ a $u'' < 0$ (u jest ściśle wklęsła).

Wyjątek stanowi funkcja $u(x) = \ln(x)$, która też jest funkcją HARA, ale nie daje się zapisać wskazanym wzorem.

Dowód. Bezpośredni rachunek daje $-\frac{u''}{u'}(x) = \left(b + \frac{x}{c}\right)^{-1}$, czyli odwrotność jest liniową funkcją x . ■

Przykład. 1.27 (Przykłady funkcji użyteczności)

1. CARA (ang. *constant absolute risk aversion*)

Dla $c \rightarrow \infty$ $ARA \rightarrow const.$ Wtedy funkcja użyteczności przyjmuje postać $u(x) = -\frac{\exp(-Ax)}{A}$. ($ARA = A > 0$)

2. CRRA (ang. *constant relative risk aversion*)

To jest bardzo popularna funkcja użyteczności. Kładąc $b = 0$ dostajemy $RRA = c$ oraz

$$u(x) = \begin{cases} \frac{x^{1-c}}{1-c} & \text{dla } c \neq 1, \\ \ln(x) & \text{dla } c = 1. \end{cases}$$

Dla $c < 1$ $u(x) \in [0, \infty)$, a dla $c > 1$ $u(x) \in (-\infty, 0]$.

3. Kwadratowa funkcja użyteczności ($c = -1$)

Wtedy $a < 0$ i funkcja jest malejąca na przedziale (b, ∞) . Czyli funkcja użyteczności jest dobrze określona na przedziale $(-\infty, b]$.

Rozdział 2

Portfele optymalne bez ograniczenia na krótką sprzedaż

2.1 Optymalizacja portfela instrumentów ryzykownych

Markowitz tworząc matematyczny model optymalizacji portfeli inwestycyjnych nie odwołał się do teorii użyteczności, ale do pojęć dobrze zrozumiałych dla praktyków rynku inwestycyjnego: korzyści z inwestycji oraz ryzyka inwestycji.

DEFINICJA. 2.1 (Model Markowitza bez instrumentu wolnego od ryzyka)

Rynek składa się d instrumentów ryzykownych. Niech R będzie d -wymiarową zmienną losową przyszłych stóp zwrotu z inwestycji w instrumenty ryzykowne. Zmienna R ma d -wymiarowy rozkład normalny $\mathcal{N}_d(\mu, \Sigma)$, gdzie μ – wektor średnich (wartości oczekiwanych) zmiennej R , a Σ – macierz kowariancji. Niech $x = (x_1, \dots, x_d)^\top$ – wektor udziałów instrumentów finansowych w portfelu. Jako miarę korzyści z inwestycji przyjmujemy średnią stopę zwrotu portfela inwestycyjnego $\mu^\top x$ a jako miarę ryzyka – wariację tego portfela $x^\top \Sigma x$.

DEFINICJA. 2.2 (Zadanie optymalizacyjne dla wariancji) Zadanie polega na znalezieniu w modelu Markowitza bez instrumentu wolnego od ryzyka portfela o minimalnym ryzyku przy założonej oczekiwanej stopie zwrotu z inwestycji ρ . Rozwiązanie tego problemu sprowadza się do następującego zadania optymalizacji kwadratowej

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x = \rho, \\ e^\top x = 1, \end{cases} \quad (2.1)$$

gdzie e to wektor złożony z d jedynek, $e = (1, \dots, 1)^\top$.

TWIERDZENIE. 2.3 Jeśli macierz Σ jest symetryczna i dodatnio określona, a wektor μ nie jest równoległy do wektora e , to zadanie (2.1) posiada jednoznaczne

rozwiązanie w przestrzeni \mathbb{R}^d . Rozwiązanie to dane jest wzorem

$$x^* = \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu), \quad \text{gdzie } \lambda_1 = \frac{\gamma - \beta \rho}{\delta}, \lambda_2 = \frac{\alpha \rho - \beta}{\delta},$$

$$\alpha = e^\top \Sigma^{-1} e, \quad \beta = e^\top \Sigma^{-1} \mu, \quad \gamma = \mu^\top \Sigma^{-1} \mu, \quad \delta = \alpha \gamma - \beta^2.$$

Dowód. Tworzymy Lagrangian dla naszego problemu

$$L(x; \lambda_1, \lambda_2) = \frac{1}{2} x^\top \Sigma x + \lambda_1 (1 - e^\top x) + \lambda_2 (\rho - \mu^\top x).$$

Z warunku koniecznego istnienia rozwiązania dostajemy

$$\begin{cases} \Sigma x - \lambda_1 e - \lambda_2 \mu = 0, \\ 1 - e^\top x = 0, \\ \rho - \mu^\top x = 0. \end{cases}$$

Stąd $x^* = \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu)$. Podstawiając ten wynik do pozostałych równań dostajemy

$$\begin{cases} \lambda_1 e^\top \Sigma^{-1} e + \lambda_2 e^\top \Sigma^{-1} \mu = 1, \\ \lambda_1 \mu^\top \Sigma^{-1} e + \lambda_2 \mu^\top \Sigma^{-1} \mu = \rho. \end{cases}$$

Czyli

$$\begin{cases} \alpha \lambda_1 + \beta \lambda_2 = 1, \\ \beta \lambda_1 + \gamma \lambda_2 = \rho. \end{cases} \Rightarrow \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \rho \end{pmatrix}.$$

Pokażemy w następnym lemacie, że przy założeniach twierdzenia macierz $\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$ jest dodatnio określona. Stąd

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ \rho \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} \gamma & -\beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} 1 \\ \rho \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} \gamma - \beta \rho \\ \alpha \rho - \beta \end{pmatrix}.$$

Ponieważ funkcja Lagrange'a jest ściśle wypukła, to otrzymane rozwiązanie jest jedyne i jest ono minimum tej funkcji (patrz twierdzenie 9.12). ■

LEMAT. 2.4 Przy założeniach twierdzenia 2.3 stałe α, γ, δ są dodatnie.

Dowód. Ponieważ macierz Σ jest dodatnio określona, więc także Σ^{-1} jest dodatnio określona. Stąd wynika dodatniość α i γ . Zauważmy, że dla $(x, y) \neq (0, 0)$

$$0 < (xe^\top + y\mu^\top)\Sigma^{-1}(xe + y\mu) = x^2\alpha + 2xy\beta + y^2\gamma$$

Niech $y \neq 0$. Wtedy $\left(\frac{x}{y}\right)^2 \alpha + 2\left(\frac{x}{y}\right)\beta + \gamma > 0$. Ta forma kwadratowa jest dodatnia, więc $\Delta < 0$ dla $\alpha > 0$. Stąd wynika, że $\alpha\gamma - \beta^2 > 0$. ■

Uwaga. 9 Korzystając z elementarnych faktów z algebry liniowej można pokazać, że

$$\alpha \in \left[\|e\|^2 \lambda_{\min}(\Sigma^{-1}), \|e\|^2 \lambda_{\max}(\Sigma^{-1}) \right] = \left[\frac{\|e\|^2}{\lambda_{\max}(\Sigma)}, \frac{\|e\|^2}{\lambda_{\min}(\Sigma)} \right],$$

$$\gamma \in \left[\frac{\|\mu\|^2}{\lambda_{\max}(\Sigma)}, \frac{\|\mu\|^2}{\lambda_{\min}(\Sigma)} \right], \quad |\beta| < \frac{\|e\| \cdot \|\mu\|}{\lambda_{\min}(\Sigma)}.$$

Portfel optymalny x^* jest kombinacją portfeli $\Sigma^{-1}e$ i $\Sigma^{-1}\mu$. Okazuje się jednak, że z twierdzenia 2.3 wynika znacznie silniejszy wniosek, który można traktować jako wersję "Two Fund Theorem" (patrz Wniosek 2.19).

WNIOSEK. 2.5 Niech na rynku bez instrumentu wolnego od ryzyka dane będą dwa dowolne portfele optymalne x_1, x_2 będące rozwiązaniem problemu (2.1) z oczekiwanymi stopami zwrotu ρ_1 i ρ_2 , takimi że $\rho_1 \neq \rho_2$. Kombinacje liniowe tych dwóch portfeli wyznaczają dowolny inny portfel optymalny.

Dowód. Przekształcając wzór na portfel optymalny z twierdzenia 2.3 otrzymujemy

$$\begin{aligned} x &= \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu) = \Sigma^{-1} \left(\frac{\gamma - \beta \rho}{\delta} e + \frac{\alpha \rho - \beta}{\delta} \mu \right) \\ &= \frac{1}{\delta} \left(\gamma \Sigma^{-1} e - \beta \rho \Sigma^{-1} e + \alpha \rho \Sigma^{-1} \mu - \beta \Sigma^{-1} \mu \right) \\ &= \frac{1}{\delta} \left((\alpha \Sigma^{-1} \mu - \beta \Sigma^{-1} e) \rho + (\gamma \Sigma^{-1} e - \beta \Sigma^{-1} \mu) \right) = A \rho + B, \end{aligned} \quad (2.2)$$

gdzie A, B to wektory niezależne od portfela x .

Niech x będzie portfelem optymalnym o zwrocie ρ . Wtedy

$$\rho = \eta \rho_1 + (1 - \eta) \rho_2, \quad \text{gdzie } \eta = \frac{\rho - \rho_2}{\rho_1 - \rho_2}.$$

Z powyższej równości oraz zależności (2.2) wynika, że portfel x jest kombinacją liniową $x = \eta x_1 + (1 - \eta) x_2$. ■

2.2 Funkcje preferencji

Problem optymalizacyjny w modelu Markowitza jest jakby z innego porządku niż omawiana w poprzednim rozdziale preferencje wyrażane przez funkcje użyteczności. Rozpatruje on dwie liczbowe miary: korzyści i ryzyka, a następnie z ich pomocą buduje portfel optymalny. Powstaje pytanie: czy optymalizację w sensie Markowitza można sprowadzić do maksymalizacji funkcji użyteczności? Odpowiedź na to pytanie jest pozytywna, ale przy pewnych dodatkowych założeniach.

Wprowadzimy w tym celu pojęcie **funkcji preferencji**. Okaże się, że optymalizacja portfela w modelu Markowitza jest szczególnym przypadkiem optymalizacji funkcji preferencji.

DEFINICJA. 2.6 Funkcję $W(\Phi, \varrho)$ klasy C^1 , taką że $\frac{\partial W}{\partial \Phi} > 0$ oraz $\frac{\partial W}{\partial \varrho} < 0$, nazywamy funkcją preferencji, gdzie

Φ – jest miarą korzyści (np. oczekiwanym zyskiem, ang. reward),
 ϱ – jest miarą ryzyka.

Poniżej podamy przykłady funkcji preferencji, które otrzymujemy z funkcji użyteczności przy pewnych dodatkowych założeniach.

Przykład. 2.7 (Kwadratowa funkcja użyteczności) Jeśli funkcja użyteczności jest kwadratowa $u(V) = V - \frac{b}{2}V^2$, $b > 0$, to $\mathbb{E}(u(V)) = \mathbb{E}(V) - \frac{b}{2}\mathbb{E}(V^2)$. Zakładając, że wartość inwestycji jest proporcjonalna do kapitału początkowego $V = (1 + \bar{R})V_0$, otrzymujemy $\mathbb{E}(V) = V_0(1 + \mu)$, gdzie $\mu = \mathbb{E}(\bar{R})$. Mamy także $\mathbb{E}(V^2) = V_0^2(1 + 2\mu + \mathbb{E}(\bar{R}^2)) = V_0^2(1 + 2\mu + \sigma^2 + \mu^2)$, gdzie $\sigma^2 = \text{Var}(\bar{R})$. Zatem

$$\begin{aligned}\mathbb{E}(u(V)) &= V_0 + V_0\mu - \frac{b}{2}V_0^2 - bV_0^2\mu - \frac{b}{2}V_0^2\sigma^2 - \frac{b}{2}V_0^2\mu^2 \\ &= V_0 - \frac{b}{2}V_0^2 + (V_0 - bV_0^2)\mu - \frac{b}{2}V_0^2\sigma^2 - \frac{b}{2}V_0^2\mu^2 \equiv W(\mu, \sigma).\end{aligned}$$

Uwaga. 10 Arrow pokazał, że kwadratowa funkcja użyteczności implikuje, że inwestorzy posiadający większy kapitał inwestują mniej w ryzykowne instrumenty, co przeczy zarówno intuicji jak też empirycznym obserwacjom.

Przykład. 2.8 Rozważmy teraz dowolną funkcję użyteczności, ale założmy, że korzyść z inwestycji mierzy się nie wartością inwestycji, ale stopą zwrotu inwestycji. Dodatkowo założmy, że stopy zwrotu instrumentów, w które inwestujemy, mają wielowymiarowy rozkład normalny. Przy tym założeniu stopa zwrotu inwestycji \bar{R} , ma jednowymiarowy rozkład normalny. Niech μ będzie średnią stopy zwrotu tej inwestycji a σ^2 – jej wariacją. Wtedy

$$\begin{aligned}\mathbb{E}(u(\bar{R})) &= \frac{1}{\sqrt{2\pi}\sigma} \int u(\bar{R}) \exp\left(-\frac{(\bar{R} - \mu)^2}{2\sigma^2}\right) d\bar{R} \\ &= \frac{1}{\sqrt{2\pi}} \int u(\sigma z + \mu) e^{-\frac{z^2}{2}} dz \equiv W(\mu, \sigma).\end{aligned}$$

Sformułujemy teraz zadanie optymalizacyjne dla pewnej funkcji preferencji wynikającej z funkcji użyteczności. W dalszej części wykładu rozwiążemy to zadanie i udowodnimy, że rozwiązanie to jest identyczne jak dla oryginalnego problemu Markowitza minimalizacji wariacji.

LEMAT. 2.9 (Zadanie optymalizacyjne dla funkcji preferencji) Załóżmy, że

1. Wartość inwestycji jest liniową funkcją kapitału początkowego.

2. Rozkład stóp zwrotu instrumentów ryzykownych jest wielowymiarowym rozkładem normalnym $\mathcal{N}(\mu, \Sigma)$.
3. Użyteczność inwestycji mierzy się funkcją typu CARA, czyli $u(x) = -\frac{1}{k}e^{-kx}$, $k > 0$.

Przy powyższych założeniach maksymalizacja użyteczności wartości inwestycji sprowadza się do następującego zadania maksymalizacji funkcji preferencji $W(\mu, \Sigma) = \mu^\top x - \frac{A}{2}x^\top \Sigma x$

$$\begin{cases} \max_x \mu^\top x - \frac{A}{2}x^\top \Sigma x, \\ e^\top x = 1, \end{cases} \quad (2.3)$$

dla pewnej stałej A nazywanej awersją do ryzyka inwestora, która zależy od współczynnika awersji do ryzyka Arrowa-Pratta k , ale nie zależy od parametrów rozkładu μ i Σ .

Dowód. Jeśli stopy zwrotu tworzą d -wymiarowy rozkład normalny $R \sim \mathcal{N}_d(\mu, \Sigma)$, to stopa zwrotu $\bar{R}_x = R^\top x$ ma 1-wymiarowy rozkład normalny $\mathcal{N}(m, \sigma^2)$, gdzie $m = \mu^\top x$, $\sigma^2 = x^\top \Sigma x$. Z założenia $V = (1 + \bar{R}_x)V_0$, gdzie V_0 jest kapitałem początkowym (patrz str.11). Korzystając ze znanego faktu, że jeśli zmienna losowa X ma rozkład normalny $\mathcal{N}(m, \sigma^2)$, to $\mathbb{E}(e^X) = \exp(m + \frac{\sigma^2}{2})$, dostajemy

$$\mathbb{E}(u(V)) = -\frac{1}{k} \exp\left(-k\mathbb{E}(V) + \frac{k^2}{2}\text{Var}(V)\right).$$

Ponieważ $\mathbb{E}(V) = V_0(1 + \mu^\top x)$, $\text{Var}(V) = V_0^2 \text{Var}(\bar{R}_x) = V_0^2 x^\top \Sigma x$, więc

$$\mathbb{E}(u(V)) = -\frac{1}{k} \exp\left(-kV_0\left(\mu^\top x - \frac{kV_0}{2}x^\top \Sigma x\right)\right)e^{-kV_0}.$$

Stąd

$$\max_V \mathbb{E}(u(V)) \Leftrightarrow \max_x \mu^\top x - \frac{kV_0}{2}x^\top \Sigma x \Leftrightarrow \max_x \mu^\top x - \frac{A}{2}x^\top \Sigma x. \quad \blacksquare$$

Uwaga. 11 Wykorzystując uwagę 8 pokażemy związek współczynnika awersji A w problemie (2.3) ze współczynnikiem awersji Arrowa-Pratta dla dowolnej funkcji użyteczności a nie tylko dla funkcji typu CARA. Równość (1.5) z rozdziału 1 można dla inwestycji V zapisać w postaci

$$C[V] \approx \mathbb{E}(V) + \frac{1}{2} \frac{u''(\mathbb{E}(V))}{u'(\mathbb{E}(V))} \text{Var}(V).$$

Niech γ_{AP} oznacza współczynnik absolutnej awersji do ryzyka Arrowa-Pratta. Wykorzystując policzone w dowodzie lematu 2.9 wartości $\mathbb{E}(V)$ i $\text{Var}(V)$ możemy pewny ekwiwalent V zapisać przybliżoną równością

$$C[V] \approx V_0(1 + \mu^\top x) - \frac{\gamma_{AP}}{2} V_0^2 x^\top \Sigma x = V_0 + V_0 \left(\mu^\top x - \frac{\gamma_{AP} V_0}{2} x^\top \Sigma x \right).$$

Z powyższego równania widać, że maksymalizacja funkcji celu w problemie (2.3) jest równoważna maksymalizacji pewnego ekwiwalentu inwestycji V (wyraz V_0 występujący we wzorze na $C[V]$ nie ma wpływu na optymalizację, ponieważ jest stały). Porównanie obu funkcji daje równość $A = V_0 \gamma_{AP}$, czyli taki sam związek współczynnika awersji do ryzyka w problemie (2.3) ze współczynnikiem awersji do ryzyka Arrowa-Pratta jak dla funkcji użyteczności typu CARA.

TWIERDZENIE. 2.10 Problem (2.3) przy założeniach twierdzenia 2.3 ma jednoznaczne rozwiązanie $x^* = \frac{1}{A} \Sigma^{-1} (\lambda e + \mu)$, gdzie $\lambda = \frac{A - \beta}{\alpha}$.

Dowód. Tworzymy funkcję Lagrange'a dla rozważanego problemu maksymalizacji funkcji preferencji

$$L(x; \lambda) = \mu^\top x - \frac{A}{2} x^\top \Sigma x - \lambda(1 - e^\top x).$$

Z warunku koniecznego istnienia rozwiązania dostajemy

$$\begin{cases} -A \Sigma x + \lambda e + \mu = 0, \\ 1 - e^\top x = 0. \end{cases}$$

Stąd $x^* = \frac{1}{A} \Sigma^{-1} (\lambda e + \mu)$. Podstawiając ten wynik do drugiego równania znajdujemy λ

$$\lambda e^\top \Sigma^{-1} e + e^\top \Sigma^{-1} \mu = A \Rightarrow \lambda = \frac{A - \beta}{\alpha}.$$

Ponieważ funkcja Lagrange'a jest ściśle wklęsła, to otrzymane rozwiązanie jest jedyne i jest ono maksimum tej funkcji (patrz twierdzenie 9.12). ■

TWIERDZENIE. 2.11 Przy założeniach twierdzenia 2.3 rozwiązanie problemu (2.1) oraz problemu (2.3) są takie same, jeśli przyjąć, że oczekiwana stopa zwrotu z inwestycji ρ w problemie (2.1) jest równa oczekiwanej stopie zwrotu z inwestycji w problemie (2.3). W tym sensie problem (2.1) oraz problem (2.3) są równoważne.

Dowód. Przypomnijmy, że problem (2.1) ma postać

$$(2.1) \begin{cases} \min_{x_{(1)}} \frac{1}{2} x_{(1)}^\top \Sigma x_{(1)}, \\ \mu^\top x_{(1)} = \rho_{(1)}, \\ e^\top x_{(1)} = 1. \end{cases}$$

Analogicznie mamy problem (2.3)

$$(2.3) \begin{cases} \max_{x(2)} \mu^\top x(2) - \frac{A}{2} x(2)^\top \Sigma x(2), \\ e^\top x(2) = 1. \end{cases}$$

W obu sformułowaniach, dla podkreślenia, który problem rozważamy, dodaliśmy indeksy do odpowiednich zmiennych.

Równoważność obu problemów oznacza, że dla dowolnego $\rho_{(1)}$ istnieje A , takie że rozwiązania obu problemów są równe $x_{(1)}^* = x_{(2)}^*$ oraz $\rho_{(2)} \equiv \mu^\top x_{(2)}^* = \rho_{(1)}$.

Korzystając z postaci rozwiązania problemu (2.3) dostajemy

$$\begin{aligned} \rho_{(2)} &= \mu^\top x_{(2)}^* = \frac{A-\beta}{A\alpha} \mu^\top \Sigma^{-1} e + \frac{1}{A} \mu^\top \Sigma^{-1} \mu \\ &= \left(1 - \frac{\beta}{A}\right) \frac{\beta}{\alpha} + \frac{\gamma}{A} = \frac{1}{\alpha} \left(\beta - \frac{\beta^2}{A} + \frac{\gamma\alpha}{A}\right) = \frac{\beta + \frac{\delta}{A}}{\alpha}. \end{aligned}$$

Ponieważ

$$x_{(1)}^* = \lambda_1 \Sigma^{-1} e + \lambda_2 \Sigma^{-1} \mu, \quad x_{(2)}^* = \frac{\lambda}{A} \Sigma^{-1} e + \frac{1}{A} \Sigma^{-1} \mu,$$

to $x_{(1)}^* = x_{(2)}^*$, jeśli $\lambda_1 = \frac{\lambda}{A}$ oraz $\lambda_2 = \frac{1}{A}$.

Z wyrażenia $\rho_{(2)} = \frac{\beta + \frac{\delta}{A}}{\alpha}$ dostajemy $\frac{1}{A} = \frac{\alpha \rho_{(2)} - \beta}{\delta}$. Natomiast z rozwiązania problemu (2.1) mamy $\lambda_2 = \frac{\alpha \rho_{(1)} - \beta}{\delta}$. Jeśli więc λ_2 ma się równać $\frac{1}{A}$, to musi zachodzić równość $\rho_{(2)} = \rho_{(1)}$.

Wtedy

$$\begin{aligned} \lambda_1 &= \frac{\gamma - \beta \rho_{(1)}}{\delta} = \frac{\alpha \gamma - \alpha \beta \rho_{(1)}}{\alpha \delta} = \frac{\delta - \alpha \beta \rho_{(1)} + \beta^2}{\alpha \delta} = \frac{1 - \frac{\alpha \rho_{(1)} - \beta}{\delta} \beta}{\alpha} \\ &= \frac{1 - \frac{\alpha \rho_{(2)} - \beta}{\delta} \beta}{\alpha} = \frac{1 - \frac{\beta}{A}}{\alpha} = \frac{A - \beta}{\alpha} = \frac{\lambda}{A}, \end{aligned}$$

czyli zachodzi równość współczynników Lagrange'a, a tym samym równość rozwiązań $x_{(1)}^* = x_{(2)}^*$. ■

2.3 Granica portfelową

Zajmiemy się teraz obrazem zbioru portfeli optymalnych na płaszczyźnie zmiennych ρ i σ . Rozpocniemy od następującego wyniku pomocniczego.

TWIERDZENIE. 2.12 Niech x^* będzie rozwiązaniem optymalnym w problemach (2.1) lub (2.3). Wtedy wariancja optymalnego portfela wynosi

$$\sigma^2 = \frac{\alpha \hat{\rho}^2 - 2\beta \hat{\rho} + \gamma}{\delta},$$

gdzie $\hat{\rho} = \mu^\top x^*$ jest średnią stopy zwrotu portfela optymalnego.

Najmniejsza wartość wariancji jest osiągana dla $\hat{\rho} = \frac{\beta}{\alpha}$. Wtedy $\sigma^2 = \frac{1}{\alpha}$. Taki portfel nazywa się **portfelem minimalnej wariancji** (portfelem minimalnego ryzyka).

Dowód.

$$\sigma^2 = x^{*\top} \Sigma x^* = (\lambda_1 e + \lambda_2 \mu)^\top \Sigma^{-1} (\lambda_1 e + \lambda_2 \mu) = \lambda_1^2 \alpha + 2\lambda_1 \lambda_2 \beta + \lambda_2^2 \gamma.$$

Podstawiając λ_1, λ_2 z twierdzenia 2.3 dostajemy

$$\begin{aligned} \sigma^2 &= \lambda_1(\lambda_1 \alpha + \lambda_2 \beta) + \lambda_2(\lambda_1 \beta + \lambda_2 \gamma) \\ &= \lambda_1 \left(\frac{\gamma \alpha - \beta \rho \alpha}{\delta} + \frac{\alpha \rho \beta - \beta^2}{\delta} \right) + \lambda_2 \left(\frac{\gamma \beta - \beta^2 \rho}{\delta} + \frac{\alpha \rho \gamma - \beta \gamma}{\delta} \right) \\ &= \lambda_1 + \lambda_2 \rho = \frac{\gamma - \beta \rho + \alpha \rho^2 - \beta \rho}{\delta} = \frac{\alpha \rho^2 - 2\beta \rho + \gamma}{\delta}. \end{aligned}$$

Reszta jest oczywista. ■

WNIOSEK. 2.13 Na płaszczyźnie (ρ, σ^2) portfele optymalne tworzą parabolę o wierzchołku $(\frac{\beta}{\alpha}, \frac{1}{\alpha})$. Parabola ta nazywana jest **granicą portfelową** (ang. portfolio frontier). Połowa tej paraboli dla $\rho \geq \frac{\beta}{\alpha}$ (powyżej wierzchołka) nazywana jest **granicą efektywną** (ang. efficient frontier) (Rys. 2.1 (a)).

Na płaszczyźnie (ρ, σ) granica portfelowa zapisuje się równaniem

$$\frac{\sigma^2}{1/\alpha} - \frac{(\rho - \beta/\alpha)^2}{\delta/\alpha^2} = 1,$$

czyli jest to gałąź hiperboli odpowiadająca $\sigma > 0$ (Rys. 2.1 (b)). Jej asymptoty dane są równaniem $\rho = \frac{\beta}{\alpha} \pm \sigma \sqrt{\frac{\delta}{\alpha}}$.

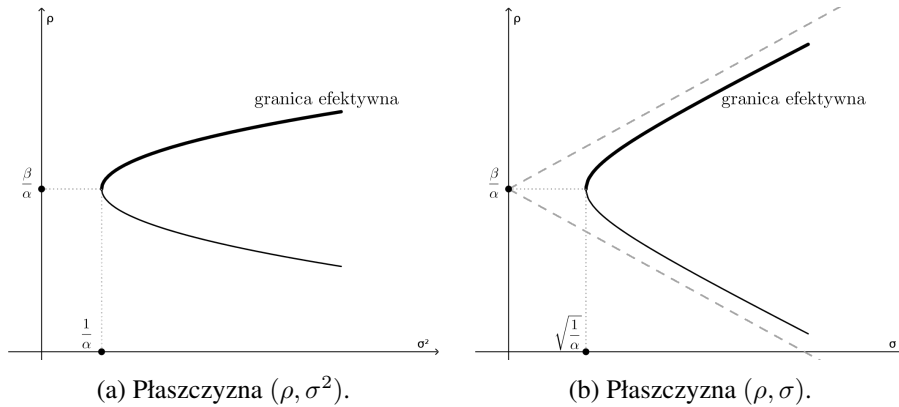
Dowód. Wzór na gałąź hiperboli bierze się z następujących przekształceń

$$\alpha \rho^2 - 2\beta \rho + \gamma = \alpha \left(\rho^2 - 2\frac{\beta}{\alpha} \rho + \frac{\beta^2}{\alpha^2} \right) + \gamma - \frac{\beta^2}{\alpha} = \alpha \left(\rho - \frac{\beta}{\alpha} \right)^2 + \frac{\delta}{\alpha}$$

Stąd

$$\sigma^2 = \frac{\alpha}{\delta} \left(\rho - \frac{\beta}{\alpha} \right)^2 + \frac{1}{\alpha} \quad \text{czyli} \quad \alpha \sigma^2 - \frac{\alpha^2}{\delta} \left(\rho - \frac{\beta}{\alpha} \right)^2 = 1,$$

a to jest równanie hiperboli z tezy. Równanie asymptot wynika ze znanych faktów z analizy i geometrii. ■



Rysunek 2.1: Granica portfelowa i granica efektywna.

WNIOSEK. 2.14 Niech $\mathbb{X} = \{x \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1\}$ będzie zbiorem portfeli dopuszczalnych. Rozważmy odwzorowanie $x \mapsto (\mu^\top x, x^\top \Sigma x)$. Niech zbiór M na płaszczyźnie (ρ, σ^2) będzie obrazem \mathbb{X} przy tym odwzorowaniu. Jeśli macierz Σ jest symetryczna i dodatnio określona a wektor μ nie jest równoległy do wektora e oraz $d \geq 3$, to zbiór M jest częścią płaszczyzny (ρ, σ^2) zawartą między ramionami paraboli tworzącej granicę portfelową. Dokładniej, jeśli $(\rho^*, (\sigma^*)^2)$ jest obrazem portfela optymalnego x^* , to półprosta $\{(\rho^*, \sigma^2) : \sigma \geq \sigma^*\}$ należy do zbioru M . Analogiczne stwierdzenie zachodzi dla obrazu portfeli dopuszczalnych \mathbb{X} przy odwzorowaniu $x \mapsto (\mu^\top x, \sqrt{x^\top \Sigma x})$ na płaszczyznę (ρ, σ) . Dla dwóch instrumentów ryzykownych ($d = 2$) zbiór M składa się wyłącznie z punktów granicy portfelowej.

Dowód. Niech

$$H_{d-2} = \{x \in \mathbb{R}^d : \mu^\top x = \rho\} \cap \{x \in \mathbb{R}^d : e^\top x = 1\}.$$

Z założenia, że wektor μ nie jest równoległy do e , wynika że H_{d-2} jest dla każdego ρ hiperpłaszczyzną o kowymiarze 2. Oznacza to, że dla każdego $\rho^* \in \mathbb{R}$ istnieją portfele dopuszczalne o stopie zwrotu ρ^* . Niech x^* będzie portfelem optymalnym odpowiadającym stopie zwrotu ρ^* . Wtedy $x^{*\top} \Sigma x^* = (\sigma^*)^2$. Jeśli $d = 2$, to H_{d-2} składa się wyłącznie z jednego punktu x^* . Wtedy zbiór M pokrywa się z granicą portfelową. Jeśli $d > 2$, to hiperpłaszczyzna H_{d-2} zawiera przynajmniej jedną prostą. Rozważmy taką prostą przechodzącą przez punkt x^* . Prosta ta dana jest równaniem $x(t) = a^* + b^*t$, gdzie a^* i b^* są wektorami w \mathbb{R}^d . Wtedy

$$x(t)^\top \Sigma x(t) = A^* + 2B^*t + C^*t^2,$$

gdzie $A^* = a^{*\top} \Sigma a^*$, $B^* = a^{*\top} \Sigma b^*$, $C^* = b^{*\top} \Sigma b^*$. Z założeń o macierzy Σ wynika, że $C^* > 0$, co oznacza nieograniczoność z góry trójmianu kwadratowego $A^* + 2B^*t + C^*t^2$. Ponieważ prosta $x(t)$ przechodzi przez punkt x^* , to minimum tego trójmianu $(A^*C^* - (B^*)^2)/C^*$ pokrywa się z $(\sigma^*)^2$ (optymalność x^*). Pokazaliśmy więc, że zbiór M zawiera całą półprostą $\{(\rho^*, \sigma^2) : \sigma \geq \sigma^*\}$. Zauważmy, że

każda prosta zawarta w H_{d-2} może być przedstawiona równaniem $x(t) = a_c + b_c t$, gdzie a_c i b_c są wektorami w \mathbb{R}^d . Obrazem takiej prostej na płaszczyźnie (ρ, σ^2) jest półprosta $\{(\rho^*, \sigma^2) : \sigma \geq \sigma_c\}$, gdzie $\sigma_c^2 = (A_c C_c - B_c^2)/C_c$ a stałe A_c, B_c, C_c są zdefiniowane analogicznie jak stałe A^*, B^*, C^* . Ponieważ $\sigma_c \geq \sigma^*$, to półprosta $\{(\rho^*, \sigma^2) : \sigma \geq \sigma_c\}$ jest podzbiorem półprostej $\{(\rho^*, \sigma^2) : \sigma \geq \sigma^*\}$. ■

Uwaga. 12 Wyróżnienie granicy efektywnej jest związane z faktem, że oryginalny problem Markowitza był sformułowany ogólniej niż zadanie (2.1). To ogólniejsze sformułowanie ma postać

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x \geq \rho, \\ e^\top x = 1. \end{cases} \quad (2.4)$$

Z postaci granicy portfelowej widać, że dla $\rho < \frac{\beta}{\alpha}$ rozwiązaniem zadania (2.4) jest portfel o większej oczekiwanej stopie zwrotu oraz mniejszej wariancji niż portfel będący rozwiązaniem zadania (2.1). Portfelem takim jest portfel minimalnej wariancji, którego obrazem jest punkt $(\beta/\alpha, 1/\alpha)$ należący do granicy efektywnej. Oznacza to, że granica efektywna jest obrazem rozwiązań optymalnych zadania (2.4).

Portfele optymalne w modelu Markowitza mogą być także otrzymane jako wynik dwukryterialnej optymalizacji Pareto, gdzie kryteriami optymalizacyjnymi są: maksymalizacja oczekiwanej stopy zwrotu oraz minimalizacja wariancji stopy zwrotu portfela. Te kryteria optymalizacji Pareto wyznaczają relację częściowego porządku w zbiorze portfeli dopuszczalnych. Tę relację można opisać następująco: niech dane będą dwa portfele $x, y \in \mathbb{X}$ o stopach zwrotu \bar{R}_x i \bar{R}_y ; portfel x jest nie gorszy niż portfel y , jeśli $\rho_x \geq \rho_y$ oraz $\sigma_x^2 \leq \sigma_y^2$, gdzie $\rho_z = \mathbb{E}[\bar{R}_z]$, $\sigma_z^2 = \text{Var}(\bar{R}_z)$. Mówimy także, że portfel x jest lepszy niż portfel y , jeśli przynajmniej jedna z powyższych nierówności jest ostra. **Portfelem efektywnym** względem tej relacji częściowego porządku nazywamy portfel, dla którego nie istnieje portfel lepszy. Zbiór punktów optymalnych w sensie Pareto nazywa się zbiorem **punktów efektywnych**. Zbiór punktów efektywnych w sensie Pareto na płaszczyźnie (ρ, σ^2) tworzy opisaną wcześniej granicę efektywną, która jest obrazem efektywnych portfeli. Działanie opisanej wyżej relacji częściowego porządku oraz związane z nią znajdowanie punktów efektywnych w sensie Pareto można łatwo zilustrować na Rys. 2.1 (a). Pamiętając, że obrazami portfeli dopuszczalnych są punkty położone między ramionami paraboli, można zauważyć, że portfel x jest nie gorszy niż y , jeśli punkt (ρ_x, σ_x^2) znajduje się w lewej górnej ćwiartce płaszczyzny wyznaczonej przez prostokątny układ współrzędnych o środku w punkcie (ρ_y, σ_y^2) .

2.4 Optymalizacja z instrumentem wolnym od ryzyka

Przejdziemy teraz do problemu znajdowania optymalnego portfela na rynku, na którym istnieje instrument wolny od ryzyka. Sprawą intuicyjnie oczywistą (ale także łatwą do formalnego udowodnienia) jest przyjęcie, że na rynku może istnieć tylko jeden instrument wolny od ryzyka, jeśli rynek ma być pozbawiony możliwości arbitrażu. Nie będziemy analizowali tego problemu zakładając od początku, że na rynku jest tylko jeden instrument wolny od ryzyka.

TWIERDZENIE. 2.15 (Zadanie optymalizacyjne z instrumentem wolnym od ryzyka)

Zakładamy następującą postać rynku finansowego:

1. Rynek składa się z d instrumentów ryzykownych, których stopy zwrotu mają łączny rozkład normalny $\mathcal{N}_d(\mu, \Sigma)$.
2. Prócz instrumentów ryzykownych na rynku jest jeden instrument wolny od ryzyka o deterministycznej stopie zwrotu μ_0 .
3. O parametrach rozkładu normalnego instrumentów ryzykownych zakładamy, że $\mu_0 e \neq \mu$ oraz Σ jest symetryczna i dodatnio określona.

Na tym rynku poszukujemy rozwiązania problemu optymalizacyjnego

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x + \mu_0 x_0 = \rho, \\ e^\top x + x_0 = 1, \end{cases} \quad (2.5)$$

Przy powyższych założeniach problem (2.5) posiada jednoznaczne rozwiązanie

$$\begin{aligned} x^* &= \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu) = \lambda_2 \Sigma^{-1}(\mu - \mu_0 e) = \frac{\rho - \mu_0}{\delta_0} \Sigma^{-1}(\mu - \mu_0 e), \\ x_0^* &= 1 - e^\top x = 1 - \lambda_2(\beta - \alpha \mu_0), \end{aligned}$$

gdzie $\lambda_1 = -\lambda_2 \mu_0$, $\lambda_2 = \frac{\rho - \mu_0}{\delta_0}$ a $\delta_0 = \mu_0^2 \alpha - 2\mu_0 \beta + \gamma > 0$.

Wariancja portfela optymalnego wynosi $\sigma^2 = \frac{(\rho - \mu_0)^2}{\delta_0}$. W szczególności portfel o minimalnej wariancji odpowiada oczekiwanej stopie zwrotu $\hat{\rho} = \mu_0$ i ma zerową wariancję. Odpowiada on inwestycji całego kapitału w instrument wolny od ryzyka $(\hat{x}, \hat{x}_0) = (0, 1)$. Dla takiej inwestycji współczynniki Lagrange'a są zerowe $\hat{\lambda}_1 = \hat{\lambda}_2 = 0$.

Dowód. Funkcja Lagrange'a dla tego problemu optymalizacyjnego ma postać

$$L(x, x_0; \lambda_1, \lambda_2) = \frac{1}{2} x^\top \Sigma x + \lambda_1 (1 - e^\top x - x_0) + \lambda_2 (\rho - \mu^\top x - \mu_0 x_0).$$

Warunek konieczny optymalności daje

$$\begin{cases} \Sigma x - \lambda_1 e - \lambda_2 x = 0, & (\text{pochodna po } x), \\ \lambda_1 + \lambda_2 \mu_0 = 0, & (\text{pochodna po } x_0), \\ e^\top x + x_0 = 1, \\ \mu^\top x + \mu_0 x_0 = \rho. \end{cases}$$

Z drugiego z tych równań dostajemy $\lambda_1 = -\lambda_2 \mu_0$. Równanie pierwsze w połączeniu z powyższym daje rozwiązanie optymalne

$$x^* = \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu) = \lambda_2 \Sigma^{-1}(\mu - \mu_0 e).$$

Po wstawieniu tego rozwiązania do trzeciego równania otrzymujemy

$$x_0^* = 1 - e^\top x^* = 1 - \lambda_2(\beta - \alpha \mu_0).$$

Wstawiając te rozwiązania do czwartego równania mamy

$$\rho = \mu^\top x^* + \mu_0 x_0^* = \lambda_2(\gamma - \mu_0 \beta) + \mu_0 - \lambda_2(\mu_0 \beta - \mu_0^2 \alpha) = \mu_0 + \lambda_2 \delta_0.$$

Otrzymujemy stąd wzór $\lambda_2 = \frac{\rho - \mu_0}{\delta_0}$.

Nierówność $\delta_0 > 0$ wynika z innego przedstawienia stałej δ_0 oraz założeń twierdzenia $\delta_0 = (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e) > 0$. Ponieważ $\sigma^2 = x^\top \Sigma x$, to dla rozwiązania optymalnego $x^* = \lambda_2 \Sigma^{-1}(\mu - \mu_0 e)$ dostajemy

$$\sigma^2 = \lambda_2^2 (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e) = \lambda_2^2 \delta_0 = \frac{(\rho - \mu_0)^2}{\delta_0}.$$

Reszta twierdzenia jest oczywista. ■

Podobnie jak w przypadku optymalizacji dla instrumentów ryzykownych można rozpatrywać zagadnienie z funkcją preferencji.

$$\begin{cases} \max_{(x, x_0)} \mu^\top x + \mu_0 x_0 - \frac{A}{2} x^\top \Sigma x \\ e^\top x + x_0 = 1 \end{cases} \quad (2.6)$$

TWIERDZENIE. 2.16 (Zadanie optymalizacyjne dla funkcji preferencji z instrumentem wolnym od ryzyka)

Przy założeniach twierdzenia 2.15 rozwiązania problemu (2.5) oraz problemu (2.6) są takie same, jeśli tylko przyjąć, że oczekiwana stopa zwrotu z inwestycji ρ w problemie (2.5) jest równa oczekiwanej stopie zwrotu z inwestycji w problemie (2.6). W tym sensie problem (2.5) oraz problem (2.6) są równoważne.

Dowód. Analogiczny jak twierdzenia 2.11 ■

Uwaga. 13 Z twierdzenia 2.15 wynika, że dla zadania optymalizacyjnego z instrumentem wolnym od ryzyka granica portfelowa na płaszczyźnie (ρ, σ) składa się z dwóch półprostych o początku w punkcie $(\mu_0, 0)$. Na płaszczyźnie (ρ, σ^2) granica jest parabolą o wierzchołku w punkcie $(\mu_0, 0)$.

TWIERDZENIE. 2.17 Wariancja portfela optymalnego w problemie (2.5) jest „prawie zawsze” mniejsza niż w problemie (2.1) dla tej samej oczekiwanej stopy zwrotu ρ .

Jeśli $\beta > \mu_0\alpha$, to granice portfelowe w problemach (2.1) i (2.5) mają jeden punkt styczności (ρ, σ) , gdzie

$$\rho = \mu_0 + \frac{\delta_0}{\beta - \mu_0\alpha} = \frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha}, \quad \sigma^2 = \frac{\delta_0}{(\beta - \mu_0\alpha)^2},$$

w którym rozwiązania optymalne są „identyczne” w tym sensie, że jeśli x^* jest rozwiązaniem problemu (2.1), to $(x^*, 0)$ jest rozwiązaniem problemu (2.5) (Rys. 2.2 (a)).

Jeśli $\beta = \mu_0\alpha$, to dla rozwiązania (x^*, x_0^*) problemu (2.5) mamy $x_0^* = 1$, $e^\top x^* = 0$ (Rys. 2.2 (b)). Wariancja rozwiązania problemu (2.1)

$$\sigma_{(1)}^2 = \frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}$$

i wariancja rozwiązania problemu (2.5)

$$\sigma_{(3)}^2 = \frac{(\rho - \mu_0)^2}{\delta_0}$$

dla tej samej oczekiwanej stopy zwrotu ρ , różnią się o $\frac{1}{\alpha}$

$$\frac{(\rho - \mu_0)^2}{\delta_0} + \frac{1}{\alpha} = \frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}.$$

Dowód. Przypadek $\beta > \mu_0\alpha$.

Problem (2.5) ma rozwiązanie odpowiadające $x_0^* = 0$ dla

$$\lambda_1 = -\frac{\mu_0}{\beta - \mu_0\alpha}, \quad \lambda_2 = \frac{1}{\beta - \mu_0\alpha}.$$

Wtedy z dowodu twierdzenia 2.15 wynika wzór

$$\rho = \mu_0 + \lambda_2\delta_0 = \mu_0 + \frac{\delta_0}{\beta - \mu_0\alpha} = \frac{\beta\mu_0 - \mu_0^2\alpha + \mu_0^2\alpha - 2\mu_0\beta + \gamma}{\beta - \mu_0\alpha} = \frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha}.$$

Podobnie

$$\sigma^2 = \frac{(\rho - \mu_0)^2}{\delta_0} = \frac{\delta_0}{(\beta - \mu_0\alpha)^2}.$$

Wstawiając λ_2 do wzoru z twierdzenia 2.15 dostajemy rozwiązanie $(x_{(3)}^*, 0)$ problemu (2.5), gdzie

$$x_{(3)}^* = \frac{1}{\beta - \mu_0\alpha} \Sigma^{-1}(\mu - \mu_0 e).$$

Wstawiając z kolei ρ do wzorów na λ_1 i λ_2 z twierdzenia 2.3 dostajemy rozwiązanie $x_{(1)}^*$ problemu (2.1)

$$\begin{aligned} x_{(1)}^* &= \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu) = \Sigma^{-1}\left(\frac{\gamma - \beta\rho}{\delta} e + \frac{\alpha\rho - \beta}{\delta} \mu\right) \\ &= \Sigma^{-1}\left(\frac{\gamma - \beta\frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha}}{\delta} e + \frac{\alpha\frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha} - \beta}{\delta} \mu\right) \\ &= \frac{1}{\delta(\beta - \mu_0\alpha)} \Sigma^{-1}\left((\gamma\beta - \mu_0\alpha\gamma - \gamma\beta + \mu_0\beta^2)e \right. \\ &\quad \left. + (\alpha\gamma - \mu_0\alpha\beta - \beta^2 + \mu_0\alpha\beta)\mu\right) \\ &= \frac{1}{\delta(\beta - \mu_0\alpha)} \Sigma^{-1}\left(-\mu_0(\alpha\gamma - \beta^2)e + (\alpha\gamma - \beta^2)\mu\right) \\ &= \frac{1}{\beta - \mu_0\alpha} \Sigma^{-1}(\mu - \mu_0 e). \end{aligned}$$

Czyli portfel $x_{(1)}^*$ będący rozwiązaniem problemu (2.1) jest identyczny z $x_{(3)}^*$ dla problemu (2.5). Łatwo można sprawdzić, że te portfele mają identyczne średnie stopy zwrotu i wariancje.

Obliczając styczną do hiperboli będącej rozwiązaniem problemu (2.1) w punkcie, w którym ma ona punkt wspólny z rozwiązaniem zagadnienia (2.5), dostajemy

$$\begin{aligned} \sigma(\rho) &= \delta^{-1/2}(\alpha\rho^2 - 2\beta\rho + \gamma)^{1/2}, \\ \frac{d\sigma}{d\rho} &= \delta^{-1/2}(\alpha\rho^2 - 2\beta\rho + \gamma)^{-1/2}(\alpha\rho - \beta). \end{aligned}$$

Funkcja $\alpha\rho^2 - 2\beta\rho + \gamma$ w punkcie $\rho = \frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha}$ ma wartość $\alpha\rho^2 - 2\beta\rho + \gamma = \frac{1}{(\beta - \mu_0\alpha)^2} \delta_0 \delta$.

Rzeczywiście

$$\begin{aligned} \alpha\rho^2 - 2\beta\rho + \gamma &= \alpha\rho \frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha} - \beta\rho + \gamma - \beta\rho \\ &= \frac{1}{\beta - \mu_0\alpha} \rho(\alpha\gamma - \mu_0\alpha\beta - \beta^2 + \mu_0\alpha\beta) + \gamma - \beta\frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha} \\ &= \frac{1}{\beta - \mu_0\alpha} \rho\delta - \frac{1}{\beta - \mu_0\alpha} (\beta\gamma - \mu_0\beta^2 - \beta\gamma + \mu_0\alpha\gamma) \\ &= \frac{1}{\beta - \mu_0\alpha} \rho\delta - \frac{1}{\beta - \mu_0\alpha} \mu_0\delta = \frac{\delta}{\beta - \mu_0\alpha} (\rho - \mu_0) \\ &= \frac{\delta}{(\beta - \mu_0\alpha)^2} (\gamma - \mu_0\beta - \mu_0\beta + \mu_0^2\alpha) = \frac{\delta\delta_0}{(\beta - \mu_0\alpha)^2}. \end{aligned}$$

Z powyższych przekształceń wynika także równość $\alpha\rho - \beta = \frac{\delta}{\beta - \mu_0\alpha}$, bo

$$\alpha\rho - \beta = \alpha \frac{\gamma - \mu_0\beta}{\beta - \mu_0\alpha} - \beta = \frac{1}{\beta - \mu_0\alpha} (\alpha\gamma - \mu_0\alpha\beta - \beta^2 + \mu_0\alpha\beta) = \frac{\delta}{\beta - \mu_0\alpha}.$$

Stąd styczna do hiperboli w punkcie $(\rho(x_{(1)}^*), \sigma(x_{(1)}^*))$ ma nachylenie

$$\frac{d\sigma}{d\rho} = \delta^{-1/2} \left(\frac{\delta\delta_0}{(\beta - \mu_0\alpha)^2} \right)^{-1/2} \frac{\delta}{\beta - \mu_0\alpha} = \delta_0^{-1/2}.$$

Z drugiej strony z twierdzenia 2.15 mamy

$$\sigma(\rho) = \delta_0^{-1/2}(\rho - \mu_0),$$

czyli prosta będąca granicą efektywną w problemie (2.5) ma nachylenie $\frac{d\sigma}{d\rho} = \delta_0^{-1/2}$.

Wynika stąd, że prosta będąca rozwiązaniem problemu (2.5) jest styczna w punkcie $(\rho(x_{(1)}^*), \sigma(x_{(1)}^*))$ do hiperboli będącej rozwiązaniem problemu (2.1).

Przypadek $\beta = \mu_0\alpha$.

Rozwiązanie jest wtedy oczywiste. Zauważmy, że w tym przypadku proste $\sigma(\rho) = \pm \frac{\rho - \mu_0}{\sqrt{\delta_0}}$ są asymptotami hiperboli. Z wzorów we wniosku 2.13 wynika

$$\sigma_{(1)}^2 = \frac{\alpha}{\delta} \left(\rho - \frac{\beta}{\alpha} \right)^2 + \frac{1}{\alpha}.$$

Jeśli zauważymy, że $\frac{\beta}{\alpha} = \mu_0$, to

$$\sigma_{(1)}^2 = \frac{\alpha}{\delta} (\rho - \mu_0)^2 + \frac{1}{\alpha}.$$

Wystarczy teraz skorzystać z definicji stałej δ_0 oraz warunku $\beta = \mu_0\alpha$, aby otrzymać

$$\delta_0 = \mu_0^2\alpha - 2\mu_0\beta + \gamma = \frac{\beta^2}{\alpha} - 2\frac{\beta^2}{\alpha} + \gamma = \frac{\alpha\gamma - \beta^2}{\alpha} = \frac{\delta}{\alpha}.$$

Stąd

$$\sigma_{(1)}^2 = \frac{(\rho - \mu_0)^2}{\delta_0} + \frac{1}{\alpha} = \sigma_{(3)}^2 + \frac{1}{\alpha}. \quad \blacksquare$$

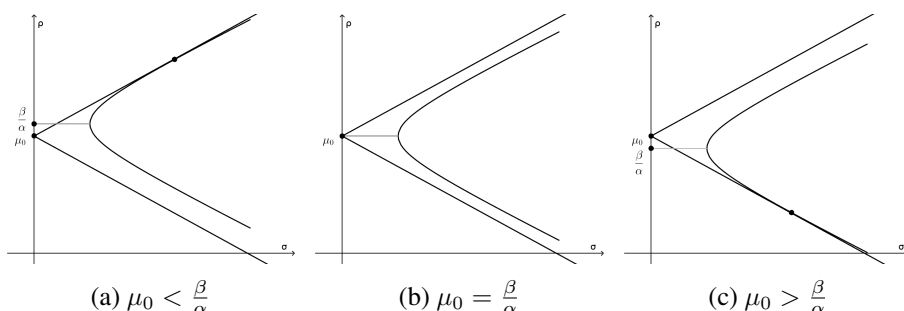
Przypadek $\beta < \mu_0\alpha$ opisuje twierdzenie, które po raz pierwszy pojawiło się w pracy Rockafellara, Uryaseva i Zabaranikina.

TWIERDZENIE. 2.18 *Jeśli $\beta < \mu_0\alpha$, to granice portfelowe w problemach (2.1) i (2.5) mają punkt styczności*

$$\rho = \mu_0 + \frac{\delta_0}{\beta - \mu_0\alpha}, \quad \sigma^2 = \frac{\delta_0}{(\beta - \mu_0\alpha)^2}.$$

Punkt ten nie leży na granicy efektywnej (Rys. 2.2 (c)). Punktem na granicy efektywnej o identycznej wariancji jest punkt odpowiadający oczekiwanej stopie zwrotu

$$\rho = \mu_0 - \frac{\delta_0}{\beta - \mu_0\alpha}.$$



Rysunek 2.2: Granice portfelowe dla problemów (2.1) i (2.5).

Dowód. Dowód pierwszej części twierdzenia wynika z dowodu twierdzenia 2.17. Punkt styczności nie leży na granicy efektywnej, bo $\rho = \mu_0 + \frac{\delta_0}{\beta - \mu_0 \alpha} < \mu_0$. Aby zrozumieć, jak wygląda rozwiązanie optymalne dla $\beta < \mu_0 \alpha$ należy sformułować problem (2.5) bardziej ogólnie

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ e^\top x + x_0 = 1, \\ \mu^\top x + \mu_0 x_0 \geq \rho. \end{cases} \quad (2.7)$$

To sformułowanie jest rzeczywiście bardziej ogólne niż oryginalne sformułowanie (2.5), bo portfel z identyczną wariancją a średnią stopy zwrotu nie mniejszą niż ρ jest nie gorszy od tego ze średnią stopy zwrotu ρ . Zauważmy przy tym, że sensowne ekonomicznie jest tylko poszukiwanie rozwiązania dla $\rho > \mu_0$. (Matematycznie rozwiązanie dla $\rho < \mu_0$ jest trywialne!)

Rozwiązanie problemu (2.7) wymaga skorzystania z twierdzenia Kuhna-Tuckera (twierdzenie 9.16). Z twierdzenia tego wynika, że jeśli rozwiązanie istnieje, to mnożnik Lagrange'a λ_2 odpowiadający ograniczeniu nierównościowemu spełnia warunek $\lambda_2 \geq 0$. Ponieważ z twierdzenia 2.15 dostajemy równość $e^\top x = \lambda_2(\beta - \mu_0 \alpha)$, to

$$\begin{aligned} e^\top x > 0 &\Leftrightarrow \mu_0 < \frac{\beta}{\alpha}, \\ e^\top x < 0 &\Leftrightarrow \mu_0 > \frac{\beta}{\alpha}. \end{aligned}$$

Przyjęcie w dowodzie twierdzenia 2.17 $x_0 = 0$ implikuje $e^\top x = 1$, czyli $\mu_0 < \frac{\beta}{\alpha}$. Jeśli chcemy rozpatrywać przypadek $\mu_0 > \frac{\beta}{\alpha}$, to należy wybrać ujemną wartość $e^\top x$. Najwygodniej jest przyjąć $e^\top x = -1$, bo odpowiada to portfelowi krótkich pozycji netto, co daje tę samą wariancję jak dla portfela spełniającego warunek $e^\top x = 1$ (odpowiada zamianie $x \mapsto -x$).

Wtedy $\lambda_2(\beta - \mu_0 \alpha) = -1$, czyli $\lambda_2 = \frac{1}{\mu_0 \alpha - \beta}$, co w połączeniu z twierdzeniem 2.15 daje

$$\rho = \mu_0 - \frac{\delta_0}{\beta - \mu_0 \alpha}, \quad \sigma^2 = \frac{\delta_0}{(\beta - \mu_0 \alpha)^2} \quad \text{oraz} \quad x_0 = 2.$$

Aby zilustrować graficznie przypadek opisany w twierdzeniu 2.18 rozważmy granicę portfelową na rynku z instrumentem wolnym od ryzyka

$$\rho = \mu_0 \pm \sigma\sqrt{\delta_0}.$$

To są dwie półproste przechodzące przez punkt $(\mu_0, 0)$ na płaszczyźnie (ρ, σ) (patrz Rys 2.3).

Niech punkt (ρ^*, σ^*) leży na prostej $\rho = \mu_0 - \sigma\sqrt{\delta_0}$. Punktowi (ρ^*, σ^*) odpowiada rozwiązanie (x^*, x_0^*) , gdzie

$$x^* = \frac{\rho^* - \mu_0}{\delta_0} \Sigma^{-1}(\mu - \mu_0 e), \quad x_0^* = 1 - \frac{\rho^* - \mu_0}{\delta_0}(\beta - \mu_0 \alpha).$$

Niech

$$e^\top x^* = \frac{\rho^* - \mu_0}{\delta_0}(\beta - \mu_0 \alpha) = \eta, \quad \text{wtedy} \quad x_0^* = 1 - \eta.$$

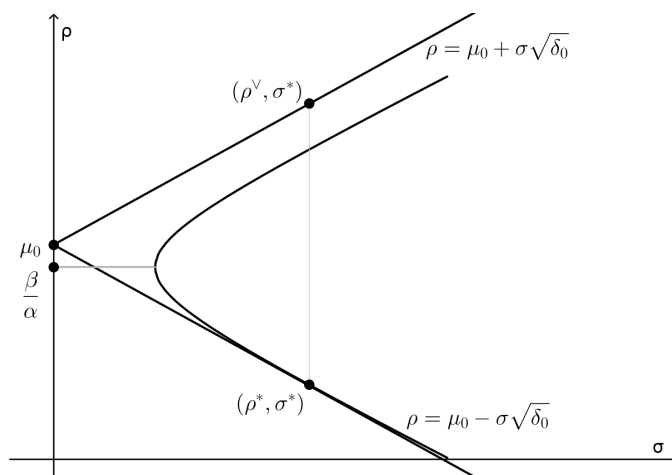
Niech (ρ^\vee, σ^\vee) będzie punktem na prostej $\rho = \mu_0 + \sigma\sqrt{\delta_0}$. Punktowi (ρ^\vee, σ^\vee) odpowiada rozwiązanie (x^\vee, x_0^\vee)

$$x^\vee = \frac{\rho^\vee - \mu_0}{\delta_0} \Sigma^{-1}(\mu - \mu_0 e), \quad x_0^\vee = 1 - \frac{\rho^\vee - \mu_0}{\delta_0}(\beta - \mu_0 \alpha),$$

dla którego

$$e^\top x^\vee = -\eta, \quad x_0^\vee = 1 + \eta,$$

bo z symetrii mamy $\rho^\vee = \mu_0 - (\rho^* - \mu_0) = 2\mu_0 - \rho^*$, więc $\rho^\vee - \mu_0 = \mu_0 - \rho^*$, czyli $x^\vee = -x^*$.



Rysunek 2.3: Granice portfelowe z twierdzenia 2.18.

Z rozważań tych wynika, że jeśli (x^*, x_0^*) jest portfelem, któremu odpowiada na płaszczyźnie (ρ, σ) punkt (ρ^*, σ^*) oraz $e^\top x^* = 1$ i $x_0^* = 0$, to punktowi (ρ^\vee, σ^\vee) na tej płaszczyźnie odpowiada portfel (x^\vee, x_0^\vee) , taki że $e^\top x^\vee = -1$ oraz $x_0^\vee = 2$.

WNIOSEK. 2.19 (Two Fund Theorem) Zakładamy, że $\mu_0 < \frac{\beta}{\alpha}$. Wtedy każdy portfel efektywny z instrumentem wolnym od ryzyka jest kombinacją liniową portfela złożonego wyłącznie z instrumentu wolnego od ryzyka i portfela złożonego wyłącznie z instrumentów ryzykownych (ten ostatni portfel nazywa się **portfelem stycznym**).

Dowód. Z twierdzenia 2.15 wynika, że jeśli (x^*, x_0^*) jest rozwiązaniem problemu (2.5), to

$$x^* = \lambda_2 \Sigma^{-1}(\mu - \mu_0 e).$$

Z twierdzenia 2.17 wiemy, że dla $x_0^* = 0$ mamy $\lambda_2 = \frac{1}{\beta - \mu_0 \alpha}$, czyli portfel $x_T = \frac{1}{\beta - \mu_0 \alpha} \Sigma^{-1}(\mu - \mu_0 e)$ jest portfelem stycznym. Odpowiada mu odchylenie standardowe $\sigma_T = \frac{\sqrt{\delta_0}}{\beta - \mu_0 \alpha}$.

Jeśli rozpatrujemy dowolny portfel efektywny (x, x_0) o odchyleniu standardowym σ , to ponieważ wszystkie takie portfele leżą na linii prostej, więc biorąc $\eta = \frac{\sigma}{\sigma_T}$ otrzymujemy $x = \eta x_T$, bo $\sigma^2 = x^\top \Sigma x = \eta^2 x_T^\top \Sigma x_T = \eta^2 \sigma_T^2$.

Ponieważ $e^\top x_T = 1$, więc $e^\top x = \eta$ oraz $x_0 = 1 - \eta$.

Otrzymujemy więc rozwiązanie w postaci kombinacji liniowej, o której mowa w tezie, $(x, x_0) = (\eta x_T, 1 - \eta)$. ■

Portfele optymalne na rynku z instrumentem wolnym od ryzyka mają ścisły związek z portfelami optymalnymi w zadaniu maksymalizacji współczynnika Sharpe'a portfela.

DEFINICJA. 2.20 Współczynnikiem Sharpe'a portfela inwestycyjnego x instrumentów ryzykownych o stopie zwrotu \bar{R}_x na rynku z instrumentem wolnym od ryzyka nazywamy iloraz

$$SR_x = \frac{\mathbb{E}[\bar{R}_x] - \mu_0}{\sigma(\bar{R}_x)},$$

gdzie μ_0 jest stopą zwrotu instrumentu wolnego od ryzyka.

Uwaga. 14 Dla $\frac{\beta}{\alpha} > \mu_0$ rozwiązanie zadania

$$\begin{cases} \max_x SR_x, \\ e^\top x = 1, \end{cases} \quad (2.8)$$

polega na znalezieniu portfela stycznego. Wynika to z obserwacji, że rozwiązanie zadania (2.8) oznacza znalezienie na płaszczyźnie (ρ, σ) prostej mającej maksymalne nachylenie i przechodzącej przez punkt $(\mu_0, 0)$ oraz punkt (ρ_x, σ_x) będący obrazem portfela dopuszczalnego x . Łatwo zauważyć, że taką prostą jest prosta przechodząca przez punkt styczności granicy efektywnej dla rynku z instrumentem wolnym od ryzyka i granicy efektywnej dla rynku bez instrumentu wolnego od ryzyka, czyli punkt będący obrazem portfela stycznego.

Rozdział 3

Model wyceny dóbr kapitałowych CAPM

W tym rozdziale będziemy w dalszym ciągu zajmować się rynkiem w modelu Markowitza. Na rynku mamy d instrumentów ryzykownych o stopach zwrotu, które są zmiennymi losowymi o wielowymiarowym rozkładzie normalnym z wektorem wartości oczekiwanej μ oraz macierzą kowariancji Σ . Zakładamy, że $\mu \not\parallel e$ a macierz Σ jest symetryczna i dodatnio określona. Poza tym na rynku jest jeden instrument wolny od ryzyka o stopie zwrotu μ_0 .

3.1 Portfele ortogonalne

DEFINICJA. 3.1 *Jeśli na rynku bez instrumentu wolnego od ryzyka mamy dwa portfele dopuszczalne x^1 i x^2 , to kowariancją portfeli x^1 i x^2 nazywamy wyrażenie $\text{Cov}(x^1, x^2) = (x^1)^\top \Sigma x^2$.*

DEFINICJA. 3.2 *Dwa dopuszczalne portfele inwestycyjne x^1 i x^2 na rynku bez instrumentu wolnego od ryzyka nazywamy ortogonalnymi, jeśli $\text{Cov}(x^1, x^2) = 0$.*

LEMAT. 3.3 *Jeśli mamy dwa portfele x^1 i x^2 optymalne, tj. ich obrazy leżą na granicy portfelowej, to ich kowariancja wyraża się wzorem*

$$\text{Cov}(x^1, x^2) = \frac{\rho(x^1)\rho(x^2)\alpha - \rho(x^1)\beta - \rho(x^2)\beta + \gamma}{\delta},$$

gdzie $\rho(x)$ oznacza stopę zwrotu portfela x .

Dowód. Z twierdzenia 2.3 mamy wzór na skład portfela x

$$x = \frac{\gamma - \beta\rho}{\delta} \Sigma^{-1} e + \frac{\alpha\rho - \beta}{\delta} \Sigma^{-1} \mu.$$

Korzystając z tego wzoru obliczamy

$$\begin{aligned}
(x^1)^\top \Sigma x^2 &= \frac{1}{\delta^2} \left(\alpha(\gamma - \beta\rho(x^1))(\gamma - \beta\rho(x^2)) + \beta(\gamma - \beta\rho(x^1))(\alpha\rho(x^2) - \beta) \right. \\
&\quad \left. + \beta(\gamma - \beta\rho(x^2))(\alpha\rho(x^1) - \beta) + \gamma(\alpha\rho(x^1) - \beta)(\alpha\rho(x^2) - \beta) \right) \\
&= \frac{1}{\delta^2} \left((\gamma - \beta\rho(x^1))(\alpha\gamma - \alpha\beta\rho(x^2) + \alpha\beta\rho(x^2) - \beta^2) \right. \\
&\quad \left. + (\alpha\rho(x^1) - \beta)(\beta\gamma - \beta^2\rho(x^2) + \alpha\gamma\rho(x^2) - \beta\gamma) \right) \\
&= \frac{1}{\delta^2} \left((\gamma - \beta\rho(x^1))\delta + (\alpha\rho(x^1) - \beta)\rho(x^2)\delta \right) \\
&= \frac{\rho(x^1)\rho(x^2)\alpha - \rho(x^1)\beta - \rho(x^2)\beta + \gamma}{\delta}.
\end{aligned}$$

■

LEMAT. 3.4 Niech x będzie portfelem optymalnym różnym od portfela minimalnego ryzyka (portfela o minimalnej wariancji). Portfelowi temu odpowiada dokładnie jeden optymalny portfel ortogonalny x^\perp dany wzorem

$$x^\perp = \frac{\rho(x)\Sigma^{-1}e - \Sigma^{-1}\mu}{\rho(x)\alpha - \beta}.$$

Stopa zwrotu tego portfela wynosi

$$\rho(x^\perp) = \frac{\rho(x)\beta - \gamma}{\rho(x)\alpha - \beta}.$$

Dowód. Warunek ortogonalności daje

$$\rho(x)\rho(x^\perp)\alpha - \rho(x)\beta - \rho(x^\perp)\beta + \gamma = 0.$$

Jeśli $\rho(x) \neq \frac{\beta}{\alpha}$, czyli x nie jest portfelem o minimalnej wariancji, to

$$\rho(x^\perp) = \frac{\rho(x)\beta - \gamma}{\rho(x)\alpha - \beta}.$$

Z twierdzenia 2.3 mamy wzór na skład portfela x

$$x = \frac{\gamma - \beta\rho}{\delta}\Sigma^{-1}e + \frac{\alpha\rho - \beta}{\delta}\Sigma^{-1}\mu.$$

Podstawiając do tego wzoru $\rho(x^\perp)$ otrzymamy wzór na portfel x^\perp . Aby otrzymać ten wzór obliczamy

$$\begin{aligned}
\gamma - \beta\rho(x^\perp) &= \gamma - \beta \frac{\rho(x)\beta - \gamma}{\rho(x)\alpha - \beta} = \frac{1}{\rho(x)\alpha - \beta} (\rho(x)\alpha\gamma - \gamma\beta - \rho(x)\beta^2 + \beta\gamma) \\
&= \frac{\rho(x)\delta}{\alpha\rho(x) - \beta},
\end{aligned}$$

$$\begin{aligned}\alpha\rho(x^\perp) - \beta &= \alpha \frac{\rho(x)\beta - \gamma}{\rho(x)\alpha - \beta} - \beta = \frac{1}{\rho(x)\alpha - \beta} (\rho(x)\alpha\beta - \alpha\gamma - \rho(x)\alpha\beta + \beta^2) \\ &= \frac{-\delta}{\alpha\rho(x) - \beta}.\end{aligned}$$

Podstawiając powyższe wyrażenia do wzoru na skład portfela dostajemy

$$x^\perp = \frac{\rho(x)\Sigma^{-1}e - \Sigma^{-1}\mu}{\rho(x)\alpha - \beta}.$$

■

TWIERDZENIE. 3.5 Niech x będzie portfelem optymalnym różnym od portfela o minimalnej wariancji, a x^\perp portfelem optymalnym do niego ortogonalnym. Niech ρ będzie średnią stopą zwrotu portfela x , a ρ^\perp portfela x^\perp . Wtedy dla i -tego instrumentu ryzykownego na rynku zachodzi wzór

$$\frac{\mu_i - \rho^\perp}{\sigma_{i,x}^2} = \frac{\rho - \rho^\perp}{\sigma^2(x)},$$

gdzie μ_i jest średnią stopy zwrotu instrumentu i -tego, $\sigma^2(x)$ jest wariancją portfela x , a $\sigma_{i,x}^2$ jest kowariancją instrumentu i -tego z portfelem x .

W przypadku gdy kowariancja instrumentu i -tego z portfelem x jest zero (instrument i -ty jest ortogonalny do portfela x) teza przyjmuje postać

$$\mu_i = \rho^\perp.$$

Dowód. Korzystając z lematu 3.4 zamieniając x na x^\perp dostajemy

$$\rho = \frac{\rho^\perp\beta - \gamma}{\rho^\perp\alpha - \beta}. \quad (3.1)$$

$$x = \frac{1}{\beta - \alpha\rho^\perp} \Sigma^{-1}(\mu - \rho^\perp e). \quad (3.2)$$

Z tego wzoru liczymy wariancję stóp zwrotu portfela x

$$\begin{aligned}\sigma^2(x) &= x^\top \Sigma x = \frac{1}{(\beta - \alpha\rho^\perp)^2} (\mu - \rho^\perp e)^\top \Sigma^{-1} (\mu - \rho^\perp e) \\ &= \frac{1}{(\beta - \alpha\rho^\perp)^2} (\gamma - 2\rho^\perp\beta + (\rho^\perp)^2\alpha) \\ &= \frac{1}{(\beta - \alpha\rho^\perp)^2} (\gamma - 2\rho^\perp\beta + (\rho^\perp)^2\alpha - \gamma + (\rho + \rho^\perp)\beta - \rho\rho^\perp\alpha) \\ &= \frac{\rho - \rho^\perp}{\beta - \alpha\rho^\perp}.\end{aligned}$$

Ze wzoru (3.2) wynika

$$\text{Cov}(e_i, x) \equiv e_i^\top \Sigma x = (\Sigma x)_i = \sigma_{i,x}^2 = \frac{\mu_i - \rho^\perp}{\beta - \alpha \rho^\perp},$$

gdzie $e_i = (0, \dots, 1, \dots, 0)^\top$, a 1 stoi na i -tym miejscu. Wykorzystując wzór na $\sigma^2(x)$ do wyeliminowania $\beta - \alpha \rho^\perp$, dostajemy

$$\mu_i - \rho^\perp = \frac{\sigma_{i,x}^2}{\sigma^2(x)} (\rho - \rho^\perp),$$

■

LEMAT. 3.6 *Rozważmy granicę portfelową na płaszczyźnie zmiennych (ρ, σ^2) . Jeśli średnia stopy zwrotu $\rho(x)$ portfela x na granicy portfelowej jest większa niż $\frac{\beta}{\alpha}$ (czyli portfel leży na granicy efektywnej), to średnia stopy zwrotu portfela ortogonalnego $\rho(x^\perp)$ jest mniejsza od $\frac{\beta}{\alpha}$ (czyli portfel ten leży na dolnej części granicy portfelowej). Dokładniej ρ^\perp jest punktem na osi 0ρ , w którym tę oś przecina prosta przechodząca przez punkt $(\rho(x), \sigma^2(x))$ odpowiadający portfelowi x oraz punkt $(\rho_{mv}, \sigma_{mv}^2)$ odpowiadający portfelowi optymalnemu o minimalnej wariancji (patrz Rys. 3.1).*

Dowód. Równanie prostej przechodzącej przez punkty $(\rho(x), \sigma^2(x))$ i $(\rho_{mv}, \sigma_{mv}^2)$ ma postać

$$\frac{\sigma^2 - \sigma^2(x)}{\sigma_{mv}^2 - \sigma^2(x)} = \frac{\rho - \rho(x)}{\rho_{mv} - \rho(x)}.$$

Prosta ta przecina oś 0ρ ($\sigma^2 \equiv 0$) w punkcie

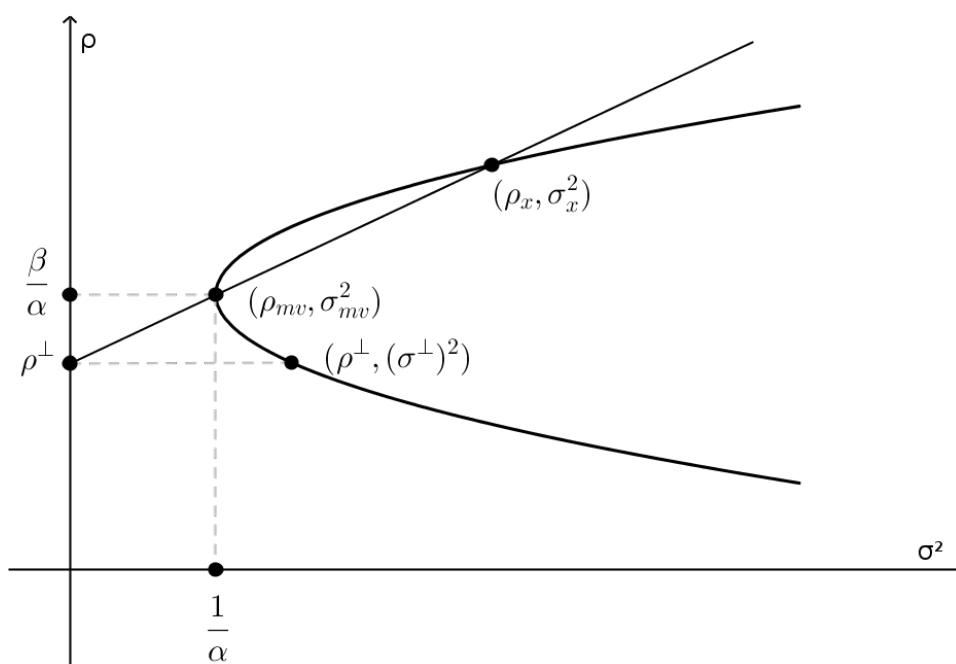
$$\rho = \frac{\rho(x)\sigma_{mv}^2 - \rho_{mv}\sigma^2(x)}{\sigma_{mv}^2 - \sigma^2(x)}.$$

Ponieważ z twierdzenia 2.12 wynika równość

$$\sigma^2(x) = \frac{\alpha\rho^2(x) - 2\beta\rho(x) + \gamma}{\delta}$$

i zachodzą równości $\sigma_{mv}^2 = \frac{1}{\alpha}$ oraz $\rho_{mv} = \frac{\beta}{\alpha}$, więc

$$\begin{aligned} \rho &= \frac{\rho(x)\frac{1}{\alpha} - \frac{\beta}{\alpha} \frac{\alpha\rho^2(x) - 2\beta\rho(x) + \gamma}{\delta}}{\frac{1}{\alpha} - \frac{\alpha\rho^2(x) - 2\beta\rho(x) + \gamma}{\delta}} = \frac{\rho(x)\delta - \beta(\alpha\rho^2(x) - 2\beta\rho(x) + \gamma)}{\delta - \alpha(\alpha\rho^2(x) - 2\beta\rho(x) + \gamma)} \\ &= \frac{-\alpha\beta\rho^2(x) + 2\beta^2\rho(x) - \gamma\beta + \alpha\gamma\rho(x) - \beta^2\rho(x)}{\alpha\gamma - \beta^2 - \alpha^2\rho^2(x) + 2\alpha\beta\rho(x) - \alpha\gamma} \\ &= \frac{-\alpha\beta\rho^2(x) + \beta^2\rho(x) + \alpha\gamma\rho(x) - \beta\gamma}{-\alpha^2\rho^2(x) + 2\alpha\beta\rho(x) - \beta^2} = \frac{-(\alpha\rho(x) - \beta)(\beta\rho(x) - \gamma)}{-(\alpha\rho(x) - \beta)^2} \\ &= \frac{\beta\rho(x) - \gamma}{\alpha\rho(x) - \beta}, \end{aligned}$$



Rysunek 3.1: Konstrukcja portfela ortogonalnego.

a to jest stopa zwrotu portfela ortogonalnego z lematu 3.4. ■

Na rynku z instrumentem wolnym od ryzyka mamy szczególną postać twierdzenia 3.5.

TWIERDZENIE. 3.7 *Rozważmy rynek z instrumentem wolnym od ryzyka o stopie zwrotu μ_0 . Załóżmy, że $\beta > \mu_0\alpha$ a x^* jest portfelem stycznym na tym rynku (x^* składa się wyłącznie z instrumentów ryzykownych). Wtedy dla i -tego instrumentu ryzykownego na rynku*

$$\frac{\mu_i - \mu_0}{\sigma_{i,x^*}^2} = \frac{\rho(x^*) - \mu_0}{\sigma^2(x^*)}.$$

Dowód. Z twierdzenia 2.17 mamy

$$\sigma^2(x^*) = \frac{\delta_0}{(\beta - \mu_0\alpha)^2} \quad \text{oraz} \quad \rho(x^*) = \mu_0 + \frac{\delta_0}{\beta - \mu_0\alpha}.$$

Z drugiej z tych równości dostajemy wzór

$$\delta_0 = (\rho(x^*) - \mu_0)(\beta - \mu_0\alpha)$$

Podstawiając to wyrażenie do wzoru na $\sigma^2(x^*)$ dostajemy

$$\sigma^2(x^*) = \frac{\rho(x^*) - \mu_0}{\beta - \mu_0\alpha}. \quad (3.3)$$

Z dowodu twierdzenia 2.17 otrzymujemy

$$x^* = \frac{1}{\beta - \mu_0\alpha} \Sigma^{-1}(\mu - \mu_0 e),$$

więc

$$\sigma_{i,x^*}^2 = e_i^\top \Sigma x^* = (\Sigma x^*)_i = \frac{\mu_i - \mu_0}{\beta - \mu_0\alpha}.$$

Łącząc tę równość z wzorem (3.3) dostajemy tezę twierdzenia

$$\frac{\mu_i - \mu_0}{\sigma_{i,x^*}^2} = \beta - \mu_0\alpha = \frac{\rho(x^*) - \mu_0}{\sigma^2(x^*)}.$$

■

3.2 Model CAPM

Rozważać teraz będziemy rynek z instrumentem wolnym od ryzyka. Zakładamy, że inwestorzy na tym rynku spełniają następujące warunki :

- Każdy inwestor zna rozkłady stóp zwrotu instrumentów na rynku, tj. zna μ_0 , μ i Σ .
- Każdy inwestor inwestuje zgodnie z funkcją preferencji z modelu Markowitza (patrz lemat 2.9) ale niekoniecznie o tej samej awersji do ryzyka.
- $\mu_0 < \frac{\beta}{\alpha}$, macierz Σ jest symetryczna i dodatnio określona, $\mu \not\parallel e$.

DEFINICJA. 3.8 *Portfelem rynkowym nazywamy portfel $x_M \in \mathbb{R}^d$, taki że*

$$x_{M,i} = \frac{W_i}{W},$$

gdzie

$$W = \sum_{i=1}^d W_i,$$

a W_i jest kapitalizacją rynkową i -tego instrumentu ryzykownego, zaś W jest kapitalizacją całego rynku instrumentów ryzykownych.

WNIOSEK. 3.9

$$x_M = x^*,$$

gdzie x^* to portfel styczny na rynku z instrumentem wolnym od ryzyka.

Dowód. Na mocy wniosku 2.19 (Two Fund Theorem) każdy inwestor dzieli swoje bogactwo pomiędzy instrument wolny od ryzyka i portfel styczny x^* . Jeśli przez U_k oznaczymy wartość inwestycji w instrumenty ryzykowne inwestora k , $k = 1, 2, \dots, K$, to sumaryczna wartość inwestycji wszystkich inwestorów w instrument i -ty wynosi

$$W_i = \sum_{k=1}^K U_k x_i^*.$$

Ponieważ $W = \sum_{k=1}^K U_k$, zatem

$$x_{M,i} = \frac{\sum_{k=1}^K U_k x_i^*}{W} = x_i^*.$$

■

TWIERDZENIE. 3.10 (CAPM) *Przy założeniach sformułowanych powyżej, dla i -tego instrumentu ryzykownego zachodzi wzór*

$$\mu_i - \mu_0 = (\rho_M - \mu_0) \frac{\sigma_{i,M}^2}{\sigma_M^2},$$

gdzie $\sigma_{i,M}^2$ jest kowariancją stopy zwrotu i -tego instrumentu ze stopą zwrotu portfela rynkowego, σ_M^2 jest wariancją stopy zwrotu portfela rynkowego, μ_i jest średnią stopą zwrotu instrumentu i -tego, a $\rho_M = \mu^\top x_M$ średnią stopą zwrotu portfela rynkowego.

Dowód. Ponieważ $x_M = x^*$, gdzie x^* jest portfelem stycznym (wniosek 3.9), to z twierdzenia 3.7 dla i -tego instrumentu ryzykownego zachodzi równość

$$\frac{\mu_i - \mu_0}{\sigma_{i,M}^2} = \frac{\rho_M - \mu_0}{\sigma^2(x_M)},$$

co po przekształceniu daje

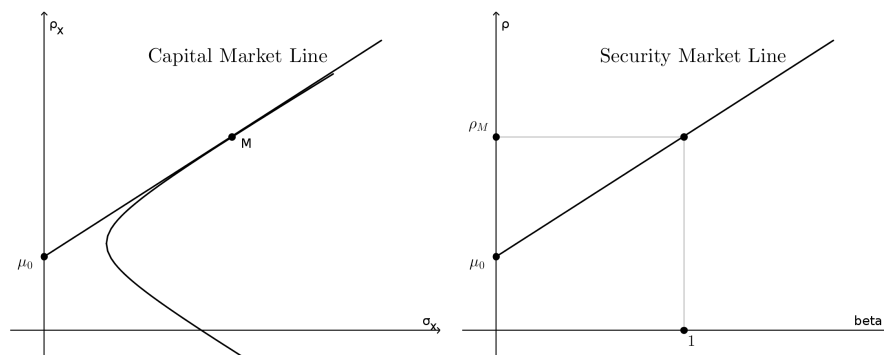
$$\mu_i - \mu_0 = (\rho_M - \mu_0) \frac{\sigma_{i,M}^2}{\sigma_M^2}.$$

■

Uwaga. 15 Wielkość $\beta_i = \frac{\sigma_{i,M}^2}{\sigma_M^2}$ nazywana jest betą i -tego instrumentu, zaś zależność

$$(\mu_i - \mu_0) = (\rho_M - \mu_0) \beta_i$$

nazywa się **security market line**.



Rysunek 3.2: Proste CML oraz SML.

TWIERDZENIE. 3.11 (Capital Market Line) Niech (x, x_0) będzie portfelem optymalnym na rynku z instrumentem wolnym od ryzyka. Wówczas

$$\rho(x, x_0) = \mu_0 + \sigma(x) \frac{\rho_M - \mu_0}{\sigma_M},$$

gdzie $\rho(x, x_0) = x_0\mu_0 + \mu^\top x$ jest średnią stopy zwrotu z portfela (x, x_0) , zaś $\sigma(x)$ jest odchyleniem standardowym tego portfela.

Dowód. Zapisując tezę twierdzenia 3.10 w postaci wektorowej otrzymujemy

$$\mu = \mu_0 e + (\rho_M - \mu_0) \frac{\Sigma x_M}{\sigma_M^2}.$$

Stąd, korzystając z równości $x_0 + e^\top x = 1$, mamy równość

$$\rho(x, x_0) = x_0\mu_0 + \mu^\top x = x_0\mu_0 + \mu_0 e^\top x + (\rho_M - \mu_0) \frac{x^\top \Sigma x_M}{\sigma_M^2}.$$

Wystarczy teraz udowodnić, że jeśli portfel (x, x_0) jest portfelem optymalnym, to $x^\top \Sigma x_M = \sigma(x)\sigma_M$.

Z twierdzenia 2.15 mamy wzór na udział instrumentów ryzykownych w dowolnym portfelu optymalnym

$$x = \frac{\rho - \mu_0}{\delta_0} \Sigma^{-1}(\mu - \mu_0 e).$$

Podobnie z twierdzenia 2.17 mamy wzór na portfel styczny, czyli portfel rynkowy,

$$x_M = \frac{1}{\beta - \alpha\mu_0} \Sigma^{-1}(\mu - \mu_0 e).$$

Łącząc te wzory możemy policzyć kowariancję stóp zwrotu portfela x i portfela rynkowego

$$\begin{aligned} x^\top \Sigma x_M &= \frac{\rho - \mu_0}{\delta_0(\beta - \alpha\mu_0)} (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e) \\ &= \frac{\rho - \mu_0}{\delta_0(\beta - \alpha\mu_0)} (\gamma - 2\beta\mu_0 + \mu_0^2\alpha) = \frac{\rho - \mu_0}{\beta - \alpha\mu_0}. \end{aligned}$$

Z drugiej strony, z twierdzeń 2.15 i 2.17 mamy

$$\sigma(x) = \frac{\rho - \mu_0}{\sqrt{\delta_0}}, \quad \sigma_M = \frac{\sqrt{\delta_0}}{\beta - \alpha\mu_0}.$$

Otrzymujemy stąd

$$\sigma(x)\sigma_M = \frac{\rho - \mu_0}{\beta - \alpha\mu_0} = x^\top \Sigma x_M.$$

■

Zauważmy na koniec, że także na rynku z instrumentem wolnym od ryzyka można otrzymać lemat analogiczny do lematu 3.4. W tym celu zdefiniujemy ortogonalność portfeli na rynku z instrumentem wolnym od ryzyka.

DEFINICJA. 3.12 Dwa portfele (x, x_0) , (\hat{x}, \hat{x}_0) na rynku z instrumentem wolnym od ryzyka nazywamy ortogonalnymi, jeśli ich kowariancja wynosi 0, tzn. $x^\top \Sigma \hat{x} = \text{Cov}(x, \hat{x}) = 0$.

LEMAT. 3.13 Niech (x, x_0) będzie portfelem optymalnym, takim że $x_0 \neq 1$. Jedynym portfelem optymalnym ortogonalnym do niego jest portfel $(0, 1)$.

Dowód. Na mocy wniosku 2.19 (Two Fund Theorem) każdy portfel optymalny jest postaci $((1 - x_0)x^\top, x_0)$, gdzie x_T jest portfelem stycznym.

Weźmy dwa portfele $((1 - x_0)x_T, x_0)$ i $((1 - \hat{x}_0)\hat{x}_T, \hat{x}_0)$. Ich kowariancja wynosi

$$(1 - x_0)(1 - \hat{x}_0)x_T^\top \Sigma x_T.$$

Ta wielkość jest zerowa wtedy i tylko wtedy, gdy $(1 - x_0)(1 - \hat{x}_0) = 0$. Ponieważ z założenia $x_0 \neq 1$, to musi być $\hat{x}_0 = 1$. ■

Rozdział 4

Portfele optymalne z ograniczeniem na krótką sprzedaż

4.1 Optymalizacja z ograniczeniem krótkiej sprzedaży

DEFINICJA. 4.1 Zadanie optymalizacyjne z ograniczeniem krótkiej sprzedaży

*Tak jak w rozdziale 2 zakładamy, że rynek spełnia założenia definicji 2.1. Problemem optymalizacyjnym Markowitza z **ograniczeniem krótkiej sprzedaży** nazywać będziemy optymalizację na rynku bez instrumentu wolnego od ryzyka, ale z ograniczeniem, że udziały instrumentów ryzykownych w portfelu muszą być nieujemne.*

Problem ten formułuje się następująco:

$$\begin{cases} \min_x \frac{1}{2}x^\top \Sigma x, \\ \mu^\top x = \rho, \\ e^\top x = 1, \\ x \geq 0. \end{cases} \quad (4.1)$$

Jego rozwiązaniem są portfele, które wyznaczają granicę portfelową.

Podobnie jak w przypadku problemu Markowitza bez ograniczenia na krótką sprzedaż, jeśli chcemy znaleźć granicę efektywną, to należy rozwiązać problem:

$$\begin{cases} \min_x \frac{1}{2}x^\top \Sigma x, \\ \mu^\top x \geq \rho, \\ e^\top x = 1, \\ x \geq 0. \end{cases} \quad (4.2)$$

W dalszym ciągu będziemy zajmowali się jedynie problemem (4.1), ponieważ znając granice portfelową można łatwo znaleźć granicę efektywną – podobnie jak dla problemu bez ograniczenia na krótką sprzedaż.

Do rozwiązania problemu optymalizacyjnego z ograniczeniem krótkiej sprzedaży potrzebne są pewne twierdzenia z optymalizacji nieliniowej zebrane w rozdziale 9. Potrzebujemy także dodatkowe założenia dotyczące rynku.

DEFINICJA. 4.2 (I założenie o braku degeneracji) *Niech R będzie d -wymiarową zmienną losową przyszłych stóp zwrotu z inwestycji w instrumenty ryzykowne. Zmienna R ma d -wymiarowy rozkład normalny $\mathcal{N}_d(\mu, \Sigma)$, gdzie μ – wektor średnich (wartości oczekiwanych) zmiennej R , a Σ – macierz kowariancji. O macierzy Σ zakładamy, jak poprzednio, że jest symetryczna i dodatnio określona. Na temat wektora μ robimy założenie silniejsze niż poprzednio. Zakładamy, że wszystkie składowe wektora μ są różne, tzn. $\mu_i \neq \mu_j, \forall i \neq j$.*

Rozpocznijmy od wprowadzenia pewnych oznaczeń, które będą pomocne w dalszym ciągu.

Przez B^d oznaczajmy granicę portfelową na płaszczyźnie (ρ, σ^2) dla problemu (2.1) z d instrumentami ryzykownymi. Krzywą B^d będziemy traktowali jako wykres funkcji zmiennej ρ .

Przez \bar{B}^d oznaczajmy granicę portfelową na płaszczyźnie (ρ, σ^2) dla problemu (4.1) z d instrumentami ryzykownymi. Krzywą \bar{B}^d będziemy traktowali jako wykres funkcji zmiennej ρ .

TWIERDZENIE. 4.3 *Granica portfelowa \bar{B}^d dla problemu (4.1) na płaszczyźnie (ρ, σ^2) jest ograniczonym łukiem, którego punktami krańcowymi są: $P_{min} = (\rho_{min}, \sigma_{min}^2)$ oraz $P_{max} = (\rho_{max}, \sigma_{max}^2)$, gdzie $\rho_{min} = \min_{1 \leq i \leq n} \mu_i$ i $\rho_{max} = \max_{1 \leq i \leq n} \mu_i$ są najmniejszą oraz największą oczekiwaną stopą zwrotu z jednego z instrumentów ryzykownych, a $\sigma_{min}^2, \sigma_{max}^2$ są odpowiednimi wariancjami.*

Dowód. Ponieważ $x \geq 0$, więc największą oczekiwaną stopę zwrotu z portfela można osiągnąć, inwestując cały kapitał w instrument o największej średniej stopy zwrotu. Analogicznie najmniejszą oczekiwaną stopę zwrotu z portfela można osiągnąć, inwestując cały kapitał w instrument o najmniejszej średniej stopy zwrotu.

Zauważmy, że warunki $e^\top x = 1, x \geq 0$, wyznaczają zwarty sympleks na hiperpłaszczyźnie $e^\top x = 1$. Ponieważ $x^\top \Sigma x$ jest funkcją ciągłą, to osiąga ona minimum na tym zwartym sympleksie. Warunek $\mu^\top x = \rho$ wyznacza odcinek w tym sympleksie, jeśli $\rho \in [\rho_{min}, \rho_{max}]$. Fakt, że \bar{B}^d jest ciągłą krzywą (a nawet krzywą kawałkami klasy C^1) zostanie pokazany później (wynika on z ogólnego twierdzenia o zależności rozwiązań optymalnych od ograniczeń). ■

Opiszemy algorytm rozwiązywania problemu optymalizacyjnego (4.1).

TWIERDZENIE. 4.4 Przy założeniach definicji 4.1 oraz 4.2 rozwiązanie algorytmiczne problemu optymalizacyjnego (4.1) jest jednoznaczne i otrzymuje się je w co najwyżej $2^d - d - 1$ krokach algorytmicznych.

Dowód. Dowód twierdzenia polega na opisie algorytmu oraz udowodnieniu, że każdy krok algorytmu jest jednoznaczny, a także, że algorytm zatrzyma się po wskazanej liczbie kroków.

Ponieważ $x^\top \Sigma x$ jest funkcją wypukłą, a funkcje $\mu^\top x - \rho$ i $e^\top x - 1$ są liniowe, więc na podstawie twierdzenia 9.17 znalezienie rozwiązania sprowadza się do znalezienia punktu spełniającego warunki I rzędu. Z wniosku 9.18 te warunki mają postać:

$$\begin{cases} \Sigma x - \lambda_1 e - \lambda_2 \mu \geq 0, \\ x_i (\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0, \quad i = 1, \dots, d, \\ e^\top x = 1, \\ \rho - \mu^\top x = 0. \end{cases}$$

Rozwiązywanie odbywa się metodą przeszukiwania wszystkich możliwości krok po kroku.

1. Zakładamy, że $\forall_i x_i > 0$. Wtedy rozwiązanie optymalne dane jest wzorem $x = \Sigma^{-1}(\lambda_1 e + \lambda_2 \mu)$. Sprawdzamy, czy $\forall_i x_i > 0$. Jeśli TAK – koniec. Jeśli NIE – przechodzimy do następnego kroku.
2. $\exists_j x_j = 0$. Zakładamy, że $\forall_{i \neq j} x_i > 0$. Wtedy dostajemy układ

$$(\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0 \quad \text{dla } i \neq j,$$

oraz nierówność $(\Sigma x - \lambda_1 e - \lambda_2 \mu)_j \geq 0$.

Układ $(\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0$ można rozwiązać, bo macierz Σ bez j -tego wiersza i j -tej kolumny jest dalej dodatnio określona. Znajdujemy $x_i, i \neq j$.

Sprawdzamy, czy $x_i > 0$. Jeśli TAK – koniec. Jeśli NIE – przechodzimy do następnego kroku.

3. $\exists_{j,k} x_j = 0 = x_k$. Zakładamy, że $\forall_{i \neq j, i \neq k} x_i > 0$. Wtedy dostajemy układ

$$(\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0 \quad \text{dla } i \neq j, i \neq k,$$

oraz nierówności

$$(\Sigma x - \lambda_1 e - \lambda_2 \mu)_j \geq 0, \quad (\Sigma x - \lambda_1 e - \lambda_2 \mu)_k \geq 0.$$

Analogicznie jak w poprzednim kroku, układ $(\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0$ można rozwiązać. Znajdujemy $x_i, i \neq j, i \neq k$. Sprawdzamy, czy $x_i > 0$. Jeśli TAK – koniec. Jeśli NIE – przechodzimy do następnego kroku. itd.

Ponieważ wiemy, że rozwiązanie istnieje, to po $2^d - d - 1$ krokach wyczerpiemy wszystkie możliwości, bo 2^d jest liczbą wszystkich podzbiorów indeksów i , dla których $x_i > 0$. Ale zbiór pusty wykluczamy. Zbiory 1-elementowe możemy też wykluczyć, bo odpowiadają one warunkom $\exists_j \forall_{i \neq j} x_i = 0$. Ale wtedy musi być $x_j = 1$, bo $e^\top x = 1$.

Założenie braku degeneracji sformułowane w definicji 4.2 jest potrzebne, aby wyznaczyć jednoznacznie współczynniki Lagrange'a λ_1 i λ_2 (a tym samym wyznaczyć jednoznacznie rozwiązanie x), tak aby spełnione były jednocześnie warunki $e^\top x = 1$ oraz $\lambda_2(\rho - \mu^\top x) = 0$. ■

TWIERDZENIE. 4.5 *Na płaszczyźnie (ρ, σ^2) krzywa \overline{B}^d jest, poza skończoną liczbą punktów, krzywą klasy C^1 złożoną z kawałków parabol, z których każda jest granicą portfelową dla portfeli inwestycyjnych w pewnym podzbiorze zbioru d instrumentów ryzykownych.*

Dowód. Dowód polega na konstrukcji krzywej \overline{B}^d .

Niech $x(\rho)$ oznacza rozwiązanie problemu (4.1) odpowiadające oczekiwanej stopie zwrotu ρ . Załóżmy, że na paraboli B^d istnieje przedział $\rho_m \leq \rho \leq \rho_M$, taki że $x(\rho) > 0$ dla $\rho \in (\rho_m, \rho_M)$, a co najmniej jedna składowa wektorów $x(\rho_m)$ i $x(\rho_M)$ jest zerowa. Niech $x_m(\rho_m) = 0$ i $x_M(\rho_M) = 0$, czyli zakładamy, że znika dokładnie jedna składowa $x(\rho_m)$ i $x(\rho_M)$.

Załóżmy, że $\rho_m \neq \rho_{min}$ i $\rho_M \neq \rho_{max}$. Weźmy punkt $(\rho_m, \sigma_m^2) \in B^d$ ($\sigma_m^2 = x^\top(\rho_m)\Sigma x(\rho_m)$). Ponieważ $x_m(\rho_m) = 0$, więc punkt (ρ_m, σ_m^2) leży na paraboli będącej granicą portfelową dla $d-1$ instrumentów, którą będziemy oznaczali B_m^{d-1} (indeks m oznacza, że jest to granica portfelowa $d-1$ instrumentów z wykluczonym instrumentem o numerze m). Analogicznie $(\rho_M, \sigma_M^2) \in B^d$ leży na paraboli będącej granicą portfelową B_M^{d-1} .

Zauważmy, że B_m^{d-1} (i analogicznie B_M^{d-1}) leży wewnątrz zbioru portfeli dopuszczalnych dla d instrumentów. Stąd wynika, że punkt $(\rho_m, \sigma_m^2) \in B^d \cap B_m^{d-1}$ musi być punktem styczności tych parabol (B_m^{d-1} leży po jednej stronie B^d). Z faktu, że tylko składowa x_m znika w punkcie (ρ_m, σ_m^2) oraz ciągłości $x(\rho)$ na paraboli B_m^{d-1} wynika, że istnieje przedział $\rho_{m_1} \leq \rho \leq \rho_m$, taki że dla $\rho \in (\rho_{m_1}, \rho_m)$ $(d-1)$ -wymiarowy wektor $x^{(d-1)}(\rho) > 0$ oraz istnieje co najmniej jedna składowa wektora $x^{(d-1)}$, która znika w punkcie ρ_{m_1} : $x_{m_1}^{(d-1)}(\rho_{m_1}) = 0$ (zakładamy ponownie, że jest tylko jedna taka składowa). Jeśli $\rho_{m_1} \neq \rho_{min}$, to kontynuujemy tę procedurę przez skończoną liczbą kroków, aż $\rho_{m_i} = \rho_{min}$. Analogicznie postępujemy dla ρ_{M_i} .

Należy jeszcze rozpatrzyć pominięte przypadki:

1. B^d nie zawiera żadnego łuku \overline{B}^d . Wtedy sprawdzamy dla wszystkich B_i^{d-1} , $i = 1, \dots, d$, czy one nie zawierają łuku \overline{B}^d . Jeśli nie, to sprawdzamy wszystkie B^{d-2} itd. Ponieważ wiadomo, że \overline{B}^d jest niepusty, istnieje taki

B^{d-i} , który zawiera fragment \overline{B}^d . Następnie postępowanie jest analogiczne, jak poprzednio.

2. W punkcie $(\rho_{m_i}, \sigma_{m_i}^2)$ więcej niż jedna składowa wektora $x(\rho_{m_i})$ zeruje się. Wtedy rozpatrujemy wszystkie granice portfelowe odpowiadające pominięciu dokładnie jednej składowej zerującej się, a jako łuk należący do \overline{B}^d wybieramy ten łuk, który dla tej samej wartości ρ ma najmniejszą wartość σ^2 .

■

WNIOSEK. 4.6 *Jeśli P_{min} i P_{max} należą do granicy portfelowej B^d , to \overline{B}^d jest fragmentem B^d zawartym między tymi punktami.*

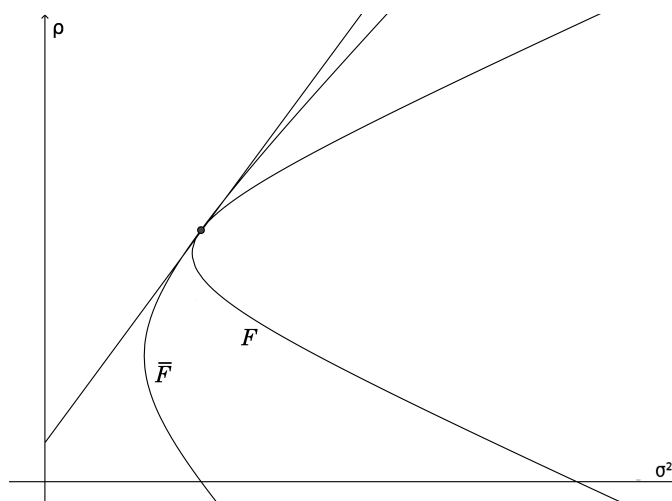
Dowód twierdzenia 4.5 jest niekompletny. Z dowodu wynika, że granica portfelowa \overline{B}^d jest kawałkami krzywą klasy C^1 , ponieważ składa się z łuków parabol połączonych w sposób styczny. Dowód nie wyjaśnia jednak skąd bierze się skończona liczba punktów, w których krzywa \overline{B}^d nie jest klasy C^1 . Wynika to z poniższego lematu.

LEMAT. 4.7 *Jeśli portfel optymalny x^* , taki że $(\rho(x^*), \sigma^2(x^*)) \in \overline{B}^d$ zawiera co najmniej 2 instrumenty ryzykowne o różnych średnich stopach zwrotu, to w otoczeniu $(\rho(x^*), \sigma^2(x^*))$ krzywa \overline{B}^d jest klasy C^1 .*

Dowód. Niech \overline{F} oznacza fragment łuku \overline{B}^d w otoczeniu $(\rho(x^*), \sigma^2(x^*))$. Weźmy granicę portfelową F bez ograniczenia na krótką sprzedaż, otrzymaną przez rozwiązanie problemu optymalizacyjnego z tymi instrumentami, które wchodzi do x^* z niezerowymi wagami. Oczywiście $(\rho(x^*), \sigma^2(x^*))$ leży także na F .

Ponieważ F jest obrazem kombinacji liniowej portfela x^* oraz pewnego innego portfela złożonego z instrumentów, które wchodzi do x^* i którego obraz leży na F (do tego potrzeba, aby x^* zawierał co najmniej dwa instrumenty o różnych średnich stopach zwrotu – patrz wniosek 2.5), to w otoczeniu $(\rho(x^*), \sigma^2(x^*))$ na F znajdują się też obrazy portfeli x o wyłącznie dodatnich wagach ($x > 0$), czyli F w otoczeniu $(\rho(x^*), \sigma^2(x^*))$ zawiera też obrazy portfeli dopuszczalnych. Z optymalności \overline{F} wynika, że w otoczeniu $(\rho(x^*), \sigma^2(x^*))$, jeśli mamy obraz portfela x , taki że $(\rho(x), \sigma^2(x)) \in F$ oraz obraz portfela \bar{x} , taki że $(\rho(\bar{x}), \sigma^2(\bar{x})) \in \overline{F}$, a jednocześnie $\rho(x) = \rho(\bar{x})$, to $\sigma^2(\bar{x}) \leq \sigma^2(x)$ (Rys. 4.1). Pokażemy niżej, że \overline{F} jest wypukłą jako funkcja ρ . Z wypukłości wynika, że \overline{F} posiada prostą podpierającą w punkcie $(\rho(x^*), \sigma^2(x^*))$. Ponieważ F jest klasy C^1 w punkcie $(\rho(x^*), \sigma^2(x^*))$, to \overline{F} zawarta między prostą podpierającą a funkcją klasy C^1 musi też być klasy C^1 .

Pozostało nam do udowodnienia, że w otoczeniu punktu $(\rho(x^*), \sigma^2(x^*))$ funkcja $\overline{F}(\rho)$ jest wypukłą. Zauważmy, że wykres tej funkcji jest złożony z punktów

Rysunek 4.1: Położenie granic portfelowych \bar{F} i F z dowodu lematu 4.7.

$(\rho, \sigma^2(\rho))$, gdzie $\sigma^2(\rho)$ jest wartością funkcji $x^\top \Sigma x$ w punkcie będącym rozwiązaniem zadania optymalizacyjnego:

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x = \rho, \\ e^\top x = 1, \\ x > 0. \end{cases} \quad (4.3)$$

Problem sprowadza się do wykazania wypukłości funkcji $\sigma^2(\rho)$. Wprowadzając przestrzeń $\mathbb{X} = \{x : e^\top x = 1, x > 0\}$ możemy zadanie optymalizacyjne (4.3) przeformułować następująco:

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x = \rho, \\ x \in \mathbb{X}. \end{cases} \quad (4.4)$$

Niech $M(\rho)$ będzie rozwiązaniem zadania optymalizacyjnego (4.4), t.j.

$$M(\rho) = \min_{x \in \mathbb{X}, \mu^\top x = \rho} \frac{1}{2} x^\top \Sigma x.$$

Korzystając z faktu, że zbiór \mathbb{X} jest wypukły i funkcja $\frac{1}{2} x^\top \Sigma x$ jest wypukła mamy

dla $\rho_1, \rho_2 \in (\rho_{min}, \rho_{max})$ oraz $\lambda \in [0, 1]$

$$\begin{aligned}
& \lambda M(\rho_1) + (1 - \lambda)M(\rho_2) \\
&= \min_{x_1 \in \mathbb{X}, \mu^\top x_1 = \rho_1} \frac{1}{2}(\lambda x_1^\top \Sigma x_1) + \min_{x_2 \in \mathbb{X}, \mu^\top x_2 = \rho_2} \frac{1}{2}((1 - \lambda)x_2^\top \Sigma x_2) \\
&= \min_{x_1, x_2 \in \mathbb{X}, \mu^\top x_1 = \rho_1, \mu^\top x_2 = \rho_2} \frac{1}{2}(\lambda x_1^\top \Sigma x_1 + (1 - \lambda)x_2^\top \Sigma x_2) \\
&\geq \min_{x_1, x_2 \in \mathbb{X}, \mu^\top x_1 = \rho_1, \mu^\top x_2 = \rho_2} \frac{1}{2}((\lambda x_1 + (1 - \lambda)x_2)^\top \Sigma (\lambda x_1 + (1 - \lambda)x_2)) \\
&= \min_{x = (\lambda x_1 + (1 - \lambda)x_2) \in \mathbb{X}, \mu^\top x = (\lambda \rho_1 + (1 - \lambda)\rho_2) = \rho} \frac{1}{2}x^\top \Sigma x = M(\rho).
\end{aligned}$$

Ponieważ $\sigma^2(\rho) = 2M(\rho)$, więc jest to też funkcja wypukła. ■

WNIOSEK. 4.8 \bar{B}^d w punkcie $(\rho(x^*), \sigma^2(x^*))$ nie musi być klasy C^1 , tylko jeśli portfel x^* złożony jest z jednego instrumentu.

Przykład. 4.9 Rozważmy rynek złożony z trzech instrumentów, których rozkład stóp zwrotu ma parametry:

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \mu = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Znajdziemy granicę portfelową dla problemu bez ograniczenia na krótką sprzedaż. Rozpoczynamy od znalezienia macierzy Σ^{-1} oraz wyliczenia stałych.

$$\Sigma^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Odpowiednie stałe mają wartości: $\alpha = \frac{11}{6}, \beta = 3, \gamma = 6, \delta = 2$.

Z twierdzeń 2.3 i 2.12 obliczamy funkcję $\sigma^2(\rho)$ oraz portfele optymalne $x(\rho)$:

$$\sigma^2 = \frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}, \quad x = \frac{\gamma - \beta\rho}{\delta}\Sigma^{-1}e + \frac{\alpha\rho - \beta}{\delta}\Sigma^{-1}\mu.$$

Otrzymujemy wtedy

$$\sigma^2 = \frac{11}{12}\rho^2 - 3\rho + 3, \quad x_1 = -\frac{7}{12}\rho + \frac{3}{2}, x_2 = \frac{1}{6}\rho, x_3 = \frac{5}{12}\rho - \frac{1}{2}.$$

Rozważmy teraz portfele, w których $x_i = 0$.

Dla portfela, w którym $x_1 = 0$, parametry rozkładu dane są wzorami

$$\Sigma^{(1)} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad \mu^{(1)} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Odpowiednie stałe mają wartości: $\alpha = \frac{5}{6}$, $\beta = 2$, $\gamma = 5$, $\delta = \frac{1}{6}$.

Obliczamy funkcję $\sigma^2(\rho)$ oraz portfele optymalne $x(\rho)$:

$$\sigma^2 = 5\rho^2 - 24\rho + 30, \quad x_2 = -\rho + 3, x_3 = \rho - 2.$$

Dla portfela, w którym $x_2 = 0$, parametry rozkładu dane są wzorami

$$\Sigma^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad \mu^{(2)} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Odpowiednie stałe mają wartości: $\alpha = \frac{4}{3}$, $\beta = 2$, $\gamma = 4$, $\delta = \frac{4}{3}$.

Obliczamy funkcję $\sigma^2(\rho)$ oraz portfele optymalne $x(\rho)$:

$$\sigma^2 = \rho^2 - 3\rho + 3 \quad x_1 = -\frac{1}{2}\rho + \frac{3}{2}, x_3 = \frac{1}{2}\rho - \frac{1}{2}.$$

Dla portfela, w którym $x_3 = 0$, parametry rozkładu dane są wzorami

$$\Sigma^{(3)} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \mu^{(3)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Odpowiednie stałe mają wartości: $\alpha = \frac{3}{2}$, $\beta = 2$, $\gamma = 3$, $\delta = \frac{1}{2}$.

Obliczamy funkcję $\sigma^2(\rho)$ oraz portfele optymalne $x(\rho)$:

$$\sigma^2 = 3\rho^2 - 8\rho + 6, \quad x_1 = -\rho + 2, x_2 = \rho - 1.$$

Kiedy portfel trzech instrumentów ma dodatnie składowe?

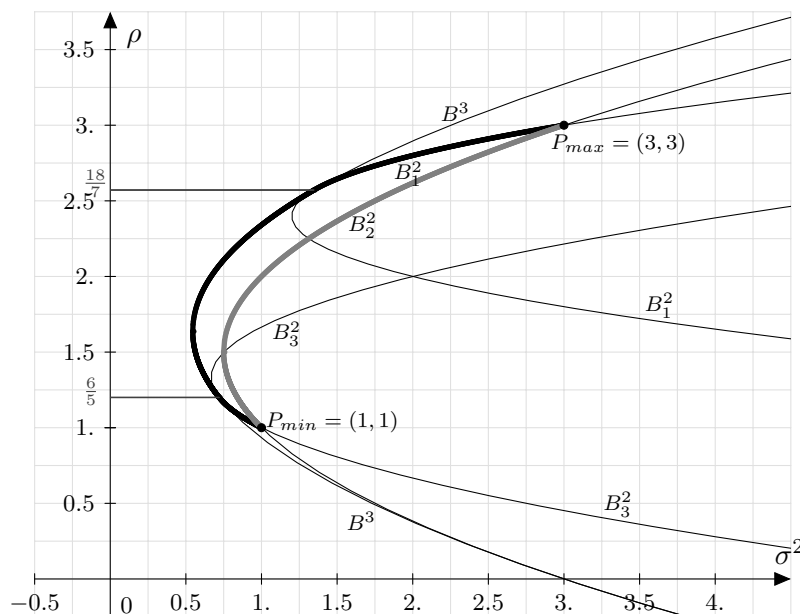
$$x_1 \geq 0 \Leftrightarrow \rho \leq \frac{18}{7}, \quad x_2 \geq 0 \Leftrightarrow \rho \geq 0, \quad x_3 \geq 0 \Leftrightarrow \rho \geq \frac{6}{5},$$

czyli dla $\frac{6}{5} \leq \rho \leq \frac{18}{7}$ mamy $x \geq 0$. Dla $\rho = \frac{6}{5}$ mamy $x_3(\frac{6}{5}) = 0$. Zmniejszając ρ przechodzimy na parabolę B_3^2 . Na niej: $x_1 \geq 0 \Leftrightarrow \rho \leq 2$, $x_2 \geq 0 \Leftrightarrow \rho \geq 1$. Czyli w przedziale $\frac{6}{5} \geq \rho \geq 1$ jesteśmy na B_3^2 . Wartość $\rho = 1$ odpowiada P_{min} , czyli inwestycji $x_1 = 1$. Teraz zwiększamy ρ od wartości $\rho = \frac{18}{7}$. Dla $\rho = \frac{18}{7}$, $x_1 = 0$, czyli przechodzimy na parabolę B_1^2 . Na niej $x_2 \geq 0 \Leftrightarrow \rho \leq 3$, $x_3 \geq 0 \Leftrightarrow \rho \geq 2$, czyli dla przedziału $\frac{18}{7} \leq \rho \leq 3$ jesteśmy na B_1^2 . Ponieważ $\rho = 3$ odpowiada P_{max} , czyli inwestycji $x_3 = 1$, to kończy to konstrukcję \bar{B}^3 .

4.2 Algorytm prostej krytycznej

Twierdzenie 4.5 daje obraz granicy portfelowej w płaszczyźnie (ρ, σ^2) , ale nie mówi, jakim portfelom odpowiadają punkty tej krzywej. Aby znaleźć odpowiednie portfele należy sięgnąć do metody opisanej w twierdzeniu 4.4. Algorytm znajdowania wszystkich portfeli odpowiadających granicy portfelowej pochodzi od Markowitza. Nazywa się go **Critical Line Algorithm** (algorytm prostej krytycznej).

Aby sprawnie opisywać realizacje algorytmu musimy wprowadzić oznaczenia ułatwiające opis działania algorytmu.



Rysunek 4.2: Wyznaczanie granicy portfelowej z Przykładu 4.9.

Uwaga. 16 Jak wiemy warunki $e^\top x = 1$, $x \geq 0$, wyznaczają w przestrzeni \mathbb{R}^d zwarty sympleks. Sympleks ten będziemy oznaczali Δ^d . k -wymiarowe ściany tego sympleksy będziemy oznaczali Δ^k , $1 \leq k \leq d$. Oczywiście Δ^k jest też sympleksem. Sympleks Δ^k będziemy utożsamiali z wierzchołkami, które go generują, a te wierzchołki z numerami niezerowych współrzędnych w tych wierzchołkach. Gdy będzie nam zależało na wskazaniu, przez jakie wierzchołki wyznaczony jest sympleks Δ^k , to będziemy go zapisywali jako odpowiedni zbiór indeksów (numerów niezerowych współrzędnych) $\{i_1, \dots, i_k\}$. Zwykle będziemy pisali Δ^k mając na uwadze jeden konkretny sympleks wymiaru k , jeśli numery wierzchołków nie będą istotne dla prowadzonych rozważań.

Przeanalizujemy realizację algorytmu z twierdzenia 4.4. Warunki I rzędu dają:

$$\begin{aligned} x_i > 0 & \quad (\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0, \quad i \in \Delta^k, \\ x_j = 0 & \quad (\Sigma x - \lambda_1 e - \lambda_2 \mu)_j \geq 0, \quad j \in \Delta^d \setminus \Delta^k. \end{aligned}$$

Markowitz zaproponował następujące podejście do rozwiązywania powyższego układu równań i nierówności:

1. Analizować portfele optymalne jako funkcję mnożnika Lagrange'a λ_2 . Ponieważ jest to mnożnik związany ze oczekiwanymi stopami zwrotu dokonujemy zmiany oznaczenia, aby podkreślić tę zależność $\lambda_2 \equiv \lambda_\rho$.
2. Zastąpić warunki I rzędu (układ równań i nierówności) układem równań o

wymiarze n :

$$\begin{cases} \Sigma_{\Delta} x - \lambda_1 e_{\Delta} = \lambda_{\rho} \mu_{\Delta}, \\ e_{\Delta}^{\top} x = 1, \end{cases} \quad (4.5)$$

gdzie:

$$(e_{\Delta})_k = \begin{cases} 1, & k = i, \\ 0, & k = j, \end{cases} \quad (\mu_{\Delta})_k = \begin{cases} \mu_i, & k = i, \\ 0, & k = j, \end{cases}$$

$$(\Sigma_{\Delta})_{k,l} = \begin{cases} \Sigma_{k,l}, & k \in \Delta^k, l \in \Delta^k \\ 1, & k = l \in \Delta^d \setminus \Delta^k \\ 0, & k \neq l, k, l \in \Delta^d \setminus \Delta^k. \end{cases}$$

Zauważmy, że układ równań (4.5) jest innym zapisem równań

$$\begin{cases} (\Sigma x - \lambda_1 e - \lambda_2 \mu)_i = 0, & i \in \Delta^k, \\ x_j = 0, & j \in \Delta^d \setminus \Delta^k, \end{cases}$$

które są formą zapisu warunków I rzędu.

Układ równań (4.5) zapisujemy w postaci macierzowej

$$\begin{pmatrix} \Sigma_{\Delta} & -e_{\Delta} \\ e_{\Delta}^{\top} & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} \lambda_{\rho} \mu_{\Delta} \\ 1 \end{pmatrix}. \quad (4.6)$$

Równanie (4.6) posiada jednoznaczne rozwiązanie, jeśli macierz Σ jest symetryczna i dodatnio określona oraz μ ma wszystkie składowe różne (spełnione są założenia definicji 4.2).

Wynika to z faktu, że macierz po lewej stronie równania jest odwracalna, czyli jest nieosobliwa. Ten ostatni fakt wynika łatwo z lematu 8.11. Zauważmy bowiem, że macierz Σ_{Δ} jest dodatnio określona, więc odwracalna. Gdyby cała macierz

$$A = \begin{pmatrix} \Sigma_{\Delta} & -e_{\Delta} \\ e_{\Delta} & 0 \end{pmatrix}$$

była osobliwa, to z lematu 8.11 wynika, że wyraz A_{dd} tej macierz (prawy dolny róg) miałby przedstawienie

$$A_{dd} = -e_{\Delta}^{\top} \Sigma_{\Delta}^{-1} e_{\Delta}.$$

Jeśli $e_{\Delta} \neq 0$, to $A_{dd} \neq 0$ (macierz Σ_{Δ} jako dodatnio określona nie ma zerowych wartości własnych), co przeczy faktowi, że w prawym dolnym rogu macierzy mamy właśnie 0.

Rozwiązanie równania (4.6) można zapisać wzorem (dokładna postać współczynników α_{Δ} i β_{Δ} nie jest istotna dla dalszych rozważań)

$$\begin{pmatrix} x \\ \lambda_1 \end{pmatrix} = \alpha_{\Delta} + \beta_{\Delta} \lambda_{\rho}.$$

Wykorzystując to rozwiązanie, wektor $\Sigma x - \lambda_1 e - \lambda_\rho \mu$ można zapisać w postaci

$$\begin{aligned}\Sigma x - \lambda_1 e - \lambda_\rho \mu &= (\Sigma, -e) \begin{pmatrix} x \\ \lambda_1 \end{pmatrix} - \lambda_\rho \mu = (\Sigma, -e)(\alpha_\Delta + \beta_\Delta \lambda_\rho) - \mu \lambda_\rho \\ &= (\Sigma, -e)\alpha_\Delta + ((\Sigma, -e)\beta_\Delta - \mu)\lambda_\rho \equiv \gamma_\Delta + \delta_\Delta \lambda_\rho.\end{aligned}$$

Teraz warunki I rzędu można zapisać we współrzędnych:

$$x_i = \alpha_{\Delta,i} + \beta_{\Delta,i} \lambda_\rho > 0, \quad i \in \Delta^k, \quad (4.7)$$

$$\gamma_{\Delta,j} + \delta_{\Delta,j} \lambda_\rho \geq 0, \quad j \in \Delta^d \setminus \Delta^k. \quad (4.8)$$

Założmy, że dla sympleksu Δ^k i ustalonej wartości λ_ρ powyższe nierówności są spełnione. Zmniejszamy wartość λ_ρ . Tylko jeśli $\beta_{\Delta,i} > 0$ i $\delta_{\Delta,j} > 0$, zmniejszając λ_ρ możemy dojść do wartości, dla której przynajmniej jedna nierówność przestanie być spełniana. Stąd wynikają ograniczenia:

$$\lambda_\rho > -\frac{\alpha_{\Delta,i}}{\beta_{\Delta,i}}, \quad i \in \Delta^k, \beta_{\Delta,i} > 0$$

$$\lambda_\rho \geq -\frac{\gamma_{\Delta,j}}{\delta_{\Delta,j}}, \quad j \in \Delta^d \setminus \Delta^k, \delta_{\Delta,j} > 0.$$

Kresem dolnym zmniejszania λ_ρ w sympleksie Δ^k jest

$$\lambda_\rho^{\Delta^k} = \max\{i \in \Delta^k, j \in \Delta^d \setminus \Delta^k, \beta_{\Delta,i} > 0, \delta_{\Delta,j} > 0: -\frac{\alpha_{\Delta,i}}{\beta_{\Delta,i}}, -\frac{\gamma_{\Delta,j}}{\delta_{\Delta,j}}\}. \quad (4.9)$$

Aby algorytm mógł jednoznacznie przechodzić z jednego sympleksu na drugi musimy zrobić dodatkowe założenie o braku degeneracji.

DEFINICJA. 4.10 (II założenie o braku degeneracji) Dla każdego niepustego sympleksu Δ^k , w zbiorze ułameków

$$\left\{i \in \Delta^k, j \in \Delta^d \setminus \Delta^k: -\frac{\alpha_{\Delta,i}}{\beta_{\Delta,i}}, -\frac{\gamma_{\Delta,j}}{\delta_{\Delta,j}}\right\}$$

istnieje dokładnie jeden ułamek o największej wartości.

Opis algorytmu.

Startujemy od wartości ρ_{max} – to odpowiada sympleksowi jednoelementowemu $\Delta^1 = \{i\}$. W tym sympleksie $x_i = 1$, czyli warunek 4.7 jest spełniony dla każdego λ_ρ . Zaczynając od $\lambda_\rho = +\infty$ i stopniowo je zmniejszając znajdujemy $\lambda_\rho^{\Delta^1}$ z warunku 4.9. Ponieważ tylko spełnienie warunku 4.8 odgrywa rolę, to $\lambda_\rho^{\Delta^1} = -\frac{\gamma_{\Delta,j}}{\delta_{\Delta,j}}$ dla pewnego j . To oznacza, że portfele optymalne z sympleksu $\Delta^1 = \{i\}$ przechodzą na sympleks $\Delta^2 = \{i, j\} = \Delta^1 \cup \{j\}$. Teraz zmniejszamy wartość λ_ρ na sympleksie Δ^2 (na tym sympleksie λ_ρ jest liniową funkcją ρ , patrz

nizej). Także portfel optymalny na sympleksie Δ^2 jest liniową funkcją ρ (patrz wniosek 2.5), czyli portfele optymalne tworzą na Δ^2 odcinek.

Założmy, że wykonaliśmy już pewną liczbę kroków algorytmu, otrzymując złożoną z odcinków łamaną portfeli optymalnych. Jesteśmy na sympleksie Δ^k i zmniejszamy wartość λ_ρ aż dojdziemy do wartości progowej $\lambda_\rho^{\Delta^k}$. Na mocy założenia o braku degeneracji jest tylko jeden indeks $i \in \Delta^k$ lub $j \in \Delta^d \setminus \Delta^k$, taki że $\lambda_\rho^{\Delta^k} = -\frac{\alpha_{\Delta,i}}{\beta_{\Delta,i}}$ lub $\lambda_\rho^{\Delta^k} = -\frac{\gamma_{\Delta,j}}{\delta_{\Delta,j}}$.

Jeśli $\lambda_\rho^{\Delta^k} = -\frac{\alpha_{\Delta,i}}{\beta_{\Delta,i}}$, to przechodzimy na sympleks $\Delta^{k-1} = \Delta^k \setminus \{i\}$.

Jeśli $\lambda_\rho^{\Delta^k} = -\frac{\gamma_{\Delta,j}}{\delta_{\Delta,j}}$, to przechodzimy na sympleks $\Delta^{k+1} = \Delta^k \cup \{j\}$.

Tak postępujemy, aż dojdziemy do wierzchołka $\{i'\}$ odpowiadającego ρ_{min} .

Co dzieje się z funkcją $\lambda_\rho = \lambda_\rho(\rho)$ przy przechodzeniu z jednego sympleksu na inny?

Z twierdzenia 2.3 wiadomo, że λ_ρ jest liniową funkcją ρ , czyli jeśli portfele optymalne zmieniają się w obrębie ustalonego sympleksu Δ^k , to λ_ρ jest funkcją liniową, czyli ciągłą.

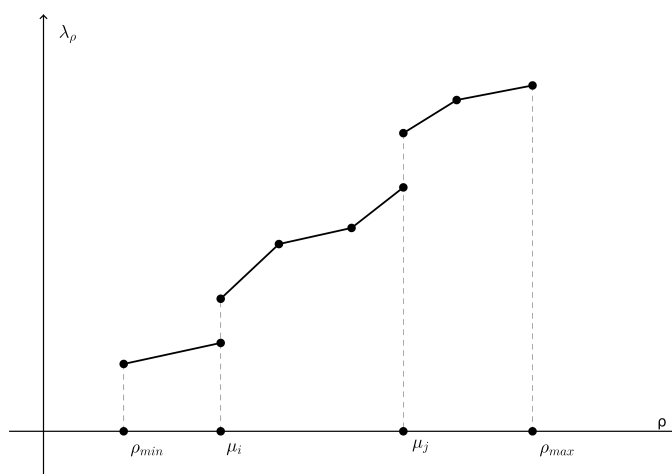
Przy przejściu $\Delta^k \rightarrow \Delta^{k'}$, $k, k' \geq 2$, funkcja liniowa zmienia parametry, ale zachowuje ciągłość. Wynika to z lematu 4.7, który zapewnia, że jeśli $x^* \in \Delta^k$, gdzie $k \geq 2$, to granica portfelowa jest funkcją klasy C^1 . Ta własność pozostaje prawdziwa, gdy przechodzimy z sympleksu Δ^k na $\Delta^{k'}$, jeśli tylko $k, k' \geq 2$. Z drugiej strony z twierdzenia 9.22 wiadomo, że granica portfelowa traktowana jako funkcja ρ jest funkcją wrażliwości dla problemu (4.1). Z twierdzenia tego wynika także, że mnożnik Lagrange'a λ_ρ jest wektorem wrażliwości dla tej funkcji. Ponieważ $\sigma^2(\rho)$ jest klasy C^1 , to wektor wrażliwości jako gradient funkcji wrażliwości jest funkcją ciągłą, co daje ciągłość $\lambda_\rho = \frac{1}{2} \frac{d\sigma^2(\rho)}{d\rho}$.

Uwaga. 17 Korzystając ze wzoru na $\sigma^2(\rho)$ w twierdzeniu 2.12 oraz wzoru na λ_2 w twierdzeniu 2.3 można sprawdzić, że w problemie optymalizacji portfela bez ograniczeń krótkiej sprzedaży zachodzi wzór $\lambda_2 = \frac{1}{2} \frac{d\sigma^2(\rho)}{d\rho}$. W przeciwieństwie do obecnie rozpatrywanego ograniczenia na krótką sprzedaż, przy braku tego ograniczenia wzór ten jest prawdziwy zawsze. Wynika to z faktu, że przy braku ograniczeń na krótką sprzedaż funkcja $\sigma^2(\rho)$ jest funkcją gładką dla wszystkich ρ .

Przy przejściu $\Delta^2 \rightarrow \{i\} \rightarrow \Delta^2$ w funkcji $\lambda_\rho(\rho)$ wystąpić może nieciągłość. Z faktu, że $\sigma^2(\rho)$ jest wypukła, wynika, że λ_ρ jest rosnąca. Nieciągłość może więc polegać wyłącznie na "skoku w górę" wartości funkcji w punkcie nieciągłości. Taki skok musi oznaczać pojawienie się "dzióbka" na granicy portfelowej, tj. punktu w którym granica portfelowa nie jest klasy C^1 .

Problemem trudnym jest znalezienie warunków, dla których przy przejściu $\Delta^2 \rightarrow \{i\} \rightarrow \Delta^2$ w funkcji $\lambda_\rho(\rho)$ wystąpi nieciągłość.

Dla ustalenia uwagi, rozważmy przejście $\{i, j\} \rightarrow \{i\} \rightarrow \{i, j'\}$, tzn. portfel dwóch instrumentów o numerach i, j przechodzi w portfel złożony tylko z instrumentu i , a ten z kolei w portfel dwóch instrumentów i, j' . Taka kolejność implikuje

Rysunek 4.3: Nieciągłości funkcji $\lambda(\rho)$.

relacje $\mu_j > \mu_i > \mu_{j'}$.

Zapiszmy warunek I rzędu $\Sigma x - \lambda_1 e - \lambda_\rho \mu \geq 0$ w wierzchołku $\{i\}$. Pamiętając, że w tym wierzchołku portfel optymalny ma postać $x = (0, \dots, 1, \dots, 0)$, gdzie 1 stoi na i -tym miejscu a poza tym są same zera, dostajemy układ

$$\begin{aligned}\Sigma_{ii} - \lambda_1 - \lambda_\rho \mu_i &= 0, \\ \Sigma_{ij} - \lambda_1 - \lambda_\rho \mu_j &\geq 0, \\ \Sigma_{ij'} - \lambda_1 - \lambda_\rho \mu_{j'} &\geq 0.\end{aligned}$$

W układzie tym wypisaliśmy tylko równania odpowiadające współrzędnym i , j oraz j' wektora $\Sigma x - \lambda_1 e - \lambda_\rho \mu \geq 0$, ponieważ te 3 składowe wystarczają do wyznaczenia ograniczeń na stałą λ_ρ

$$\frac{\Sigma_{ii} - \Sigma_{ij'}}{\mu_i - \mu_{j'}} \leq \lambda_\rho \leq \frac{\Sigma_{ii} - \Sigma_{ij}}{\mu_i - \mu_j}. \quad (4.10)$$

Nierówność (4.10) daje warunek konieczny i dostateczny na to, aby w wierzchołku $\{i\}$ następował skok wartości λ_ρ . Jeśli ułamek po lewej stronie nierówności jest mniejszy niż ułamek po prawej stronie, to jak pokażemy niżej, musi wystąpić nieciągłość funkcji λ_ρ w punkcie $\rho = \mu_i$. Oczywiście dla pewnych specjalnych wartości parametrów μ i Σ , może się zdarzyć, że oba ułamki we wzorze (4.10) będą równe sobie, nawet jeśli spełnione są warunki braku degeneracji. W takim przypadku, funkcja λ_ρ w punkcie $\rho = \mu_i$ będzie ciągła a granica portfelową klasy C^1 .

Przeanalizujmy portfele optymalne na sympleksie $\{i, j\}$, czyli krawędzi łączącej wierzchołek $\{j\}$ z wierzchołkiem $\{i\}$. Na tej krawędzi niezerowe są tylko składowe x_i oraz x_j portfela x . Taki portfel jest także optymalny jako portfel dwóch instrumentów bez ograniczenia krótkiej sprzedaży (obie składowe są nieujemne). Dla takiego problemu łatwo znaleźć rozwiązanie z twierdzenia 2.3. Rozwiązanie

to wyraża się wzorami

$$x_i = \frac{\rho - \mu_j}{\mu_i - \mu_j}, \quad x_j = \frac{\mu_i - \rho}{\mu_i - \mu_j}.$$

Ponieważ macierz kowariancji dla tego problemu z 2 instrumentami ma postać

$$\begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix},$$

to wariancja portfela optymalnego dana jest wzorem

$$\sigma^2(\rho) = x_i^2 \Sigma_{ii} + 2x_i(1 - x_i) \Sigma_{ij} + (1 - x_i)^2 \Sigma_{jj}.$$

Ponieważ rozpatrujemy problem bez ograniczenia na krótką sprzedaż, więc możemy skorzystać z uwagi 17. Znając postać rozwiązania x_i obliczamy λ_ρ

$$\lambda_\rho(\rho) = \frac{1}{2} \frac{d\sigma^2(\rho)}{d\rho} = \frac{1}{(\mu_i - \mu_j)^2} (\Sigma_{ii}(\rho - \mu_j) + \Sigma_{ij}(\mu_i + \mu_j - 2\rho) + \Sigma_{jj}(\mu_i - \rho)).$$

Przechodząc do granicy $\rho \searrow \mu_i$ otrzymujemy

$$\lim_{\rho \searrow \mu_i} \lambda_\rho(\rho) = \frac{\Sigma_{ii} - \Sigma_{ij}}{\mu_i - \mu_j},$$

czyli prawy z ułamków we wzorze (4.10).

Podobnie rozpatrując zachowanie rozwiązania na krawędzi $\{i, j'\}$ otrzymujemy

$$\lim_{\rho \nearrow \mu_i} \lambda_\rho(\rho) = \frac{\Sigma_{ii} - \Sigma_{ij'}}{\mu_i - \mu_{j'}},$$

czyli lewy z ułamków we wzorze (4.10).

Wynika stąd, że jeśli te dwa ułamki są różne, to granica lewostronna i prawostronna λ_ρ w wierzchołku $\{i\}$ są różne, co oznacza nieciągłość funkcji λ_ρ w punkcie $\rho = \mu_i$.

Poniższy przykład pokazuje, że przy przejściu $\Delta^2 \rightarrow \{i\} \rightarrow \Delta^2$ w funkcji $\lambda_\rho(\rho)$ nie musi wystąpić nieciągłość.

Przykład. 4.11 Przykład ten pochodzi z pracy Dybviga.

Rozważmy rynek złożony z trzech instrumentów, których rozkład stóp zwrotu ma parametry:

$$\Sigma = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1\frac{1}{3} & 2 \\ 0 & 2 & 4\frac{1}{3} \end{pmatrix}, \quad \mu = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}.$$

Dla portfela, w którym $x_1 = 0$, parametry rozkładu dane są wzorami

$$\Sigma^{(1)} = \begin{pmatrix} 1\frac{1}{3} & 2 \\ 2 & 4\frac{1}{3} \end{pmatrix}, \quad \mu^{(1)} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Odpowiednie stałe mają wartości: $\alpha = \frac{15}{16}$, $\beta = \frac{39}{16}$, $\gamma = \frac{111}{16}$, $\delta = \frac{9}{16}$.

Obliczamy funkcję $\sigma^2(\rho)$ oraz portfele optymalne $x(\rho)$:

$$\sigma^2 = \frac{5}{3}\rho^2 - \frac{26}{3}\rho + \frac{37}{3} \quad x_2 = -\rho + 4, x_3 = \rho - 3.$$

Dla portfela, w którym $x_3 = 0$, parametry rozkładu dane są wzorami

$$\Sigma^{(3)} = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1\frac{1}{3} \end{pmatrix}, \quad \mu^{(3)} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Odpowiednie stałe mają wartości: $\alpha = \frac{15}{4}$, $\beta = \frac{21}{4}$, $\gamma = \frac{39}{4}$, $\delta = 9$.

Obliczamy funkcję $\sigma^2(\rho)$ oraz portfele optymalne $x(\rho)$:

$$\sigma^2 = \frac{5}{12}\rho^2 - \frac{7}{6}\rho + \frac{13}{12}, \quad x_1 = \frac{1}{2}(-\rho + 3), x_2 = \frac{1}{2}(\rho - 1).$$

Łatwo zauważyć, że na paraboli B_3^2 : $x_1 \geq 0 \Leftrightarrow \rho \leq 3$, $x_2 \geq 0 \Leftrightarrow \rho \geq 1$. Czyli w przedziale $3 \geq \rho \geq 1$ jesteśmy na B_3^2 . Wartość $\rho = 1$ odpowiada P_{min} , czyli inwestycji $x_1 = 1$. Teraz zwiększamy ρ od wartości $\rho = 3$. Dla $\rho = 3$, $x_1 = 0$, czyli przechodzimy na parabolę B_1^2 . Na niej $x_2 \geq 0 \Leftrightarrow \rho \leq 4$, $x_3 \geq 0 \Leftrightarrow \rho \geq 3$, czyli dla przedziału $3 \leq \rho \leq 4$ jesteśmy na B_1^2 . Ponieważ $\rho = 4$ odpowiada P_{max} , czyli inwestycji $x_3 = 1$, to kończy konstrukcję \bar{B}^3 .

Dla $\rho = 3$ portfel optymalny $x = (0, 1, 0)$, czyli jest złożony wyłącznie z instrumentu x_2 . Licząc pochodną $\frac{d\sigma^2}{d\rho}$ otrzymujemy

$$\lim_{\rho \nearrow 3} \frac{d\sigma^2}{d\rho} = \frac{4}{3},$$

podobnie

$$\lim_{\rho \searrow 3} \frac{d\sigma^2}{d\rho} = \frac{4}{3}.$$

Równość pochodnych jednostronnych oznacza, że w punkcie $\rho = 3$ granica portfelową jest klasy C^1 , mimo że portfel optymalny w tym punkcie składa się wyłącznie z jednego instrumentu.

4.3 Optymalizacja z instrumentem wolnym od ryzyka i ograniczeniem krótkiej sprzedaży

Zajmiemy się teraz optymalizacją portfela inwestycyjnego w przypadku obecności instrumentu wolnego od ryzyka przy nałożeniu ograniczeń na krótką sprzedaż.

DEFINICJA. 4.12 *W przypadku obecności instrumentu wolnego od ryzyka mamy dwa różne przypadki odpowiadające ograniczeniu krótkiej sprzedaży:*

A. Brak krótkiej sprzedaży instrumentów ryzykownych, co odpowiada rozwiązaniu problemu:

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x + \mu_0 x_0 = \rho, \\ e^\top x + x_0 = 1, \\ x \geq 0. \end{cases}$$

B. Brak krótkiej sprzedaży instrumentów ryzykownych i pożyczek instrumentu wolnego od ryzyka, co odpowiada problemowi:

$$\begin{cases} \min_x \frac{1}{2} x^\top \Sigma x, \\ \mu^\top x + \mu_0 x_0 = \rho, \\ e^\top x + x_0 = 1, \\ x \geq 0, x_0 \geq 0. \end{cases}$$

Z twierdzenia 2.15 i wniosku 2.19 wynika natychmiast wniosek dający częściowe rozwiązanie rozważanego problemu optymalizacyjnego.

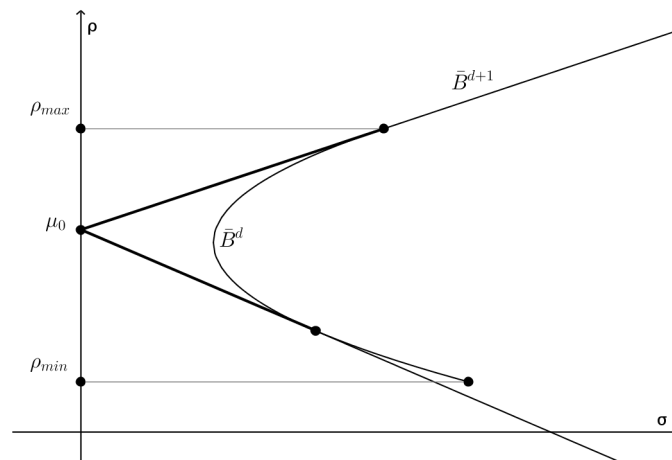
WNIOSEK. 4.13 Niech macierz Σ będzie symetryczna i dodatnio określona oraz wektor μ nie będzie równoległy do e . Rozważmy na płaszczyźnie (ρ, σ) prostą łączącą punkt (ρ_v, σ_v) z punktem odpowiadającym inwestycji wolnej od ryzyka $(\mu_0, 0)$. Prosta ta dana jest równaniem

$$\rho = \mu_0 + \frac{\rho_v - \mu_0}{\sigma_v} \sigma. \quad (4.11)$$

Jeśli optymalny portfel inwestycyjny $x_v = x(\rho_v)$ opisujący inwestycję w instrumenty ryzykowne odpowiadający punktowi (ρ_v, σ_v) jest nieujemny ($x \geq 0$), to każdy portfel na prostej (4.11) spełnia warunek A definicji 4.12.

Dowód. Rozważmy portfel o wariancji σ^2 . Niech $\eta = \frac{\sigma}{\sigma_v}$. Wariancji σ^2 odpowiada portfel inwestycyjny $(\eta x_v, 1 - \eta e^\top x_v)$, bo dla $x = \eta x_v$ mamy $\sigma^2 = x^\top \Sigma x = \eta^2 \sigma_v^2$. Wynika stąd, że jeśli $x_v \geq 0$, to $x = \eta x_v \geq 0$, bo $\eta \geq 0$. ■

TWIERDZENIE. 4.14 Granica portfelowa \bar{B}^{d+1} na płaszczyźnie (ρ, σ) w przypadku obecności instrumentu wolnego od ryzyka i ograniczeniu krótkiej sprzedaży, opisanego w definicji 4.12 punkt A, składa się z dwóch półprostych przechodzących przez punkt $(\mu_0, 0)$ i mających punkt kontaktu z granicą portfelową \bar{B}^d od góry i od dołu. W przypadku ograniczenia na krótką sprzedaż oraz pożyczki, opisanego w definicji 4.12 punkt B, granica portfelowa składa się z odcinków wyżej opisanych półprostych znajdujących się między punktem $(\mu_0, 0)$ a punktem kontaktu odpowiedniej półprostej ze zbiorem \bar{B}^d . W twierdzeniu mowa jest wyłącznie o punkcie kontaktu ze zbiorem \bar{B}^d , ponieważ punkt ten nie musi być punktem styczności, np. gdy jest to punkt P_{max} lub P_{min} .



Rysunek 4.4: Granica portfelowa przy ograniczeniach na krótką sprzedaż. Pogrubione linie to granica portfelowa w przypadku opisanym w definicji 4.12 punkt B.

Dowód. Dowód wynika bezpośrednio z dowodu wniosku 4.13 oraz obserwacji, że jeśli $(\rho_v, \sigma_v) \in \bar{B}^d$, to $e^\top x_v = 1$. ■

Uwaga. 18 W przypadku portfeli optymalnych zawierających instrument wolny od ryzyka oraz ograniczeń na krótką sprzedaż, granica portfelowa nie jest symetryczna względem prostej $\rho = \mu_0$, jak było to w przypadku braku ograniczeń na krótką sprzedaż. Jednocześnie w obecności tych ograniczeń granica efektywna będąca częścią granicy portfelowej \bar{B}^{d+1} ma zawsze punkt wspólny z granicą efektywną będącą częścią granicy portfelowej \bar{B}^d dla problemu bez instrumentu wolnego od ryzyka.

Rozdział 5

Estymatory portfeli optymalnych Markowitza

Model Markowitza w swojej oryginalnej postaci zakładał, że rozkład przyszłych stóp zwrotu instrumentów ryzykownych jest wielowymiarowym rozkładem normalnym, którego momenty są znane inwestorowi. Takie założenie jest oczywiście nierealistyczne praktycznie. Inwestor nie ma żadnej wiedzy o przyszłości. Nawet gdyby przyjąć *a priori*, że rozkład przyszłych stóp zwrotu jest normalny, to inwestor nie zna momentów tego rozkładu. Jedyną informacją jaką dysponuje inwestor, to znajomość cen instrumentów ryzykownych z przeszłości. Kalibracja modelu Markowitza do danych rynkowych polega na wyznaczeniu estymatorów momentów rozkładu stóp zwrotu na podstawie notowań cen z przeszłości przy założeniu, że stopy te mają rozkład normalny. W rozdziale tym opiszemy typowe estymatory oraz konsekwencje, jakie dla znajdowanych portfeli optymalnych ma zastąpienie prawdziwych momentów rozkładu przez ich wartości estymowane.

5.1 Rynek finansowy

Notowane na rynku ceny instrumentów ryzykownych można traktować jako realizacje pewnych procesów stochastycznych. Aby przedstawić poprawny model matematyczny ograniczymy się na początku do cen jednego instrumentu ryzykowanego. Niech (S_t) będzie procesem stochastycznym opisującym cenę tego instrumentu (S_t to cena instrumentu w chwili t).

DEFINICJA. 5.1 *Logarytmiczną stopą zwrotu instrumentu o cenie S_t w okresie (t_0, t_1) nazywamy $\log(S_{t_1}/S_{t_0})$. (Należy pamiętać, że na rynku finansowym czas liczymy w latach. Odnosi się to także do liczb t_0, t_1 liczonych w latach od pewnego arbitralnie ustalonego momentu początkowego.)*

DEFINICJA. 5.2 (Model rynku finansowego) *Zakładamy, że logarytmiczne stopy zwrotu spełniają następujące warunki:*

1. Jeśli $(t_0, t_1) \cap (s_0, s_1) = \emptyset$, to zmienna losowa $\log(\frac{S_{t_1}}{S_{t_0}})$ jest niezależna od $\log(\frac{S_{s_1}}{S_{s_0}})$.
2. $\log(\frac{S_{t_1}}{S_{t_0}}) \sim \mathcal{N}(\mu(t_1 - t_0), \sigma^2(t_1 - t_0))$ dla pewnych stałych μ i σ .

LEMAT. 5.3 *Stopy zwrotu spełniające warunki definicji 5.2 poprawnie skalują się z czasem w tym sensie, że jeśli mamy 3 momenty czasu $t_0 < t_1 < t_2$, to stopa zwrotu w okresie $[t_0, t_2]$ jest sumą stóp zwrotu z okresów $[t_0, t_1]$ i $[t_1, t_2]$ oraz*

$$\log\left(\frac{S_{t_1}}{S_{t_0}}\right) + \log\left(\frac{S_{t_2}}{S_{t_1}}\right) \sim \mathcal{N}(\mu(t_2 - t_0), \sigma^2(t_2 - t_0))$$

Dowód. Rzeczywiści

$$\log\left(\frac{S_{t_2}}{S_{t_0}}\right) = \log\left(\frac{S_{t_2}}{S_{t_1}} \frac{S_{t_1}}{S_{t_0}}\right) = \log\left(\frac{S_{t_2}}{S_{t_1}}\right) + \log\left(\frac{S_{t_1}}{S_{t_0}}\right),$$

co dowodzi addytywności logarytmicznych stóp zwrotu.

Aby znaleźć rozkład stóp zwrotu $\log(\frac{S_{t_2}}{S_{t_0}})$ skorzystamy z punktu 1 definicji 5.2. Mamy wtedy

$$\begin{aligned} \log\left(\frac{S_{t_2}}{S_{t_1}}\right) + \log\left(\frac{S_{t_1}}{S_{t_0}}\right) &\sim \mathcal{N}(\mu(t_2 - t_1), \sigma^2(t_2 - t_1)) + \mathcal{N}(\mu(t_1 - t_0), \sigma^2(t_1 - t_0)) \\ &\sim \mathcal{N}(\mu(t_2 - t_0), \sigma^2(t_2 - t_0)), \end{aligned}$$

co wynika z niezależności stóp zwrotu $\log(\frac{S_{t_2}}{S_{t_1}})$ i $\log(\frac{S_{t_1}}{S_{t_0}})$. ■

Uwaga. 19 *Z definicji 5.2 wynika, że $\log(S_t)$ ma przyrosty niezależne i stacjonarne (tzn. o rozkładzie zależnym tylko od długości okresu). Gdybyśmy dodatkowo założyli ciągłość trajektorii S_t , to $\log S_t$ byłby procesem Wienera z dryfem $\log S_t = \log S_0 + \mu t + \sigma W_t$, gdzie W_t jest standardowym procesem Wienera.*

Przejdziemy teraz do problemu kalibracji modelu Markowitza, czyli wyznaczenia momentów rozkładu stóp zwrotu. W dalszym ciągu ograniczamy się do jednego instrumentu ryzykownego. Niech dane będą obserwacje ceny instrumentu S_t w momentach $t_0, t_0 + \Delta, t_0 + 2\Delta, \dots, t_0 + n\Delta$ dla pewnego Δ . Oznacza to, że ograniczamy się do jednej realizacji (trajektorii) procesu S_t . Stopy zwrotu wyznaczamy ze wzoru $r_i = \log(\frac{S_{t_0+i\Delta}}{S_{t_0+(i-1)\Delta}})$, $i = 1, \dots, n$. Z założeń definicji 5.2 wynika, że (r_i) są niezależnymi próbkami z rozkładu $\mathcal{N}(\mu\Delta, \sigma^2\Delta)$.

Naszym celem jest estymacja momentów μ i σ^2 rozkładu stóp zwrotu. Te estymatory zostaną użyte zamiast nieznanymi prawdziwych momentów rozkładu do wyznaczenia portfela optymalnego Markowitza.

Dokonując estymacji z próbki (patrz definicja 8.3) otrzymujemy

$$\begin{aligned}\hat{\mu}_\Delta &= \frac{1}{n} \sum_{i=1}^n r_i, \\ \hat{\sigma}_\Delta^2 &= \frac{1}{n-1} \sum_{i=1}^n (r_i - \hat{\mu}_\Delta)^2.\end{aligned}\tag{5.1}$$

Stąd

$$\begin{aligned}\hat{\mu} &= \frac{\hat{\mu}_\Delta}{\Delta}, \\ \hat{\sigma}^2 &= \frac{\hat{\sigma}_\Delta^2}{\Delta}.\end{aligned}\tag{5.2}$$

LEMAT. 5.4 *Otrzymane w równaniu (5.2) estymatory mają następujące rozkłady:*

1. $\hat{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n\Delta})$,
2. $\frac{n-1}{\sigma^2} \hat{\sigma}^2 \sim \chi_{n-1}^2$.

Dowód. Dla rozkładu estymatora $\hat{\mu}$ dowód wynika z własności rozkładu normalnego. Jeśli bowiem $\hat{\mu} = \frac{\hat{\mu}_\Delta}{\Delta}$, gdzie $\hat{\mu}_\Delta$ ma rozkład $\mathcal{N}(\mu\Delta, \frac{\sigma^2\Delta}{n})$, to $\hat{\mu}$ ma rozkład $\frac{1}{\Delta} \mathcal{N}(\mu\Delta, \frac{\sigma^2\Delta}{n}) = \mathcal{N}(\mu, \frac{\sigma^2}{n\Delta})$.

Do dowodu punktu 2 wykorzystujemy własności rozkładu χ^2 oraz twierdzenie 8.8. Z twierdzenia tego wynika rozkład $\frac{n-1}{\sigma^2\Delta} \cdot \hat{\sigma}_\Delta^2 \sim \chi_{n-1}^2$. Po przekształceniach otrzymujemy $\frac{n-1}{\sigma^2} \frac{\hat{\sigma}_\Delta^2}{\Delta} = \frac{n-1}{\sigma^2} \hat{\sigma}^2 \sim \chi_{n-1}^2$. ■

Uwaga. 20 (Ważny wniosek) *Załóżmy, że możemy obserwować ceny S_t k razy częściej, tzn. co okres $\frac{\Delta}{k}$. Oznaczmy przez $\hat{\mu}_k, \hat{\sigma}_k^2$ estymatory μ i σ^2 otrzymane z tego częstszego próbkowania. Wówczas*

$$\hat{\mu}_k \sim \mathcal{N}(\mu, \frac{\sigma^2}{n\Delta}) \sim \hat{\mu} \quad \text{oraz} \quad \frac{kn-1}{\sigma^2} \hat{\sigma}_k^2 \sim \chi_{kn-1}^2.$$

Z postaci rozkładu $\hat{\mu}_k$ wynika, że

$$\text{Var}(\hat{\mu}_k) = \text{Var}(\hat{\mu}),$$

czyli wariancja rozkładu estymatora $\hat{\mu}_k$ nie ulega zmianie w stosunku do wariancji estymatora $\hat{\mu}$ (brak poprawy).

Aby obliczyć wariancję estymatora $\hat{\sigma}_k^2$ wykorzystamy fakt, że $\text{Var}(\chi_{n-1}^2) = 2(n-1)$. Stąd $\text{Var}(\frac{n-1}{\sigma^2} \hat{\sigma}^2) = 2(n-1)$. Z drugiej strony mamy $\text{Var}(\frac{n-1}{\sigma^2} \hat{\sigma}^2) = \frac{(n-1)^2}{\sigma^4} \text{Var}(\hat{\sigma}^2)$. Łącząc te dwie równości dostajemy $\text{Var}(\hat{\sigma}^2) = \frac{2\sigma^4}{n-1}$. Wykorzystując ostatnią równość do obliczenia wariancji estymatora $\hat{\sigma}_k^2$ otrzymujemy

$$\text{Var}(\hat{\sigma}_k^2) = \frac{2\sigma^4}{kn-1} = \frac{n-1}{kn-1} \text{Var}(\hat{\sigma}^2).$$

Oznacza to zmniejszenie wariancji w przybliżeniu k razy.

Uwaga. 21 Łatwo można zauważyć, że $\hat{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^2}{t_0+n\Delta-t_0})$. Zatem równie dobrym estymatorem μ są stopy zwrotu $\log(\frac{S_{t_0+n\Delta}}{S_{t_0}}) \cdot \frac{1}{n\Delta}$. Nie jest to zaskakujące, gdyż

$$\sum_{i=1}^n r_i = \sum_{i=1}^n \log\left(\frac{S_{t_0+i\Delta}}{S_{t_0+(i-1)\Delta}}\right) = \log\left(\frac{S_{t_0+n\Delta}}{S_{t_0}}\right).$$

Obserwacja ta ma bardzo istotne konsekwencje, ponieważ pokazuje ona, że jedynym sposobem na poprawę statystycznych własności (zmniejszenie wariancji) estymatora μ jest wydłużenie okresu obserwacji rynku.

5.2 Portfel optymalny dla jednego instrumentu ryzykownego

Rozważmy inwestycję o czasie trwania h . Niech x będzie wagą instrumentu ryzykownego w portfelu a $x_0 = (1-x)$ wagą instrumentu wolnego od ryzyka. Chcemy rozwiązać klasyczne zadanie Markowitza optymalizacji portfela, czyli znaleźć optymalną wartość x wykorzystując wyestymowane wartości średniej stopy zwrotu $\hat{\mu}$ oraz wariancji $\hat{\sigma}^2$. W klasycznym sformułowaniu zadanie optymalizacyjne wygląda następująco:

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^2 \hat{\sigma}^2 \cdot h, \\ (1-x) \cdot \mu_0 h + x \hat{\mu} h &= \rho h, \end{aligned}$$

gdzie ρ jest oczekiwaną stopą zwrotu z inwestycji.

Łatwo zauważyć, że rozwiązanie tego problemu optymalizacyjnego nie zależy od długości okresu inwestycyjnego. Portfel optymalny jest więc rozwiązaniem zadania

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^2 \hat{\sigma}^2, \\ (1-x) \cdot \mu_0 + x \hat{\mu} &= \rho. \end{aligned}$$

Zauważmy, że z warunku na oczekiwaną stopę zwrotu dostajemy jeden punkt dopuszczalny $\hat{x} = \frac{\rho - \mu_0}{\hat{\mu} - \mu_0}$. Dla jednego instrumentu ryzykownego problem ten jest zdegenerowany. Co więcej, dla tak wyznaczonego portfela nie istnieje wartość oczekiwana \hat{x} , bo całka

$$\mathbb{E}\hat{x} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{\rho - \mu_0}{u - \mu_0} \exp\left(-\frac{(u - \mu)^2}{2\sigma^2/n\Delta}\right) du$$

nie istnieje (to oznacza, że dokładność wyznaczenia portfela optymalnego jest „bardzo mała”).

Okazuje się jednak, że sformułowanie zadania optymalizacyjnego przez funkcję preferencji prowadzi do rozwiązania niezdegenerowanego. W rozpatrywanym

przypadku zadanie optymalizacyjne z funkcją preferencji ma postać

$$\max_{x \in \mathbb{R}} h \cdot \mu_0(1 - x) + h \cdot \mu x - \frac{A}{2} x^2 \sigma^2 h,$$

gdzie A jest współczynnikiem awersji do ryzyka inwestora.

Ten problem jest równoważny następującemu

$$\max_{x \in \mathbb{R}} (\mu - \mu_0)x - \frac{A}{2} x^2 \sigma^2. \quad (5.3)$$

Rozwiązaniem problemu jest portfel optymalny

$$x^* = \frac{\mu - \mu_0}{A\sigma^2}.$$

Będziemy teraz modyfikować zadanie optymalizacyjne wstawiając zamiast dokładnych wartości μ lub σ^2 wielkości wyestymowane z próbki rynkowej. Naszym celem będzie skonstruowanie nieobciążonych estymatorów portfela optymalnego x^* oraz zbadanie ich wariancji.

TWIERDZENIE. 5.5 *Jeśli w funkcji preferencji (5.3) przyjmiemy, że średnia stopa zwrotu μ jest znana dokładnie a wariancję zastąpimy wielkością wyestymowaną $\hat{\sigma}^2$, to portfel optymalny dany jest wzorem*

$$\hat{x}_\mu = \frac{n - 3}{n - 1} \frac{\mu - \mu_0}{A\hat{\sigma}^2}.$$

Wtedy estymator \hat{x}_μ ma następujące momenty

$$\begin{aligned} \mathbb{E}(\hat{x}_\mu) &= x^*, \\ \text{Var}(\hat{x}_\mu) &= \frac{2}{n - 5} (x^*)^2. \end{aligned}$$

Jeśli w funkcji preferencji wariancja σ^2 jest znana dokładnie a średnią stopę zwrotu zastąpimy wielkością wyestymowaną $\hat{\mu}$, to portfel optymalny dany jest wzorem

$$\hat{x}_\sigma = \frac{\hat{\mu} - \mu_0}{A\sigma^2}.$$

Wtedy estymator \hat{x}_σ ma następujące momenty

$$\begin{aligned} \mathbb{E}(\hat{x}_\sigma) &= x^*, \\ \text{Var}(\hat{x}_\sigma) &= \frac{1}{A^2 \sigma^2 n \Delta}. \end{aligned}$$

Jeśli w funkcji preferencji średnią stopę zwrotu oraz wariancję zastąpimy wielkościami wyestymowanymi, to portfel optymalny dany jest wzorem

$$\hat{x} = \frac{n - 3}{n - 1} \frac{\hat{\mu} - \mu_0}{A\hat{\sigma}^2}.$$

Wtedy estymator \hat{x} ma następujące momenty

$$\begin{aligned}\mathbb{E}(\hat{x}) &= x^*, \\ \text{Var}(\hat{x}) &= \frac{n-3}{A^2\sigma^2(n-5)n\Delta} + \frac{2}{n-5}(x^*)^2.\end{aligned}$$

Dowód. W przypadku estymatora \hat{x}_μ proste rachunki dają

$$\begin{aligned}\mathbb{E}(\hat{x}_\mu) &= \frac{n-3}{n-1} \frac{\mu - \mu_0}{A} \mathbb{E}\left(\frac{1}{\hat{\sigma}^2}\right) = \frac{n-3}{n-1} \frac{\mu - \mu_0}{A} \mathbb{E}\left(\frac{\sigma^2}{(n-1)\hat{\sigma}^2}\right) \frac{n-1}{\sigma^2} = \\ &= (n-3) \frac{\mu - \mu_0}{A\sigma^2} \cdot \frac{1}{n-3} = x^*.\end{aligned}$$

Skorzystalismy przy tym z faktu, że zmienna $\frac{\sigma^2}{(n-1)\hat{\sigma}^2}$ ma rozkład χ_{n-1}^2 oraz twierdzenia 8.7.

Dla wariancji mamy

$$\begin{aligned}\text{Var}(\hat{x}_\mu) &= \mathbb{E}(\hat{x}_\mu^2) - (\mathbb{E}\hat{x}_\mu)^2 = \frac{(n-3)^2}{(n-1)^2} \frac{(\mu - \mu_0)^2}{A^2} \mathbb{E}\left(\frac{1}{\hat{\sigma}^2}\right)^2 - (x^*)^2 = \\ &= \frac{(n-3)^2}{(n-1)^2} \frac{(\mu - \mu_0)^2}{A^2} \mathbb{E}\left(\frac{\sigma^2}{(n-1)\hat{\sigma}^2}\right)^2 \frac{(n-1)^2}{\sigma^4} - (x^*)^2 = \\ &= (x^*)^2 (n-3)^2 \frac{1}{(n-3)(n-5)} - (x^*)^2 = \\ &= (x^*)^2 \left(\frac{n-3}{n-5} - 1\right) = \frac{2}{n-5}(x^*)^2.\end{aligned}$$

Dla estymatora \hat{x}_σ rachunki są prostsze

$$\begin{aligned}\mathbb{E}(\hat{x}_\sigma) &= \mathbb{E}\left(\frac{\hat{\mu} - \mu_0}{A\sigma^2}\right) = \frac{\mu - \mu_0}{A\sigma^2} = x^*. \\ \text{Var}(\hat{x}_\sigma) &= \text{Var}\left(\frac{\hat{\mu} - \mu_0}{A\sigma^2}\right) = \frac{1}{A^2\sigma^4} \text{Var}(\hat{\mu}) = \frac{\sigma^2}{A^2\sigma^4 n\Delta} = \frac{1}{A^2\sigma^2 n\Delta}.\end{aligned}$$

Dowód braku obciążenia dla estymatora \hat{x} jest analogiczny do dowodu dla estymatora \hat{x}_μ

$$\begin{aligned}\mathbb{E}(\hat{x}) &= \frac{n-3}{n-1} \mathbb{E}\left(\frac{\hat{\mu} - \mu_0}{A\hat{\sigma}^2}\right) = \frac{n-3}{n-1} \mathbb{E}(\hat{\mu} - \mu_0) \frac{1}{A} \mathbb{E}\left(\frac{1}{\hat{\sigma}^2}\right) = \\ &= \frac{n-3}{n-1} \frac{(\mu - \mu_0)}{A} \frac{n-1}{n-3} \frac{1}{\sigma^2} = x^*.\end{aligned}$$

Przy czym w drugiej równości korzystamy z niezależności estymatorów $\hat{\mu}$ i $\hat{\sigma}^2$.

Wariancję estymatora \hat{x} liczymy podobnie jak poprzednio

$$\text{Var}(\hat{x}) = \mathbb{E}(\hat{x}^2) - (x^*)^2 = \left(\frac{n-3}{n-1}\right)^2 \mathbb{E}(\hat{\mu} - \mu_0)^2 \frac{1}{A^2} \mathbb{E}\left(\frac{1}{\hat{\sigma}^2}\right)^2 - (x^*)^2.$$

Wartość oczekiwaną $\mathbb{E}(\hat{\mu} - \mu_0)^2$ obliczamy korzystając z przedstawienia

$$\mathbb{E}(\hat{\mu} - \mu_0)^2 = \text{Var}(\hat{\mu} - \mu_0) + (\mathbb{E}(\hat{\mu} - \mu_0))^2.$$

Ponieważ $\text{Var}(\hat{\mu} - \mu_0) = \frac{\sigma^2}{n\Delta}$, więc wykorzystując twierdzenie 8.7 otrzymujemy

$$\begin{aligned} \text{Var}(\hat{x}) &= \\ & \frac{(n-3)^2}{(n-1)^2 A^2} \left(\frac{\sigma^2}{n\Delta} + (\mu - \mu_0)^2 \right) \mathbb{E} \left(\frac{\sigma^2}{(n-1)\hat{\sigma}^2} \right)^2 \frac{(n-1)^2}{\sigma^4} - (x^*)^2 = \\ & \frac{(n-3)}{(n-5)A^2\sigma^4} \left(\frac{\sigma^2}{n\Delta} + (\mu - \mu_0)^2 \right) - (x^*)^2 = \\ & \frac{(n-3)}{(n-5)A^2\sigma^2 n\Delta} + \frac{n-3}{n-5} (x^*)^2 - (x^*)^2 = \\ & \frac{n-3}{(n-5)A^2\sigma^2 n\Delta} + \left(\frac{n-3}{n-5} - 1 \right) (x^*)^2 = \frac{n-3}{(n-5)A^2\sigma^2 n\Delta} + \frac{2}{n-5} (x^*)^2. \end{aligned}$$

Uwaga. 22 Estymatory z twierdzenia 5.5 są nie tylko nieobciążone, ale także są one mocno zgodne, tzn. przy $n \rightarrow \infty$ zbiegają p.n. do x^* .

Rozważmy jeszcze w przypadku jednego instrumentu ryzykowanego problem, który można nazwać problemem ceny informacji.

Niech $U(x) = (1-x)\mu_0 + x\mu - \frac{A}{2}\sigma^2 x^2$ będzie funkcją preferencji dla zadania optymalizacyjnego. Portfel optymalny x^* odpowiada maksimum tej funkcji. Jeśli \hat{x} jest portfelem optymalnym uzyskanym przez zastąpienie dokładnych momentów rozkładu ich estymatorami, to $U(\hat{x}) \leq U(x^*)$, przy czym nierówność ta jest zwykle ostra. Powstaje pytanie: ile maksymalnie jesteśmy skłonni zapłacić (jako rezygnację z pewnej części zysku czyli oczekiwanej stopy zwrotu portfela), aby uzyskać pełną wiedzę o parametrach rozkładu przyszłych zwrotów?

Aby odpowiedzieć na to pytanie zdefiniujemy zmodyfikowaną funkcję preferencji

$$U_\delta(x) = (1-x)(\mu_0 - \delta) + x(\mu - \delta) - \frac{A}{2}\sigma^2 x^2 = U(x) - \delta$$

gdzie $\delta > 0$.

Odpowiedź na nasze pytanie oznacza poszukiwanie maksymalnego δ , takiego że $U_\delta(x^*) \geq \mathbb{E}U(\hat{x})$.

LEMAT. 5.6 Niech δ^* będzie największą wartością δ , dla której zachodzi nierówność $U_\delta(x^*) \geq \mathbb{E}U(\hat{x})$. Wtedy

$$\delta^* = \frac{A\sigma^2}{n-5} (x^*)^2 + \frac{n-3}{2An(n-5)\Delta}.$$

Dowód. Oczywiście $\delta^* = U(x^*) - \mathbb{E}U(\hat{x})$. Korzystając z dowodu twierdzenia 5.5 dostajemy

$$\begin{aligned} \mathbb{E}U(\hat{x}) &= \mu_0 + (\mu - \mu_0)\mathbb{E}(\hat{x}) - \frac{A\sigma^2}{2}\mathbb{E}(\hat{x}^2) = \\ &= \mu_0 + (\mu - \mu_0)x^* - \frac{A\sigma^2}{2}\left(\frac{n-3}{(n-5)A^2\sigma^2n\Delta} + \frac{n-3}{n-5}(x^*)^2\right) = \\ &= U(x^*) + \frac{A\sigma^2}{2}(x^*)^2 - \frac{(n-3)}{2A(n-5)n\Delta} - \frac{A\sigma^2}{2}(x^*)^2\frac{n-3}{n-5} = \\ &= U(x^*) - \frac{n-3}{2A(n-5)n\Delta} - \frac{A\sigma^2}{2}(x^*)^2\left(\frac{n-3}{n-5} - 1\right) = \\ &= U(x^*) - \frac{n-3}{2A(n-5)n\Delta} - \frac{A\sigma^2}{2}(x^*)^2\frac{2}{n-5}. \end{aligned}$$

■

Podobnie możemy policzyć δ dla przypadku, gdy znamy prawdziwe μ lub σ^2 a tylko drugi z momentów rozkładu jest estymowany.

5.3 Portfele optymalne dla wielu instrumentów ryzykownych

Rozszerzmy teraz wyniki uzyskane dla $d = 1$ na większą liczbę instrumentów ryzykownych. W przypadku wielu instrumentów musimy rozszerzyć definicję rynku finansowego. Będziemy teraz rozpatrywali wielowymiarowe procesy stochastyczne S_t^i , $i = 1, \dots, d$, gdzie S_t^i to cena i -tego instrumentu w chwili t .

DEFINICJA. 5.7 Zakładamy, że rynkowe logarytmiczne stopy zwrotu dla każdego instrumentu spełniają następujące warunki, które są naturalnym rozszerzeniem na wiele wymiarów warunków z definicji 5.2

1. Jeśli $(t_0, t_1) \cap (s_0, s_1) = \emptyset$, to $\log\left(\frac{S_{t_1}^i}{S_{t_0}^i}\right)$ jest niezależny od $\log\left(\frac{S_{s_1}^j}{S_{s_0}^j}\right)$, $i, j = 1, \dots, d$.
2. $\left(\log\left(\frac{S_{t_1}^1}{S_{t_0}^1}\right), \dots, \log\left(\frac{S_{t_1}^d}{S_{t_0}^d}\right)\right) \sim \mathcal{N}(\mu(t_1 - t_0), \Sigma(t_1 - t_0))$ dla pewnego wektora μ oraz macierzy Σ .

Obecnie przejdziemy do badania własności estymatorów wag portfeli optymalnych dla wielu instrumentów ryzykownych. Okaże się, że zwiększenie liczby instrumentów zwiększy błąd ponad liniowo, tzn. zaczną odgrywać rolę kowariancje rozkładów poszczególnych instrumentów.

Zauważmy także, że w przypadku, gdy próbki pochodzą z rozkładów ciągłych (dla problemu Markowitza jest to rozkład normalny), to wyestymowana z próbki macierz kowariancji jest nieosobliwa (twierdzenie 8.14). Nieosobliwość macierzy

kowariancji jest konieczna do stosowania wzorów na rozwiązanie problemu Markowitza, ponieważ wzory te otrzymywaliśmy przy założeniu nieosobliwości Σ .

Przejdziemy teraz do rozwiązania problemu optymalizacji portfela złożonego z wielu instrumentów ryzykownych. Podobnie jak w przypadku jednowymiarowym, będziemy rozwiązywali zadanie optymalizacyjne z funkcją preferencji. Portfele optymalne zapisujemy w postaci (x, x_0) , gdzie $x \in \mathbb{R}^d$ jest wektorem wag instrumentów ryzykownych a x_0 wagą instrumentu wolnego od ryzyka. Problem optymalizacyjny polega na znalezieniu portfela, który rozwiązuje zadanie (przypomnijmy, że rozwiązanie nie zależy od okresu inwestycji)

$$\max_{x \in \mathbb{R}^d} (\mu - \mu_0 e)^\top x - \frac{A}{2} x^\top \Sigma x \quad (5.4)$$

z ograniczeniem

$$x_0 + e^\top x = 1.$$

Jak wiemy z twierdzenia 2.10 rozwiązanie dane jest wzorem

$$x^* = \frac{1}{A} \Sigma^{-1} (\mu - \mu_0 e).$$

Na początku sformułujmy pomocniczy lemat, który wykorzystamy przy dowodzie głównego twierdzenia.

LEMAT. 5.8 *Niech Z będzie d -wymiarową zmienną losową, dla której istnieje $\text{Cov}(Z)$. Wówczas $\mathbb{E}Z^\top Z = \text{tr}(\text{Cov}(Z)) + \|\mathbb{E}Z\|^2$*

Dowód. Dowód wynika bezpośrednio z następujących dwóch równości:

$$\begin{aligned} \mathbb{E}(Z - \mathbb{E}Z)^\top (Z - \mathbb{E}Z) &= \mathbb{E}Z^\top Z - \|\mathbb{E}Z\|^2, \\ \mathbb{E}(Z - \mathbb{E}Z)^\top (Z - \mathbb{E}Z) &= \sum_{i=1}^d \mathbb{E}(Z_i - (\mathbb{E}Z)_i)^2 = \text{tr}(\text{Cov}(Z)). \end{aligned}$$

■

TWIERDZENIE. 5.9 *Jeśli w funkcji preferencji (5.4) średnia stopa zwrotu μ jest znana dokładnie a macierz kowariancji zastąpimy macierzą wyestymowaną, to*

$$\hat{x}_\mu = \frac{1}{A} \frac{n-d-2}{n-1} \hat{\Sigma}^{-1} (\mu - \mu_0 e).$$

Wtedy dla estymatora \hat{x}_μ otrzymujemy

$$\begin{aligned} \mathbb{E}(\hat{x}_\mu) &= x^*, \\ \mathbb{E}\|\hat{x}_\mu - x^*\|^2 &= \frac{n-d}{(n-d-1)(n-d-4)} \|x^*\|^2 \\ &\quad + \frac{1}{A^2} \frac{n-d-2}{(n-d-1)(n-d-4)} \text{tr}(\Sigma^{-1}) (\mu - \mu_0 e)^\top \Sigma^{-1} (\mu - \mu_0 e). \end{aligned}$$

Jeśli w funkcji preferencji średnią stopę zwrotu μ oraz macierz kowariancji Σ zastąpimy wielkościami wyestymowanymi, to

$$\hat{x} = \frac{1}{A} \frac{n-d-2}{n-1} \hat{\Sigma}^{-1} (\hat{\mu} - \mu_0 e).$$

Wtedy dla estymatora \hat{x} otrzymujemy

$$\begin{aligned} \mathbb{E}(\hat{x}) &= x^*, \\ \mathbb{E}\|\hat{x} - x^*\|^2 &= \mathbb{E}\|\hat{x}_\mu - x^*\|^2 \\ &+ \frac{1}{A^2} \frac{(n-d-2)^2}{n\Delta(n-d-1)(n-d-4)} \text{tr}(\Sigma^{-1}) \left(1 + \frac{d}{n-d-2}\right). \end{aligned}$$

Dowód. Udowodnimy tylko punkt 2 twierdzenia bo punkt 1 dowodzi się analogicznie. Zaczniemy od dowodu braku obciążenia estymatora \hat{x} . Z niezależności $\hat{\Sigma}$ i $\hat{\mu}$ dostajemy

$$\begin{aligned} \mathbb{E}(\hat{x}) &= \frac{1}{A} \frac{n-d-2}{n-1} \mathbb{E}(\hat{\Sigma}^{-1}) \mathbb{E}(\hat{\mu} - \mu_0 e) = \\ &\frac{1}{A} \frac{n-d-2}{n-1} (n-1) \mathbb{E}\left(\left((n-1)\hat{\Sigma}\right)^{-1}\right) (\mu - \mu_0 e) = \\ &\frac{1}{A} \Sigma^{-1} (\mu - \mu_0 e) = x^*, \end{aligned}$$

ponieważ $\mathbb{E}\left(\left((n-1)\hat{\Sigma}\right)^{-1}\right) = \frac{1}{n-1-d-1} \Sigma^{-1}$ (twierdzenie 8.20).

Korzystając z braku obciążenia estymatora \hat{x} dla $\mathbb{E}\|\hat{x} - x^*\|^2$ dostajemy wyrażenie

$$\begin{aligned} \mathbb{E}\|\hat{x} - x^*\|^2 &= \mathbb{E}\left((\hat{x} - x^*)^\top (\hat{x} - x^*)\right) = \mathbb{E}\hat{x}^\top \hat{x} - 2\mathbb{E}\hat{x}^\top x^* + (x^*)^\top x^* = \\ &\mathbb{E}\hat{x}^\top \hat{x} - (x^*)^\top x^*. \end{aligned}$$

Z definicji \hat{x} mamy

$$\mathbb{E}(\hat{x}^\top \hat{x}) = \frac{1}{A^2} \frac{(n-d-2)^2}{(n-1)^2} \mathbb{E}\left((\hat{\mu} - \mu_0 e)^\top \hat{\Sigma}^{-2} (\hat{\mu} - \mu_0 e)\right).$$

Na mocy niezależności estymatorów $(\hat{\mu} - \mu_0 e)$ i $\hat{\Sigma}$ dostajemy

$$\begin{aligned} \mathbb{E}\left((\hat{\mu} - \mu_0 e)^\top \hat{\Sigma}^{-2} (\hat{\mu} - \mu_0 e)\right) &= \\ &\mathbb{E}\left(\mathbb{E}\left((\hat{\mu} - \mu_0 e)^\top \hat{\Sigma}^{-2} (\hat{\mu} - \mu_0 e) \mid (\hat{\mu} - \mu_0 e)\right)\right) = \\ &\mathbb{E}\left((\hat{\mu} - \mu_0 e)^\top \mathbb{E}(\hat{\Sigma}^{-2}) (\hat{\mu} - \mu_0 e)\right). \end{aligned}$$

Twierdzenie 8.20 implikuje, że

$$\begin{aligned} & (\hat{\mu} - \mu_0 e)^\top \mathbb{E}(\hat{\Sigma}^{-2})(\hat{\mu} - \mu_0 e) = \\ & \frac{(n-1)^2}{(n-d-1)(n-d-2)(n-d-4)} \text{tr}(\Sigma^{-1})(\hat{\mu} - \mu_0 e)^\top \Sigma^{-1}(\hat{\mu} - \mu_0 e) \\ & + \frac{(n-1)^2}{(n-d-1)(n-d-4)} (\hat{\mu} - \mu_0 e)^\top \Sigma^{-2}(\hat{\mu} - \mu_0 e). \end{aligned}$$

Równość wynika z przedstawienia $\hat{\Sigma}^{-2} = (n-1)^2((n-1)\hat{\Sigma})^{-2}$ oraz faktu, że $((n-1)\hat{\Sigma})^{-2}$ ma rozkład Wisharta $W_d(n-1, \Sigma)$ (twierdzenie 8.19).

Pierwszy składnik powyższej równości przekształcamy wykorzystując fakt, że $(\hat{\mu} - \mu_0 e) \sim \mathcal{N}((\mu - \mu_0 e), \frac{\Sigma}{n\Delta})$. Ponieważ Σ^{-1} jest macierzą symetryczną i dodatnio określoną, więc istnieje macierz nieosobliwa C , taka że $C^\top C = \Sigma^{-1}$. Biorąc $Y = C(\hat{\mu} - \mu_0 e)$ dostajemy $(\hat{\mu} - \mu_0 e)^\top \Sigma^{-1}(\hat{\mu} - \mu_0 e) = Y^\top Y$. Wtedy z lematu 8.16 wynika, że

$$Y \sim \mathcal{N}\left(C(\mu - \mu_0 e), \frac{C\Sigma C^\top}{n\Delta}\right) = \mathcal{N}\left(C(\mu - \mu_0 e), \frac{I}{n\Delta}\right),$$

ponieważ

$$C\Sigma C^\top = C(C^\top C)^{-1}C^\top = CC^{-1}(C^\top)^{-1}C^\top = I.$$

Korzystając z lematu 5.8 dostajemy

$$\mathbb{E}Y^\top Y = \text{tr}\left(\frac{I}{n\Delta}\right) + \|C(\mu - \mu_0 e)\|^2 = \frac{d}{n\Delta} + (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e),$$

ponieważ $\|C(\mu - \mu_0 e)\|^2 = (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e)$.

W efekcie otrzymujemy równość

$$\mathbb{E}((\hat{\mu} - \mu_0 e)^\top \Sigma^{-1}(\hat{\mu} - \mu_0 e)) = \frac{d}{n\Delta} + (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e).$$

Podobnie postępujemy z drugim składnikiem $(\hat{\mu} - \mu_0 e)^\top \Sigma^{-2}(\hat{\mu} - \mu_0 e)$.

Bierzemy $Y = \Sigma^{-1}(\hat{\mu} - \mu_0 e)$ i zauważamy, że $Y \sim \mathcal{N}(\Sigma^{-1}(\mu - \mu_0 e), \frac{\Sigma^{-1}}{n\Delta})$. Zatem z lematu 5.8,

$$\begin{aligned} \mathbb{E}((\hat{\mu} - \mu_0 e)^\top \Sigma^{-2}(\hat{\mu} - \mu_0 e)) &= \mathbb{E}Y^\top Y = \text{tr}\left(\frac{\Sigma^{-1}}{n\Delta}\right) + \|\Sigma^{-1}(\mu - \mu_0 e)\|^{-2} = \\ &= \text{tr}\left(\frac{\Sigma^{-1}}{n\Delta}\right) + A^2\|x^*\|^2. \end{aligned}$$

Wstawiając policzone wyrażenia do początkowego wzoru otrzymujemy

$$\begin{aligned} \mathbb{E}\|\hat{x} - x^*\|^2 &= \frac{1}{A^2} \frac{(n-d-2)^2}{(n-1)^2} \left(\frac{(n-1)^2}{(n-d-1)(n-d-2)(n-d-4)} \times \right. \\ & \quad \times \text{tr}(\Sigma^{-1}) \left(\frac{d}{n\Delta} + (\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e) \right) \\ & \quad \left. + \frac{(n-1)^2}{(n-d-1)(n-d-4)} \left(\frac{1}{n\Delta} \text{tr}(\Sigma^{-1}) + A^2\|x^*\|^2 \right) \right) - \|x^*\|^2. \end{aligned}$$

Pozostaje nam jeszcze uprościć to wyrażenie.

Przy $\|x^*\|^2$ mamy współczynnik

$$\begin{aligned} & \frac{1}{A^2} \frac{(n-d-2)^2}{(n-1)^2} \frac{(n-1)^2}{(n-d-1)(n-d-4)} A^2 - 1 = \\ & \frac{(n-d-2)^2 - (n-d-1)(n-d-4)}{(n-d-1)(n-d-4)} = \\ & \frac{n-d}{(n-d-1)(n-d-4)}. \end{aligned}$$

Przy $\text{tr}(\Sigma^{-1})(\mu - \mu_0 e)^\top \Sigma^{-1}(\mu - \mu_0 e)$ mamy

$$\begin{aligned} & \frac{1}{A^2} \frac{(n-d-2)^2}{(n-1)^2} \frac{(n-1)^2}{(n-d-1)(n-d-2)(n-d-4)} = \\ & \frac{1}{A^2} \frac{n-d-2}{(n-d-1)(n-d-4)}. \end{aligned}$$

Przy $\frac{1}{n\Delta} \text{tr}(\Sigma^{-1})$ dostajemy

$$\begin{aligned} & \frac{1}{A^2} \frac{(n-d-2)^2}{(n-1)^2} \frac{(n-1)^2}{(n-d-1)(n-d-2)(n-d-4)} d \\ & + \frac{1}{A^2} \frac{(n-d-2)^2}{(n-1)^2} \frac{(n-1)^2}{(n-d-1)(n-d-4)} = \\ & \frac{1}{A^2} \frac{(n-d-2)^2}{(n-d-1)(n-d-4)} \left(\frac{d}{n-d-2} + 1 \right). \end{aligned}$$

■

Uwaga. 23 Zauważmy, że podobnie jak w przypadku jednego instrumentu ryzykownego, przy $n \rightarrow \infty$

$$\mathbb{E}\|\hat{x} - x^*\| \rightarrow \frac{1}{A^2 T} \text{tr}(\Sigma^{-1}),$$

gdzie $T = \lim_{n \rightarrow \infty} n\Delta$ jest długością przedziału czasu, w którym obserwujemy ceny instrumentów ryzykownych.

Oznacza to, że $\mathbb{E}\|\hat{x} - x^*\|$ nie dąży do zera gdy $n \rightarrow \infty$. W przypadku wielu instrumentów ryzykownych sytuacja może być jednak znacznie gorsza niż dla jednego instrumentu. Wynika to obecności wyrazu $\text{tr}(\Sigma^{-1})$. Dla rzeczywistych danych rynkowych macierz kowariancji Σ ma zwykle wartości własne, które potrafią się różnić nawet o kilka rzędów wielkości, co powoduje, że $\text{tr}(\Sigma^{-1})$ może być bardzo duży a tym samym bardzo duży będzie średni błąd kwadratowy dla wag portfela $\mathbb{E}\|\hat{x} - x^*\|$.

Rozdział 6

Model Blacka–Littermana

6.1 Założenia modelu

Model Markowitza, któremu był poświęcony dotychczasowy wykład, ma kilka istotnych mankamentów. Jeden z tych mankamentów analizowaliśmy w rozdziale 5. Zauważyliśmy wtedy, że jeśli parametry rozkładu używane w modelu Markowitza estymujemy z danych historycznych, to otrzymane estymatory wag instrumentów w portfelu obarczone są bardzo dużym błędem. Stawia to pod znakiem zapytania celowość używania modelu Markowitza do znajdowania optymalnych portfeli inwestycyjnych.

Zarządzający portfelami inwestycyjnymi widzieli jeszcze inny mankament modelu. Związany jest on z faktem, że w modelu nie ma miejsca na indywidualną ocenę rynku przez zarządzającego portfelem. Model daje takie same wyniki każdemu użytkownikowi. Tymczasem praktycy w znacznie większym stopniu niż z wyników optymalizacji Markowitza korzystali z własnych przewidywań na temat przyszłego zachowania rynku. Postępowanie takie było czysto intuicyjne i nie miało żadnego teoretycznego uzasadnienia.

Zmiana tej sytuacji nastąpiła w momencie wprowadzenia modelu Blacka-Littermana. Model został opracowany przez F. Blacka i R. Littermana w firmie Goldman Sachs i od 1991 roku zrobił ogromną karierę w firmach Asset Management. To jest skomplikowany model, który jest używany zarówno w optymalizacji portfela na rynku krajowym (instrumenty w jednej walucie) jak i rynku międzynarodowym (instrumenty w wielu walutach). Zrozumienie funkcjonowania modelu na rynku międzynarodowym jest dość trudne, dlatego ograniczymy się do opisu modelu w wersji inwestycji w jednej walucie.

Podstawowy pomysł Blacka i Littermana polegał na rozdzieleniu opisu rynku na dwa elementy:

1. Informacji o długofalowym zachowaniu rynku.
2. Opinii ekspertów na temat przyszłych stóp zwrotu (niektórych) instrumentów na rynku.

Te dwa obiekty są definiowane jako odpowiednie rozkłady prawdopodobieństwa a następnie łączone przy pomocy wzoru Bayesa w celu otrzymania informacji o przyszłym zachowaniu rynku.

Założenie. *Podstawowe założenia modelu Blacka-Littermana można sformułować w dwóch punktach:*

- *rozkłady prawdopodobieństwa opisujące długookresowe zachowanie rynku oraz opinie ekspertów są wielowymiarowymi rozkładami normalnymi,*
- *portfel optymalny jest znajdowany jako rozwiązanie zadania minimalizacji wariancji portfela przy założonej oczekiwanej stopie zwrotu, czyli klasycznego zadania Markowitza.*

Konsekwencją tych założeń jest zredukowanie problemu do znalezienia z wzoru Bayesa nie całego rozkładu a jedynie wektora wartości oczekiwanych oraz macierzy kowariancji stóp zwrotu.

Rozkład równowagowy

Aby opisać długofalowe zachowanie rynku zakładamy, że

1. Wszyscy inwestorzy na rynku podejmują decyzje inwestycyjne na podstawie rozwiązania tego samego zadania optymalizacyjnego

$$\max_x \mu^\top x - \frac{A}{2} x^\top \Sigma x,$$

choć być może z innym współczynnikiem awersji do ryzyka A .

2. W wyniku działań inwestorów na rynku ustala się długookresowy stan równowagi, a portfel rynkowy odpowiadający temu stanowi równowagi jest portfelem efektywnym.

W celu znalezienia rozkładu równowagowego czynimy dodatkowe założenia techniczne:

1. Znamy skład portfela rynkowego, tzn. znamy instrumenty, które do niego wchodzi oraz ich udziały w portfelu rynkowym.
2. Dysponujemy dostatecznie długim szeregiem czasowym historycznych stóp zwrotu instrumentów tworzących portfel rynkowy (dane rynkowe).
3. Dane rynkowe dają zły estymator średnich stóp zwrotu, co prowadzi do otrzymania wag instrumentów w portfelu, które są dalekie od optymalnych i sprzecznych z intuicją inwestora (duże ujemne i dodatnie wagi, niestabilność).

4. Dane rynkowe dają całkiem niezły estymator macierzy kowariancji. Własności statystyczne tego estymatora można poprawić zwiększając częstotliwość próbkowania (ale tylko do pewnej granicy – problem niezależności różnych próbek).

Jeśli portfel rynkowy jest portfelem efektywnym, to na rynku z instrumentem wolnym od ryzyka jest on rozwiązaniem następującego zagadnienia z funkcją preferencji

$$\min_x (\mu - \mu_0 e)^\top x - \frac{\gamma}{2} x^\top \Sigma x,$$

gdzie γ jest współczynnikiem "rynkowej" awersji do ryzyka.

Rozwiązując ten problem optymalizacyjny otrzymujemy

$$x_M = \frac{1}{\gamma} \Sigma^{-1} (\mu - \mu_0 e) = \frac{1}{\gamma} \Sigma^{-1} \mu_{eq}, \quad (6.1)$$

gdzie wprowadziliśmy nowe oznaczenie μ_{eq} na średnie stopy zwrotu ($\mu - \mu_0 e$) (termin angielski *excess returns* nie ma dobrego polskiego odpowiednika).

Znamy skład portfela rynkowego x_M . Uważamy, że estymator $\hat{\Sigma}$ ma dostatecznie dobre własności, czyli zastępując nim macierz kowariancji Σ we wzorze (6.1) nie popełniamy znaczącego błędu. Możemy więc odwrócić wzór (6.1), aby wyznaczyć średnie stopy zwrotu z inwestycji o portfelu rynkowym

$$\mu_{eq} = \gamma \hat{\Sigma} x_M.$$

WNIOSEK. 6.1 Rozkład równowagowy dany jest wzorem $X \sim \mathcal{N}(\mu_{eq}, \hat{\Sigma})$.

Uwaga. 24 (Jak wyznaczyć współczynnik "rynkowej" awersji do ryzyka?)

Aby znać rozkład równowagowy musimy wyznaczyć równowagowe średnie stopy zwrotu instrumentów w portfelu rynkowym, a to wymaga znajomości rynkowego współczynnika awersji do ryzyka.

Sposób proponowany przez Blacka i Littermana: wybrać taką awersję do ryzyka, aby współczynnik Sharpe'a portfela rynkowego miał wartości między 0.5 a 1.0. Takie współczynniki Sharpe'a są rzeczywiście obserwowane dla dobrze zdywersyfikowanych portfeli rynkowych w długim przedziale czasu.

Z twierdzeń 2.10 oraz 2.15 dostajemy $\gamma \equiv A = \frac{1}{\lambda_2}$, gdzie $\lambda_2 = \frac{\rho - \mu_0}{\delta_0}$, zaś $\delta_0 = \frac{(\rho - \mu_0)^2}{\sigma^2}$, a σ^2 to wariancja portfela optymalnego. Stąd $\gamma = \frac{\rho - \mu_0}{\sigma^2}$. Ponieważ $\frac{\rho - \mu_0}{\sigma}$ jest współczynnikiem Sharpe'a portfela optymalnego, więc $\gamma = \frac{SR}{\sigma}$.

Uwaga. 25 (O portfelu rynkowym) Portfel rynkowy to nie koniecznie ogólnowiatowy indeks. Dla każdego (dostatecznie dużego) fragmentu rynku można wybrać portfel rynkowy dla tego fragmentu rynku. Instrumenty w portfelu rynkowym to nie pojedyncze akcje czy obligacje, ale bardziej specyficzne indeksy. Na przykład dla rynku WGPW portfel rynkowy to WIG a instrumenty to indeksy sektorowe tego rynku. Można też tworzyć własne indeksy sektorowe na podstawie danych rynkowych szczegółowych (patrz strona webowa Kennetha Frencha). Dla rynku obligacji sektory to indeksy obligacji o zapadalnościach 1–3, 3–5, 5–7, 7–10, 10–15 itd.

Rozkład prognoz

Opinie ekspertów (prognozy) sformułowane są w postaci warunków liniowych ograniczających średnie stopy zwrotu niektórych instrumentów. Istnieje macierz P oraz zmienna losowa Q takie, że $Q|X \sim \mathcal{N}(PX, \Omega)$, gdzie Ω odpowiada za niepewność prognoz

$$Q|X = PX + \varepsilon,$$

gdzie ε jest realizacją zmiennej o rozkładzie $\mathcal{N}(0, \Omega)$.

Jakie prognozy są dopuszczalne:

1. Opinie absolutne (ang. *absolute view*): instrument X_1 będzie miał stopę zwrotu q_1 procent.
2. Opinie względne (ang. *relative view*): instrument X_2 będzie miał stopę zwrotu większą o q_2 procent niż instrument X_3 .

Przykład. 6.2 Jak tworzyć macierz P opiszemy na prostym przykładzie wykorzystującym powyższe dane.

Mamy 3 instrumenty X_1, X_2, X_3 oraz 2 opinie: pierwsza opinia mówi, że instrument X_1 będzie miał stopę zwrotu q_1 , druga, że instrument X_2 będzie miał stopę zwrotu większą o q_2 procent niż instrument X_3 . Wtedy

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

Są możliwe bardziej wyrafinowane opinie:

1. Niech instrumenty X_2 i X_3 mają zwroty większe o q procent od instrumentów X_4 i X_5 . Niech k będzie numerem wiersza macierzy P odpowiadającym tej opinii, wtedy $P_k = \begin{pmatrix} 0 & 0.5 & 0.5 & -0.5 & -0.5 \end{pmatrix}$.
2. Opinie są identyczne jak w poprzednim przypadku, ale zamiast równego podziału stosujemy podział proporcjonalny do pewnych informacji o instrumentach (np. ich kapitalizacji). Dla kapitalizacji $X_2 = 0.9, X_3 = 0.1$ a $X_4 = 0.8, X_5 = 0.2$ mamy $P_k = \begin{pmatrix} 0 & 0.9 & 0.1 & -0.8 & -0.2 \end{pmatrix}$.

Reguła: Dla opinii absolutnych suma wag w wierszu macierzy P równa się 1. Dla opinii względnych suma wag w wierszu macierzy P równa się 0.

Wyznaczanie macierzy kowariancji Ω

Black i Litterman proponowali aby macierz kowariancji $\Omega = \tau^{-1}\hat{\Omega}$, gdzie τ – mierzy nasze zaufanie do wyrażanych opinii a $\hat{\Omega}$ – prezentuje rzeczywiste zależności kowariancji opinii.

O postaci macierzy $\hat{\Omega}$ zakładamy, że:

1. Opinie nie są skorelowane – macierz $\hat{\Omega}$ jest diagonalna.

2. Opinie dziedziczą swoje wariancje z wariancji instrumentów – jeśli instrument ma dużą wariancję, to opinia na jego temat też musi mieć dużą wariancję.

Przy tych założeniach otrzymujemy $\hat{\Omega} = \text{diag}(\omega_i^2)$, gdzie $\omega_i^2 = P_i \hat{\Sigma} P_i^\top$, zaś P_i jest i -tym wierszem macierzy P .

Wyznaczanie współczynnika τ

Jest jasne, że małe τ czyni wariancję opinii większą, a duże τ czyni tę wariancję małą. Parametr τ charakteryzuje ogólne zaufanie do wszystkich opinii. Można oczywiście wprowadzić indywidualne współczynniki zaufania do poszczególnych opinii – wtedy nie można specyfikować τ jako mnożnika macierzy $\hat{\Omega}$.

Ale jaka wartość τ jest właściwa? Tu można opierać się wyłącznie na doświadczeniu ekspertów. W literaturze pojawiają się liczby w przedziale 0.01–0.1, Black i Litterman używali $\tau = 0.025$.

6.2 Wzór Blacka-Littermana

TWIERDZENIE. 6.3 (Wzór Blacka-Littermana) Zakładamy, że równowagowe stopy zwrotu instrumentów inwestycyjnych mają rozkład $X \sim \mathcal{N}(\mu_{eq}, \hat{\Sigma})$ a prognozy wyrażone są w postaci liniowych ograniczeń na oczekiwane przyszłe stopy zwrotu $Q|X \sim \mathcal{N}(PX, \Omega)$. Wtedy rozkład a posteriori oczekiwanych stóp zwrotu dany jest wyrażeniem

$$(X|Q = q) \sim \mathcal{N}(\mu_{BL}, \Sigma_{BL}),$$

gdzie

$$\begin{aligned} \mu_{BL} &= (\hat{\Sigma}^{-1} + P^\top \Omega^{-1} P)^{-1} (\hat{\Sigma}^{-1} \mu_{eq} + P^\top \Omega^{-1} q), \\ \Sigma_{BL} &= (\hat{\Sigma}^{-1} + P^\top \Omega^{-1} P)^{-1}. \end{aligned}$$

Uwaga. 26 Jeśli mamy 100% zaufania do opinii ($\tau = +\infty$), to

$$\mu_{BL} = \mu_{eq} + \hat{\Sigma} P^\top (P \hat{\Sigma} P^\top)^{-1} (q - P \mu_{eq}).$$

Uwaga. 27 Black i Litterman w swojej oryginalnej pracy ograniczyli wyznaczanie rozkładu a posteriori do estymacji punktowej rozkładu stóp zwrotu, czyli policzenia wartości oczekiwanej tego rozkładu. Takie rozwiązanie prowadzi do wyznaczenia rozkładu ze średnią μ_{BL} , ale macierzą kowariancji $\hat{\Sigma}$.

Do dowodu twierdzenia 6.3 wykorzystamy poniższy lemat.

LEMAT. 6.4 (wzór Bayesa dla rozkładów) Niech (X, Y) będzie wektorem losowym o gęstości łącznej $f_{XY}(x, y)$ i gęstościach brzegowych $f_X(x)$, $f_Y(y)$. Wtedy

$$f_X(x|Y = y) = \begin{cases} \frac{f_{XY}(x, y)}{f_Y(y)} & \text{dla } f_Y(y) > 0, \\ g(x) & \text{dla } f_Y(y) = 0, \end{cases}$$

gdzie $g(x)$ – dowolna gęstość.

Analogiczny wzór zachodzi dla $f_Y(y|X = x)$. Więc

$$f_X(x|Y = y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{f_Y(y|X = x)f_X(x)}{f_Y(y)}.$$

$f_X(x)$ to rozkład a priori, rozkład $f_Y(y|X = x)$ to rozkład z próby, zaś rozkład $f_X(x|Y = y)$ to rozkład a posteriori.

Dowód. (twierdzenia 6.3) Znając parametry rozkładu a priori równowagowej stopy zwrotu X oraz warunkowego rozkładu $Q|X$, można wyznaczyć, korzystając z twierdzenia Bayesa, rozkład a posteriori warunkowej zmiennej losowej $X|Q = q$.

Zgodnie z wcześniejszymi wynikami rozkład zmiennej $Q|X$ jest dany wzorem

$$Q|X \sim \mathcal{N}(PX, \Omega). \quad (6.2)$$

Z prawa Bayesa, gęstość rozkładu a posteriori oczekiwanej zmiennej losowej dana jest wzorem

$$f_{X|Q=q}(x) = \frac{f_{Q|X}(q)f_X(x)}{\int f_{Q|X}(q)f_X(x)dx}.$$

Całkowanie mianownika daje wyrażenie zależne co najwyżej od q . We wzorze na gęstość zmiennej losowej X składnik ten będzie pełnił rolę stałej, zatem dla uproszczenia rachunków można go pominąć w obliczeniach. Stałą tą będzie można wyznaczyć przyrównując całkę z otrzymanego wzoru do jedynki (ponieważ całka z gęstości rozkładu a posteriori po przestrzeni parametru x wynosi 1).

Ostatecznie warto skupić się na wyznaczeniu postaci wyrażenia $f_{Q|X}(q)f_X(x)$.

$$f_{X|Q=q}(x) \propto f_{Q|X}(q)f_X(x). \quad (6.3)$$

Gęstość n -wymiarowego rozkładu normalnego Z o wektorze średnich m i macierzy kowariancji S dana jest wzorem

$$f_Z(z) = \frac{1}{(2\pi)^{\frac{n}{2}}|S|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(z - m)^\top S^{-1}(z - m)\right).$$

Podstawiając parametry rozkładów we wzorze (6.3) oraz pomijając stałe dostajemy

$$f_{X|Q=q}(x) \propto \exp\left(-\frac{1}{2}\left[(x - \mu_{eq})^\top \hat{\Sigma}^{-1}(x - \mu_{eq}) + (q - Px)^\top \Omega^{-1}(q - Px)\right]\right).$$

Przekształcając wyrażenie w nawiasie kwadratowym otrzymujemy

$$\begin{aligned}
& (x - \mu_{eq})^\top \hat{\Sigma}^{-1} (x - \mu_{eq}) + (q - Px)^\top \Omega^{-1} (q - Px) \\
&= x^\top \hat{\Sigma}^{-1} x - \mu_{eq}^\top \hat{\Sigma}^{-1} x - x^\top \hat{\Sigma}^{-1} \mu_{eq} + \mu_{eq}^\top \hat{\Sigma}^{-1} \mu_{eq} + q^\top \Omega^{-1} q \\
&\quad - q^\top \Omega^{-1} Px - (Px)^\top \Omega^{-1} q + (Px)^\top \Omega^{-1} Px \\
&= x^\top \hat{\Sigma}^{-1} x - 2x^\top \hat{\Sigma}^{-1} \mu_{eq} + \mu_{eq}^\top \hat{\Sigma}^{-1} \mu_{eq} + q^\top \Omega^{-1} q - 2(Px)^\top \Omega^{-1} q \\
&\quad + (Px)^\top \Omega^{-1} Px \\
&= x^\top (\hat{\Sigma}^{-1} + P^\top \Omega^{-1} P) x - 2x^\top (\hat{\Sigma}^{-1} \mu_{eq} + P^\top \Omega^{-1} q) \\
&\quad + \mu_{eq}^\top \hat{\Sigma}^{-1} \mu_{eq} + q^\top \Omega^{-1} q.
\end{aligned}$$

Oznaczając

$$\begin{aligned}
H &:= \hat{\Sigma}^{-1} + P^\top \Omega^{-1} P, \\
C &:= \hat{\Sigma}^{-1} \mu_{eq} + P^\top \Omega^{-1} q, \\
A &:= \mu_{eq}^\top \hat{\Sigma}^{-1} \mu_{eq} + q^\top \Omega^{-1} q,
\end{aligned}$$

gdzie H jest macierzą $d \times d$, C jest wektorem $d \times 1$, a A jest liczbą, możemy ostatnie wyrażenie zapisać w postaci

$$\begin{aligned}
& x^\top (\hat{\Sigma}^{-1} + P^\top \Omega^{-1} P) x - 2x^\top (\hat{\Sigma}^{-1} \mu_{eq} + P^\top \Omega^{-1} q) + \mu_{eq}^\top \hat{\Sigma}^{-1} \mu_{eq} + q^\top \Omega^{-1} q \\
&= x^\top H x - 2x^\top C + A.
\end{aligned}$$

Otrzymane wyrażenie jest postaci "trójmianu kwadratowego" wektora x . Dalsze rachunki prowadzą do otrzymania wyrażenia analogicznego do wyrażenia występującego we wzorze na gęstość rozkładu normalnego

$$(z - m)^\top S^{-1} (z - m) = z^\top S^{-1} z - 2z^\top S^{-1} m + m^\top S^{-1} m.$$

Z postaci otrzymanego wzoru widać, że

$$S^{-1} = H, \quad C = Hm.$$

Zatem $m = H^{-1}C$. Ponieważ

$$(x - H^{-1}C)^\top H (x - H^{-1}C) = x^\top H x - 2x^\top C + C^\top H^{-1}C,$$

więc

$$x^\top H x - 2x^\top C + A = (x - H^{-1}C)^\top H (x - H^{-1}C) + A - C^\top H^{-1}C.$$

Zauważmy, że wyrażenie $A - C^\top H^{-1}C$ jest niezależne od x , zatem $\exp(A - C^\top H^{-1}C)$ jest stałą przez którą mnożymy gęstość *a posteriori*. Dla uproszczenia rachunków składnik ten może zatem być pominięty

$$f_{X|Q=q}(x) \propto \exp\left(-\frac{1}{2}(x - H^{-1}C)^\top H (x - H^{-1}C)\right).$$

Otrzymany rozkład *a posteriori* jest zatem rozkładem normalnym o średniej $H^{-1}C$ oraz macierzy kowariancji H^{-1}

$$(X|Q = q) \sim \mathcal{N}(H^{-1}C, H^{-1}) \quad (6.4)$$

czyli:

$$(X|Q = q) \sim \mathcal{N}\left(\left((\hat{\Sigma}^{-1}) + P^{\top}\Omega^{-1}P\right)^{-1}\left((\hat{\Sigma})^{-1}\mu_{eq} + P^{\top}\Omega^{-1}q\right), \left((\hat{\Sigma})^{-1} + P^{\top}\Omega^{-1}P\right)^{-1}\right).$$

■

Mając wyznaczone momenty rozkładu *a posteriori* znajdujemy optymalny portfel inwestycyjny rozwiązując zadanie optymalizacyjne

$$\begin{cases} \min_x \frac{1}{2}x^{\top}\Sigma_{BL}x, \\ \mu_{BL}^{\top}x = \rho, \\ e^{\top}x = 1, \end{cases}$$

lub zadanie z funkcją preferencji

$$\begin{cases} \max_x \mu_{BL}^{\top}x - \frac{A}{2}x^{\top}\Sigma_{BL}x, \\ e^{\top}x = 1, \end{cases}$$

gdzie A jest współczynnikiem awersji do ryzyka inwestora.

Jak to zostało zaznaczone w uwadze 27 w oryginalnym modelu Blacka-Littermana nie używano macierzy kowariancji Σ_{BL} , ale macierz estymowaną z danych rynkowych $\hat{\Sigma}$.

Model Blacka-Littermana ma już wieloletnią historię. Można więc zasadnie pytać, jak model sprawdza się w praktyce rynkowej. Odpowiedź na to pytanie nie jest łatwa. Model został opracowany w firmie Goldman Sachs i był przez wiele lat używany do taktycznego zarządzania portfelami tej firmy. Także wiele innych firm używało i używa tego modelu. Jednak fakt jego intensywnego wykorzystywania w praktyce rynkowej sprawia, że ukazujące się publikacje sygnowane przez pracowników tych firm zajmują się raczej opisywaniem zalet modelu niż porównywaniem wyników otrzymywanych z jego pomocą z wynikami, jakie można byłoby osiągnąć stosując inne metody zarządzania portfelem. Oczywiście fakt stosowania modelu Blacka-Littermana przemawia na jego korzyść. Firmy zarządzające aktywami klientów nie stosowałyby metodologii, której wyniki są gorsze niż innych dostępnych metod zarządzania portfelem.

Ta ogólna argumentacja przemawiająca za wyższością metody Blacka-Littermana nad innymi metodami zarządzania, nie zwalnia nas jednak z próby obiektywnego ilościowego porównania. Liczba dostępnych w literaturze wyników na

ten temat jest bardzo ograniczona. Interesujące porównanie wyników uzyskiwanych tradycyjną metodą Markowitza z wynikami otrzymywanymi metodą Blacka-Littermana znajduje się w pracy Mertensa i Zimmermanna. Autorzy ci porównali portfele Markowitza otrzymywane przy estymacji parametrów z 12-letniego okna czasowego z wynikami modelu Blacka-Littermana. Jako metodę porównania wybrano współczynnik Sharpe'a odpowiednich portfeli. W przypadku inwestycji na rynku akcji model Blacka-Littermana daje tylko nieznacznie lepsze wyniki. Współczynnik Sharpe'a wzrasta od wartości 0.53 dla portfeli Markowitza do 0.55 dla portfeli Blacka-Littermana. W przypadku inwestycji na rynku obligacji różnica wyników jest już znacznie bardziej wyraźna: dla portfela Markowitza $SR=1.07$, a dla portfela Blacka-Littermana $SR=1.15$. Różnice w wynikach portfela, choć wyraźnie przemawiające na korzyść modelu Blacka-Littermana, nie są więc bardzo duże. To co jest uderzające to różnice w stabilności portfeli. Udziały poszczególnych instrumentów w portfelu Markowitza potrafiły się zmieniać w czasie procedur dostosowywania portfela o kilka tysięcy procent (patrz uwaga 23 na końcu rozdziału 5). Tymczasem wahania portfeli Blacka-Littermana wynosiły najwyżej kilka punktów procentowych (zmiany względne nie przekraczały 50%). Ta stabilność składu portfeli Blacka-Littermana stanowi największą zaletę modelu.

Rozdział 7

Portfele minimalizujące ryzyko

W rozdziale tym zajmiemy się poszukiwaniem portfeli, które minimalizują ryzyko inwestora. W zależności od tego jak mierzymy ryzyko będziemy otrzymywali różne portfele. Najprostsze kryteria tego typu pojawiły się równoległe z modelem Markowitza. Okazuje się przy tym, że dla normalnych rozkładów stóp zwrotu instrumentów ryzykownych te kryteria optymalności nie prowadzą do nowych rozwiązań. Ich przewaga ujawnia się przy zarządzaniu inwestycjami na rynku, na którym stopy zwrotu nie mają rozkładu normalnego.

7.1 Kryteria bezpieczeństwa

Kryteriami bezpieczeństwa (ang. *safety first*) nazywane są strategie inwestycyjne, w których inwestor zwraca uwagę na możliwe małe prawdopodobieństwo poniesienia strat.

W przypadkach omówionym niżej przyjmujemy, że znana jest granica efektywna. Kryteria bezpieczeństwa wskazują na obszar na płaszczyźnie (ρ, σ) będący obrazem portfeli dopuszczalnych, które spełniają dane kryterium. W połączeniu z warunkami granicy efektywnej daje to portfel efektywny spełniający jeden z warunków bezpieczeństwa. Na początku zajmiemy się kryteriami Roya, Telsera i Kataoki na rynku bez instrumentu wolnego od ryzyka.

Kryterium Roya

DEFINICJA. 7.1 Niech zmienna losowa $R(x)$ będzie stopą zwrotu portfela x a ρ_{\min} minimalną stopą zwrotu z portfela, jaką pragnie uzyskać inwestor (albo została ustalona w konkretnym kryterium inwestycyjnym). Kryterium Roya polega na znalezieniu portfela x , który spełnia warunek

$$\min_x \mathbb{P}(R(x) \leq \rho_{\min}).$$

Zacniemy od normalnego rozkładu stóp zwrotu. Ponieważ wektor stóp zwrotu R ma rozkład normalny (wielowymiarowy), to stopa zwrotu portfela $R(x) = R^\top x$

ma jednowymiarowy rozkład normalny. Warunek

$$\min_x \mathbb{P}(R(x) \leq \rho_{\min})$$

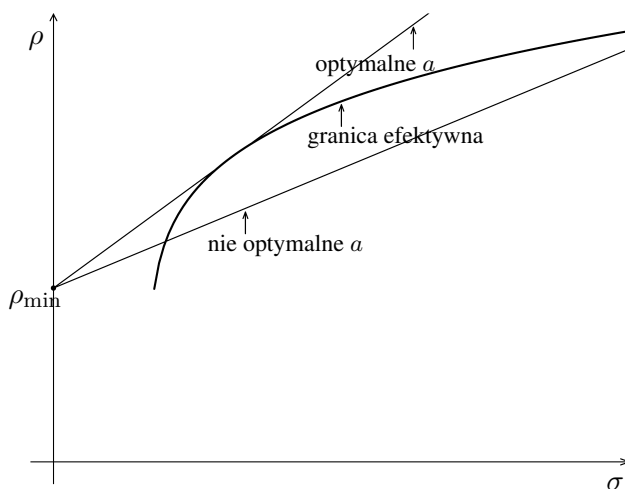
jest równoważny warunkowi

$$\min_x \mathbb{P}\left(\frac{R(x) - \rho(x)}{\sigma(x)} \leq \frac{\rho_{\min} - \rho(x)}{\sigma(x)}\right),$$

gdzie $\rho(x) = \mathbb{E}(R(x))$, $\sigma(x) = \sigma(R(x))$.

Zmienna $\frac{R(x) - \rho(x)}{\sigma(x)}$ ma rozkład $\mathcal{N}(0, 1)$. Aby zminimalizować analizowane prawdopodobieństwo, należy więc zminimalizować $\frac{\rho_{\min} - \rho(x)}{\sigma(x)}$ lub zmaksymalizować $\frac{\rho(x) - \rho_{\min}}{\sigma(x)}$.

Rozpatrzmy to zadanie na płaszczyźnie (ρ, σ) . Niech $a = \frac{\rho(x) - \rho_{\min}}{\sigma(x)}$, czyli $\rho(x) = a\sigma(x) + \rho_{\min}$. Geometrycznie oznacza to prostą na płaszczyźnie (ρ, σ) o nachyleniu a przechodzącą przez punkt $(\rho_{\min}, 0)$. Rozwiązanie optymalne odpowiada maksymalnej wartości a , przy której na tej prostej znajdują się portfele efektywne.



Rysunek 7.1: Znajdowanie portfela optymalnego wg kryterium Roya.

W przypadku dowolnego rozkładu stóp zwrotu portfela, w przeciwieństwie do normalnych rozkładów stóp zwrotu, kiedy analizę mogliśmy rozpocząć od wielowymiarowego rozkładu stóp zwrotu instrumentów ryzykownych, musimy założyć, że znamy rozkład stóp zwrotu portfela (rozkład 1-wymiarowy). Jest to jednak tylko ograniczenie rachunkowe, bo jeśli znamy wielowymiarowy rozkład stóp zwrotu instrumentów, to możemy policzyć rozkład stóp zwrotu portfela.

Do znalezienia portfela spełniającego kryterium Roya wykorzystamy nierówność Czebyszewa

$$\mathbb{P}\left(\left|\frac{R - \mathbb{E}(R)}{\sigma(R)}\right| \geq k\right) \leq \frac{1}{k^2},$$

która w tej wersji jest prawdziwa dla dowolnej 1-wymiarowej zmiennej losowej mającej skończoną średnią i wariancję.

Ponieważ w kryterium Roya interesuje nas jedynie przypadek górnego ograniczenia na $\frac{R - \mathbb{E}(R)}{\sigma(R)}$, więc z nierówności Czebyszewa wybieramy tylko warunek na górne ograniczenie

$$\mathbb{P}\left(\frac{R(x) - \rho(x)}{\sigma(x)} \leq -k\right) \leq \frac{1}{k^2}.$$

Z poprzedniego rozumowania wiemy, że dokonując podstawienia

$$-k = \frac{\rho_{\min} - \rho(x)}{\sigma(x)}$$

kryterium Roya sprowadzamy do warunku

$$\mathbb{P}\left(\frac{R(x) - \rho(x)}{\sigma(x)} \leq \frac{\rho_{\min} - \rho(x)}{\sigma(x)}\right) \leq \frac{1}{k^2}.$$

Kryterium Roya prowadzi więc do maksymalizacji k , tj. do identycznego rozwiązania jak dla przypadku rozkładu normalnego (stałą a zastępuje stała k)

$$\rho(x) = k\sigma(x) + \rho_{\min}.$$

Kryterium Telsera

DEFINICJA. 7.2 Niech zmienna losowa $R(x)$ będzie stopą zwrotu portfela x , ρ_{\min} minimalną stopą zwrotu z portfela, jaką pragnie uzyskać inwestor a ε prawdopodobieństwem, z jakim tej założonej stopy zwrotu inwestor może nie osiągnąć (poziomem ufności). Portfel x spełnia kryterium Telsera, jeśli jest rozwiązaniem następującego problemu

$$\begin{cases} \max_x \mathbb{E}(R(x)), \\ \mathbb{P}(R(x) \leq \rho_{\min}) \leq \varepsilon, \end{cases}$$

dla podanego ρ_{\min} i ε .

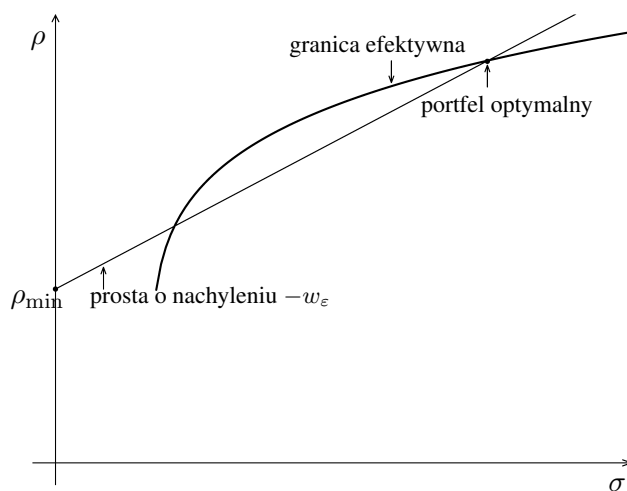
Dla normalnego rozkładu stóp zwrotu, podobnie jak w przypadku kryterium Roya, warunek $\mathbb{P}(R(x) \leq \rho_{\min})$ zamieniamy na

$$\mathbb{P}\left(\frac{R(x) - \rho(x)}{\sigma(x)} \leq \frac{\rho_{\min} - \rho(x)}{\sigma(x)}\right).$$

Ponieważ $\frac{R(x) - \rho(x)}{\sigma(x)} \sim \mathcal{N}(0, 1)$, więc zadanie sprowadza się do $\frac{\rho_{\min} - \rho(x)}{\sigma(x)} \leq w_\varepsilon$, gdzie w_ε jest kwantylem rozkładu normalnego na poziomie ε . Otrzymujemy wtedy rozwiązanie

$$\rho(x) \geq -w_\varepsilon \sigma(x) + \rho_{\min}.$$

Oznacza to, że portfel spełniający kryterium Telsera, to portfel którego obraz leży na przecięciu prostej o nachyleniu $-w_\varepsilon$ przechodzącej przez punkt $(\rho_{\min}, 0)$ z granicą efektywną, co odpowiada maksymalnej możliwej wartości $\rho(x)$ na tej prostej.



Rysunek 7.2: Portfel optymalny wg kryterium Telsera.

W przypadku dowolnego rozkładu stóp zwrotu nierówność Czebyszewa daje

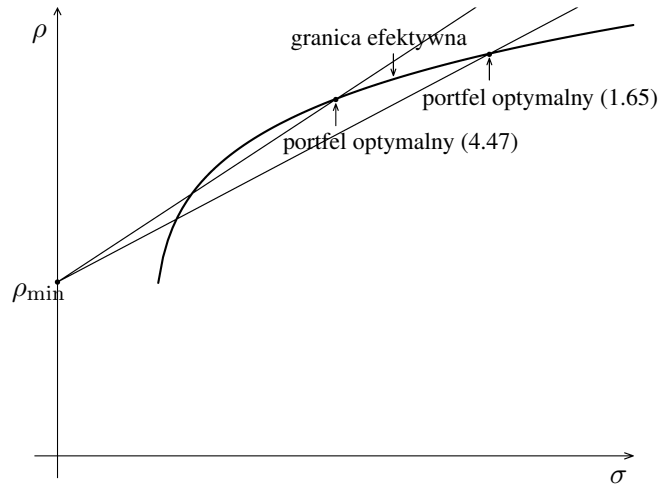
$$\mathbb{P}\left(\frac{R(x) - \rho(x)}{\sigma(x)} \leq \frac{\rho_{\min} - \rho(x)}{\sigma(x)}\right) \leq \frac{1}{k^2},$$

gdzie $k = \frac{\rho(x) - \rho_{\min}}{\sigma(x)}$.

Ponieważ $\frac{1}{k^2} = \varepsilon$ to $k = \frac{1}{\sqrt{\varepsilon}}$. Stąd

$$\rho(x) \geq \frac{1}{\sqrt{\varepsilon}} \sigma(x) + \rho_{\min}.$$

Czyli "prawie" tak samo jak dla rozkładu normalnego. "Prawie" bo oczywiście $-w_\varepsilon \neq \frac{1}{\sqrt{\varepsilon}}$. Np. $\varepsilon = 0.05 \Rightarrow -w_\varepsilon = 1.65$, a $\frac{1}{\sqrt{\varepsilon}} = 4.47$.



Rysunek 7.3: Portfele optymalny wg kryterium Telsera dla rozkładu normalnego ($-w_\varepsilon = 1.65$) i dowolnego rozkładu ($\frac{1}{\sqrt{\varepsilon}} = 4.47$).

Kryterium Kataoki

DEFINICJA. 7.3 Niech zmienna losowa $R(x)$ będzie stopą zwrotu portfela x , ρ_{\min} minimalną stopą zwrotu z portfela, jaką pragnie uzyskać inwestor a ε prawdopodobieństwem, z jakim tej założonej stopy zwrotu inwestor może nie osiągnąć (poziomem ufności). Portfel x spełnia kryterium Kataoki, jeśli jest rozwiązaniem następującego problemu

$$\begin{cases} \max \rho_{\min}, \\ \mathbb{P}(R(x) \leq \rho_{\min}) \leq \varepsilon. \end{cases}$$

dla podanego ρ_{\min} i ε .

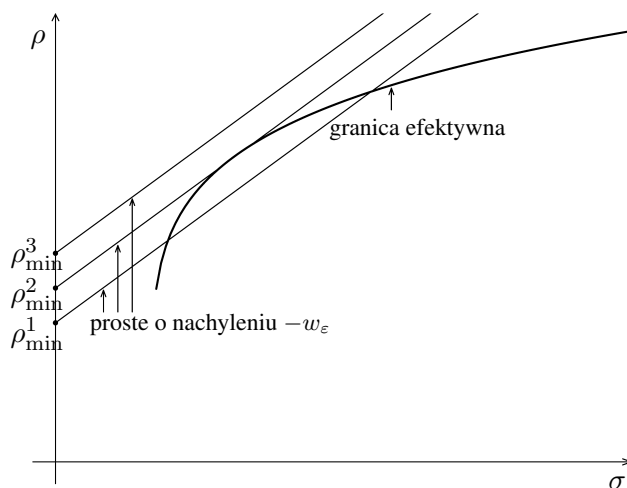
Dla normalnego rozkładu stóp zwrotu zadanie redukuje się do

$$\begin{cases} \max \rho_{\min}, \\ \frac{\rho(x) - \rho_{\min}}{\sigma(x)} \geq -w_\varepsilon, \end{cases}$$

czyli $\rho(x) \geq -w_\varepsilon \sigma(x) + \rho_{\min}$.

Oznacza to, że portfel spełniający kryterium Kataoki, to portfel którego obraz leży na granicy efektywnej w punkcie styczności z prostą $\rho(x) + w_\varepsilon \sigma(x) = \rho_{\min}$ i odpowiada to największej wartości ρ_{\min} , przy której prosta o nachyleniu $-w_\varepsilon$ przechodząca przez punkt $(\rho_{\min}, 0)$ jest styczna do granicy efektywnej.

Dla dowolnego rozkładu stóp zwrotu, korzystając z wyników dla kryterium Telsera, zauważamy, że wystarczy zastąpić stałą $(-w_\varepsilon)$ przez stałą $\frac{1}{\sqrt{\varepsilon}}$.



Rysunek 7.4: Portfel optymalny wg kryterium Kataoki. Optymalny jest portfel odpowiadający ρ_{\min}^2 .

Kryteria bezpieczeństwa w obecności instrumentu wolnego od ryzyka

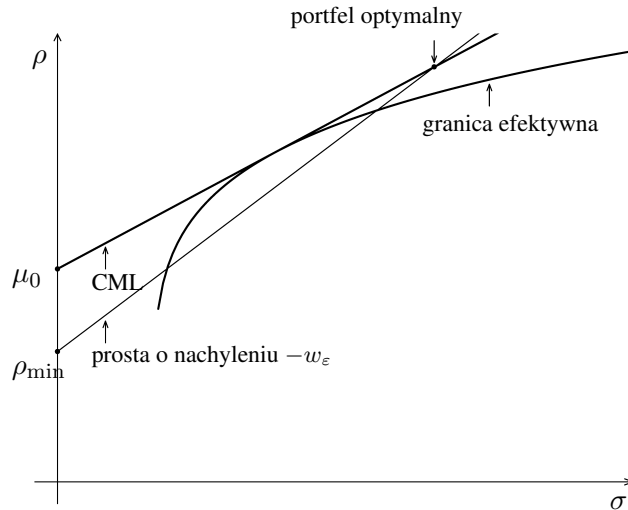
Obecnie zajmiemy się analizą omówionych wcześniej kryteriów bezpieczeństwa na rynku, na którym jest instrument wolny od ryzyka. W analizie tej ograniczymy się do przypadku, gdy stopy zwrotu mają rozkłady normalne. Przypadek dowolnych rozkładów stóp zwrotu różni się jedynie występującymi stałymi.

Rozpocznijmy do analizy kryterium Telsera. Zauważmy, że część probabilistyczna rozwiązania jest identycznie jak przy braku instrumentu wolnego od ryzyka, czyli stopa zwrotu portfela musi spełniać nierówność

$$\rho(x) \geq -w_\varepsilon \sigma(x) + \rho_{\min}. \quad (7.1)$$

Portfel optymalny w sensie kryterium Telsera, to portfel efektywny (leżący na granicy efektywnej), którego stopa zwrotu spełnia jednocześnie nierówność (7.1). W obecności instrumentu wolnego od ryzyka portfele efektywne tworzą półprosta (*capital market line*), więc portfel optymalny w sensie Telsera, to portfel leżący na przecięciu CML z prostą opisaną równaniem (7.1). Rozwiązanie istnieje jeśli $\rho_{\min} < \mu_0$ i nachylenie $-w_\varepsilon$ jest większe niż nachylenie CML lub jeśli $\rho_{\min} > \mu_0$ i nachylenie $-w_\varepsilon$ jest mniejsze niż nachylenie CML.

W przypadku kryterium Roya problem sprowadza się do znalezienia maksymalnego a , takiego że prosta $\rho(x) = a\sigma(x) + \rho_{\min}$ ma punkt wspólny z granicą efektywną, czyli CML. Jeśli $\rho_{\min} < \mu_0$, to $a = +\infty$ a portfelem optymalnym jest portfel złożony wyłącznie z inwestycji w instrument wolny od ryzyka. Jeśli



Rysunek 7.5: Portfel optymalny wg kryterium Telsera na rynku z instrumentem wolnym od ryzyka.

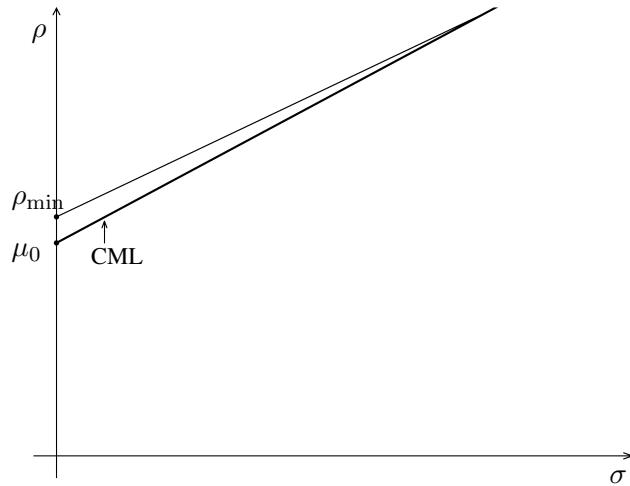
$\rho_{\min} > \mu_0$, to optymalne a odpowiada nachyleniu prostej CML, ale wtedy portfel optymalny ma nieograniczone σ , czyli nieograniczone ryzyko.

W przypadku kryterium Kataoki kluczowa jest informacja o nachyleniu prostej $\rho(x) = -w_\varepsilon \sigma(x) + \rho_{\min}$. Jeśli nachylenie tej prostej $-w_\varepsilon$ jest większe niż CML, to rozwiązaniem optymalnym jest $\rho_{\min} = \mu_0$. Jeśli nachylenie tej prostej $-w_\varepsilon$ jest mniejsze niż CML, to rozwiązaniem optymalnym jest $\rho_{\min} = +\infty$.

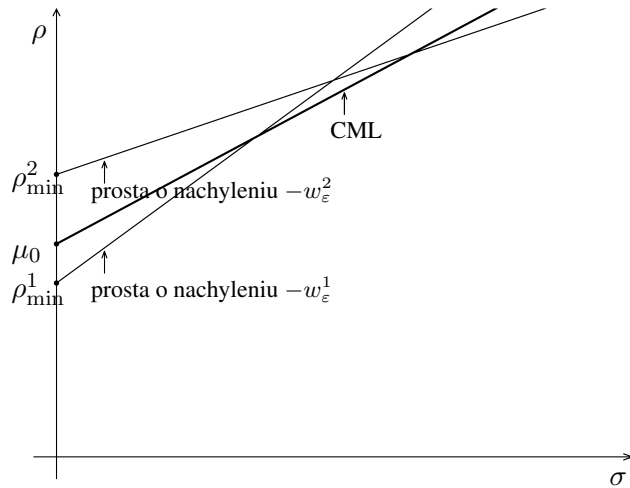
7.2 Miary ryzyka

Dotychczas zajmowaliśmy się minimalizacją strat inwestycji ograniczając lub minimalizując prawdopodobieństwo wystąpienia strat. Obecnie zajmiemy się problemem minimalizacji strat w sposób systemowy wprowadzając funkcję, która będzie mierzyła wielkość poniesionej straty.

Inwestycję X będziemy traktowali jako zmienną losową na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$. Interpretujemy X jako wartość końcową zainwestowanego kapitału. Przyjmujemy, że X należy do pewnego zbioru inwestycji dopuszczalnych \mathcal{X} , który jest podzbiorem zbioru zmiennych losowych na przestrzeni $(\Omega, \mathcal{F}, \mathbb{P})$. Oznaczmy przez $\varrho(X)$ funkcję, która mierzy ryzyko inwestycji X . Zadanie poszu-



Rysunek 7.6: Portfel optymalny wg kryterium Roya na rynku z instrumentem wolnym od ryzyka.



Rysunek 7.7: Portfel optymalny wg kryterium Kataoki na rynku z instrumentem wolnym od ryzyka. $\rho_{\text{optim}}^1 = \mu_0$ dla prostej o nachyleniu $-w_{\varepsilon}^1$, $\rho_{\text{optim}}^2 = +\infty$ dla prostej o nachyleniu $-w_{\varepsilon}^2$.

kiwania inwestycji minimalizującej ryzyko można wtedy sformułować następująco

$$\begin{cases} \varrho(X) \rightarrow \min, \\ \mathbb{E}(X) \geq \Delta, \\ X \in \mathcal{X}. \end{cases} \quad (7.2)$$

Aby rozwiązać problem (7.2) musimy zrobić pewne założenia na temat funkcji ϱ .

DEFINICJA. 7.4 *Odwzorowanie $\varrho : \mathcal{X} \rightarrow \mathbb{R}$ nazywamy miarą ryzyka, jeśli:*

1. $\varrho(X + c) = \varrho(X) - c$, $c \in \mathbb{R}$ (*translacyjna niezmienniczość*),
2. $X \geq Y \Rightarrow \varrho(X) \leq \varrho(Y)$, dla każdego $X, Y \in \mathcal{X}$ (*monotoniczność*),

gdzie nierówność $X \geq Y$ oznacza, że $X(\omega) \geq Y(\omega)$ dla wszystkich $\omega \in \Omega$.

Zawarte w powyższej definicji warunki mają naturalną interpretacją ekonomiczną: translacyjna niezmienniczość oznacza, że jeśli do inwestycji X dołożymy kapitał gotówkowy c to ryzyko inwestycji zmniejszy się o wielkość dołożonego kapitału; podobnie jeśli mamy dwie inwestycje, z których pierwsza dla każdego zdarzenia losowego ma większą wartość niż druga, to jej ryzyko jest mniejsze (monotoniczność).

Zdefiniowana wyżej miara ryzyka nie ma jednak pewnych własności, które są pożądane przez rynek finansowy. Lepsze własności ma klasa miar ryzyka, które zostały nazwane koherentnymi miarami ryzyka.

DEFINICJA. 7.5 *Odwzorowanie $\varrho : \mathcal{X} \rightarrow \mathbb{R}$ nazywamy koherentną miarą ryzyka, jeśli:*

1. $\varrho(X + c) = \varrho(X) - c$, $c \in \mathbb{R}$ (*translacyjna niezmienniczość*).
2. $X \geq 0 \Rightarrow \varrho(X) \leq 0$.
3. $\varrho(\lambda X) = \lambda \varrho(X)$, $\lambda \geq 0$ (*dodatnia jednorodność*).
4. $\varrho(X + Y) \leq \varrho(X) + \varrho(Y)$ (*podaddytywność*).

TWIERDZENIE. 7.6 *Jeśli $\varrho : \mathcal{X} \rightarrow \mathbb{R}$ jest koherentną miarą ryzyka, to*

- (a) $X \geq Y \Rightarrow \varrho(X) \leq \varrho(Y)$,
- (b) $\varrho(\lambda X + (1 - \lambda)Y) \leq \lambda \varrho(X) + (1 - \lambda)\varrho(Y)$, $\lambda \in [0, 1]$,
- (c) $\varrho(c) = -c$, $\forall c \in \mathbb{R}$,
- (d) $a \leq X \leq b \Rightarrow -b \leq \varrho(X) \leq -a$.
- (e) $\varrho(X + \varrho(X)) = 0$.

Dowód. Własność (a) wynika z nierówności

$$\varrho(X) = \varrho(X - Y + Y) \leq \varrho(X - Y) + \varrho(Y) \leq \varrho(Y),$$

ponieważ $\varrho(X - Y) \leq 0$ z warunku 2. definicji 7.5.

Własność (b) wynika z warunków 3 i 4 definicji 7.5 oraz nierówności

$$\varrho(\lambda X + (1 - \lambda)Y) \leq \varrho(\lambda X) + \varrho((1 - \lambda)Y) = \lambda\varrho(X) + (1 - \lambda)\varrho(Y).$$

Własność (c) wynika z obserwacji $\varrho(0) = 0$, która jest konsekwencją warunku 3. definicji 7.5. Wtedy $\varrho(0 + c) = \varrho(0) - c = -c$.

Własność (d) wynika z (a) a własność (e) z warunku 1. definicji 7.5. ■

Przykład. 7.7 Jeśli $(\Omega, \mathcal{F}, \mathbb{P})$ jest przestrzenią probabilistyczną, na której zdefiniowane są zmienne losowe z \mathcal{X} , to dla dowolnego zbioru \mathcal{Q} miar probabilistycznych, które są absolutnie ciągłe względem \mathbb{P} ,

$$\varrho_{\mathcal{Q}}(X) = \sup\{\mathbb{E}_{\mathbb{Q}}(-X) : \mathbb{Q} \in \mathcal{Q}\}$$

jest koherentną miarą ryzyka (przy pewnych słabych założeniach każda miara koherentna ma takie przedstawienie, ale dowód tego faktu wykracza poza materiał tego wykładu).

Uwaga. 28 Jeśli $\mathcal{X} = L^2(\Omega)$, to wariancja X nie jest koherentną miarą ryzyka.

Uwaga. 29 W poprzednich rozdziałach opisywaliśmy inwestycje ryzykowne przez podawanie ich stopy zwrotu, a nie wartości inwestycji, jak to zrobiliśmy w definicjach 7.4 i 7.5. Okazuje się jednak, że jeśli miara ryzyka jest translacyjnie niezmiennicza i dodatnio jednorodna, to przy pewnych oczywistych założeniach problem optymalizacyjny (7.2) można sprowadzić do problemu dla stopy zwrotu z inwestycji.

Załóżmy w tym celu, że wartość inwestycji jest liniową funkcją stóp zwrotu instrumentów. Niech V_0 będzie początkową wartością zainwestowanego kapitału. Kapitał ten inwestujemy w instrument wolny od ryzyka o stopie zwrotu μ_0 oraz instrumenty ryzykowne o stopach zwrotu R będących zmiennymi losowymi, takimi że $\mathbb{E}(R) = \mu$. Niech (x, x_0) będzie składem portfela inwestycyjnego (x to inwestycja w instrumenty ryzykowne, x_0 – w instrument wolny od ryzyka). Wówczas

$$\begin{aligned} V_T &= V_0 x_0 (1 + \mu_0) + V_0 x^\top (e + R) = (V_0 x_0 + V_0 x^\top e) + V_0 x_0 \mu_0 + V_0 x^\top R \\ &= V_0 + V_0 x_0 \mu_0 + V_0 x^\top R. \end{aligned}$$

Korzystając z translacyjnej niezmienniczości i dodatniej jednorodności ϱ dostajemy

$$\begin{aligned} \varrho(V_T) &= \varrho(V_0 x^\top R) - V_0 - V_0 x_0 \mu_0 = V_0 (\varrho(x^\top R) - 1 - x_0 \mu_0) \\ &= V_0 (\varrho(x^\top (R - \mu_0 e)) - x^\top e \mu_0 - 1 - x_0 \mu_0) \\ &= V_0 (\varrho(x^\top (R - \mu_0 e)) - 1 - \mu_0). \end{aligned}$$

Podobnie

$$\mathbb{E}(V_T) = V_0((1 + \mu_0) + x^\top \mathbb{E}(R - \mu_0 e)) = V_0((1 + \mu_0) + x^\top (\mu - \mu_0 e)).$$

Oznacza to, że problem optymalizacyjny (7.2), w którym X jest zastąpiony przez $V_T^{(x)}$ dla podkreślenia zależności od składu portfela inwestycyjnego x ,

$$\begin{cases} \varrho(V_T^{(x)}) \rightarrow \min, \\ \mathbb{E}(V_T^{(x)}) \geq \Delta, \\ x \in \mathbb{R}^d \end{cases} \quad (7.3)$$

jest równoważny problemowi dla stopy zwrotu z inwestycji

$$\begin{cases} \varrho(x^\top (R - \mu_0 e)) \rightarrow \min, \\ x^\top (\mu - \mu_0 e) \geq \frac{\Delta}{V_0} - (1 + \mu_0) \equiv \rho. \end{cases} \quad (7.4)$$

Wartość narażona na ryzyko (Value at Risk)

DEFINICJA. 7.8 Niech $X \in \mathcal{X}$ będzie inwestycją ryzykowną. Ustalmy $\alpha \in (0, 1)$. Wartość narażoną na ryzyko (ang. Value at Risk) dla inwestycji X na poziomie α definiujemy wzorem

$$\text{VaR}_\alpha(X) = -\sup\{x \in \mathbb{R} : \mathbb{P}(X < x) \leq \alpha\}$$

Uwaga. 30 VaR można zdefiniować równoważnym wzorem

$$\text{VaR}_\alpha = \inf\{m \in \mathbb{R} : \mathbb{P}(X + m < 0) \leq \alpha\}.$$

W definicji tej wielkość m interpretujemy jako dodatkowy kapitał, jaki należy dodać do inwestycji X , aby prawdopodobieństwo poniesienia straty było mniejsze niż α .

TWIERDZENIE. 7.9 Niech X będzie inwestycją ze zbioru \mathcal{X} . Wtedy:

1. Miara ryzyka $\text{VaR}_\alpha(X)$ spełnia warunki 1–3 definicji 7.5 koherentnej miary ryzyka.
2. VaR_α nie jest podaddytywny.
3. Jeśli zmienne (X, Y) mają łączny rozkład normalny, to

$$\text{VaR}_\alpha(X + Y) \leq \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y), \quad \forall \alpha \leq \frac{1}{2}.$$

Dowód. Dowód punktu 1. wynika bezpośrednio z definicji $\text{VaR}_\alpha(X)$.

Aby udowodnić punkt 2. skonstruujemy dwie inwestycje X i Y , dla których nie jest spełniony warunek podaddytywności. Niech będzie dana dyskretna przestrzeń probabilistyczna $\Omega = \{\omega_1, \omega_2\}$, $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \frac{1}{2}$. Inwestycje X i Y definiujemy na tej przestrzeni przy pomocy następującej tabeli

	ω_1	ω_2
X	-1	-2
Y	-2	-1
$X + Y$	-3	-3

Jeśli rozważymy VaR na poziomie $\alpha = 60\%$, to $\text{VaR}_{60\%}(X) = 1$, $\text{VaR}_{60\%}(Y) = 1$ a $\text{VaR}_{60\%}(X + Y) = 3$. Ponieważ $3 > 1 + 1$ więc nie zachodzi warunek podaddytywności.

Aby udowodnić punkt 3. zauważmy, że jeśli $X \sim \mathcal{N}(\mu, \sigma^2)$, to $\text{VaR}_\alpha(X) = -\Phi^{-1}(\alpha)\sigma - \mu$, gdzie Φ jest dystrybuantą standardowego rozkładu normalnego. Rzeczywiście, korzystając z punktu 1. otrzymujemy

$$\text{VaR}_\alpha\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma}\text{VaR}_\alpha(X - \mu) = \frac{1}{\sigma}(\text{VaR}_\alpha(X) + \mu).$$

Ponieważ $\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$, więc

$$\text{VaR}_\alpha\left(\frac{X - \mu}{\sigma}\right) = -\Phi^{-1}(\alpha),$$

co daje poszukiwany wzór.

Rozważmy teraz parę inwestycji o łącznym rozkładzie normalnym

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right), \quad \sigma_1, \sigma_2 > 0.$$

Korzystając z wzoru na VaR dla inwestycji o rozkładzie normalnym dostajemy $\text{VaR}_\alpha(X) = -\mathcal{A}\sigma_1 - \mu_1$, $\text{VaR}_\alpha(Y) = -\mathcal{A}\sigma_2 - \mu_2$, gdzie $\mathcal{A} = \Phi^{-1}(\alpha)$. Ponieważ

$$X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\sigma_{12}),$$

więc

$$\text{VaR}_\alpha(X + Y) = -\mathcal{A}\sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}} - \mu_1 - \mu_2.$$

Wynika z tych rachunków, że

$$\text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y) - \text{VaR}_\alpha(X + Y) = -\mathcal{A}\left(\sigma_1 + \sigma_2 - \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}}\right).$$

Ponieważ $-\mathcal{A} \geq 0$ dla $\alpha \leq \frac{1}{2}$, wystarczy sprawdzić, że

$$\sigma_1 + \sigma_2 - \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}} \geq 0.$$

Jeśli nierówność ta jest prawdziwa, to przenosząc pierwiastek na drugą stronę nierówności dostajemy

$$\sigma_1 + \sigma_2 \geq \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}}.$$

Po podniesieniu do kwadratu obu stron nierówności mamy

$$\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2 \geq \sigma_1^2 + \sigma_2^2 + 2\sigma_{12},$$

co daje po skróceniu

$$\sigma_1\sigma_2 \geq \sigma_{12}.$$

Ostatnia nierówność jest oczywiście prawdziwa, co wynika z nieujemnej określoności macierzy kowariancji. ■

Uwaga. 31 *Pomimo braku własności podaddytywności VaR jest stosowany w bankowości. Takie postępowanie jest dopuszczalne przy ocenie bieżącego ryzyka portfela akcji (rozkłady stóp zwrotu akcji giełdowych są dość bliskie rozkładowi normalnemu). Jednak stosowanie VaR jest niedopuszczalne dla rynku kredytowego, gdzie występują rozkłady, dla których VaR nie jest podaddytywny.*

Istotnym mankamentem VaR jest brak informacji, jaka jest wielkość straty ponad VaR. Ten mankament VaR doprowadził do konstrukcji innych miar ryzyka, które są koherentne (czyli podaddytywne) a ponadto informują o możliwej wielkości straty, jeśli ta strata przekroczyła poziom VaR.

Warunkowa wartość narażona na ryzyko (Conditional VaR)

DEFINICJA. 7.10 *Niech $X \in \mathcal{X}$ będzie zmienną losową całkowalną i o ciągłym rozkładzie. Warunkową wartość narażoną na ryzyko definiujemy wzorem*

$$CVaR_\alpha(X) = \mathbb{E}(-X | -X \geq VaR_\alpha(X)) = -\mathbb{E}(X | X \leq -VaR_\alpha(X)).$$

Uwaga. 32 *W literaturze istnieje wiele miar ryzyka, które różnią się nieznacznie definicjami a także są różnie nazywane. Jednak dla zmiennych losowych całkowalnych, o ciągłym rozkładzie te różne definicje pokrywają się. Dlatego w tym wykładzie zrezygnowaliśmy z ogólnej definicji dla rozkładów nieciągłych, aby w ten sposób uniknąć rozważania różnych definicji.*

TWIERDZENIE. 7.11 *Miara $CVaR_\alpha$ jest koherentną miarą ryzyka na zbiorze zmiennych losowych $X \in \mathcal{X}$ całkowalnych i o ciągłym rozkładzie.*

W dowodzie tego twierdzenia wykorzystamy następujący lemat.

LEMAT. 7.12 *Jeśli $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, $x \in \mathbb{R}$ oraz $\mathbb{P}(X \leq x) > 0$, to $\forall A \in \mathcal{F}$ zachodzi implikacja $\mathbb{P}(A) \geq \mathbb{P}(X \leq x) \Rightarrow \mathbb{E}(X|A) \geq \mathbb{E}(X|X \leq x)$.*

Dowód. (twierdzenia 7.11) Warunki 1–3 definicji 7.5 wynikają z definicji $CVaR$ oraz analogicznych własności VaR.

Aby udowodnić podaddytywność CVaR zauważmy, że

$$\begin{aligned} \text{CVaR}_\alpha(X + Y) &= -\mathbb{E}(X + Y | X + Y \leq -\text{VaR}_\alpha(X + Y)) \\ &= -\mathbb{E}(X | X + Y \leq -\text{VaR}_\alpha(X + Y)) - \mathbb{E}(Y | X + Y \leq -\text{VaR}_\alpha(X + Y)). \end{aligned}$$

Z lematy 7.12 mamy oszacowania

$$\begin{aligned} \mathbb{E}(X | X + Y \leq -\text{VaR}_\alpha(X + Y)) &\geq \mathbb{E}(X | X \leq -\text{VaR}_\alpha(X)), \\ \mathbb{E}(Y | X + Y \leq -\text{VaR}_\alpha(X + Y)) &\geq \mathbb{E}(Y | Y \leq -\text{VaR}_\alpha(Y)), \end{aligned}$$

które wynikają z ciągłości rozkładów oraz równości

$$\begin{aligned} \mathbb{P}(X + Y \leq -\text{VaR}_\alpha(X + Y)) &= \alpha = \mathbb{P}(X \leq -\text{VaR}_\alpha(X)) \\ &= \mathbb{P}(Y \leq -\text{VaR}_\alpha(Y)). \end{aligned}$$

Dostajemy więc

$$\text{CVaR}_\alpha(X + Y) \leq \text{CVaR}_\alpha(X) + \text{CVaR}_\alpha(Y).$$

■

LEMAT. 7.13 Niech $X \sim \mathcal{N}(\mu, \sigma^2)$. Wówczas $\text{CVaR}_\alpha(X) = -\mu - \sigma T_\alpha$, gdzie

$$T_\alpha = \frac{1}{\alpha} \int_{-\infty}^{\Phi^{-1}(\alpha)} \frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} dx.$$

Dowód. Niech $\tilde{X} = X - \mu$, wtedy $\tilde{X} \sim \mathcal{N}(0, \sigma^2)$. Stąd

$$\text{CVaR}_\alpha(X) = \text{CVaR}_\alpha(\tilde{X} + \mu) = \text{CVaR}_\alpha(\tilde{X}) - \mu.$$

Dla \tilde{X} dostajemy

$$\begin{aligned} \text{CVaR}_\alpha(\tilde{X}) &= -\mathbb{E}(\tilde{X} | \tilde{X} \leq -\text{VaR}_\alpha(\tilde{X})) = -\mathbb{E}(\tilde{X} | \tilde{X} \leq \Phi^{-1}(\alpha)\sigma) \\ &= -\sigma \mathbb{E}\left(\frac{\tilde{X}}{\sigma} \mid \frac{\tilde{X}}{\sigma} \leq \Phi^{-1}(\alpha)\right) = -\sigma T_\alpha. \end{aligned}$$

W obliczeniach tych wykorzystaliśmy równość $\text{VaR}_\alpha(\tilde{X}) = -\Phi^{-1}(\alpha)$ oraz fakt, że $\tilde{X}/\sigma \sim \mathcal{N}(0, 1)$. ■

7.3 Portfele minimalizujące VaR lub CVaR

Zajmiemy się teraz problemem znajdowania portfeli optymalnych, dla których optymalność rozumiana jest jako minimalizacja odpowiedniej miary ryzyka. Tak ogólnie postawiony problem jest trudno rozwiązać, dlatego ograniczymy się do rozpatrywania jedynie dwóch najpopularniejszych miar ryzyka: VaR oraz CVaR. Nasze rozważania rozpoczniemy od udowodnienia ogólnego twierdzenia pochodzącego od Krokhmała, Palmquista i Uryaseva, które pokazuje równoważność różnych sformułowań problemu optymalizacyjnego. Jest to uogólnienie wyniku zawartego w twierdzeniu 2.11 z rozdziału 2.

TWIERDZENIE. 7.14 Niech $\mathbb{X} \subset \mathbb{R}^d$ będzie zbiorem portfeli dopuszczalnych na pewnym rynku finansowym. Na zbiorze tych portfeli zdefiniowane są dwie funkcje:

$\varrho : \mathbb{X} \rightarrow \mathbb{R}$ – mierząca ryzyko portfela,

$\Phi : \mathbb{X} \rightarrow \mathbb{R}$ – mierząca zysk portfela.

Rozważmy następujące problemy optymalizacyjne:

$$(P1) \quad \begin{cases} \Phi(x) - \frac{A}{2}\varrho(x) \rightarrow \max, & A > 0 \\ x \in \mathbb{X}. \end{cases}$$

$$(P2) \quad \begin{cases} \varrho(x) \rightarrow \min, \\ \Phi(x) \geq \rho, \\ x \in \mathbb{X}. \end{cases}$$

$$(P3) \quad \begin{cases} \Phi(x) \rightarrow \max, \\ \varrho(x) \leq c, \\ x \in \mathbb{X} \end{cases}$$

Założmy, że zbiór \mathbb{X} jest wypukły, funkcja ϱ jest funkcją wypukłą, a funkcja Φ – funkcją wklęsłą. Przy tych założeniach prawdziwe są następujące stwierdzenia:

1. Niech x^* będzie rozwiązaniem problemu (P1). Wówczas x^* jest rozwiązaniem problemu (P2) dla $\rho = \Phi(x^*)$ oraz rozwiązaniem problemu (P3) dla $c = \varrho(x^*)$.
2. Niech x^* będzie rozwiązaniem problemu (P2). Założmy, że x^* nie jest portfelem minimalnego ryzyka (nie jest rozwiązaniem problemu $\varrho(x) \rightarrow \min, x \in \mathbb{X}$) oraz istnieje $x \in \mathbb{X}$, taki że $\Phi(x) > \rho$. Wówczas x^* jest rozwiązaniem problemu (P1) dla pewnego $A > 0$. Ponadto $\rho = \Phi(x^*)$.
3. Niech x^* będzie rozwiązaniem problemu (P3). Założmy, że x^* nie jest portfelem największego zwrotu (nie jest rozwiązaniem problemu $\Phi(x) \rightarrow \max, x \in \mathbb{X}$) oraz istnieje $x \in \mathbb{X}$, taki że $\varrho(x) < c$. Wówczas x^* jest rozwiązaniem problemu (P1) dla pewnego $A > 0$. Ponadto $\varrho(x^*) = c$.

Dowód. Rozpocznijmy od dowodu punktu 1. twierdzenia. Niech x^* będzie rozwiązaniem problemu (P1). Wówczas

$$\Phi(x^*) - \frac{A}{2}\varrho(x^*) \geq \Phi(x) - \frac{A}{2}\varrho(x), \quad \forall x \in \mathbb{X}. \quad (7.5)$$

Warunek dostateczny dla problemu (P2) ma postać (patrz twierdzenie 9.20)

$$\begin{cases} \varrho(x^*) + \mu_2(\rho - \Phi(x^*)) \leq \varrho(x) + \mu_2(\rho - \Phi(x)), \quad \forall x \in W, \\ \mu_2(\rho - \Phi(x^*)) = 0, \end{cases}$$

gdzie $W = \{x \in \mathbb{X}: \Phi(x) \geq \rho\}$.

Łatwo zauważyć, że wynika on z nierówności (7.5) przy $\mu_2 = \frac{2}{A}$ i $\rho = \Phi(x^*)$. Warunek dostateczny dla problemu (P3) ma postać

$$\begin{cases} -\Phi(x^*) + \mu_3((\varrho(x^*) - c) \leq -\Phi(x) + \mu_3(\varrho(x) - c), \quad \forall x \in W, \\ \mu_3(\varrho(x^*) - c) = 0, \end{cases}$$

gdzie $W = \{x \in \mathbb{X}: \varrho(x) \leq c\}$.

Warunek ten wynika z nierówności (7.5) przy $\mu_3 = \frac{A}{2}$ i $c = \varrho(x^*)$.

Aby udowodnić punkt 2. twierdzenia zauważmy, że na mocy twierdzenia 9.21 istnieje $\mu_2 \geq 0$, takie że

$$\begin{cases} \varrho(x^*) + \mu_2(\rho - \Phi(x^*)) \leq \varrho(x) + \mu_2(\rho - \Phi(x)), \quad \forall x \in \mathbb{X}, \\ \mu_2(\rho - \Phi(x^*)) = 0, \end{cases}$$

co można zapisać równoważnie

$$\begin{cases} \varrho(x^*) - \mu_2\Phi(x^*) \leq \varrho(x) - \mu_2\Phi(x), \quad \forall x \in \mathbb{X} \\ \mu_2(\rho - \Phi(x^*)) = 0. \end{cases}$$

Nie może być $\mu_2 = 0$, bo x^* nie jest portfelem najmniejszego ryzyka. Zatem $\rho = \Phi(x^*)$ i x^* rozwiązuje problem (P1) dla $A = \frac{2}{\mu_2}$.

Dowód punktu 3. twierdzenia przebiega analogicznie do dowodu punktu 2. ■

Rozważmy teraz rynek finansowy, na którym stopy zwrotu instrumentów ryzykownych mają d -wymiarowy rozkład normalny $R \sim \mathcal{N}_d(\mu, \Sigma)$. Jeśli $x \in \mathbb{X}$ jest portfelem dopuszczalnym, to stopa zwrotu tego portfela ma rozkład $x^\top R \sim \mathcal{N}(x^\top \mu, x^\top \Sigma x)$. Rozważmy problem optymalizacyjny poszukiwania portfela inwestycyjnego, który minimalizuje miarę ryzyka ϱ

$$\begin{cases} \varrho(x^\top R) \rightarrow \min, \\ x^\top \mu \geq \rho_0, \\ x \in \mathbb{X}. \end{cases} \quad (7.6)$$

TWIERDZENIE. 7.15 *Na rynku finansowym, na którym stopy zwrotu instrumentów ryzykownych mają rozkład normalny a miarą ryzyka jest VaR lub CVaR, problem optymalizacyjny (7.6) ma identyczne rozwiązanie jak problem optymalizacyjny Markowitza*

$$\begin{cases} x^\top \Sigma x \rightarrow \min, \\ x^\top \mu \geq \rho_0, \\ x \in \mathbb{X}. \end{cases}$$

Dowód. Niech $X = x^\top R$ i $X \sim \mathcal{N}(\rho, \sigma^2)$, gdzie $\rho = x^\top \mu$ a $\sigma^2 = x^\top \Sigma x$. Z twierdzenia 7.9 mamy równość $\text{VaR}_\alpha(X) = -\Phi^{-1}(\alpha)\sigma - \rho$, a z lematu 7.13 równość $\text{CVaR}_\alpha(X) = -T_\alpha\sigma - \rho$. Obie te równości możemy zapisać jako warunek spełniany przez miarę ryzyka VaR lub CVaR: $\varrho(X) = \mathcal{A}\sigma - \rho$ ze stałą dodatnią \mathcal{A} .

Z twierdzenia 7.14 wynika, że problem optymalizacyjny (7.6) jest równoważny problemowi

$$\begin{cases} x^\top \mu - \frac{\Gamma}{2}\varrho(x^\top R) \rightarrow \max, \\ x \in \mathbb{X}, \end{cases} \quad (7.7)$$

dla pewnego $\Gamma > 0$. Korzystając z wzoru na $\varrho(X)$ dla miar ryzyka VaR lub CVaR przekształcamy funkcję celu w problemie optymalizacyjnym (7.7)

$$x^\top \mu - \frac{\Gamma}{2}\varrho(x^\top R) = x^\top \mu - \frac{\Gamma}{2}(\mathcal{A}\sigma - x^\top \mu) = x^\top \mu \left(1 + \frac{\Gamma}{2}\right) - \frac{\Gamma}{2}\mathcal{A}\sigma.$$

Problem optymalizacyjny (7.7) można więc zapisać jako

$$\begin{cases} x^\top \mu - \frac{\Gamma}{2+\Gamma}\mathcal{A}\sigma \rightarrow \max, \\ x \in \mathbb{X}, \end{cases}$$

lub równoważnie

$$\begin{cases} x^\top \mu - \frac{\gamma}{2}\sigma \rightarrow \max, \\ x \in \mathbb{X}, \end{cases}$$

gdzie $\gamma = 2\frac{\Gamma}{2+\Gamma}\mathcal{A}$.

Z twierdzenia 7.14 ostatni problem jest równoważny następującej wersji problemu (P2)

$$\begin{cases} \sigma \rightarrow \min, \\ x^\top \mu \geq \rho_0, \\ x \in \mathbb{X}, \end{cases}$$

a ten jest oczywiście równoważny problemowi

$$\begin{cases} \sigma^2 = x^\top \Sigma x \rightarrow \min, \\ x^\top \mu \geq \rho_0, \\ x \in \mathbb{X}, \end{cases}$$

czyli klasycznemu problemowi optymalizacyjnemu Markowitza. ■

Uwaga. 33 W przypadku klasycznego problemu Markowitza

$$\begin{cases} \min_x x^\top \Sigma x, \\ x^\top \mu \geq \rho_0, \\ x \in \mathbb{X}, \end{cases}$$

jeśli $\rho_0 \geq \frac{\beta}{\alpha}$ (α w tej uwadze oznacza $e^\top \Sigma^{-1} e$, a nie poziom ufności dla VaR lub CVaR), to ograniczenie $x^\top \mu \geq \rho_0$ jest aktywne dla portfela optymalnego x^* , czyli $\mu^\top x^* = \rho_0$. W problemie (7.6) równość $\mu^\top x^* = \rho_0$ zachodzi dla znacznie większych wartości ρ_0 . Wynika to z obserwacji, że obrazem portfela minimalnego ryzyka, gdy ρ jest równe VaR $_\alpha$ lub CVaR $_\alpha$, jest punkt styczności prostej $\mathcal{A}\sigma - \rho = c$ z granicą efektywną problemu Markowitza.

Rzeczywiście, równanie punktu przecięcia prostej $\mathcal{A}\sigma - \rho = c$ z granicą efektywną $\sigma^2 = \frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}$ ma postać

$$\mathcal{A}\sqrt{\frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}} - \rho = c.$$

Różniczkując to wyrażenie po ρ i przyrównując pochodną do zera mamy

$$-1 + \mathcal{A}\frac{\rho\alpha - \beta}{\sqrt{\delta(\alpha\rho^2 - 2\beta\rho + \gamma)}} = 0.$$

Rozwiązując to równanie względem ρ otrzymujemy

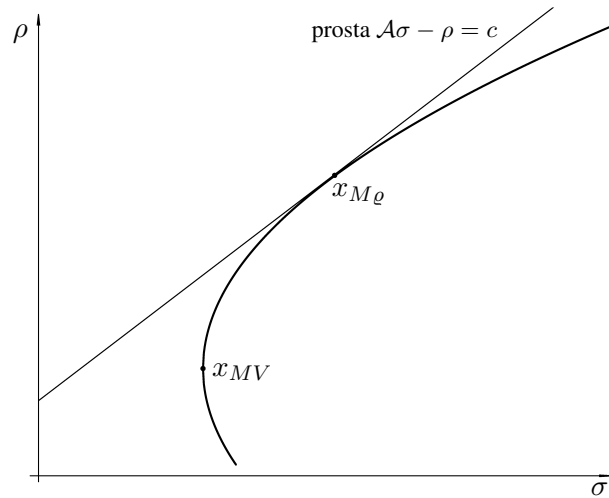
$$\rho_{\min} = \frac{\beta}{\alpha} + \frac{\delta}{\alpha\sqrt{\alpha\mathcal{A}^2 - \delta}}.$$

To jest minimalna wartość ρ w problemie (7.6), przy której ograniczenie $\rho \equiv x^\top \mu \geq \rho_0$ jest aktywne. Tej wartości ρ odpowiada odchylenie standardowe

$$\sigma_{\min} = \sqrt{\frac{\alpha\rho_{\min}^2 - 2\beta\rho_{\min} + \gamma}{\delta}} = \frac{\mathcal{A}}{\sqrt{\alpha\mathcal{A}^2 - \delta}}.$$

Rozważania te można podsumować następującym wnioskiem: dla problemu Markowitza granica efektywna to gałąź hiperboli na płaszczyźnie (ρ, σ) zaczynająca się w punkcie minimalnej wariancji $(\frac{\beta}{\alpha}, \frac{1}{\sqrt{\alpha}})$. Dla zadania optymalizacyjnego (7.6) z miarą ryzyka VaR lub CVaR granica efektywna jest częścią tej gałęzi hiperboli począwszy od punktu $(\rho_{\min}, \sigma_{\min})$, który leży wyżej na tej gałęzi niż punkt minimalnej wariancji.

Uwaga. 34 Dla optymalizacji z instrumentem wolnym od ryzyka rozwiązanie jest bardzo skomplikowane. Niech μ_0 oznacza stopę zwrotu instrumentu wolnego od ryzyka. W dalszym ciągu będziemy analizowali jedynie przypadek $\mu_0 \leq \frac{\beta}{\alpha}$. Wtedy granica efektywna dla problemu Markowitza, to półprosta na płaszczyźnie (ρ, σ)



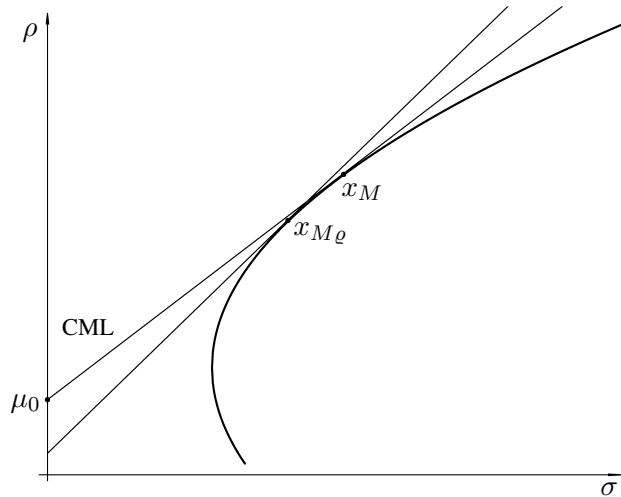
Rysunek 7.8: Granica efektywna dla miary ryzyka ϱ zaczyna się w punkcie minimalnego ryzyka $x_{M\varrho}$ położonym wyżej niż punkt minimalnej wariancji x_{MV} .

zaczynająca się w punkcie $(\mu_0, 0)$ i przechodząca przez obraz portfela stycznego x_T . Jeśli obraz portfela stycznego leży na gałęzi hiperboli powyżej punktu $(\rho_{\min}, \sigma_{\min})$, czyli na granicy efektywnej problemu (7.6), to granica efektywna dla problemu z instrumentem wolnym od ryzyka

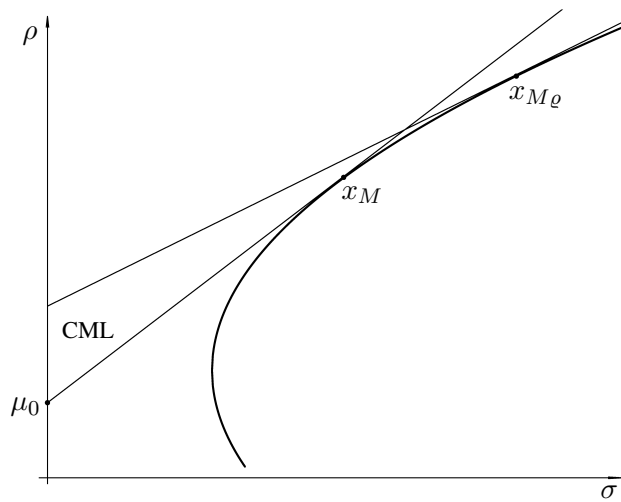
$$\begin{cases} \varrho(x^\top R) \rightarrow \min, \\ x^\top \mu + x_0 \mu_0 \geq \rho_0, \\ (x, x_0) \in \mathbb{X}, \end{cases} \quad (7.8)$$

ma dokładnie takie samo rozwiązanie jak dla problemu Markowitza, czyli jest półprostą zaczynającą się w punkcie $(\mu_0, 0)$ i przechodzącą przez obraz portfela stycznego x_T .

Jeśli obraz portfela stycznego leży na hiperboli poniżej punktu $(\rho_{\min}, \sigma_{\min})$, czyli nie leży na granicy efektywnej problemu (7.6), to granica efektywna dla problemu (7.8) z instrumentem wolnym od ryzyka nie istnieje. Geometrycznie odpowiada to sytuacji, kiedy nachylenie prostej będącej granicą efektywną dla problemu Markowitza z instrumentem wolnym od ryzyka jest większe niż nachylenie prostej $A\sigma - \rho = c$.



Rysunek 7.9: Granica efektywna dla miary ryzyka ρ na rynku z instrumentem wolnym od ryzyka. Portfel minimalnego ryzyka $x_{M\rho}$ ma mniejszą wariancję niż portfel rynkowy x_M . Granicą efektywną jest CML.



Rysunek 7.10: Granica efektywna dla miary ryzyka ρ na rynku z instrumentem wolnym od ryzyka. Portfel minimalnego ryzyka $x_{M\rho}$ ma większą wariancję niż portfel rynkowy x_M . Granicą efektywną jest zbiorem pustym.

Rozdział 8

Wiadomości uzupełniające ze statystyki

W rozdziale tym zebrane zostały fakty ze statystyki potrzebne przy estymacji parametrów modelu Markowitza. Część tych faktów jest powtórzeniem wiedzy z wykładu *Statystyka I* (wtedy są one zwykle podawane bez dowodów). Część jednak wykracza poza materiał tego wykładu. Wtedy odpowiednie twierdzenia są podawane wraz z dowodami.

8.1 Własności rozkładów empirycznych

Rozpocznijmy od podania podstawowej dla rozkładów empirycznych definicji czym jest próbka z danego rozkładu.

DEFINICJA. 8.1 *Rozkładem prawdopodobieństwa D na $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ nazywamy miarę dodatnią D spełniającą warunek $D(\mathbb{R}^d) = 1$.*

DEFINICJA. 8.2 *Próbką z rozkładu prawdopodobieństwa D o dystrybuancie F_D nazywamy ciąg niezależnych zmiennych losowych X_1, \dots, X_n o jednakowym rozkładzie, $\mathbb{P}(X_i \leq x) = F_D(x)$ dla $i = 1, \dots, n$. Aby zaznaczyć, że ciąg zmiennych X_i jest próbką z rozkładu D będziemy używali oznaczenia*

$$X_1, \dots, X_n \sim D.$$

DEFINICJA. 8.3 *Średnią z próbki X_1, \dots, X_n nazywamy zmienną losową*

$$\bar{X}_{(n)} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Wariancją z próbki X_1, \dots, X_n nazywamy zmienną losową

$$S_{(n)}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_{(n)})^2.$$

Zauważmy, że równoważnie wariancję można zdefiniować wzorem

$$S_{(n)}^2 = \frac{1}{n-1} \sum_{i=1}^n X_i^2 - \frac{n}{n-1} \bar{X}_{(n)}^2$$

TWIERDZENIE. 8.4 Niech X_1, \dots, X_n będą próbką z pewnego rozkładu (czyli zmiennymi losowymi i.i.d.) oraz $\mathbb{E}X_1 < \infty$ i $\mathbb{E}X_1^2 < \infty$. Wówczas:

1. $\mathbb{E}\bar{X}_{(n)} = \mathbb{E}X_1$.
2. $\mathbb{E}S_{(n)}^2 = \text{Var}X_1$.
3. $\lim_{n \rightarrow \infty} \bar{X}_{(n)} = \mathbb{E}X_1$ p.n.
4. $\lim_{n \rightarrow \infty} S_{(n)}^2 = \text{Var}X_1$ p.n.

Dowód. Punkt 1. twierdzenia wynika z ciągu równości

$$\mathbb{E}\bar{X}_{(n)} = \frac{1}{n} \sum_{i=1}^n \mathbb{E}X_i = \frac{n \mathbb{E}X_1}{n} = \mathbb{E}X_1.$$

Aby udowodnić punkt 2. twierdzenia wprowadźmy oznaczenia $\mu = \mathbb{E}X_1$, $a^2 = \mathbb{E}X_1^2$. Wtedy

$$\begin{aligned} \mathbb{E}\bar{X}_{(n)}^2 &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}(X_i X_j) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{E}(X_i^2) + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^{i-1} 2\mathbb{E}(X_i X_j) = \\ &= \frac{1}{n^2} n a^2 + \frac{1}{n^2} \sum_{i=1}^n 2(i-1)\mu^2 = \frac{a^2}{n} + 2 \frac{\mu^2 n(n-1)}{n^2} = \frac{a^2}{n} + \frac{n-1}{n} \mu^2. \end{aligned}$$

Korzystając z definicji 8.3 otrzymujemy

$$\begin{aligned} \mathbb{E}S_{(n)}^2 &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}X_i^2 - \frac{n}{n-1} \mathbb{E}(\bar{X}_{(n)})^2 = \\ &= \frac{1}{n-1} n a^2 - \frac{n}{n-1} \left(\frac{a^2}{n} + \frac{n-1}{n} \mu^2 \right) = \\ &= \frac{n}{n-1} a^2 - \frac{a^2}{n-1} - \mu^2 = a^2 - \mu^2 = \text{Var}X_1. \end{aligned}$$

Punkt 3. twierdzenia wynika z mocnego prawa wielkich liczb.

Aby udowodnić punkt 4. skorzystamy z mocnego prawa wielkich liczb dla dwóch ciągów:

$$\begin{aligned} \lim_{n \rightarrow \infty} \bar{X}_{(n)} &= \mu \text{ p.n.} \\ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i^2 &= a^2 \text{ p.n.} \end{aligned} \tag{8.1}$$

Przypomnijmy, że dla zmiennych losowych i.i.d mocne prawo wielkich liczb wymaga tylko istnienia pierwszego momentu. Dla powyższych dwóch ciągów musi być zatem $\mathbb{E}|X_1| < \infty$ oraz $\mathbb{E}X_1^2 < \infty$, a te oszacowania są spełnione na mocy naszych założeń.

Korzystając z definicji 8.3 mamy $S_{(n)}^2 = \frac{1}{n-1} \sum_i X_i^2 - \frac{n}{n-1} (\bar{X}_{(n)})^2$. Stąd wynika przejście graniczne

$$\lim_{n \rightarrow \infty} S_{(n)}^2 = a^2 - 2\mu^2 + \mu^2 = a^2 - \mu^2 = \text{Var}(X_1).$$

Zauważmy, że w powyższym ciągu równości pierwsza równość zachodzi prawie na pewno. ■

W dalszym ciągu chcemy opisać własności próbki z rozkładu normalnego. Potrzebny nam do tego będzie rozkład χ^2 , dlatego zaczniemy od przypomnienia podstawowych własności tego rozkładu.

DEFINICJA. 8.5 Zmienna losowa y ma rozkład χ_n^2 (chi kwadrat o n stopniach swobody), jeśli

$$y \sim \sum_{i=1}^n X_i^2,$$

gdzie X_i są i.i.d. oraz mają rozkład $N(0, 1)$.

TWIERDZENIE. 8.6 Gęstość rozkładu χ_n^2 dana jest wzorem

$$f_n(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (8.2)$$

Wartość średnia χ_n^2 wynosi n , zaś wariancja $2n$.

TWIERDZENIE. 8.7 Niech $y \sim \chi_n^2$. Wówczas:

1. $\mathbb{E}\left(\frac{1}{y}\right) = \frac{1}{n-2}$.
2. $\mathbb{E}\left(\frac{1}{y^2}\right) = \frac{1}{(n-2)(n-4)}$.

Dowód. Dowód pierwszej równości otrzymujemy przez całkowanie

$$\begin{aligned} \mathbb{E}\left(\frac{1}{y}\right) &= \int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} \frac{1}{x} dx = \\ &= \int_0^\infty \frac{1}{2^{\frac{n-2}{2}}\Gamma(n-2/2)} e^{-\frac{x}{2}} x^{\frac{n-2}{2}-1} dx \cdot \frac{2^{\frac{n-2}{2}}}{2^{\frac{n}{2}}} \cdot \frac{\Gamma(n-2/2)}{\Gamma(n/2)} = \\ &= 1 \cdot \frac{1}{2} \cdot \frac{1}{\frac{n-2}{2}} = \frac{1}{n-2}, \end{aligned}$$

gdź $\Gamma(z+1) = z\Gamma(z)$. (Przypomnijmy, że $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$.)

Dowód drugiej równości otrzymujemy analogicznie. ■

Możemy teraz przejść do opisanego własności momentów z próbki z rozkładu normalnego.

TWIERDZENIE. 8.8 Załóżmy, że $X_i \sim N(\mu, \sigma^2)$, $i = 1, \dots, n$. Wówczas:

1. $\bar{X}_{(n)} \sim N(\mu, \frac{\sigma^2}{n})$.
2. $\frac{n-1}{\sigma^2} S_{(n)}^2 \sim \chi_{n-1}^2$.
3. Zmienna losowa $\bar{X}_{(n)}$ jest niezależna od $S_{(n)}^2$.

Dowód. Pierwszy punkt twierdzenia jest dowodzony na wykładzie statystyki, dlatego pomijamy ten dowód. Pozostałe dwa punkty są też dowodzone na tym wykładzie, ale w tym przypadku przytaczamy dowód, ponieważ będziemy się na niego powoływać w przypadku wielowymiarowego rozkładu normalnego.

Zauważmy na początku, że wystarczy dowieść punkty 2 i 3 dla $\mu = 0$ oraz $\sigma = 1$. Definiując $z_i = \frac{(X_i - \mu)}{\sigma}$ sprowadzamy problem do takiego przypadku, bo wówczas $z_i \sim N(0, 1)$.

Zakładamy teraz, że $X_i \sim N(0, 1)$, czyli

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N(0, I).$$

Oznaczmy przez v_1, \dots, v_{n-1} bazę ortonormalną podprzestrzeni prostopadłej do wektora $e = (1, \dots, 1)$. Wówczas

$$C = \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ \frac{1}{\sqrt{n}} e \end{pmatrix}$$

jest macierzą ortogonalną: $C^\top C = C C^\top = I$. Połóżmy $y = CX$. Wtedy $y \sim N(0, C^\top C)$. Oznaczmy przez y_1, \dots, y_n współrzędne wektora y . Zauważmy, że $y_n = \sqrt{n} \bar{X}_n$. Ponadto

$$y^\top y = X^\top C^\top C X = X^\top X$$

oraz

$$y^\top y = \sum_{i=1}^{n-1} y_i^2 + y_n^2$$

Stąd

$$X^\top X = \sum_{i=1}^{n-1} y_i^2 + n \bar{X}_{(n)}^2,$$

a zatem

$$(n-1)S_{(n)}^2 = (X^\top X - n \bar{X}_{(n)}^2) = \sum_{i=1}^{n-1} y_i^2 \sim \chi_{n-1}^2,$$

co dowodzi 2.

Zauważmy też, że y_n jest niezależne od y_1, \dots, y_{n-1} . Zatem $S_{(n)}^2$ jest niezależne od $y_n = n\bar{X}_{(n)}^2$, co dowodzi 3. ■

8.2 Własności próbek wielowymiarowych

Rozszerzymy teraz wyniki uzyskane dla $d = 1$ na próbki z rozkładów wielowymiarowych. Oczywiście definicja średniej z próbki nie ulega zmianie, ponieważ jest ona identyczna w każdym wymiarze. Wariancję z próbki musimy w przypadku wielowymiarowym zastąpić bardziej skomplikowanym obiektem

DEFINICJA. 8.9 *Macierz kowariancji z próbki nazywamy macierz $\hat{\Sigma}_{(n)} \in \mathbb{R}^{d \times d}$ zadaną wzorem*

$$(\hat{\Sigma}_{(n)})_{ij} = \frac{1}{n-1} \sum_{k=1}^n ((X_k)_i - (\bar{X}_{(n)})_i)((X_k)_j - (\bar{X}_{(n)})_j).$$

Podobnie jak w przypadku jednowymiarowym wzór ten można przekształcić do następującej postaci

$$(\hat{\Sigma}_{(n)})_{ij} = \frac{1}{n-1} \sum_{k=1}^n (X_k)_i (X_k)_j - \frac{n}{n-1} (\bar{X}_{(n)})_i (\bar{X}_{(n)})_j$$

lub zapisać w postaci macierzowej

$$\hat{\Sigma}_{(n)} = \frac{1}{n-1} X^\top X - \frac{n}{n-1} \bar{X}_{(n)} \bar{X}_{(n)}^\top,$$

gdzie próbkę X_i traktujemy jako wektor kolumnowy a

$$X = \begin{pmatrix} X_1^\top \\ \vdots \\ X_n^\top \end{pmatrix}.$$

Uwaga. 35 *W symbolu $\hat{\Sigma}_{(n)}$ będziemy pomijać indeks n , jeśli nie będzie to prowadzić do niejednoznaczności.*

TWIERDZENIE. 8.10 *Załóżmy, że X_1, \dots, X_n jest próbką z pewnego rozkładu (próbką i.i.d.) oraz $\mu = \mathbb{E}X_1$ i $\Sigma = \text{Cov}(X_1)$. Wówczas:*

1. $\mathbb{E} \bar{X}_{(n)} = \mu$.
2. $\mathbb{E} \hat{\Sigma}_{(n)} = \Sigma$.
3. $\lim_{n \rightarrow \infty} \bar{X}_{(n)} = \mu$ p.n.

$$4. \lim_{n \rightarrow \infty} \hat{\Sigma}_{(n)} = \Sigma \text{ p.n.}$$

Dowód. Dowód jak w przypadku jednowymiarowym. ■

Uwaga. 36 Macierz $\hat{\Sigma}_{(n)}$ jest symetryczna. Jeśli $n \leq d$, to macierz $\hat{\Sigma}_{(n)}$ jest osobliwa. Jeśli $n > d$, to nie wiadomo – zależy to od rozkładu X_1 . Pokażemy dalej, że dla rozkładu ciągłego X_1 oraz $n > d$ macierz $\hat{\Sigma}_{(n)}$ p.n. jest nieosobliwa.

LEMAT. 8.11 Niech $M \in \mathbb{R}^{k \times k}$ będzie macierzą postaci

$$M = \begin{pmatrix} m_{1,1} & m_{1,2}^\top \\ m_{2,1} & M_{2,2} \end{pmatrix},$$

gdzie $m_{1,1} \in \mathbb{R}$, $m_{1,2}, m_{2,1} \in \mathbb{R}^{k-1}$, $M_{2,2} \in \mathbb{R}^{(k-1) \times (k-1)}$.

Jeśli $\det(M_{2,2}) \neq 0$ i $\det(M) = 0$, to $m_{1,1} = m_{1,2}^\top M_{2,2}^{-1} m_{2,1}$.

Dowód. Z faktu, że M jest osobliwa oraz $M_{2,2}$ jest nieosobliwa wynika, że pierwsza kolumna może być przedstawiona jako kombinacja liniowa pozostałych, tzn. istnieje $u \in \mathbb{R}^{k-1}$ takie, że

$$\begin{pmatrix} m_{1,1} \\ m_{2,1} \end{pmatrix} = \begin{pmatrix} m_{1,2}^\top \\ M_{2,2} \end{pmatrix} u.$$

Stąd dostajemy: $m_{1,1} = m_{1,2}^\top u$, $m_{2,1} = M_{2,2} u$. Z drugiej równości otrzymujemy $u = M_{2,2}^{-1} m_{2,1}$. Wstawiając to u do pierwszej równości mamy poszukiwaną tezę $m_{1,1} = m_{1,2}^\top M_{2,2}^{-1} m_{2,1}$. ■

LEMAT. 8.12 Niech $Z \in \mathbb{R}^{k \times k}$ będzie macierzą losową o rozkładzie ciągłym. Wówczas $\mathbb{P}(\det(Z) = 0) = 0$.

Dowód. Dowód prowadzimy przez indukcję po k . Dla $k = 1$ teza jest oczywista. Załóżmy, że teza jest prawdziwa dla $k - 1$. Macierz Z zapisujemy następująco

$$Z = \begin{pmatrix} z_{1,1} & z_{1,2}^\top \\ z_{2,1} & Z_{2,2} \end{pmatrix},$$

gdzie $z_{1,1} \in \mathbb{R}$, $z_{1,2}, z_{2,1} \in \mathbb{R}^{k-1}$, $Z_{2,2} \in \mathbb{R}^{(k-1) \times (k-1)}$.

Wówczas

$$\begin{aligned} \mathbb{P}(\det(Z) = 0) &= \\ &= \mathbb{P}(\det(Z) = 0, \det(Z_{2,2}) \neq 0) + \mathbb{P}(\det(Z) = 0, \det(Z_{2,2}) = 0) = \\ &= \mathbb{P}(z_{1,1} = z_{1,2}^\top Z_{2,2}^{-1} z_{2,1}, \det(Z_{2,2}) \neq 0) \end{aligned}$$

bo $\mathbb{P}(\det(Z) = 0, \det(Z_{2,2}) = 0) = 0$ z założenia indukcyjnego.

Do zakończenia dowodu wystarczy zauważyć, że zbiór macierzy M , takich że

$$M = \begin{pmatrix} m_{1,1} & m_{1,2}^\top \\ m_{2,1} & M_{2,2} \end{pmatrix} \in \mathbb{R}^{k \times k} : \det(M_{2,2}) \neq 0 \text{ i } m_{1,1} = m_{1,2}^\top M_{2,2}^{-1} m_{2,1}$$

ma miarę Lebesgue'a zero. Istotnie tak jest, gdyż $m_{1,1}$ jest jednoznacznie wyznaczone przez pozostałe współrzędne. ■

Następujący pomocniczy lemat z algebry liniowej będzie nam potrzebny w dowodzie następnego twierdzenia.

LEMAT. 8.13 *Jeśli A jest macierzą wymiaru $n \times d$, to $\text{rz}(A) = \text{rz}(A^\top A)$*

Dowód. $\text{rz}(A) = d - \dim \ker A$. Pokażemy, że jeśli $x \notin \ker A$, to $x \notin \ker A^\top A$. Jeśli $Ax \neq 0$, to $0 \neq \|Ax\|^2 = (Ax)^\top Ax = x^\top A^\top Ax = x^\top (A^\top Ax)$. Stąd $A^\top Ax \neq 0$. Oczywiście jest, że jeśli $x \in \ker A$, to $x \in \ker A^\top A$. ■

TWIERDZENIE. 8.14 *Jeśli zmienne losowe X_i , $i = 1, \dots, n$ są i.i.d. a ich rozkład jest ciągły oraz $n > d$, to macierz $\hat{\Sigma}_{(n)}$ jest nieosobliwa p.n.*

Dowód. Oznaczmy:

$$\tilde{x} = \begin{bmatrix} X_1^\top - \bar{X}_{(n)}^\top \\ \vdots \\ X_d^\top - \bar{X}_{(n)}^\top \end{bmatrix}, \quad x = \begin{bmatrix} X_1^\top - \bar{X}_{(n)}^\top \\ \vdots \\ X_n^\top - \bar{X}_{(n)}^\top \end{bmatrix}.$$

Zauważmy, że $\hat{\Sigma}_{(n)} = \frac{1}{n-1} x^\top x$. Zatem

$$\text{rz} \hat{\Sigma}_{(n)} = \text{rz}(x^\top x) = \text{rz}(x) \geq \text{rz}(\tilde{x}).$$

Jeśli $n \geq d + 1$ i X_i mają ciągły rozkład, to wektory X_1, \dots, X_d są liniowo niezależne p.n. Ponieważ rozkład $\bar{X}_{(n)}$ pod warunkiem X_1, \dots, X_d jest ciągły (tutaj używamy $n \geq d + 1$), to również wektory $X_1 - \bar{X}_{(n)}, \dots, X_d - \bar{X}_{(n)}$ są liniowo niezależne (tylko skończona liczba wartości $\bar{X}_{(n)}$ mogłaby spowodować powstanie liniowej zależności a miara takiego zbioru jest zero). A zatem macierz \tilde{x} ma pełny rząd p.n. ■

Uwaga. 37 *Z dowodu poprzedniego twierdzenia łatwo zauważyć, że macierz $\hat{\Sigma}_{(n)}$ jest nieujemnie określona p.n. Jeśli weźmiemy wektor $v \neq 0$, to*

$$v^\top \hat{\Sigma}_{(n)} v = \frac{1}{n-1} v^\top x^\top x v = \frac{1}{n-1} \|xv\|^2 \geq 0.$$

Przejdziemy teraz do badania własności próbek z wielowymiarowego rozkładu normalnego.

DEFINICJA. 8.15 Zmienna losowa X ma rozkład normalny d -wymiarowy o wektorze średnich $\mu \in \mathbb{R}^d$ i macierzy kowariancji $\Sigma \in \mathbb{R}^{d \times d}$, jeśli ma gęstość

$$f(x) = \frac{1}{2^{\frac{d}{2}} \sqrt{\det(\Sigma)}} e^{-\frac{(x-\mu)^\top \Sigma^{-1} (x-\mu)}{2}}.$$

Zmienną taką oznaczamy $X \sim N_d(\mu, \Sigma)$.

Następujący lemat jest znany z rachunku prawdopodobieństwa.

LEMAT. 8.16 Jeśli X jest zmienną losową o d -wymiarowym rozkładzie normalnym $X \sim N_d(\mu, \Sigma)$, to

1. $\mathbb{E}X = \mu$, $\text{Cov}(X) = \Sigma$.
2. Jeśli C jest macierzą wymiaru $r \times d$ o rzędzie r , $r \leq d$, a b jest wektorem z \mathbb{R}^r , to $b + CX \sim N_r(C\mu + b, C\Sigma C^\top)$.
3. Jeśli macierz Σ jest symetryczna, dodatnio określona, to istnieje macierz kwadratowa C , taka że $CC^\top = \Sigma$. To znaczy, że $\mu + CY \sim N_d(\mu, \Sigma)$, jeśli $Y \sim N_d(0, I)$.

DEFINICJA. 8.17 Zmienna losowa U ma rozkład Wisharta $W_d(m, \Sigma)$, jeśli

$$U \sim Y^\top Y = \sum_{k=1}^m Y_k Y_k^\top,$$

gdzie

$$Y = \begin{pmatrix} Y_1^\top \\ \vdots \\ Y_m^\top \end{pmatrix}$$

a zmienne losowe Y_i są i.i.d o rozkładzie $N_d(0, \Sigma)$ na \mathbb{R}^d . (Zakładamy oczywiście nieosobliwość Σ , aby rozkład $N_d(0, \Sigma)$ miał gęstość.)

Uwaga. 38 Rozkład Wisharta jest uogólnieniem rozkładu χ^2 , ponieważ

$$W_1(m, 1) \sim \chi_m^2$$

LEMAT. 8.18 Zmienna losowa o rozkładzie Wisharta ma następujące własności:

1. Jeśli $U \sim W_d(m, \Sigma)$ i $m \geq d$, to macierz U jest nieosobliwa p.n.
2. Jeśli macierz $C \in \mathbb{R}^{d \times d}$ jest nieosobliwa i $U \sim W_d(m, \Sigma)$, to

$$CUC^\top \sim W_d(m, C\Sigma C^\top).$$

W szczególności U może być przedstawiona jako

$$U = CVC^\top,$$

gdzie $CC^\top = \Sigma$ i $V \sim W_d(m, I)$.

Dowód. Punkt pierwszy twierdzenia dowodzi się analogicznie jak twierdzenie 8.14.

Aby udowodnić punkt drugi, korzystamy z definicji rozkładu Wisharta. Wtedy $U \sim \sum_{k=1}^m Y_k Y_k^\top$, gdzie $Y_k \sim N_d(0, \Sigma)$. Zatem

$$\begin{aligned} CUC^\top &\sim C\left(\sum_{k=1}^m Y_k Y_k^\top\right)C^\top = \sum_{k=1}^m C Y_k Y_k^\top C^\top = \\ &= \sum_{k=1}^m (C Y_k)(C Y_k)^\top = \sum_{k=1}^m \tilde{Y}_k \tilde{Y}_k^\top, \end{aligned}$$

gdzie $\tilde{Y}_k \sim N(0, C\Sigma C^\top)$. ■

TWIERDZENIE. 8.19 Załóżmy, że X_1, \dots, X_n jest próbką i.i.d. z rozkładu normalnego $N_d(\mu, \Sigma)$. Wówczas:

1. $\bar{X}_{(n)} \sim N_d(\mu, \frac{\Sigma}{n})$.
2. $(n-1)\hat{\Sigma}_{(n)} \sim W_d(n-1, \Sigma)$.
3. Zmienne losowe $\bar{X}_{(n)}$ i $\hat{\Sigma}_{(n)}$ są niezależne.

Dowód. Dowód jest podobny do dowodu twierdzenia 8.8. Punkt 1. dowodzi się analogicznie jak w przypadku jednowymiarowym.

Dla dowodu punktów 2. i 3. zauważamy, że istnieje taka macierz nieosobliwa A , że $AA^\top = \Sigma$. Połóżmy $Z_k = A^{-1}(X_k - \mu)$. Wówczas, na mocy lematu 8.16, $Z_k \sim N(0, A^{-1}\Sigma(A^{-1})^\top) = N(0, I_d)$. Korzystając z faktu, że $AZ_k = (X_k - \mu)$ dostajemy

$$A\bar{Z}_{(n)} = \bar{X}_{(n)} - \mu.$$

Wtedy

$$\begin{aligned} \hat{\Sigma}_{(n)} &= \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_{(n)})(X_k - \bar{X}_{(n)})^\top = \\ &= \frac{1}{n-1} \sum_{k=1}^n A(Z_k - \bar{Z}_{(n)})(Z_k - \bar{Z}_{(n)})^\top A^\top = A\hat{\Sigma}^Z A^\top, \end{aligned} \tag{8.3}$$

gdzie $\hat{\Sigma}^Z$ jest macierzą kowariancji z próbki Z_1, \dots, Z_n .

Weźmy macierz $C \in \mathbb{R}^{n \times n}$ zdefiniowaną wzorem

$$C = \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ \frac{1}{\sqrt{n}}e \end{pmatrix},$$

gdzie v_1, \dots, v_{n-1} jest bazą ortonormalną w podprzestrzeni prostopadłej do wektora $e = (1, \dots, 1)$. Niech

$$Z = \begin{pmatrix} Z_1^\top \\ \vdots \\ Z_n^\top \end{pmatrix}.$$

Wiemy, że współrzędne tej macierzy (należącej do $\mathbb{R}^{n \times d}$) są próbką i.i.d. z rozkładu $N(0, 1)$. Stąd kolumny są niezależne o rozkładzie $N_n(0, I_n)$. Dokonując zamiany zmiennych $Y = CZ$ otrzymujemy macierz Y , której kolumny mają rozkład $N(0, CI_nC^\top) = N(0, I_n)$ a ostatni wiersz jest równy $\sqrt{n}\bar{Z}_{(n)}^\top$. Z tożsamości

$$\begin{aligned} Y^\top Y &= (CZ)^\top CZ = Z^\top C^\top CZ = Z^\top Z, \\ Y^\top Y &= \sum_{k=1}^n Y_k Y_k^\top = \sum_{k=1}^{n-1} Y_k Y_k^\top + n\bar{Z}_{(n)}\bar{Z}_{(n)}^\top, \end{aligned}$$

wynika, że $(n-1)\hat{\Sigma}^Z = Z^\top Z - n\bar{Z}_{(n)}\bar{Z}_{(n)}^\top = \sum_{k=1}^{n-1} Y_k Y_k^\top \sim W_d(n-1, I_d)$. Ze wzoru 8.3 otrzymujemy wtedy

$$(n-1)\hat{\Sigma}_{(n)} = (n-1)A\hat{\Sigma}^Z A^\top \sim W_d(n-1, AI_d A^\top) = W_d(n-1, \Sigma).$$

Pozostaje jeszcze wykazać niezależność $\bar{X}_{(n)}$ i $\hat{\Sigma}_{(n)}$. Ponieważ

$$\hat{\Sigma}_{(n)} = A \left(\sum_{k=1}^{n-1} Y_k Y_k^\top \right) A^\top \cdot \frac{1}{n-1}, \quad \bar{X}_{(n)} = AY_n \cdot \frac{1}{\sqrt{n}} + \mu,$$

więc teza wynika z niezależności wierszy Y , co jest konsekwencją faktu, że X_1, \dots, X_n jest próbką i.i.d. ■

TWIERDZENIE. 8.20 (Dalsze własności rozkładu Wisharta) *Rozkład Wisharta ma następujące dodatkowe własności:*

1. *Jeśli $U \sim W_d(m, \Sigma)$, $A \in \mathbb{R}^{r \times d}$ oraz $\text{rz}(A) = r$, gdzie $r \leq d$, to $AUA^\top \sim W_r(m, A\Sigma A^\top)$.*
2. *Niech U będzie macierzą o rozkładzie Wisharta. Przedstawmy U w postaci*

$$U = \begin{pmatrix} U_{1,1} & U_{1,2}^\top \\ U_{2,1} & U_{2,2} \end{pmatrix},$$

gdzie $U_{1,1} \in \mathbb{R}^{r \times r}$, $U_{2,2} \in \mathbb{R}^{s \times s}$ i $r + s = d$. Jeśli $U \sim W_d(m, \Sigma)$, to zastosujmy analogiczne przedstawienie do macierzy $\Sigma \in \mathbb{R}^{d \times d}$. Wtedy $U_{1,1} \sim W_r(m, \Sigma_{1,1})$.

3. Jeśli $U_i \sim W_d(m_i, \Sigma)$, $i = 1, \dots, k$ są niezależne, to

$$\sum_{i=1}^k U_i \sim W_d\left(\sum_{i=1}^k m_i, \Sigma\right).$$

4. Niech $U \sim W_d(m, \Sigma)$. Wówczas

a. $\mathbb{E}U^{-1} = \frac{1}{m-d-1}\Sigma^{-1}$, jeśli $m > d + 1$.

b. Jeśli C jest nieujemnie określona macierzą wymiaru $d \times d$, to

$$\begin{aligned} \mathbb{E}(U^{-1}CU^{-1}) &= \frac{1}{(m-d)(m-d-1)(m-d-3)} \operatorname{tr}(\Sigma^{-1}C)\Sigma^{-1} \\ &+ \frac{1}{(m-d)(m-d-3)}\Sigma^{-1}C\Sigma^{-1}, \quad \text{dla } m > d + 3. \end{aligned}$$

Dowód. Aby udowodnić punkt 1. twierdzenia zauważmy, że macierz U można przedstawić w postaci $U \sim \sum_{i=1}^m X_i X_i^\top$, gdzie X_i są zmiennymi i.i.d. z rozkładu $N_d(0, \Sigma)$. Wtedy

$$AUA^\top \sim \sum_{i=1}^m AX_i X_i^\top A^\top = \sum_{i=1}^m AX_i (AX_i)^\top, \quad \text{gdzie } AX_i \sim N_r(0, A\Sigma A^\top).$$

Dowód punktu 2. otrzymujemy przez zastosowanie powyższego rozumowania do macierzy $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Punkt 3. twierdzenia wynika bezpośrednio z definicji rozkładu Wisharta. Natomiast punkt 4. jest rachunkowy oraz trudny. Nie będziemy tutaj przytaczać tego skomplikowanego rachunku bazującego na postaci funkcji gęstości rozkładu Wisharta oraz twierdzeniu Stokesa. Zainteresowanego czytelnika odsyłamy do twierdzenia 3.2 w pracy L. R. Haff – An identity for the Wishart distribution with applications, *Journal of Multivariate Analysis*, **9** (1979), 531–544. ■

Rozdział 9

Podstawowe fakty z optymalizacji nieliniowej

W rozdziale tym przypomnimy podstawowe fakty z optymalizacji nieliniowej. Fakty te pochodzą w większości ze skryptu "Optymalizacja II" autorstwa J. Palczewskiego. Twierdzenia będą podawane bez dowodów. Dowody, często przy znacznie słabszych założeniach, można znaleźć we wspomnianym skrypcie.

Niech W będzie niepustym podzbiorem \mathbb{R}^n , zaś $f: W \rightarrow \mathbb{R}$. Rozważmy problem minimalizacji funkcji f na zbiorze W :

$$\begin{cases} \min_x f(x), \\ x \in W. \end{cases} \quad (9.1)$$

DEFINICJA. 9.1 Elementy zbioru W nazywamy **punktami dopuszczalnymi**. **Rozwiązaniem globalnym** nazywamy taki punkt $\bar{x} \in W$, że $f(\bar{x}) \leq f(x)$ dla każdego $x \in W$. **Rozwiązaniem lokalnym** jest taki punkt $\bar{x} \in W$, że istnieje $\varepsilon > 0$, dla którego $f(\bar{x}) \leq f(x)$, jeśli $x \in W \cap K(\bar{x}, \varepsilon)$. **Rozwiązanie jest ścisłe**, jeśli $f(\bar{x}) < f(x)$ dla $x \in W \cap K(\bar{x}, \varepsilon)$ i $x \neq \bar{x}$.

9.1 Optymalizacja z ograniczeniami równościowymi

Przypomnijmy podstawowe fakty z analizy II dotyczące minimalizacji funkcji f na zbiorze W zdefiniowanym przy pomocy zestawu równości. Rozpatrywany problem optymalizacyjny ma wtedy postać:

$$\begin{cases} \min_x f(x), \\ h_j(x) = 0, \quad j = 1, \dots, l, \\ x \in \mathbb{R}^n. \end{cases} \quad (9.2)$$

Z tego sformułowania wynika następująca postać zbioru punktów dopuszczalnych

$$W = \{x \in \mathbb{R}^n: h_j(x) = 0, j = 1, \dots, l\}.$$

Aby sformułować warunki konieczne i dostateczne istnienia rozwiązania problemu optymalizacyjnego (9.1) musimy wprowadzić kilka nowych pojęć.

DEFINICJA. 9.2 *Stożkiem kierunków stycznych* $T(\bar{x})$ do W w punkcie $\bar{x} \in \overline{W}$ nazywamy zbiór wektorów $d \in \mathbb{R}^n$, takich że

$$d = \lim_{k \rightarrow \infty} \lambda_k (x_k - \bar{x})$$

dla pewnych $\lambda_k > 0$, $x_k \in W$ i $x_k \rightarrow \bar{x}$.

DEFINICJA. 9.3 Niech $\bar{x} \in W$, h_j , $j = 1, \dots, l$, będą różniczkowalne w \bar{x} . *Stożkiem kierunków stycznych dla ograniczeń zlinearyzowanych* nazywamy zbiór

$$T_{lin}(\bar{x}) = \{d \in \mathbb{R}^n: \forall j = 1, \dots, l \quad Dh_j(\bar{x})d = 0\}$$

gdzie $Dh(x)$ oznacza gradient funkcji h w punkcie x .

DEFINICJA. 9.4 Punkt $\bar{x} \in W$ nazywa się **regularnym punktem ograniczeń**, lub po prostu **punktem regularnym**, jeśli

$$T(\bar{x}) = T_{lin}(\bar{x}).$$

Sformułowanie warunków, przy których punkt $\bar{x} \in W$ jest punktem regularnym, stanowi ważny i skomplikowany problem optymalizacji nieliniowej. Dla naszych celów wystarczający jest następujący lemat.

LEMAT. 9.5 *Jeśli w punkcie $\bar{x} \in W$ funkcje h_j , $j = 1, \dots, l$, są funkcjami afinicznymi, to punkt \bar{x} jest regularnym punktem ograniczeń.*

Dysponujemy teraz niezbędnym aparatem, aby sformułować warunki konieczne i dostateczne rozwiązalności zadania (9.2).

TWIERDZENIE. 9.6 (Warunek konieczny I rzędu) *Niech funkcja f oraz funkcje h_j będą różniczkowalne w punkcie $\bar{x} \in W$ a punkt \bar{x} jest punktem regularnym ograniczeń. Jeśli przy tych założeniach punkt \bar{x} jest lokalnym minimum funkcji f na zbiorze W , to istnieje wektor $\lambda \in \mathbb{R}^l$, $\lambda \neq 0$, że*

$$Df(\bar{x}) + \sum_{j=1}^l \lambda_j Dh_j(\bar{x}) = 0.$$

Przypomnijmy, że funkcję

$$L(x; \lambda) = f(x) + \sum_{j=1}^l \lambda_j h_j(x)$$

nazywamy **funkcją Lagrange'a** problemu (9.2), a składowe wektora λ **współczynniki Lagrange'a**.

TWIERDZENIE. 9.7 (Warunek dostateczny II rzędu) Niech funkcja f oraz funkcje h_j będą funkcjami klasy C^2 w punkcie $\bar{x} \in W$ a punkt \bar{x} jest punktem regularnym ograniczeń. Jeśli istnieje wektor $\lambda \in \mathbb{R}^l$, taki że w punkcie \bar{x} spełniony jest warunek I rzędu

$$D_x L(\bar{x}; \lambda) = 0,$$

oraz

$$d^\top D_x^2 L(\bar{x}; \lambda) d > 0,$$

dla każdego $d \neq 0$ spełniającego warunek

$$Dh_j(\bar{x})d = 0,$$

to \bar{x} jest ścisłym rozwiązaniem lokalnym problemu (9.2).

WNIOSEK. 9.8 Jeśli funkcja f w otoczeniu punktu \bar{x} jest klasy C^2 i jest ściśle wypukła a funkcje h_j są afiniczne, to warunek I rzędu jest warunkiem wystarczającym, aby punkt \bar{x} był ścisłym rozwiązaniem lokalnym problemu (9.2).

9.2 Optymalizacja wypukła

DEFINICJA. 9.9 Funkcję $f: W \rightarrow \mathbb{R}$, gdzie $W \subset \mathbb{R}^n$ wypukły, nazwiemy:

- **wypukłą**, jeśli dla każdego $x, y \in W$ i $\lambda \in (0, 1)$ zachodzi

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

- **ściśle wypukłą**, jeśli dla każdego $x, y \in W$, $x \neq y$ i $\lambda \in (0, 1)$ zachodzi

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

Przypomnijmy kilka ważnych własności funkcji wypukłych.

TWIERDZENIE. 9.10 Jeśli funkcja f jest wypukła, to zbiór poziomicowy

$$W_\alpha(f) = \{x \in W : f(x) \leq \alpha\}, \quad \alpha \in \mathbb{R}.$$

jest wypukły dla dowolnego α .

TWIERDZENIE. 9.11 (Twierdzenie o hiperpłaszczyźnie podpierającej) Jeśli f jest wypukła, to w każdym punkcie $\bar{x} \in \text{int } W$ istnieje hiperpłaszczyzna podpierająca, tzn. istnieje $\xi \in \mathbb{R}^n$, takie że

$$f(x) \geq f(\bar{x}) + \xi^\top (x - \bar{x}), \quad \forall x \in W.$$

Jeśli f jest ściśle wypukła, to

$$f(x) > f(\bar{x}) + \xi^\top (x - \bar{x}), \quad \forall x \in W \setminus \{\bar{x}\}.$$

Jeśli f jest różniczkowalna w \bar{x} , to w obu powyższych nierównościach możemy przyjąć $\xi = Df(\bar{x})$.

Dla funkcji wypukłych rozwiązanie problemu (9.1) ma bardzo prostą postać.

TWIERDZENIE. 9.12 Niech $W \subset \mathbb{R}^n$ wypukły oraz $f: W \rightarrow \mathbb{R}$ wypukła. Jeśli f jest różniczkowalna w punkcie $\bar{x} \in W$, to mamy następującą równoważność: \bar{x} jest rozwiązaniem lokalnym (9.1) wtedy i tylko wtedy, gdy $Df(\bar{x})(x - \bar{x}) \geq 0$ dla każdego $x \in W$.

Jeśli $\bar{x} \in W$ jest rozwiązaniem lokalnym (9.1), to:

- \bar{x} jest rozwiązaniem globalnym,
- jeśli f jest ściśle wypukła, to \bar{x} jest jedynym rozwiązaniem globalnym.

9.3 Optymalizacja z ograniczeniami mieszanymi

DEFINICJA. 9.13 Problem optymalizacyjny z **ograniczeniami mieszanymi** ma postać:

$$\begin{cases} \min_x f(x), \\ g_i(x) \leq 0, \quad i = 1, \dots, m, \\ h_j(x) = 0, \quad j = 1, \dots, l, \\ x \in \mathbb{X}, \end{cases} \quad (9.3)$$

gdzie $\mathbb{X} \subset \mathbb{R}^n$ jest zbiorem otwartym i $f, g_1, \dots, g_m, h_1, \dots, h_l : \mathbb{X} \rightarrow \mathbb{R}$. Zbiór punktów dopuszczalnych zadany jest następująco

$$W = \{x \in \mathbb{X}: g_1(x) \leq 0, \dots, g_m(x) \leq 0, h_1(x) = 0, \dots, h_l(x) = 0\}.$$

Funkcje g_i nazywane są **ograniczeniami nierównościami** a funkcje h_j , **ograniczeniami równościami**.

Dla optymalizacji z ograniczeniami mieszanymi musimy zmodyfikować definicję stożka kierunków stycznych dla ograniczeń zlinearyzowanych.

DEFINICJA. 9.14 Niech $\bar{x} \in W$, g_i różniczkowalne w \bar{x} dla ograniczeń aktywnych $i \in I(\bar{x})$ oraz h_j są różniczkowalne w \bar{x} dla $j = 1, \dots, l$. **Stożkiem kierunków stycznych dla ograniczeń zlinearyzowanych** nazywamy zbiór

$$T_{lin}(\bar{x}) = \{d \in \mathbb{R}^n: \forall i \in I(\bar{x}) \quad Dg_i(\bar{x})d \leq 0, \quad \forall j = 1, \dots, l \quad Dh_j(\bar{x})d = 0\},$$

gdzie zbiorem **ograniczeń aktywnych** w punkcie $\bar{x} \in W$ nazywamy zbiór

$$I(\bar{x}) = \{i \in \{1, 2, \dots, m\}: g_i(\bar{x}) = 0\}.$$

Definicja regularnego punktu ograniczeń pozostaje taka sama jak dla ograniczeń równościowych, tzn. \bar{x} jest regularnym punktem ograniczeń, jeśli $T(\bar{x}) = T_{lin}(\bar{x})$. Prawdziwy pozostaje także lemat 9.5.

LEMAT. 9.15 *Jeśli w punkcie $\bar{x} \in W$ funkcje g_i , $i = 1, \dots, m$ oraz funkcje h_j , $j = 1, \dots, l$, są funkcjami afinicznymi, to punkt \bar{x} jest regularnym punktem ograniczeń.*

TWIERDZENIE. 9.16 (Twierdzenia Kuhna-Tuckera) *Niech \bar{x} będzie rozwiązaniem lokalnym (9.3). Jeśli funkcje f , g_i , $i \in I(\bar{x})$, oraz h_j , $j = 1, \dots, l$, są różniczkowalne w \bar{x} oraz $T(\bar{x}) = T_{lin}(\bar{x})$, to istnieją $\mu \in [0, \infty)^m$ oraz $\lambda \in \mathbb{R}^l$, takie że*

$$\begin{cases} Df(\bar{x}) + \sum_{i \in I(\bar{x})} \mu_i Dg_i(\bar{x}) + \sum_{j=1}^l \lambda_j Dh_j(\bar{x}) = 0, \\ \mu_i g_i(\bar{x}) = 0, \quad i = 1, 2, \dots, m. \end{cases} \quad (9.4)$$

TWIERDZENIE. 9.17 (Warunek dostateczny) *Niech $\bar{x} \in W$. Załóżmy, że występujące w problemie (9.3) funkcje spełniają warunki:*

1. *funkcje g_i są różniczkowalne w \bar{x} i wypukłe,*
2. *funkcje h_j są różniczkowalne w \bar{x} i afiniczne,*
3. *funkcja f jest różniczkowalna w \bar{x} i wypukła.*

Jeśli istnieją wektory $\mu \in [0, \infty)^m$ i $\lambda \in \mathbb{R}^l$ spełniające warunek I rzędu

$$\begin{cases} Df(\bar{x}) + \sum_{i \in I(\bar{x})} \mu_i Dg_i(\bar{x}) + \sum_{j=1}^l \lambda_j Dh_j(\bar{x}) = 0, \\ \mu_i g_i(\bar{x}) = 0, \quad i = 1, 2, \dots, m, \end{cases}$$

to punkt \bar{x} jest rozwiązaniem globalnym problemu (9.3).

Uwaga. 39 *Przytoczone twierdzenie nie jest najogólniejszym, ale takie sformułowanie twierdzenia w zupełności wystarczy do rozwiązania problemu optymalizacji portfela.*

Poniższy wniosek z twierdzenia Kuhna-Tuckera jest bardzo przydatny przy optymalizacji portfela inwestycyjnego z ograniczeniami na krótką sprzedaż.

WNIOSEK. 9.18 *Niech dane będzie zadanie optymalizacyjne:*

$$\begin{cases} \min_x f(x), \\ g_i(x) \leq 0, \quad i = 1, \dots, m, \\ h_j(x) = 0, \quad j = 1, \dots, l, \\ x \geq 0, \quad x \in \mathbb{R}^n. \end{cases}$$

Jeśli \bar{x} jest jego rozwiązaniem lokalnym, to warunek I rzędu przyjmuje w punkcie \bar{x} postać:

$$\begin{cases} D_x L(\bar{x}; \mu, \lambda) \geq 0, \\ x_k D_{x_k} L(\bar{x}; \mu, \lambda) = 0, \quad k = 1, \dots, n, \\ \mu_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m, \\ h_j(\bar{x}) = 0, \quad j = 1, \dots, l, \\ \mu \geq 0, \end{cases}$$

gdzie funkcja Lagrange'a dana jest wzorem

$$L(x; \mu, \lambda) = f(x) + \sum_{i=1}^m \mu_i g_i(x) + \sum_{j=1}^l \lambda_j h_j(x).$$

9.4 Punkty siodłowe funkcji Lagrange'a

DEFINICJA. 9.19 Niech A, B będą dowolnymi zbiorami, zaś $h : A \times B \rightarrow \mathbb{R}$ funkcją. Punkt $(\bar{x}, \bar{\mu}) \in A \times B$ nazywamy punktem siodłowym funkcji h , jeśli

$$h(\bar{x}, \mu) \leq h(\bar{x}, \bar{\mu}) \leq h(x, \bar{\mu}), \quad \forall x \in A, \mu \in B.$$

Okazuje się, że punkt siodłowy funkcji Lagrange'a jest związany z rozwiązaniem globalnym problemu optymalizacji z ograniczeniami nierównościami:

$$\begin{cases} f(x) \rightarrow \min, \\ g_i(x) \leq 0, \quad i = 1, \dots, m, \\ x \in \mathbb{X}. \end{cases} \quad (9.5)$$

Przypomnijmy, że przez W oznaczamy zbiór punktów dopuszczalnych, tj.

$$W = \{x \in \mathbb{X} : g_1(x) \leq 0, \dots, g_m(x) \leq 0\}.$$

Poniższe twierdzenia pokazują, że przy pewnych dodatkowych założeniach spełnianie tylko jednej z nierówności definiującej punkt siodłowy wystarczy do istnienia rozwiązania globalnego problemu (9.5).

TWIERDZENIE. 9.20 Jeśli istnieje $\bar{\mu} \in [0, \infty)^m$ oraz $\bar{x} \in W$, takie że

$$\begin{cases} f(\bar{x}) + \sum_{i=1}^m \bar{\mu}_i g_i(\bar{x}) \leq f(x) + \sum_{i=1}^m \bar{\mu}_i g_i(x) \quad \forall x \in W, \\ \bar{\mu}_i g_i(\bar{x}) = 0 \quad \text{dla } i = 1, \dots, m, \end{cases}$$

to \bar{x} jest rozwiązaniem (9.5).

TWIERDZENIE. 9.21 Niech $\bar{x} \in W$ będzie rozwiązaniem (9.5) oraz istnieje $x \in \mathbb{X}$, taki że $g_i(x) < 0$ dla $i = 1, \dots, m$, to istnieje $\bar{\mu} \in [0, \infty)^m$, takie że

$$\begin{cases} f(\bar{x}) + \sum_{i=1}^m \bar{\mu}_i g_i(\bar{x}) \leq f(x) + \sum_{i=1}^m \bar{\mu}_i g_i(x) \quad \forall x \in \mathbb{X}, \\ \bar{\mu}_i g_i(\bar{x}) = 0 \quad \text{dla } i = 1, \dots, m. \end{cases}$$

9.5 Funkcja wrażliwości

Rozważmy problem optymalizacyjny z ograniczeniami równościowymi:

$$\begin{cases} f(x) \rightarrow \min, \\ h_j(x) = 0, \quad j = 1, \dots, l, \\ x \in \mathbb{X}, \end{cases} \quad (9.6)$$

gdzie $\mathbb{X} \subset \mathbb{R}^n$ jest zbiorem otwartym i $f, h_1, \dots, h_l : \mathbb{X} \rightarrow \mathbb{R}$. Dla uproszczenia notacji połączmy $h(x) = [h_1(x), \dots, h_l(x)]^\top$. Rozważmy problem zaburzony, tzn.

$$\begin{cases} f(x) \rightarrow \min, \\ h(x) = z, \\ x \in \mathbb{X}, \end{cases} \quad (9.7)$$

gdzie $z \in \mathbb{R}^l$.

TWIERDZENIE. 9.22 Niech \bar{x} będzie rozwiązaniem lokalnym problemu (9.6), zaś $\bar{\lambda}$ wektorem jego mnożników Lagrange'a. Załóżmy, że funkcje f, h_1, \dots, h_l są klasy C^2 na otoczeniu \bar{x} , gradienty funkcji h_1, \dots, h_l są liniowo niezależne oraz

$$d^\top D_x^2 L(\bar{x}, \bar{\lambda}) d > 0 \quad (9.8)$$

dla $d \in \mathbb{R}^n \setminus 0$ spełniających $Dh_j(\bar{x})d = 0$, $j = 1, \dots, l$. Wówczas istnieje otoczenie \tilde{O} punktu $0 \in \mathbb{R}^l$ oraz funkcje $x(z)$ i $\lambda(z)$ działające z \tilde{O} do \mathbb{X} , klasy C^1 , takie że $x(0) = \bar{x}$, $\lambda(0) = \bar{\lambda}$. Ponadto funkcja $p(z) = f(x(z))$ jest ściśłym rozwiązaniem lokalnym problemu zaburzonego (9.7) i dla wszystkich $z \in \tilde{O}$ mamy

$$Dp(z) = -\lambda(z)^\top.$$

Funkcję $p(z)$ nazywamy **funkcją perturbacji** dla problemu z ograniczeniami równościowymi. Wartość pochodnej w zerze $Dp(0) = -\bar{\lambda}^\top$ nazywa się **wektorem wrażliwości** problemu (9.6).

Literatura

- [1] Black, F., Litterman, R. – Global asset allocation with equities, bonds and currencies, Goldman Sachs, Fixed Income Research, October 1991; przedruk jako – Global portfolio optimization, *Financial Analysts J.*, **48** (1992), 28–43.
- [2] Dybvig, P. H. – Short sales restrictions and kinks on the mean variance frontier, *Journal of Finance*, **39** (1984), 239–244.
- [3] Föllmer, H., Schied, A. – *Stochastic Finance*, Walter de Gruyter, 2002.
- [4] Krokmal, P., Palmquist, J., Uryasev, S. – Portfolio optimization with conditional value-at-risk objective and constraints, *Journal of Risk*, **4** (2002), 43–68.
- [5] Markowitz, H. – Portfolio selection, *Journal of Finance*, **7** (1952), 77–91.
- [6] Mertens, E., Zimmermann, H. – International TAA strategies and currencies, Working Paper No. 1/02 WWZ/Department of Finance 2002.
- [7] Palczewski, A., Palczewski, J. – Theoretical and empirical estimates of mean-variance portfolio sensitivity, *European Journal of Operational Research*, **234** (2014), 402–410.
- [8] Prigent, J-L. – *Portfolio Optimization and Performance Analysis*, Chapman and Hall 2007.
- [9] Rockafellar, R. T., Uryasev, S., Zabarankin, M. – Master funds in portfolio analysis with general deviation measures, *Journal of Banking & Finance*, **30** (2006), 743–778.
- [10] Steinbach, M. C. – Markowitz revisited: Mean-variance models in financial portfolio analysis, *SIAM Review*, **43** (2001), 31–85.
- [11] Szegö, G. P. – *Portfolio Theory with Application to Bank Asset Management*, Academic Press, 1980.