



## Sławomir Jarek

Uniwersytet Ekonomiczny w Katowicach  
Wydział Informatyki i Komunikacji  
Katedra Badań Operacyjnych  
slawomir.jarek@ue.katowice.pl

# ANALIZA PORTFELA O MAKSYMALNEJ PRZEWIDYWALNOŚCI

**Streszczenie:** W nowoczesnej teorii portfelowej jest rozważany problem takiego doboru aktywów, aby wybrane statystyki portfela, takie jak np. oczekiwana wartość czy wariancja, przyjmowały określone wartości. W procesie tym explicite są pomijane własności predyktorów stóp zwrotu w takim sensie, że skład portfela nie zależy bezpośrednio od czynników mających pośredni wpływ na zmienność portfela. W artykule zaproponowano trzy algorytmy tworzenia portfela, które kładą główny nacisk na zmienność wyjaśnianą przez model prognostyczny użyty do szacowania przyszłych wartości stóp zwrotu. W algorytmach tych wariancja wyjaśniana przez model prognostyczny jest użyta do wyznaczania poszczególnych wag aktywów w portfelu.

**Słowa kluczowe:** portfel papierów wartościowych, portfele własne, algorytmy.

## Wprowadzenie

Celem artykułu jest opracowanie algorytmów umożliwiających wyznaczenie portfela o maksymalnej przewidywalności mierzonej z wykorzystaniem współczynnika determinacji modelu prognostycznego. Zakładamy przy tym, że model ten będzie wyrażony w postaci liniowej. Po oszacowaniu parametrów struktury modelu możliwe staje się oszacowanie macierzy wariancji-kowariancji prognoz stóp zwrotu. Macierz ta pełni istotną rolę w opisanych w kolejnych punktach pracy algorytmach wyznaczania portfela papierów wartościowych.

Konstruując model prognostyczny, stajemy przed problemem wyboru czynników mających wpływ na kształtowanie się stóp zwrotu oraz formułowanie prognoz tych czynników. W tym kontekście pojawia się teoria błędzenia losowego oraz hipoteza efektywnego rynku.

## 1. Teoria błędzenia losowego

Budując portfel inwestycyjny, kierujemy się pewnymi oczekiwaniami odnośnie do przyszłych stóp zwrotu. W ubiegłym wieku powstała hipoteza efektywnego rynku (EMH – ang. *efficient market hypothesis*) starająca się powiązać bieżące ceny akcji z dostępnymi informacjami w momencie formułowania prognozy. W ramach prowadzonych badań [Fama, 1970] nad EMH badana jest także hipoteza błędzenia losowego [Samuelson, 1972] (RWH – ang. *random walk hypothesis*), w ramach której zakłada się, że ceny akcji zachowują się w sposób losowy, który można wyrazić w postaci sumy dwóch składników, przy czym pierwszy z nich to wartość stopy zwrotu w poprzednim okresie, a drugi jest równy pewnej zmiennej losowej o wartości oczekiwanej zero. Jeśli RWH nie zostanie odrzucona, to nie ma także podstaw do odrzucenia EMH. Początkowo EMH i RWH cieszyły się przychylnością i w niewielkim stopniu podawano je w wątpliwość. Szczególnie intensywne badania [Fama, 1970] tych hipotez przeprowadzono na podstawie analiz stóp zwrotu aktywów notowanych na giełdach w USA oraz Wielkiej Brytanii. Niestety coraz częściej obserwowano na rynkach finansowych nietypowe reakcje inwestorów i wiązano je z określonymi anomaliami (efekt stycznia, efekt poniedziałków, *Sell in May*, *Santa Claus rally* oraz inne anomalie kalendarzowe, związane np. z prezentowaniem wyników przez fundusze inwestycyjne). Takie nietypowe zachowania inwestorów, które przeczą EMH, są obecnie opisywane w ramach badań nad finansami behawioralnymi.

## 2. Ogólne założenia dotyczące formułowania prognoz stóp zwrotu oraz struktury macierzy wariancji-kowariancji

Wśród krytyków RWH są autorzy Lo i MacKinlay [2002]. Opisują oni ciekawe obserwacje dotyczące własności predyktywnych akcji spółek notowanych na giełdach papierów wartościowych. W ramach prowadzonych badań nad wyceną aktywów kapitałowych istotną rolę odgrywają modele [Fama, French, 1992, 1993, 1998] starające się wyjaśnić zmienność badanych aktywów. Na tym obszarze szczególnie wkład wnieśli Fama i French. W swoich pracach proponowali modele wieloczynnikowe, których konstrukcja uwzględniała rozmiar badanych spółek (SMB), czynnik HML bazujący na wskaźniku analizy fundamentalnej (P/BV) oraz stopę zwrotu z portfela rynkowego, a także czynniki bazujące na stopach zwrotu z obligacji. Autorzy tych prac wykazywali, że tak wybrane czynniki dobrze wyjaśniają zmienność stóp zwrotu oraz są w niewielkim stopniu ze sobą skorelowane.

Bazując na powyższej obserwacji, zakładamy, że stopy zwrotu akcji są generowane przez pewien proces stochastyczny opisany zależnością (1):

$$R_i = \alpha_i + \beta_i \mathbf{F} + \varepsilon_i \quad (1)$$

gdzie:

$R_i$  – stopa zwrotu  $i$ -tej akcji,

$\mathbf{F}$  – wektor czynników o określonym rozkładzie,

$\alpha_i, \beta_i$  – nieznane współczynniki modelu (w wersji populacyjnej),

$\varepsilon_i$  – składniki losowe  $i$ -tej akcji (zasób zmienności specyficznej).

Zakładamy przy tym, że  $E(\varepsilon_i) = 0$ , czynniki  $F_j$  są nieskorelowane ze składnikiem losowym  $\text{cov}(\varepsilon_i, F_j) = 0$  oraz składniki losowe stóp zwrotu różnych akcji nie są ze sobą skorelowane  $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ . Przy tych założeniach łatwo możemy dostrzec, że struktura [Perold, 1984] macierzy wariancji-kowariancji może być opisana za pomocą zależności (2), w której w diagonalnej macierzy  $\mathbf{E}$  na głównej przekątnej są wariancje składników losowych:

$$\mathbf{V}^* = \text{cov}(\mathbf{R}) = \beta^T \text{cov}(\mathbf{F})\beta + \mathbf{E} \quad (2)$$

Niestety nie znamy rzeczywistych współczynników modelu (1) oraz macierzy wariancji-kowariancji  $\mathbf{V}^*$  w całej populacji. Jednakże na podstawie  $n$  obserwacji w próbie możemy oszacować macierz wariancji-kowariancji  $\mathbf{V}$  oraz wartości współczynników w modelu (1), a także sformułować prognozy czynników  $\mathbf{F}$  oznaczone jako  $\hat{\mathbf{F}}$  i wyznaczyć prognozy  $\hat{\mathbf{R}}$  stóp zwrotu. Ponadto na podstawie wartości historycznych  $\hat{\mathbf{R}}$  możemy wyznaczyć macierz wariancji-kowariancji  $\hat{\mathbf{V}}$  dla prognoz stóp zwrotu.

Osobna dyskusja powinna dotyczyć stałości w czasie parametrów występujących w równaniach (1) i (2). W niniejszym artykule zakładamy, że w chwili formułowania prognozy parametry te się nie zmieniają i są wyznaczone na podstawie dostępnych w tym momencie informacji.

### 3. Parametry konstruowanych portfeli

W artykule badamy portfele jednoznacznie wyznaczone przez ustalone wagi poszczególnych akcji  $\mathbf{x}$ , które cechują się własnością (3):

$$\sum_i |x_i| = 1 \quad (3)$$

Jeśli zależność (3) nie jest spełniona, to będziemy skalować otrzymany portfel  $\mathbf{x}$  zgodnie z przekształceniem (4):

$$x_i = \frac{x_i}{\sum_i |x_i|} \quad (4)$$

Znając wagi  $\mathbf{x}$  poszczególnych akcji w portfelu, możemy wyznaczyć stopy zwrotu z portfela (5) oraz prognozę stopy zwrotu portfela (6), a także odpowiednio wariancję stopy zwrotu portfela (7) oraz wariancję prognozy stopy zwrotu portfela (8):

$$R_r = \mathbf{x}^T \mathbf{R} \quad (5)$$

$$R_p = \mathbf{x}^T \hat{\mathbf{R}} \quad (6)$$

$$V_r = \mathbf{x}^T \mathbf{V} \mathbf{x} \quad (7)$$

$$V_p = \mathbf{x}^T \hat{\mathbf{V}} \mathbf{x} \quad (8)$$

W powyższych zależnościach, zgodnie z przyjętymi w poprzednim punkcie oznaczeniami, przez  $\hat{\mathbf{R}}$  oznaczono prognozy stóp zwrotu, a przez  $\hat{\mathbf{V}}$  macierz wariancji-kowariancji prognoz stóp zwrotu.

## 4. Wyznaczanie portfela o maksymalnej przewidywalności

### 4.1. Motywacja

Na podstawie zdefiniowanych wcześniej wielkości na gruncie nowoczesnej teorii portfelowej są wyznaczone wagi  $\mathbf{x}$  portfela akcji. Zbiór rozwiązań dopuszczalnych w tym podejściu jest rozpięty w przestrzeni decyzyjnej na zmiennych  $x_i$ , natomiast właściwości portfela, takie jak oczekiwana stopa zwrotu czy wariancja/odchylenie standardowe są badane w przestrzeni kryterialnej. W nowoczesnej teorii portfelowej są wyznaczone portfele sprawne, czyli takie, którym w przestrzeni kryterialnej odpowiadają punkty niezdominowane. O ile przestrzeń decyzyjna nie budzi większych wątpliwości, o tyle na kształt obrazu rozwiązań dopuszczalnych w przestrzeni kryterialnej istotny wpływ ma łączny rozkład stóp zwrotu z akcji. W wielu pracach przyjmuje się, że stopy zwrotu cechuje wielowymiarowy rozkład normalny. Niestety w praktyce najczęściej nie znamy rzeczywistego rozkładu stóp zwrotu, a w szczególności nie znamy parametrów tego rozkładu dla całej populacji. Trudno nawet wykazać, że rozkłady stóp zwrotu są symetryczne. W związku z tym szacunki tych parametrów na podstawie pobranej próby są

obarczone błędami. Wyniki otrzymywane na podstawie analiz modeli bazujących na tych założeniach bywają w praktycznych zastosowaniach niezadowolające.

W niniejszym artykule omawiamy metodę konstruowania wektora  $x$  w taki sposób, aby prognozy stóp zwrotu portfela były systematycznie jak najbliższe rzeczywistym realizacjom tej zmiennej losowej. Taki portfel będziemy nazywać portfelem o maksymalnej przewidywalności (MPP – ang *Maximally Predictable Portfolio*<sup>1</sup>). W pracy Lo i MacKinlay [2002] wprawdzie opisano możliwość konstruowania MPP, jednakże nie zdefiniowano metod umożliwiających efektywne wyznaczanie MPP. W rozdziale 9 tej pracy podano załączek równania (26) bez wyjaśnienia przesłanek stojących u podstaw tej zależności. W tym kontekście autorzy tamtej pracy odwołali się jedynie do pracy Gantmachera z 1959 r. (dotyczącej teorii macierzy) oraz pracy Boxa i Tiao z 1977 r. opisującej wielowymiarowe szeregi czasowe, w których nie rozwijano zagadnienia dotyczącego budowy MPP. Ta istotna luka w literaturze skłoniła autora tego opracowania do przeprowadzenia analizy, której wyniki zawarto w dalszej części tego rozdziału.

## 4.2. Założenia MPP

Punktem wyjścia wyznaczania MPP jest kowariancja w wersji populacyjnej (9):

$$\max_x Cov(R_r, R_p) \quad (9)$$

$$|E(R_r - ER_r)^T (R_p - ER_p)| \leq \sigma_{Rr} \sigma_{Rp} \quad (10)$$

Jak wiadomo, kowariancja występująca w (9) przyjmuje wartości spełniające zależność (10). Rozwiązując zatem problem (9) oraz biorąc pod uwagę własność (10), trudno jest uznać, na ile dobry jest MPP. Jednakże dzieląc warunek (10) przez jego prawą stronę, uzyskamy miarę (11) przyjmującą wartości w przedziale od -1 do 1:

$$\rho = \frac{E(R_r - ER_r)^T (R_p - ER_p)}{\sigma_{Rr} \sigma_{Rp}} \quad (11)$$

$$|\rho| \leq 1 \quad (12)$$

Ponieważ nie znamy ani wartości oczekiwanych stóp zwrotu, ani wariancji obu portfeli, w dalszej części będziemy stosowali miarę (17) wyznaczoną na podstawie  $n$  obserwacji w próbie, zwaną współczynnikiem korelacji liniowej Pearsona:

<sup>1</sup> Nazwa zaproponowana w pracy Lo, MacKinlay [2002].

$$\bar{R}_r = \frac{1}{n} \sum_{i=1}^n Rr_i \quad (13)$$

$$s_{Rr} = \sqrt{\frac{1}{n-1} \sum (Rr_i - \bar{R}_r)^2} \quad (14)$$

$$\bar{R}_p = \frac{1}{n} \sum_{i=1}^n Rp_i \quad (15)$$

$$s_{Rp} = \sqrt{\frac{1}{n-1} \sum (Rp_i - \bar{R}_p)^2} \quad (16)$$

$$r = \frac{1}{s_{Rr}s_{Rp}(n-1)} \sum_{i=1}^n (Rr_i - \bar{R}_r)(Rp_i - \bar{R}_p) \quad (17)$$

W dalszej części artykułu będziemy bazować na następujących założeniach:

- istnieje co najwyżej liniowy związek pomiędzy  $R_r$  i  $R_p$  (nie jest to zależność nieliniowa),
- $E(R_r - R_p) = 0$ ,
- zachodzi warunek (18).

Warunek (18) zakłada brak zależności pomiędzy resztami losowymi modelu prognozy a wartościami prognozy stóp zwrotu portfela.

$$\text{Cov}(R_r - R_p, R_p) = 0 \quad (18)$$

### 4.3. Konstrukcja MPP

Dążąc do wyznaczenia MPP, dodamy do pierwszego czynnika w znaku sumy w (17) wartość  $R_p - R_p$  i otrzymamy  $r$  w postaci (19). Następnie korzystając z zależności (18), możemy zredukować w (19) pierwszy składnik występujący w ostatniej sumie oraz podstawiając wielkości zdefiniowane w (7) i (8), możemy otrzymać zależność (20):

$$\begin{aligned} r &= \frac{1}{s_{Rr}s_{Rp}(n-1)} \sum_{i=1}^n (Rr_i - \bar{R}_r + Rp_i - Rp_i)(Rp_i - \bar{R}_p) = \\ &= \frac{1}{s_{Rr}s_{Rp}(n-1)} \sum_{i=1}^n [(Rr_i - Rp_i)(Rp_i - \bar{R}_p) + (Rp_i - \bar{R}_p)^2] \end{aligned} \quad (19)$$

$$r = \frac{1}{s_{Rr}s_{Rp}(n-1)} \sum_{i=1}^n (Rp_i - \bar{R}_p)^2 = \frac{s_{Rp}^2}{s_{Rr}s_{Rp}} = \frac{s_{Rp}}{s_{Rr}} = \sqrt{\frac{V_p}{V_r}} \quad (20)$$

Ze względów obliczeniowych wygodnie jest pozbyć się w (20) pierwiastka, obustronnie podnosząc to równanie do kwadratu. Ponownie korzystając z zależności (7) i (8), otrzymujemy w (21) dobrze znany z modelu regresji liniowej współczynnik determinacji  $r^2$ . Wartości tego współczynnika mieszczą się w przedziale od 0 do 1 i można je interpretować jako tę część wariancji portfela, która wynika z jej zależności od uwzględnionych w modelu prognostycznym informacji zawartych w czynnikach  $\mathbf{F}$ . Im wartość  $r^2$  jest większa, tym MPP lepiej pokrywa całkowitą zmienność stopy zwrotu portfela, czyli że prognozowane stopy zwrotu z portfela przyjmują wartości bliskie ich rzeczywistym realizacjom.

$$r^2 = \frac{V_p}{V_r} = \frac{\mathbf{x}^T \hat{\mathbf{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{V} \mathbf{x}} \quad (21)$$

W celu wyznaczenia MPP z zależności (21) przyjmujemy, że znana jest optymalna wartość  $r^2$  i oznaczamy tę ustaloną wartość jako  $\lambda$ . Problem wyznaczenia MPP został zdefiniowany w (22). Jak łatwo zauważyć ze względu na ułamek, w którym  $\mathbf{x}$  występuje zarówno w liczniku, jak i mianowniku, składowe wektora  $\mathbf{x}$  występujące w (21) mogą być pomnożone przez dowolną niezerową wartość  $z$ , nie zmieniając wartości współczynnika determinacji  $r^2$ . W szczególności dopuszczalne jest przekształcenie zdefiniowane w (4):

$$\max_{\mathbf{x}} r^2 = \max_{\mathbf{x}} \frac{\mathbf{x}^T \hat{\mathbf{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{V} \mathbf{x}} = \lambda \quad (22)$$

Skoro mamy ustaloną wartość współczynnika determinacji, to możemy przekształcić (22) do postaci funkcji (23), która przyjmuje tym większe wartości, im wartość  $r^2$  jest większa:

$$\max_{\mathbf{x}} \mathbf{x}^T \hat{\mathbf{V}} \mathbf{x} - \lambda \mathbf{x}^T \mathbf{V} \mathbf{x} \quad (23)$$

Wartość funkcji (23) przyjmuje wartość największą w punkcie  $\mathbf{x}$ , w którym jej gradient przyjmuje wartości zerowe. Wykonując odpowiednie przekształcenia, otrzymujemy w (24) układ  $n$  równań spełniający ten warunek. Przenosząc składniki z  $\lambda$  na drugą stronę równań, otrzymujemy w (25) *uogólniony problem własny*. Funkcja (23) będzie przyjmowała wartość maksymalną w punkcie  $\mathbf{x}$  spełniającym warunek (24), jeśli hesjan funkcji (23)  $\mathbf{H} = 2\hat{\mathbf{V}} - 2\lambda\mathbf{V}$  będzie niedodatnio określony, a zatem będzie spełniona zależność  $\Lambda_{\mathbf{y}} 2\mathbf{y}^T (\hat{\mathbf{V}} - \lambda\mathbf{V}) \mathbf{y} \leq 0$ :

$$\nabla_x f(\mathbf{x}) = 2\hat{\mathbf{V}}\mathbf{x} - 2\lambda\mathbf{V}\mathbf{x} = \mathbf{0} \quad (24)$$

$$\hat{\mathbf{V}}\mathbf{x} = \lambda\mathbf{V}\mathbf{x} \quad (25)$$

Mnożąc lewostronnie układ równań (25) przez macierz odwrotną do macierzy wariancji-kowariancji  $\mathbf{V}$ , otrzymujemy w (26) *standardowy problem własny*, w którym  $\lambda$  jest wartością własną, a  $\mathbf{x}$  wektorem własnym macierzy  $\mathbf{V}^{-1}\hat{\mathbf{V}}$ :

$$\mathbf{V}^{-1}\hat{\mathbf{V}}\mathbf{x} = \lambda\mathbf{x} \quad (26)$$

W celu wyznaczenia wartości  $\mathbf{x}$  spełniających układ równań własnych (26) należy wyznaczyć wartości własne macierzy  $\mathbf{V}^{-1}\hat{\mathbf{V}}$  i odpowiadające im wektory własne. Poszukiwanym rozwiązaniem  $\mathbf{x}$  jest wektor własny związany z największą wartością własną  $\lambda$  macierzy  $\mathbf{V}^{-1}\hat{\mathbf{V}}$ .

Niestety opisane powyżej postępowanie związane z wykorzystaniem odwrotności macierzy  $\mathbf{V}$  może prowadzić do utraty stabilności w trakcie wykonywania obliczeń. Jeśli macierz  $\mathbf{V}$  jest macierzą symetryczną i dodatnio określoną, to istnieje inny wariant rozwiązywania problemu (25) bazujący na rozkładzie Choleskiego (27). Zakładamy przy tym, że macierz  $\mathbf{L}$  nie jest macierzą osobliwą:

$$\mathbf{V} = \mathbf{L}\mathbf{L}^T \quad (27)$$

Podstawiając za  $\mathbf{V}$  w (25) wielkość określoną w (27), otrzymujemy *uogólniony problem własny* w postaci (28), który po lewostronnym pomnożeniu przez macierz odwrotną do  $\mathbf{L}$  może być zapisany w postaci (29):

$$\hat{\mathbf{V}}\mathbf{x} = \lambda\mathbf{L}\mathbf{L}^T\mathbf{x} \quad (28)$$

$$\mathbf{L}^{-1}\hat{\mathbf{V}}\mathbf{x} = \lambda\mathbf{L}^{-1}\mathbf{L}\mathbf{L}^T\mathbf{x} = \lambda\mathbf{L}^T\mathbf{x} \quad (29)$$

Wygodnie jest oznaczyć wektor znajdujący się po prawej stronie (29) jako nową zmienną. Odpowiednie definicje podano w (30) i (31), w których przez  $\mathbf{L}^{-T}$  oznaczono macierz odwrotną do transpozycji macierzy  $\mathbf{L}$ .

$$\mathbf{y} = \mathbf{L}^T\mathbf{x} \quad (30)$$

$$\mathbf{x} = \mathbf{L}^{-T}\mathbf{y} \quad (31)$$

Podstawiając odpowiednio wartości z (30) i (31) do (29), otrzymujemy *standardowy problem własny* w postaci (32) zachowującej wartość  $\lambda$  oraz z innym wektorem własnym  $\mathbf{y}$ :

$$\mathbf{L}^{-1}\hat{\mathbf{V}}\mathbf{L}^{-T}\mathbf{y} = \lambda\mathbf{y} \quad (32)$$



Po wyznaczeniu wektora własnego  $\mathbf{y}$  dla macierzy  $\mathbf{L}^{-1}\hat{\mathbf{V}}\mathbf{L}^{-T}$  można odtworzyć poszukiwany wektor  $\mathbf{x}$  z zależności (31).

#### 4.4. Algorytm wyznaczania MPP

Problem wyznaczania MPP sprowadza się zatem do rozwiązania *standardowego problemu własnego* (26) lub (32), a następnie przeskalowania otrzymanego wektora  $\mathbf{x}$  zgodnie z przekształceniem (4). Otrzymane w ten sposób rozwiązanie jest równoważne rozwiązaniom problemów (9) oraz (22). Zadanie to nieco się komplikuje ze względu na fakt, iż macierz  $\mathbf{V}^{-1}\hat{\mathbf{V}}$  nie jest macierzą symetryczną, pomimo tego, że obie macierze  $\mathbf{V}$  i  $\hat{\mathbf{V}}$  są symetryczne. Wartości własne macierzy symetrycznych przyjmują wartości rzeczywiste, natomiast dla macierzy niesymetrycznych mogą one przyjmować wartości zespolone. Z przeprowadzonych eksperymentów numerycznych wynika jednak, że stosunkowo często  $\mathbf{x}$  i  $\lambda$  przyjmują wartości rzeczywiste, a w przypadku otrzymania rozwiązań w zbiorze liczb zespolonych dla największych wartości własnych część urojona nie występowała.

Wyznaczanie wartości własnych może wymagać wielu nakładów obliczeniowych i, jak to wcześniej zostało pokazane, nie zachodzi potrzeba wyznaczania wszystkich wartości własnych i odpowiadających im wektorów własnych. Skoro z konstrukcji tego zagadnienia wynika, że maksymalna wartość własna wynosi 1, to można zastosować metody numeryczne wyznaczające wartość własną najbliższą 1 i odpowiadający jej wektor własny. Jedną z metod umożliwiających wyznaczenie największej co do modułu wartości własnej jest metoda potęgowa. Polega ona na sekwencyjnym i wielokrotnym mnożeniu wektora startowego  $\mathbf{x}$  przez macierz, dla której poszukujemy wartości i wektora własnego. W trakcie kolejnych iteracji wektor  $\mathbf{x}$  zbliża się do poszukiwanego wektora własnego związanego z największą wartością własną. Wartość własną można wyznaczyć w tej metodzie z ilorazu Rayleigha, ale – jak to pokazano wcześniej – nie ma takiej potrzeby, gdyż poszukiwaną wartość można wyznaczyć z zależności (21).

Proponowana metoda wyznaczania MPP składa się zatem z kroków opisanych w *Algorytmie 1*.

##### Algorytm 1

1. Wyznaczyć macierz wariancji-kowariancji próbkowej  $\mathbf{V}$ .
2. Sformułować prognozy stóp zwrotu  $\hat{\mathbf{R}}$  oraz ich macierz wariancji-kowariancji  $\hat{\mathbf{V}}$ .
3. Wyznaczyć macierz  $\mathbf{V}^{-1}\hat{\mathbf{V}}$ .
4. Przyjąć początkowe wartości dla wektora  $\mathbf{x} = 1/n$ .

5. Wyznaczyć wartości  $\mathbf{y}$  z formuły  $\mathbf{y} = \mathbf{V}^{-1}\widehat{\mathbf{V}}\mathbf{x}$ .
6. Podstawić za  $\mathbf{x}$  otrzymane w poprzednim kroku  $\mathbf{y}$  i wykonać normalizację wektora  $\mathbf{x}$  zgodnie z formułą (4).
7. Wyznaczyć  $\mathbf{x}^T\mathbf{x}$ .
8. Jeśli różnica wartości wyznaczonych w kroku 7 w dwóch kolejnych iteracjach jest mniejsza niż wymagana dokładność, to przejść do kolejnego kroku, w przeciwnym przypadku wrócić do kroku 5.
9. Wyznaczyć wartość  $r^2$  na podstawie (21).
10. Sprawdzić, czy  $\Lambda_{\mathbf{y}} 2\mathbf{y}^T(\widehat{\mathbf{V}} - \lambda\mathbf{V})\mathbf{y} \leq 0$ . Jeśli warunek ten nie jest spełniony, MPP nie udało się uzyskać.

Powyższy algorytm może być usprawniony poprzez zastosowanie dla macierzy  $\mathbf{V}$  rozkładu  $\mathbf{QR}$ , gdzie  $\mathbf{Q}$  spełnia warunki (33) i (34), a  $\mathbf{R}$  jest macierzą górną trójkątną, w której poniżej głównej przekątnej występują jedynie zera:

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{I} \quad (33)$$

$$\mathbf{Q}^T = \mathbf{Q}^{-1} \quad (34)$$

Macierz w postaci trójkątnej jest szczególnie przydatna przy rozwiązywaniu układu równań liniowych, gdyż w pierwszym równaniu tego układu występuje jedynie jedna zmienna, w drugim tylko dwie zmienne, w tym ta występująca w pierwszym równaniu, a w trzecim występują trzy zmienne, przy czym dwie z nich występują w poprzednich równaniach. Powyższy schemat powtarza się dla kolejnych równań, umożliwiając proste wyznaczenie rozwiązania. Rozkład  $\mathbf{QR}$  można wyznaczyć na podstawie odbić Householdera, obrotów Givensa lub poprzez proces ortogonalizacji Grama-Schmidta. Każda z tych metod ma swoje specyficzne właściwości.

Zarysowane powyżej postępowanie możemy zastosować do kroku 5 *Algorytmu 1*, mnożąc wcześniej lewostronnie to równanie przez  $\mathbf{V}$  i podstawiając za  $\mathbf{V} = \mathbf{QR}$ . Otrzymujemy wtedy układ równań w postaci (35):

$$\mathbf{QR}\mathbf{y} = \widehat{\mathbf{V}}\mathbf{x} \quad (35)$$

Ze względu na własność (34) możemy lewostronnie pomnożyć układ (35) przez  $\mathbf{Q}^T$ , otrzymując nowy w postaci (36):

$$\mathbf{R}\mathbf{y} = \mathbf{Q}^T\widehat{\mathbf{V}}\mathbf{x} \quad (36)$$

**Algorytm 2**

1. Wyznaczyć macierz wariancji-kowariancji próbkowej  $\mathbf{V}$  oraz wykonać jej faktoryzację do postaci:  $\mathbf{V} = \mathbf{QR}$ .
2. Sformułować prognozy stóp zwrotu  $\hat{\mathbf{R}}$  oraz ich macierz wariancji-kowariancji  $\hat{\mathbf{V}}$ .
3. Wyznaczyć macierz  $\mathbf{Q}^T \hat{\mathbf{V}}$ .
4. Przyjąć początkowe wartości dla wektora  $\mathbf{x} = 1/n$ .
5. Wyznaczyć wartości  $\mathbf{y}$  z układu równań  $\mathbf{R}\mathbf{y} = \mathbf{Q}^T \hat{\mathbf{V}} \mathbf{x}$ .
6. Podstawić za  $\mathbf{x}$  otrzymany w poprzednim kroku wektor  $\mathbf{y}$  i wykonać normalizację  $\mathbf{x}$  zgodnie z formułą (4).
7. Wyznaczyć  $\mathbf{x}^T \mathbf{x}$ .
8. Jeśli różnica wartości wyznaczonych w kroku 7 w dwóch kolejnych iteracjach jest mniejsza niż wymagana dokładność, to przejść do kolejnego kroku, w przeciwnym przypadku wrócić do kroku 5.
9. Wyznaczyć wartość  $r^2$  na podstawie (21).
10. Sprawdzić, czy  $\Lambda_y 2\mathbf{y}^T (\hat{\mathbf{V}} - \lambda \mathbf{V}) \mathbf{y} \leq 0$ . Jeśli warunek ten nie jest spełniony, MPP nie udało się uzyskać.

**4.5. Portfele własne**

Poprzez analogię do otrzymanych wcześniej wyników istnieje możliwość wyznaczenia portfela poprzez analizę spektralną macierzy  $\mathbf{V}$ . To zadanie odpowiada problemowi minimalizacji wariancji portfela, gdyż macierz  $2\mathbf{V}$  jest nieujemnie określona i jest jednocześnie hesjanem wariancji portfela  $\mathbf{x}^T \mathbf{V} \mathbf{x}$ . Punktem wyjścia tych rozważań może być iloraz Rayleigha. Ze względu na to, że poszukujemy portfela cechującego się najmniejszą wariancją, czyli poszukujemy najmniejszej wartości własnej, to poszukiwany *Algorytm 3* można wyprowadzić bazując na odwrotnej metodzie potęgowej.

**Algorytm 3**

1. Wyznaczyć macierz wariancji-kowariancji próbkowej  $\mathbf{V}$  oraz wyznaczyć jej rozkład  $\mathbf{V} = \mathbf{QR}$ .
2. Przyjąć początkowe wartości dla wektora  $\mathbf{x} = 1/n$ .
3. Wyznaczyć  $\mathbf{y}$  z następującego układu równań:  $\mathbf{R}\mathbf{y} = \mathbf{Q}^T \mathbf{x}$ .
4. Podstawić za  $\mathbf{x}$  wektor  $\mathbf{y}$  otrzymany w poprzednim kroku i wykonać normalizację wektora  $\mathbf{x}$  zgodnie z formułą (4).
5. Wyznaczyć  $\mathbf{x}^T \mathbf{x}$ .

6. Jeśli różnica wartości wyznaczonych w kroku 5 w dwóch kolejnych iteracjach jest mniejsza niż wymagana dokładność, to przejść do kolejnego kroku, w przeciwnym przypadku wrócić do kroku 3.
7. Wyznaczyć wartość  $V_r$  na podstawie (7).

W literaturze przedmiotu portfele wyznaczone w ten sposób są określane mianem portfeli własnych (ang. *eigenportfolio*) [Roll, 1980; Steele, 1995], przy czym badane są nie tylko portfele  $x$  odpowiadające najmniejszej wartości własnej, ale także pozostałe. W pracy Steele [1995] badano właściwości portfeli własnych.

Niestety opisane wyżej za pomocą algorytmu podejście, które maksymalizuje wariancję największego głównego czynnika, może powodować, że czynniki, dla których można zbudować dobry model prognostyczny, zostaną zredukowane, a pozostaną jedynie te, które charakteryzują się zerową wartością oczekiwaną, co dramatycznie obniża wartość wskaźnika determinacji  $r^2$ . Więcej informacji o podobnych problemach można znaleźć w pracy Lo, MacKinlay [2002].

## Podsumowanie

W artykule zdefiniowano trzy algorytmy służące do budowy portfeli o maksymalnej przewidywalności, które w pewien sposób nawiązują do teorii eigenportfeli i portfeli ortogonalnych. Wprawdzie nie można zakładać, że otrzymane MPP będą tworzyć zgodnie z nowoczesną teorią portfelową portfele sprawne, niemniej jednak od lat budzą one zainteresowanie wśród badaczy. Niestety właściwości numeryczne danych występujących w opisywanych modelach mogą stanowić pewną barierę w ich stosowaniu. Postać zaproponowanych algorytmów stara się obniżyć te bariery, umożliwiając wyznaczanie MPP.

W punkcie 4.4 artykułu zaproponowano trzy algorytmy służące do wyznaczania MPP. Pierwszy z nich oznaczony jako *Algorytm 1* umożliwia konstrukcję MPP bez stosowania faktoryzacji macierzy wariancji-kowariancji. W punkcie 10 tego algorytmu zaproponowano warunek weryfikujący, czy otrzymane rozwiązanie  $x$  spełnia założenia wstępne badanej metody. Dla macierzy  $V$  o dużych rozmiarach mogą się pojawić problemy numeryczne związane z tym, że wartość wyznacznika tej macierzy może przyjmować wartości znajdujące się bardzo blisko zera, co sprawia, że macierz ta staje się macierzą prawie osobliwą i trudno jest wyznaczyć numerycznie jej odwrotność, która jest niezbędna do wyznaczenia MPP. W rozdziale 4.3 opisano zastosowanie faktoryzacji Choleskiego, której zastosowanie w *Algorytmie 1* może poprawić jego właściwości numeryczne, jednakże celem nie była weryfikacja jakości otrzymanych metod, co będzie za-

pewne stanowić punkt wyjścia do kolejnych badań prowadzonych przez autora. W kolejnym algorytmie zastosowano faktoryzację  $QR$ , która jest znana ze swoich stabilnych właściwości numerycznych. Zastosowano szeroki rozkład  $QR$ , który wykorzystuje wszystkie wiersze macierzy  $R$ . W kolejnym etapie badań jest planowane zastosowanie wąskiego rozkładu  $QR$ , który uwzględnia jedynie niezerowe wiersze macierzy  $R$ , co w przypadku osobliwych macierzy wariancji-kowariancji może przynieść wymierne korzyści. Potwierdzenie tej hipotezy wymaga dalszych badań. Jako ostatni zaproponowano *Algorytm 3*, który bazując na szerokim rozkładzie  $QR$ , służy do wyznaczania portfeli własnych, które nawiązują swoją konstrukcją do MPP. Tym samym udało się osiągnąć wyznaczone cele. Badanie właściwości zaproponowanych algorytmów wymaga przeprowadzenia dalszych prac o charakterze empirycznym.

Ze względu na ograniczoną objętość artykułu nie uwzględniono rozdziału zawierającego empiryczne zastosowania proponowanych metod. Kolejnym etapem planowanych prac jest przeprowadzenie badań dotyczących zastosowania omawianych metod dla aktywów notowanych na Giełdzie Papierów Wartościowych w Warszawie.

## Literatura

- Fama E. (1970), *Efficient Capital Markets: A Review of the Theory and Empirical Work*, "Journal of Finance", 25, s. 383-417.
- Fama E., French K. (1992), *The Cross-section of Expected Stock Returns*, "Journal of Finance", 47, s. 427-465.
- Fama E., French K. (1993), *Common Risk Factors in the Returns on Stock and Bonds*, "Journal of Financial Economics", 33, s. 3-56.
- Fama E., French K. (1996), *Multifactor Portfolio Efficiency and Multifactor Asset Pricing*, "Journal of Financial and Quantitative Analysis", 31, s. 441-465.
- Lo A.W., MacKinlay A.C. (2002), *A Non-Random Walk Down Wall Street*, Princeton University Press, Princeton and Oxford.
- Perold A.F. (1984), *Large-Scale Portfolio Optimization*, "Management Science", 10, s. 1143-1160.
- Roll R. (1980), *Orthogonal Portfolios*, "Journal of Financial and Quantitative Analysis", 15, s. 1005-1023.
- Samuelson P. (1972), *Proof that Properly Anticipated Prices Fluctuate Randomly*, "Industrial Management Review", 6(2), s. 41-49.
- Steele A. (1995), *On the Eigen Structure of the Mean Variance Efficient Set*, "Journal of Business Finance and Accounting", 22, s. 245-255.

### ANALYSIS OF THE MAXIMALLY PREDICTABLE PORTFOLIO

**Summary:** In this study was described a three algorithms which can be used to construction of the Maximally Predictable Portfolio (MPP). It in some way is related to the theory of orthogonal portfolios and eigenportfolios. Although we cannot assume that the resulting MPP will be create the efficient portfolio, according to modern portfolio theory, but for years this direction of research aroused interest among researchers. Unfortunately, the properties of used numerical methods can create a barrier to their use. The construction of the proposed algorithms trying to reduce these barriers by allowing determination of the MPP. This work has a theoretical character and it will be need to carry out a number of empirical studies which will verify the properties of the proposed algorithms.

**Keywords:** stock portfolio, eigenportfolio, algorithms.