

V Krakowska Konferencja Matematyki Finansowej: Symulacje w R

Fundacja Rozwoju Zawodowego Quantitative Finance: Maciej Nasinski

April 12, 2016

R Markdown

Spis Tresci:

1. Interesujace dane Finansowe i nie tylko
2. Symulacje modeli VAR
3. Symulacje handlu algorytmicznego na HFT

Fundacja Quantitative Finance - Kim jesteśmy?

Fundacja Rozwoju Zawodowego Quantitative Finance powstała w maju 2012 jako odpowiedź na rozwijające się ilościowe podejście do rynków finansowych nazywane quantitative finance, opierające się o wykorzystanie najnowszych osiągnięć ekonometrii, inżynierii finansowej oraz informatyki przy podejmowaniu decyzji inwestycyjnych oraz rozwiązywaniu problemów biznesowych. Celem Fundacji jest wspieranie rozwoju rynku quantitative finance w Polsce oraz zwiększanie świadomości tej branży wśród studentów. Budujemy współpracę pomiędzy najbardziej innowacyjnymi pracodawcami w Polsce, którzy poszukują osób kształcących się w quantitative finance, oraz studentami i absolwentami.



Quantitative Finance
FUNDACJA ROZWOJU ZAWODOWEGO

Prowadzone projekty

Cykl **spotkan z praktykami rynku:**

umożliwia poznanie osób zawodowo związanych z rynkiem. Studenci dowiadują się o obecnych trendach w branży finansowej oraz najbardziej pożądanym umiejętnościach. Każdy ma możliwość uczestnictwa w dyskusji, a po spotkaniach nasi goście z chęcią odpowiadają na pytania zadane w kuluarach. Dotychczas gościliśmy m.in. osoby odpowiedzialne za wdrożenie systemu UTP na polskiej giełdzie czy zarządzających algorytmicznymi funduszami inwestycyjnymi.

Projekt **L(a)unch your career:**

to cykl spotkań w formie kameralnego lunchu, w którym bierze udział kilku studentów oraz przedstawiciel top managementu firmy. W przeciwieństwie do publicznych wystąpień, rozmowy mają kameralną formę, dając studentom szansę na zadanie dowolnego pytania. Jest to również możliwość na zaprezentowanie się jako potencjalny kandydat do pracy kandydat do pracy.

Knowledge Team

W ramach struktur Fundacji działa zespół *Knowledge Team*, budujący zaplecze merytoryczne naszej działalności. Dotychczas, zespół przeprowadzał analizy rynku handlu algorytmicznego, projekcje inflacji, analizy strategii pasywnego zarządzania portfelem, wskaźników kontrariańskich i wielu innych. Obecnie rozszerzamy działalność o consulting w pozostałych obszarach quantitative finance.

1. Interesujące dane Finansowe i nie tylko

Notowania:

- ▶ Google Finance (Reuters?) - wysokie limity, dane o wysokiej częstotliwości - max 15 dni
- ▶ Stooq - niskie limity

Naturalny popyt (modelowanie sprzedaży):

- ▶ Google Trends - skalowana przez Google liczba wyszukiwań danej frazy
- ▶ Wunderground - dane pogodowe

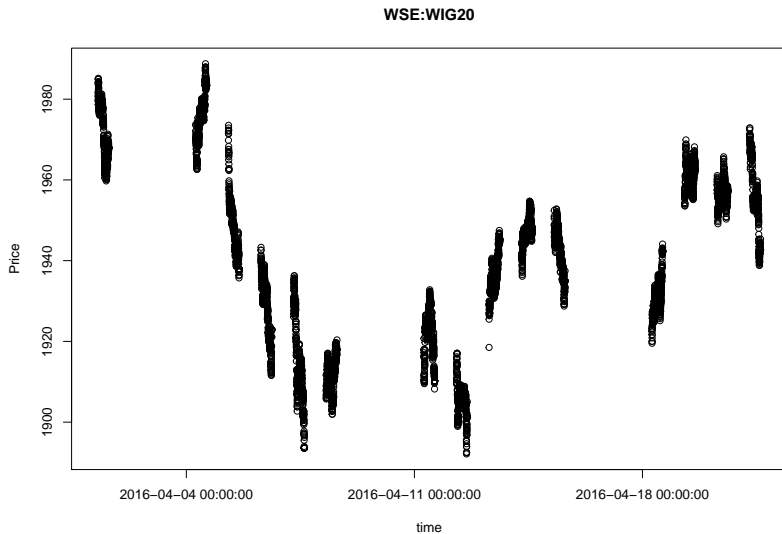
1.1.a. Google Finance

```
source("GF.R")
names<-c("WSE:WIG20")
for(f in names){GoogleFinance(f,60,15)}
```

Google Finance WIG20

CZAS UNIX - strefy czasowe

1.1.b. Dane Google Finance

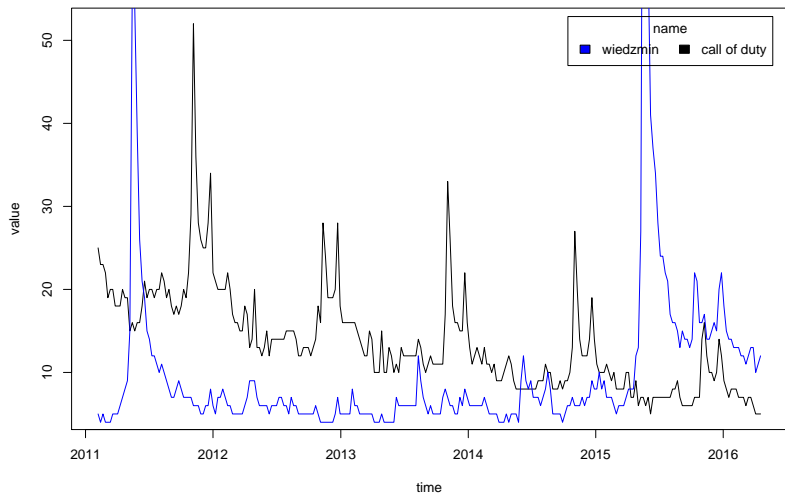


1.2.a. Google Trends

```
gconnect(usr, psw)
t<-gtrends(c("wiedzmin", "call of duty"),
           start_date =(as.Date(Sys.time())-1900),
           geo="PL")

#source("GoogleTday.R")
#GoogleTrendDaily(nam=c("wiedzmin", "call of duty"),
#type="relative", geo="PL", res="day")
```

1.2.b. Dane Google Trends



1.3.a. Wunderground

```
source("Weather_day.R")
#source("Weather_week.R")
n<-c('london', 'birmingham', 'glasgow')
city<-c("EGLC", "EGBB", "EGPF")
source("Weather_week.R")
```

Wunderground LONDON

```
for(i in 1:length(city)){
  repeat{
    url<-paste0("https://www. ...")
    webpage<-try(read.csv(url))
    if(class(webpage)=="try-error")
      {cat("connection problem: restart downloading current iteration"
res_city[[i]]<-webpage
    if(class(webpage)!="try-error") break}}}
```

1.3.b. Dane Wunderground

```
##
## GMT "2013-02-02" "2013-02-03"
## Max.TemperatureC_london "5" "9"
## Mean.TemperatureC_london "3" "4"
## Min.TemperatureC_london "2" "1"
## Dew.PointC_london "2" "7"
## MeanDew.PointC_london "-1" " 2"
## Min.DewpointC_london "-3" "-4"
## Max.Humidity_london "81" "87"
## Mean.Humidity_london "71" "80"
## Min.Humidity_london "57" "70"
## Max.Sea.Level.PressurehPa_london "1024" "1024"
## Mean.Sea.Level.PressurehPa_london "1016" "1018"
## Min.Sea.Level.PressurehPa_london "1003" "1013"
## Max.VisibilityKm_london "10" "10"
## Mean.VisibilityKm_london "10" "10"
```

1.4.a. Ściąganie danych - API

Bazowanie na pustych sekwencjach czasowych

- ▶ eliminuje problem z brakującymi obserwacjami
- ▶ ułatwia łącznie danych (cbind vs merge)
- ▶ umożliwia łatwe skalowanie czasowe np. dzienne na tygodniowe

Wykorzystanie funkcji typu try() oraz repeat() badz while()

- ▶ ogranicza błędy w komunikacji

2.1.a Trójkąt Pascala

```
## [1] "      1"
## [1] "     1 1"
## [1] "    1 2 1"
## [1] "   1 3 3 1"
## [1] "  1 4 6 4 1"
## [1] " 1 5 10 10 5 1"
## [1] "1 6 15 20 15 6 1"
## [1] "1 7 21 35 35 21 7 1"
## [1] "1 8 28 56 70 56 28 8 1"
```

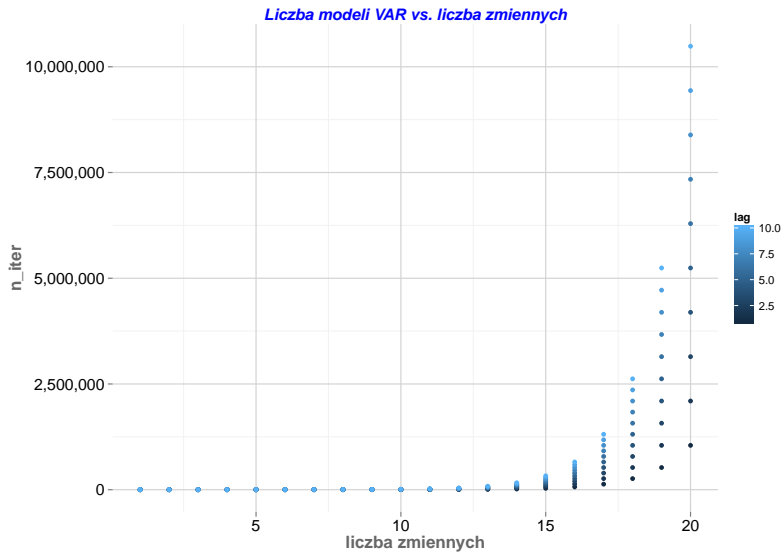
2.1.b Trójkąt Pascala

```
## [1] "1      1"
## [1] "2      1 1"
## [1] "4      1 2 1"
## [1] "8      1 3 3 1"
## [1] "16     1 4 6 4 1"
## [1] "32     1 5 10 10 5 1"
## [1] "64     1 6 15 20 15 6 1"
## [1] "128    1 7 21 35 35 21 7 1"
## [1] "256    1 8 28 56 70 56 28 8 1"
## [1] "512    1 9 36 84 126 126 84 36 9 1"
## [1] "1024   1 10 45 120 210 252 210 120 45 10 1"
```

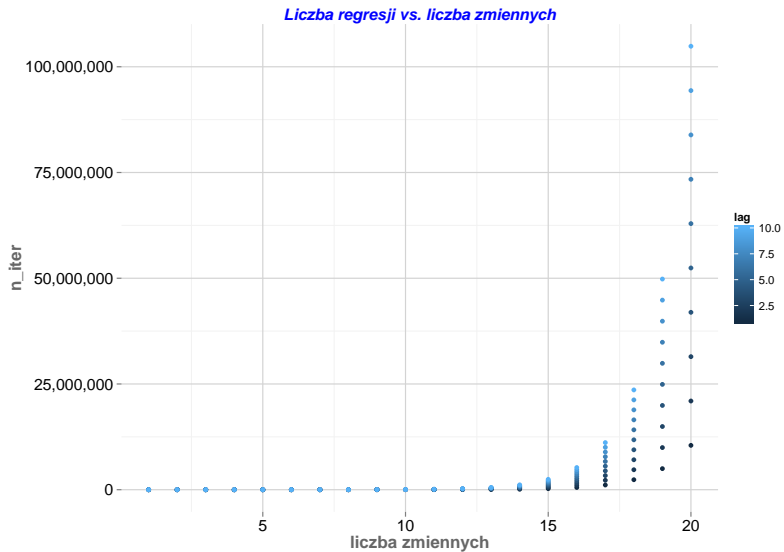
2.2.a. Model VAR

- ▶ szczególny przypadek modelu SUR (Seemingly Unrelated Regressions).
- ▶ identyczne zmienne objaśniające na przestrzeni równań \rightarrow MNK.
- ▶ możliwość implementacji IRF (Impulse Response Function) i FEVD.
- ▶ model dla zmiennych stacjonarnych - zmienne skointegrowane VECM.
- ▶ metoda zaliczana do Data Mining.
- ▶ brak wnioskowania ekonomicznego.

2.2.b. Liczba iteracji - model VAR



2.2.b.2 Liczba regresji - model VAR



2.2.c. Duża ilość zmiennych objaśniających

Redukcja wymiaru macierzy

- ▶ PCA (Principal Components Analysis)
- ▶ LDA (Linear discriminan analysis)

PCA jest czysto techniczna operacja, w przeciwieństwie do LDA, które zachowuje informacje służącą klasyfikacji zmiennych.

Problemy:

- ▶ Interpretacja zmiennych
- ▶ Niestandardowy rozkład zmiennych

2.2.d. MODEL VAR - Data Mining

```
col_names<-c("CPIUS_MM", "PKBPL_PROC_RET1", "Commodity.Index_RET1"  
            "M3_RET1", "USD_RET1", "UR_RET1",  
            "SALDO_RET1", "FED_RET1", "BEZROB_RET1",  
            "BRENT_RET1", "LIBOR3mEUR_RET1", "ZYWN_A_RET1",  
            "PKBUS_RET1", "WIBOR3M_RET1")  
com<-(combn(length(col_names),2,simplify = FALSE))  
for(i in 3:14){  
  com<-c(com,combn(length(col_names),i,simplify = FALSE))  
}  
length(com)*10
```

```
## [1] 163690
```

Transformacje: średnia ruchoma, stopy zwrotu.

2.2.e. MODEL VAR - Data Mining 2

```
all_R2<-array(NA,c(length(com),2,nr_lag))

for(g in 1:nr_lag){

  R2<-vector("numeric",length(com))
  SERIAL<-vector("numeric",length(com))
  for(i in 1:length(com)){

    model1<-VAR(data_noLAG_NA[,c("CPIPL_MM",col_names[com[[i]])]),

    sum_VAR<-summary(model1$varresult$CPIPL_MM)

    R2[i]<-sum_VAR$adj.r.squared
    SERIAL[i]<-serial.test(model1)$serial$p.value
    print(paste(i/length(com),g))

  }
  all_R2[,1,g]<-R2
  all_R2[,2,g]<-SERIAL
}
```

3.1.a. Dane Akcyjne HFT

▶ DLC

DLC (Dual Listed Companies) są efektem połączenia dwóch spółek. Spółki posiadają osobną osobowość prawną, dzięki czemu mogą być jednocześnie, jednak oddzielnie, notowane na swoim macierzystym rynku. Dochody dzielone są zgodnie z equalization agreement ratio- najczęściej dzielą się zyskami po równo.

▶ GRS

GRS klasyczne akcje notowane pod jednym numerem ISIN na całym świecie.

▶ DR

DR można powiązać z polską definicją kwitu depozytowego.

3.1.b. Przykład prostego kodu dla symulacji : repeat break

Start: Premia+

$$\text{diffcc} = (\text{price1} - \text{price2} * \text{currency}) / \text{price1}$$

```
repeat{
  repeat{
    i<-i+1
    if (dane[i,diffcc]>0.15){
      short<-dane[i,price1]
      long<-dane[i,price2]
      break}
    print(dane[i,c(price1,price2,currency,diffcc)])
  }
  b<-i
  repeat{
    i<-i+1
    if ( dane[i,diffcc]<(-0.15)){
      long2<-dane[i,price1]
      short2<-dane[i,price2]
      break}
    print(dane[i,c(price1,price2,currency,diffcc)])}
```

3.1.c. SYMULACJE ARBITRAŻU

- ▶ market frictions (krotka sprzedaz, koszty transakcyjne, koszt alternatywny)
- ▶ prawdop. zrealizowania zlecenia
- ▶ ilość volumenu z danej minuty - czekamy na częściową realizację
- ▶ poziom premii inicjujący strategię - EVT
- ▶ jak długo czekamy z zamknięciem pozycji

3.1.d. Wyniki symulacji

SYMULACJA W PLIKU SymulacjaStrategyFrompos.R

KONIEC

Co zyskujesz angażując się w pracę Fundacji QF?

Wiedzę z zakresu finansów ilościowych i nauk pokrewnych! Doświadczenie zawodowe i unikalne możliwości networkingowe! Realizację projektów dla największych instytucji finansowych w Polsce! Swobodę w organizowaniu swojej pracy i doskonalenie umiejętności organizacyjnych! Miłą atmosferę i satysfakcję! Zapraszamy do kontaktu!

www.quantitativefinance.org.pl

fundacja@quantitativefinance.org.pl

Kontakt:

Maciej Nasiński Lead Analyst:

nasinski.maciej@gmail.com

Piotr Cymcyk Prezes Zarządu:

piotrcymcyk@quantitativefinance.org.pl

Wojciech Zdunkiewicz Członek Zarządu:

wzdunkiewicz@quantitativefinance.org.pl

Bibliografia

Philippe Massicotte and Dirk Eddelbuettel (2016). gttrendsR: R Functions to Perform and Display Google Trends Queries. R package version 1.3.3. <https://CRAN.R-project.org/package=gtrendsR>

Hadley Wickham and Romain Francois (2015). dplyr: A Grammar of Data Manipulation. R package version 0.4.3. <https://CRAN.R-project.org/package=dplyr>

Hadley Wickham (2015). tidyr: Easily Tidy Data with spread() and gather() Functions. R package version 0.3.1. <https://CRAN.R-project.org/package=tidyr>

Bernhard Pfaff (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. Journal of Statistical Software 27(4). URL <http://www.jstatsoft.org/v27/i04/>.

<https://mran.microsoft.com/open/>

<https://www.wunderground.com/>

<https://www.google.pl/trends/?hl=pl>