



UNIWERSYTET  
MIKOŁAJA KOPERNIKA  
W TORUNIU

Wydział Nauk Ekonomicznych  
i Zarządzania

Piotr Fiszeder

Witold Orzeszko

# Prognozowanie kowariancji stóp zwrotu za pomocą Maszyny Wektorów Nośnych

Piernikowe Wirtualne Seminarium z Badań Operacyjnych, 21.12.2020 r.



## Cel pracy

---

Zaproponowanie nowej procedury prognozowania macierzy warunkowej kowariancji, opartej na:

- ⊙ estymatorze macierzy kowariancji uwzględniającymienne ceny minimalne i maksymalne,
- ⊙ modelu SVR (Support Vector Regression) jako metody prognozowania z uwzględnieniem dekompozycji Choleskiego (w celu zapewnienia dodatniej określoności prognozowanej macierzy).



## Plan prezentacji

---

- ⊙ Motywacje
- ⊙ Macierz kowariancji zakresu cen
- ⊙ Model SVR
- ⊙ Procedura zastosowania modelu SVR do prognozowania macierzy kowariancji
- ⊙ Badanie empiryczne – prognozowanie macierzy kowariancji dla wybranych kursów walut



## Motywacje (1)

---

- ⊙ Poszukiwanie narzędzi wspierających podejmowanie decyzji inwestycyjnych na rynkach finansowych.
- ⊙ Potrzeba wyznaczania dokładnych prognoz w zakresie wariacji (analiza ryzyka pojedynczych instrumentów) i kowariancji (analiza portfeli inwestycyjnych).
- ⊙ Metody uczenia maszynowego są obiecującym narzędziem prognozowania finansowych szeregów czasowych.



## Motywacje (2)

---

- ⦿ Estymatory wariancji/kowariancji oparte na cenach minimalnych i maksymalnych wykorzystują więcej informacji o zmienności stóp zwrotu i o zależnościach między stopami różnych aktywów niż estymatory oparte wyłącznie na cenach zamknięcia. Ceny minimalne i maksymalne są na ogół dostępne na równi z dziennymi cenami zamknięcia dla wielu aktywów finansowych.



## Macierz kowariancji zakresu cen

---

- ⊙ Estymator kowariancji:

$$\text{cov}(X, Y) = [\text{var}(X + Y) - \text{var}(X) - \text{var}(Y)]/2, \quad (1)$$

gdzie wariancje są szacowane na podstawie cen minimalnych i maksymalnych.

- ⊙ Estymator kowariancji dla stóp zmian kursów walutowych (zastosowanie kursu krzyżowego):

$$\text{cov}(\Delta \ln A/C, \Delta \ln B/C) = [\text{var}(\Delta \ln A/C) + \text{var}(\Delta \ln B/C) - \text{var}(\Delta \ln A/B)]/2$$



## Estymator wariancji

---

⊙ Estymator wariancji (Parkinson, 1980):

$$\sigma_t^2 = [\ln(H_t/L_t)]^2 / (4\ln 2), \quad (2)$$

gdzie  $H_t$  i  $L_t$  są dziennymi cenami: maksymalnymi i minimalnymi.



# Maszyna Wektorów Nośnych

---

Maszyna Wektorów Nośnych (*Support Vector Machine*; Vapnik, 1995) – metoda uczenia maszynowego stosowana do:

- ⊙ klasyfikacji (*Support Vector Classification* – SVC, SVM)
- ⊙ regresji (*Support Vector Regression* – SVR).





## Maszyna Wektorów Nośnych – zalety i wady

---

### ⊙ Zalety:

- Rozwiązuje problemy liniowe, jak i nieliniowe
- Wysokie dopasowanie do danych, z możliwością ograniczenia efektu przeuczenia modelu
- Odporność na występowanie szumu w zbiorze uczącym
- Odporność na małą liczbę danych i dużą liczbę atrybutów.

### ⊙ Wady:

- Trudność w interpretacji modelu (*black box*).



## Model SVR – istota

---

Model regresji:

$$y = r(\mathbf{x}) + \delta, \quad (3)$$

gdzie:  $r(\mathbf{x})$  funkcja regresji,  $y$  – zmienna zależna,  $\mathbf{x}$  – wektor regresorów,  $\delta$  – składnik losowy.

Cel: wyznaczenie funkcji aproksymującej  $f(\mathbf{x})$  tak, aby jej wartości nie różniły się o więcej niż  $\varepsilon$  od obserwacji  $y_t$  i jednocześnie nie była nadmiernie dopasowana do zbioru treningowego  $\{(\mathbf{x}_t, y_t)\}_{t=1, \dots, n}$ .



## Model SVR – konstrukcja

---

Wektory  $\mathbf{x}$  przekształca się poprzez odwzorowania  $\varphi_i(\mathbf{x})$  do przestrzeni o większym wymiarze –  $d$ , a następnie wyznacza się:

$$f(\mathbf{x}) = \sum_{i=1}^d \omega_i \varphi_i(\mathbf{x}) + b, \quad (4)$$

gdzie  $\omega_i$  oraz  $b$  są parametrami modelu.

W procesie wyznaczania parametrów stosuje się funkcję straty:

$$L_\varepsilon(y, f(\mathbf{x})) = \begin{cases} 0 & , \quad |y - f(\mathbf{x})| \leq \varepsilon, \\ |y - f(\mathbf{x})| - \varepsilon, & \text{w p. w.}, \end{cases} \quad (5)$$

11 według której błędy poniżej poziomu  $\varepsilon$  nie są penalizowane.



## Estymacja parametrów modelu SVR

---

Zadanie programowania kwadratowego – minimalizacja funkcjonau:

$$\Phi(\boldsymbol{\omega}, \boldsymbol{\xi}) = \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{t=1}^n (\xi_t + \xi_t^*), \quad (6)$$

gdzie:  $C > 0$ ;  $\xi_t, \xi_t^*$  – nieujemne zmienne spełniające warunki:

$$y_t - f(\mathbf{x}_t) \leq \varepsilon + \xi_t^*, \quad (7)$$

$$f(\mathbf{x}_t) - y_t \leq \varepsilon + \xi_t, \quad (8)$$

dla każdego  $t = 1, 2, \dots, n$ .

Problem praktyczny: duża liczba optymalizowanych zmiennych, złożoność obliczeniowa.



## Estymacja parametrów modelu SVR – problem dualny

---

Istota: Zastosowanie funkcji jądrowych (*kernel trick*):

$$f(\mathbf{x}) = \sum_{t=1}^{n_{SV}} (\alpha_t - \alpha_t^*) K(\mathbf{x}_t, \mathbf{x}); \quad 0 \leq \alpha_t \leq C, \quad 0 \leq \alpha_t^* \leq C, \quad (9)$$

gdzie:  $\alpha_t$  i  $\alpha_t^*$  – mnożniki Lagrange'a,  $n_{SV}$  – liczba wektorów nośnych,  $K$  – funkcja jądrowa, spełniająca warunek:

$$K(\mathbf{x}_t, \mathbf{x}) = \sum_{i=1}^d \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}_t). \quad (10)$$

Zaleta: Nie trzeba wprost wyznaczać postaci analitycznej funkcji  $\varphi_i$  (wyznacza się tylko iloczyn skalarny wektora  $\varphi(\mathbf{x})$ ), co znacząco redukuje złożoność obliczeniową algorytmu.



## SVR – najczęściej stosowane funkcje jądrowe

---

- ⊙ Liniowa (iloczyn skalarny):  $K(\mathbf{x}_t, \mathbf{x}) = \mathbf{x}_t^T \mathbf{x}$ ,
- ⊙ Gaussowska:  $K(\mathbf{x}_t, \mathbf{x}) = \exp(-\|\mathbf{x}_t - \mathbf{x}\|^2)$ ,
- ⊙ Wielomianowa:  $K(\mathbf{x}_t, \mathbf{x}) = (1 + \mathbf{x}_t^T \mathbf{x})^p$ ;  $p = 2, 3, \dots$

Zastosowanie jądra liniowego prowadzi do liniowej SVR, a pozostałych – do nieliniowej SVR.



## Badanie empiryczne – cel badania

---

Ocena dokładności prognoz z proponowanej procedury, zarówno w zakresie całej macierzy kowariancji, jak i poszczególnych kowariancji/wariancji.

Modele benchmarkowe: DCC, GARCH.



## Badanie empiryczne – przedmiot badania

---

Trzy najbardziej płynne na rynku FOREX pary walutowe: EUR/USD, USD/JPY i GBP/USD (macierze kowariancji warunkowej  $G_t$  mają wymiar  $3 \times 3$ ).

Kowariancje/wariancje warunkowe szacowane na podstawie dziennych cen minimalnych i maksymalnych.

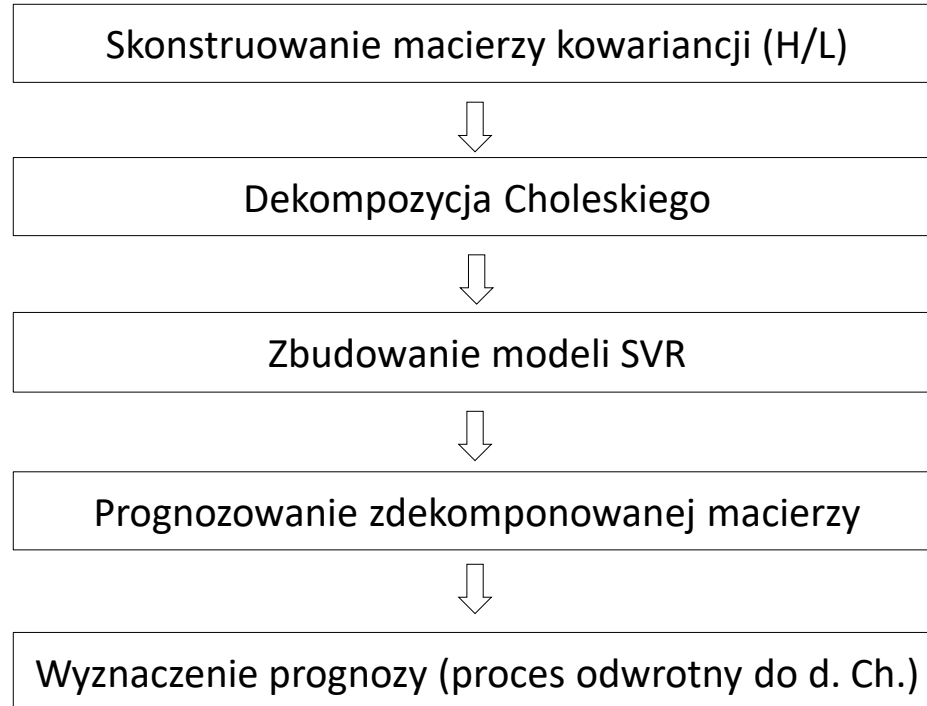
Dane z okresu 2.01.2006 – 30.12.2016.

Horyzont prognozy = 1 dzień (prognozowana macierz:  $\hat{G}_{n+1}$ ).





## Procedura zastosowania modelu SVR





## Procedura zastosowania modelu SVR

---

1. Dekompozycja Choleskiego:  $\mathbf{G}_t = \mathbf{P}'_t \mathbf{P}_t$ , gdzie  $\mathbf{P}_t$  – macierz górnotrójkątna.
2. Każdy z elementów  $p_{n+1}^{ij}$  (gdzie  $i \leq j$ ) macierzy  $\mathbf{P}_{n+1}$ , prognozowany jest oddzielnie przy zastosowaniu autoregresyjnego modelu SVR:

$$\hat{p}_{n+1}^{ij} = f\left(p_n^{ij}, p_{n-1}^{ij}, \dots, p_{n-l+1}^{ij}\right), \quad (11)$$

gdzie  $l$  jest rzędem opóźnienia.

3. Wyznaczenie prognozy dla macierzy kowariancji:

$$\hat{\mathbf{G}}_{n+1} = \hat{\mathbf{P}}'_{n+1} \hat{\mathbf{P}}_{n+1}. \quad (12)$$



## Badanie empiryczne – zastosowane modele

---

- 1) SVR z jądrem liniowym i  $l = 1$  (ozn. SVR\_lin\_1),
- 2) SVR z jądrem liniowym i  $l = 15$  (ozn. SVR\_lin\_15),
- 3) SVR z jądrem gaussowskim i  $l = 1$  (ozn. SVR\_gauss\_1),
- 4) SVR z jądrem gaussowskim i  $l = 15$  (ozn. SVR\_gauss\_15),
- 5) DCC(1,1) (ozn. DCC).

Modele budowane na podstawie przesuwanego okna złożonego z 500 poprzedzających obserwacji.

Prognozy dla okresu 2.01.2008 – 30.12.2016 (2336 prognoz).



## Modele SVR – wyznaczenie metaparametrów

---

Metaparametry do wyznaczenia:  $\varepsilon$ ,  $C$ .

Metoda: *grid search* – uwzględnienie wszystkich kombinacji metaparametrów:

$$C = 2^{-5}, 2^{-4}, \dots, 2^4, 2^5 \text{ oraz } \varepsilon = 2^{-5}, 2^{-4}, \dots, 2^4, 2^5$$

i porównanie modeli metodą 10-krotnej walidacji krzyżowej (*cross validation*).

Metaparametry aktualizowano co 100 obserwacji.



## Weryfikacja dokładności prognoz

---

Zastosowano wariancje i kowariancje zrealizowane, obliczone na podstawie danych *intraday* (15-minutowe stopy zmian):

- ⊙ wariancja: suma kwadratów stóp zmian,
- ⊙ kowariancja: suma iloczynów stóp zmian.

Uwaga: Informacje o danych *intraday* nie były wykorzystane w procesie wyznaczania prognozy.



## Miary trafności prognoz

---

- ⊙ Dla całej macierzy: funkcja straty Frobeniusa (uśredniona odległość Frobeniusa macierzy):

$$LF = (1/m) \sum_{T=1}^m \text{Tr}[(\mathbf{cov}_{Tp} - \mathbf{cov}_{Tr})'(\mathbf{cov}_{Tp} - \mathbf{cov}_{Tr})], \quad (13)$$

gdzie:  $m = 2336$  to liczba prognoz,  $\text{Tr}$  – ślad macierzy,  $\mathbf{cov}_{Tp}$  – prognoza macierzy kowariancji,  $\mathbf{cov}_{Tr}$  – macierz zrealizowanych kowariancji.

- ⊙ Dla poszczególnych szeregów wariancji/kowariancji:  $MSE$ ,  $MAE$ .



## Weryfikacja istotności różnic miar trafności – metoda MCS

---

Metoda zbioru ufności modeli (*Model Confidence Set*, Hansen i in. (2011)) pozwala wskazać podzbiór modeli prognostycznych istotnie lepszych od pozostałych.

Modele są oceniane w oparciu o zadaną funkcję straty ( $LF$ ,  $MSE$ ,  $MAE$ ). Polega na zastosowaniu ciągu testów istotności, w wyniku których modele istotnie gorsze od pozostałych są eliminowane.

Hipoteza zerowa głosi, że modele mają jednakowe (co do wartości oczekiwanej) funkcje straty.

23 Wartości  $p$  obliczone techniką *bootstrap*.



## Wyniki prognozowania a dekompozycja Choleskiego

---

**Uwaga:** Dekompozycja Choleskiego zależy od kolejności zmiennych (ich miejsca w macierzy kowariancji).

Dla trzech szeregów walut istnieje sześć permutacji. Dla każdej z nich szeregi wariancji/kowariancji po dekompozycji Choleskiego są inne, więc prognozy też są inne.

Najgorsza permutacja: EUR/USD, GBP/USD i USD/JPY.





## Trafność prognozowania pełnej macierzy kowariancji

Metoda prognozowania	<i>LF</i>
DCC	2,870
SVR_lin_1	2,952
SVR_lin_15	2,659*
SVR_gauss_1	3,043
SVR_gauss_15	3,335

Symbol \* oznacza, że dany model należy do zbioru ufności modeli (MCS) na poziomie ufności  $1 - \alpha = 0,90$ .



## Trafność prognozowania poszczególnych kowariancji

Metoda prognozowania	EUR/USD, GBP/USD		EUR/USD, JPY/USD		JPY/USD, GBP/USD	
	MSE	MAE	MSE	MAE	MSE	MAE
DCC	0,123	0,131	0,097	0,140	0,220	0,119
SVR_lin_1	0,150	0,149	0,096	0,135	0,206	0,113
SVR_lin_15	0,116*	0,121*	0,084*	0,119*	0,195*	0,102*
SVR_gauss_1	0,144	0,145	0,106	0,141	0,212	0,116
SVR_gauss_15	0,180	0,167	0,113	0,148	0,225	0,126



## Trafność prognozowania poszczególnych wariancji

Metoda prognozowania	EUR/USD		GBP/USD		JPY/USD	
	MSE	MAE	MSE	MAE	MSE	MAE
GARCH	0,168*	0,206	1,190	0,230	0,631*	0,313
SVR_lin_1	0,203	0,222	1,091	0,245	0,751	0,318
SVR_lin_15	0,164*	0,185*	0,999*	0,198*	0,706	0,295*
SVR_gauss_1	0,209	0,226	1,087	0,239	0,822	0,327
SVR_gauss_15	0,264	0,258	1,195	0,271	0,839	0,337



## Podsumowanie wyników badań

---

- ⊙ Dokładność modeli SVR zależy od przyjętego jądra i rzędu autoregresji.
- ⊙ Nawet dla najgorszej permutacji, model SVR\_lin\_15 dał istotnie dokładniejsze prognozy macierzy kowariancji niż model DCC.
- ⊙ Ten sam wniosek dotyczy też poszczególnych kowariancji (zarówno dla MSE, jak i MAE).
- ⊙ W przypadku wariancji, SVR\_lin\_15 dał lepsze prognozy niż GARCH według MAE. Według MSE, SVR\_lin\_15 znalazł się w zbiorze MCS dla EUR/USD i GBP/USD, a model GARCH dla EUR/USD i JPY/USD.



Dziękuję za uwagę.