

oi-lab4

January 14, 2024

###Obrada informacija

#4. Laboratorijska vježba: Multivarijatni financijski vremenski nizovi

Prosinac 2023.

##Upute

U ovoj bilježnici dana je priprema sa svim uputama za 4. laboratorijsku vježbu iz predmeta Obrada informacija - uz bilježnicu su dostupni i podatci u datoteci `prices.csv`.

Vaš zadatak je u bilježnicu na odgovarajuća mjesta dopisati kod Vašeg rješenja.,

Riješenu bilježnicu potrebno je predati kao izvještaj u .pdf formatu na Moodle najkasnije do 14.1.2024. u 23:59h. Datoteka koju predajete se mora zвати *PrezimeImeJMBAG.pdf*.

0.1 Uvod

U laboratorijskoj vježbi razmatra se dinamika cijena vrijednosnica na financijskim tržištima. Dane su povijesne dnevne cijene 24 ETF-a (eng. *exchange traded fund*) koji prate određene dioničke, obvezničke ili druge indekse.

Oznaka	Naziv	Klasa imovine
SPY	SPDR S&P 500 ETF Trust	Equity: U.S. - Large Cap
IEFA	iShares Core MSCI EAFE ETF	Equity: Developed Markets
VWO	Vanguard FTSE Emerging Markets ETF	Ex-U.S. - Total Market
EWJ	iShares MSCI Japan ETF	Equity: Emerging Markets - Total Market
XLF	Financial Select Sector SPDR Fund	Equity: Japan - Total Market
XLK	Technology Select Sector SPDR Fund	Equity: U.S. Financials
XLV	Health Care Select Sector SPDR Fund	Equity: U.S. Technology
XLY	Consumer Discretionary Select Sector SPDR Fund	Equity: U.S. Health Care
XLP	Consumer Staples Select Sector SPDR Fund	Equity: U.S. Consumer Cyclical

Oznaka	Naziv	Klasa imovine
XLU	Utilities Select Sector SPDR Fund	Equity: U.S. Utilities
XLI	Industrial Select Sector SPDR Fund	Equity: U.S. Industrials
XLE	Energy Select Sector SPDR Fund	Equity: U.S. Energy
XLC	Communication Services Select Sector SPDR Fund	Equity: U.S. Telecommunications
XLRE	Real Estate Select Sector SPDR Fund	Equity: U.S. Real Estate
XLB	Materials Select Sector SPDR Fund	Equity: U.S. Basic Materials
BND	Vanguard Total Bond Market ETF	Fixed Income: U.S. - Broad Market
LQD	iShares iBoxx USD Investment Grade Corporate Bond ETF	Fixed Income: U.S. - Corporate
BNDX	Vanguard Total International Bond ETF	Fixed Income: Global Ex-U.S. - Broad Market
SHV	iShares Short Treasury Bond ETF	Fixed Income: U.S. - Government
HYG	iShares iBoxx USD High Yield Corporate Bond ETF	Fixed Income: U.S. - Corporate
GLD	SPDR Gold Trust	Commodities: Precious Metals Gold
SLV	iShares Silver Trust	Commodities: Precious Metals Silver
PDBC	Invesco Optimum Yield Diversified Commodity Strategy No K-1 ETF	Commodities: Broad Market
USO	United States Oil Fund LP	Commodities: Energy Crude Oil

Pri modeliranju zajedničkog kretanja i rizika vrijednosnica, koristit ćemo aritmetičke povrate:

$$R(t) = \frac{S(t) - S(t-1)}{S(t-1)},$$

gdje je $S(t)$ cijena vrijednosnice u danu t . U sklopu ove laboratorijske vježbe cilj je analizirati kretanje danih ETF-ova i izračunati glavne komponente (PCA) koje utječu na njihovu dinamiku. Laboratorijsku vježbu je potrebno riješiti unutar ove bilježnice i predati riješenu bilježnicu kao izvještaj.

[]: `import pandas as pd`

```
prices = pd.read_csv('prices.csv')
prices.set_index('Time', inplace=True)
```

```

prices.index = pd.to_datetime(prices.index)

prices.head()

```

```

[ ]:          SPY      IEFA      VWO      EWJ      XLF      XLK  \
Time
2019-01-02  231.492233  48.000053  33.417080  47.497765  21.776472  59.029892
2019-01-03  225.968170  47.684727  32.893154  47.227894  21.286907  56.050797
2019-01-04  233.537125  49.129986  33.932262  48.828522  21.994057  58.534962
2019-01-07  235.378525  49.366474  34.002121  48.949490  22.021255  59.058445
2019-01-08  237.589920  49.716839  34.194221  49.182148  22.039383  59.553371

          XLV      XLY      XLP      XLU    ...      XLB  \
Time
2019-01-02  78.483353  95.259102  44.591763  44.784126  ...  46.068901
2019-01-03  76.889748  93.196533  44.335598  44.775520  ...  44.762505
2019-01-04  79.183434  96.280815  45.280781  45.438545  ...  46.522518
2019-01-07  79.487411  98.457954  45.218948  45.128563  ...  46.685822
2019-01-08  80.104607  99.546509  45.634129  45.688255  ...  47.175724

          BND      LQD      BNDX      SHV      HYG  \
Time
2019-01-02  69.980576  97.039024  48.517841  101.234894  63.253788
2019-01-03  70.253906  97.081886  48.517841  101.271584  63.285011
2019-01-04  70.042305  97.021873  48.473194  101.299133  64.347198
2019-01-07  69.936508  97.030441  48.357071  101.299133  64.971985
2019-01-08  69.848312  97.244827  48.321350  101.289948  65.393723

          GLD      SLV      PDBC      USO
Time
2019-01-02  121.330002  14.56  8.545244  78.800003
2019-01-03  122.430000  14.75  8.607411  79.599998
2019-01-04  121.440002  14.73  8.731748  81.440002
2019-01-07  121.860001  14.67  8.799566  82.320000
2019-01-08  121.529999  14.69  8.873038  84.000000

```

[5 rows x 24 columns]

1 Zadatak 1 - Računanje korelacijske matrice i matrice kovarijance povrata

1.1. U prvom zadatku ove laboratorijske vježbe potrebno je prvo iz danih cijena (gore učitanih u Pandas DataFrame) izračunati dnevne povrate za sve pojedine vrijednosnice (prateći formulu danu u uvodu).

Izračunajte srednje povrate i volatilnost (standardnu devijaciju povrata) za svaku pojedinu vrijednosnicu. Pri analizi srednjih povrata i volatilnosti, te se brojke često *anualiziraju* - to znači da se

srednji povrati pomnože s 252 (cca. broj trgovinskih dana u godini), a volatilnost s $\sqrt{252}$.

Izračunajte anualizirane srednje povrate i volatilnosti. Sve ETF-ove prikažite u dijagramu raspršenja s volatilnošću na x-osi i srednjim povratom na y-osi.

Razmislite - koji se ETF-ovi ističu po odnosu povrata i rizika (posebno dobri ili posebno loši kao investicije)?

[1]: #Vaš kod ide ovdje

1.2. Kovarijancu i korelaciju moguće je iz podataka izračunati koristeći Pandas, ali i NumPy ili neke druge biblioteke.

Koristeći dnevne povrate, izračunajte matricu kovarijance Σ i matricu korelacije C povrata svih ETF-ova. Matrice ispišite u konzolu ili vizualizirajte.

Proučite strukturu matrice i razmislite o tome koje zajedničke komponente u podatcima možete očekivati.

[2]: #Vaš kod ide ovdje

2 Zadatak 2 - Analiza glavnih komponenti

2.1. Za analizu glavnih komponenti potrebno je izračunati svojstvenu dekompoziciju, koju možete pronaći u sklopu biblioteke NumPy <https://numpy.org/doc/stable/reference/generated/numpy.linalg.eig.html>.

Izračunajte svojstvene vektore i pripadajuće svojstvene vrijednosti matrice kovarijance povrata Σ . Poredajte komponente padajući po svojstvenim vrijednostima i prikažite svojstvene vrijednosti grafički.

[3]: #Vaš kod ide ovdje

2.2. Izračunajte koliki udio varijance objašnjavaju prve tri komponente?

[4]: #Vaš kod ide ovdje

2.3. Komponente PCA će u financijama često opisivati neke zajedničke faktore u podatcima, što je moguće analizirati promatranjem pojedinih elemenata svojstvenih vektora. Ako je neki element određenog svojstvenog vektora velik po magnitudi (pozitivan ili negativan), to znači da ta komponenta opisuje odgovarajuću vrijednosnicu i objašnjava njenu varijancu, za razliku od slučaja kad je element blizu 0, što znači da razmatrana vrijednosnica ne ovisi previše o toj komponenti.

Prikažite grafički (npr. stupčastim dijagramom za svaku komponentu posebno) koeficijente prve 3 glavne komponente (elemente prva tri svojstvena vektora).

S obzirom na to koje vrijednosnice opisuju prve tri komponente, razmislite možete li zaključiti kakve zajedničke faktore u tržištu opisuju razmatrane komponente?

NAPOMENA: pripazite na to što vraća funkcija koju koristite i u kojoj se dimenziji (stupac ili red) nalaze svojstveni vektori.

[5] : #Vaš kod ide ovdje

2.4. Ponovite prethodnu analizu za matricu korelacije povrata C (prikažite svojstvene vrijednosti, udio varijance i koeficijente pojedinih komponenti za prve tri komponente).

Usporedite rezultate - mijenjaju li se interpretacije komponenti?

[6] : #Vaš kod ide ovdje

3 Zadatak 3 - Svojstveni portfelji

U primjeni PCA i svojstvenoj dekompoziciji kovarijance u financijama, svojstveni vektori se često zovu i tzv. svojstveni portfelji.

Općenito, portfelj je vektor $w = [w_1, \dots, w_N]$ u kojem svaki element predstavlja težinu ili udio kapitala u određenoj vrijednosnici. Same težine svojstvenih portfelja mogu biti rotirane i skalirane u odnosu na elemente svojstvenih vektora.

U ovoj analizi ćemo pomnožiti njihove težine s predznakom njihove sume - na taj način zapravo samo "okrećemo" predznak svojstvenog vektora tako da mu je suma pozitivna (konačni PCA rastav je i dalje isti ako svojstveni vektor pomnožimo s -1). Također, dobro je i skalirati svojstvene portfelje sa sumom njihovih apsolutnih vrijednosti:

$$\tilde{w}_i = \frac{w_i}{\sum_j^N |w_j|}.$$

Na taj način se osigurava da visoke magnitude pojedinih elemenata ne uzrokuju velike razlike u volatilnostima svojstvenih portfelja.

Ukoliko znamo povrate $R \in \mathbb{R}^{T \times N}$ (gdje je $R_i \in \mathbb{R}^T$ vektor povrata za vrijednosnicu i) za N vrijednosnica u nekom vremenskom periodu od T dana, povrate portfelja w u tom istom periodu možemo izračunati kao:

$$R_p = \sum R_i w_i = R \cdot w.$$

Izračunajte skalirane svojstvene portfelje \tilde{w} koji proizlaze iz prve tri glavne komponente dobivene iz matrice kovarijance Σ . Za ta tri svojstvena portfelja izračunajte povjesne povrate kroz razmatrani period. Grafički prikažite vremensko kretanje njihovih vrijednosti (njihove povrate "vratite" natrag u cijene, s tim da početna cijena bude jednak za oba portfelja, npr. 100).

[8] : #Vaš kod ide ovdje

Ako usporedite dobivene rezultate s kretanjem cijena originalnih vrijednosnica, vidjet ćete sličnosti između vrijednosnica koje pripadaju određenim klasama imovina i pojedinih svojstvenih portfelja. Svojstveni portfelji dakle predstavljaju niže-dimenzionalan prostor tzv. sintetičkih vrijednosnica (u našem slučaju 3 umjesto originalnih 24) koje najbolje opisuju cijeli razmatrani skup podataka. Dobra procjena tih komponenti je ključna u razumijevanju zajedničkog kretanja većih skupova dionica i upravljanju finansijskim rizikom.

Razmislite što to znači za tržiste koje smo analizirali - koji su glavni izvori rizika prevladavali u razmatranom periodu?