

Dylematy związane ze stosowaniem Filtru Kalmana w prognozowaniu zjawisk ekonomicznych

Mateusz Stachnik, Zygmunt Zawadzki

Otwarte seminarium Katedry Statystyki

May 12, 2014

Wprowadzenie

Wprowadzenie

Rys historyczny

Przestrzeń stanów

Filtr Kalmana

Faza predykcji

Faza uaktualnienia

Zestawienie filtrów

Przykład ekonomiczny

Dylematy?

Bibliografia

Materiały dodatkowe

Prezentacja powstała z wykorzystaniem R, Matlaba, \LaTeX 'a i knitr'a ¹.
Na uwagę zasługuje szczególnie knitr, który pozwolił w łatwy sposób zintegrować R i \LaTeX 'a.

Kod R użyty do wygenerowania przykładów w prezentacji można znaleźć w internecie pod adresem <https://github.com/zzawadz/KalmanSem>.

¹<http://yihui.name/knitr/>

Rudolf Kalman

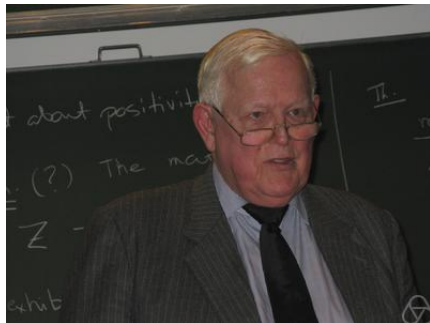


Figure: Źródło: Wikipedia

- ▶ Urodzony w 1930 na Węgrzech.
- ▶ 1960 publikacja Filtru Kalmana.
- ▶ Do dzisiaj wykłada.

Przestrzeń stanów - wprowadzenie - rys historyczny

W klasycznej teorii sterowania można było sterować układami o jednym wejściu i jednym wyjściu. Wraz z programem Apollo pojawiała się potrzeba sterowania wieloma obiektami jednocześnie. Problem polegał na tym, że występowały oddziaływania między obiektami.

$$Y = G \cdot X$$

$$G = \frac{1}{s + 1}$$

gdzie G to transmitancja operatorowa układu, Y to wyjście, X wejście, s - operator Laplace'a.

Przestrzeń stanów - wprowadzenie

Problem pojawiał się w opisie prostego układu o dwóch wejściach i dwóch wyjściach.

$$Y_1 = G_{11} \cdot X_1 + G_{12} \cdot X_2$$

$$Y_2 = G_{21} \cdot X_1 + G_{22} \cdot X_2$$

W podejściu przestrzeni stanów rozdziela się układ na dwie części. Część stanu, oraz część wyjścia. Powyższy układ można dzięki temu zapisać dwoma równaniami.

$$\dot{x}(t) = A_c x(t) + B_c u(t)$$

$$y(t) = H_c x(t) + D_c u(t)$$

gdzie A_c - macierz stanu, B_c - macierz sterowania, H_c - macierz wyjścia, D_c - macierz transmisji.

Oczywiste wydaje się, że zapis w przestrzeni stanów jest dużo prostszy i łatwo można go wykorzystać w dużych systemach. W dodatku jest bardziej ogólny niż zapis transmitancyjny.

Przestrzeń stanów - układ dyskretny

Model dla czasu ciągłego postaci:

$$\dot{x}(t) = A_c x(t) + B_c u(t)$$

$$y(t) = H_c x(t) + D_c u(t)$$

można również przedstawić w dziedzinie dyskretniej:

$$x(t+1) = Ax(t) + Bu(t)$$

$$y(t) = Hx(t)$$

Istnieją metody pozwalające na przejście z czasu ciągłego do dyskretnego i z powrotem.

Przykład - estymacja pozycji pojazdu

Równania opisujące dynamikę układu:

$$v_{t+1} = v_t + a_t T_p + w_v$$

$$s_{t+1} = s_t + v_t T_p + \frac{1}{2} a_t T_p^2 + w_s$$

Układ zapisany w przestrzeni stanów:

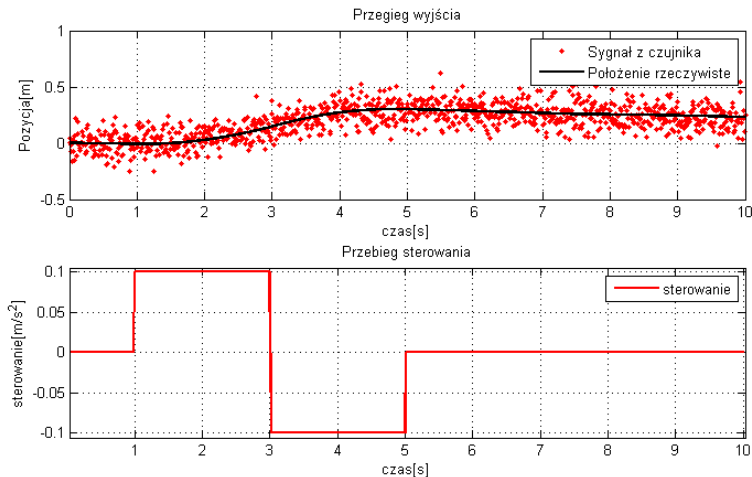
$$\begin{bmatrix} s_{t+1} \\ v_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & T_p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_t \\ v_t \end{bmatrix} + \begin{bmatrix} w_{st} \\ w_{vt} \end{bmatrix}$$

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s_t \\ v_t \end{bmatrix} + z_t$$

T_p - czas próbkowania, równy 0.05s

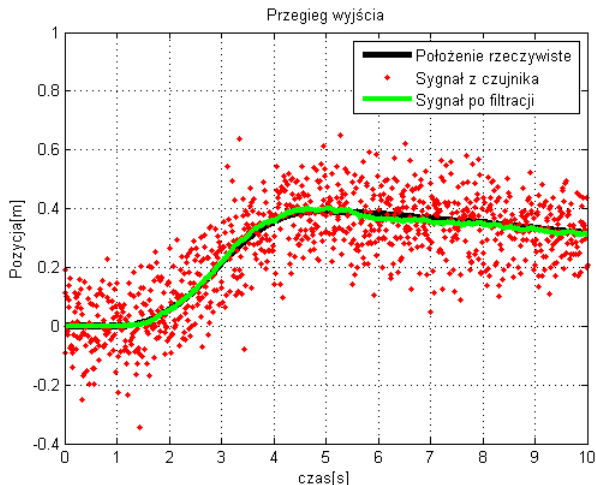
Przykład - estymacja pozycji pojazdu

$$\sigma_a = 0.1, \quad \sigma_s = 0.1$$



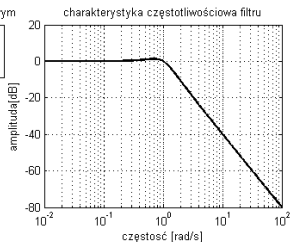
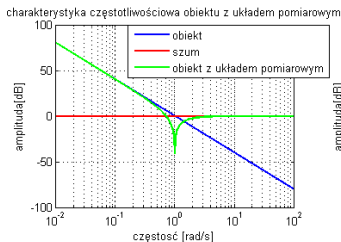
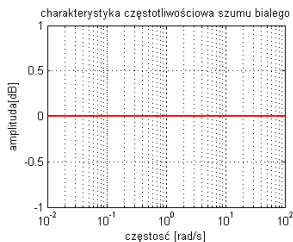
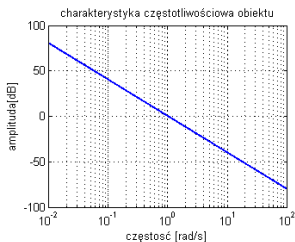
Cel filtracji

Głównym celem filtracji jest poznanie rzeczywistej wartości sygnału. Czasem problemem jest też fakt, że sygnał mocno zaszumiony jest trudny do przetwarzania, szczególnie dotyczy się to różniczkowania takiego sygnału.

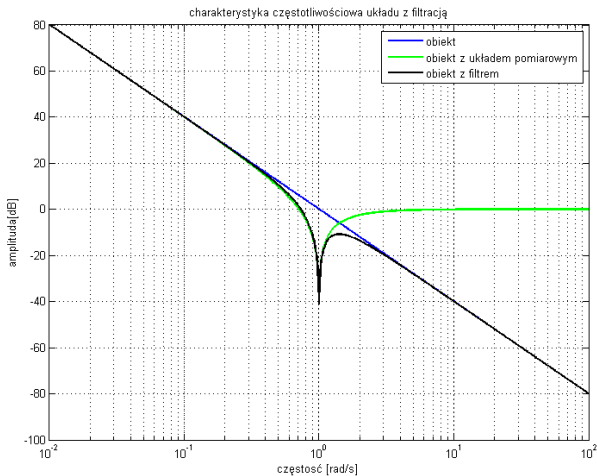


Filtr analogowy

Filtry analogowe projektuje się zazwyczaj w domenie częstotliwości.



Filtr analogowy - dolnoprzepustowy drugiego rzędu



Realizacja filtru dolnoprzepustowego jako filtru cyfrowego

$$G = \frac{1}{s^2 + s + 1}$$

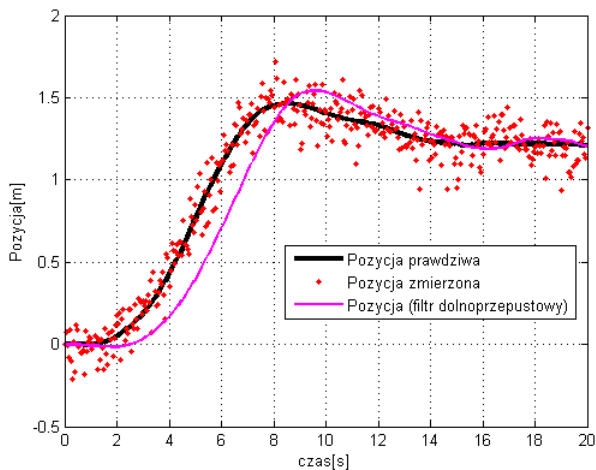
przechodząc z transmitancji ciągłej do dyskretnej metodą ZOH (ZERO ORDER HOLD) z czasem próbkowania 0.05s otrzymujemy:

$$G = \frac{0.0012z^{-1} + 0.0012z^{-2}}{1 - 1.949z^{-1} + 0.095z^{-2}}$$

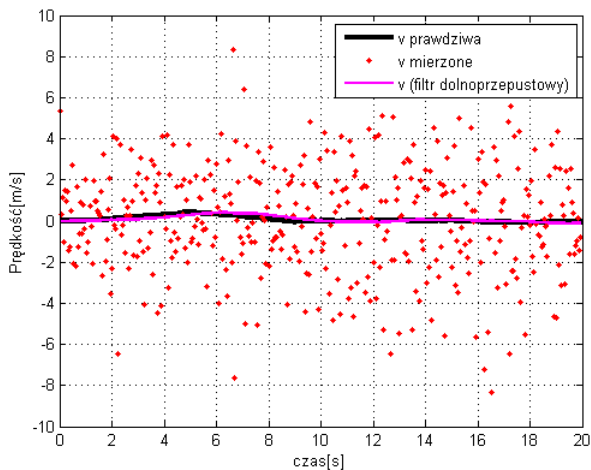
Z transmitancji dyskretnej można w prosty sposób przejść na równanie różnicowe:

$$y_t = 0.0012x_{t-1} + 0.0012x_{t-2} + 1.949y_{t-1} - 0.095y_{t-2}$$

Charakterystyka czasowa - położenie



Charakterystyka czasowa - prędkość



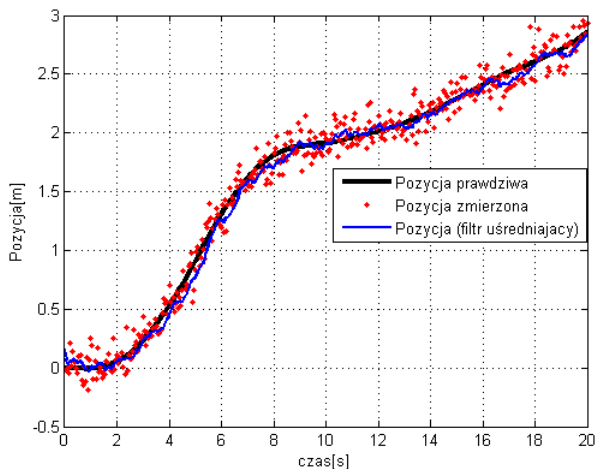
Filtr cyfrowy - średnia ruchoma

Prosta średnia ruchoma jest klasycznym filtrem cyfrowym typu MA:

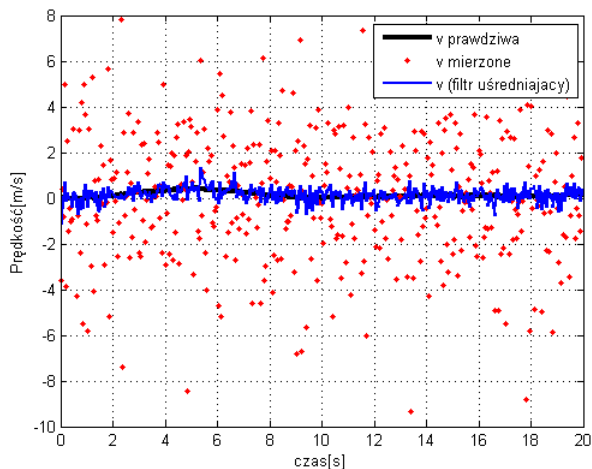
$$y_t = \frac{1}{n} \sum_{i=t-n+1}^t x_{t-i}$$

W rozważanym przypadku przyjęto $n = 10$.

Filtr cyfrowy - charakterystyka czasowa - położenie



Filtr cyfrowy - charakterystyka czasowa - prędkość



Filtr Kalmana

Zadaniem filtracji Kalmana jest estymacja wartości wektora stanu x przy niepewnych parametrach obiektu na podstawie zmierzonych próbek obarczonych szumem.

Cały proces przebiega dwuetapowo - najpierw następuje faza predykcji na chwilę t , na podstawie chwili $t - 1$, następnie na podstawie obserwacji (już w chwili t), następuje uaktualnienie estymowanych wartości.

Dla ustalenia uwagi:

$$\begin{aligned}x_t &= Ax_{t-1} + Bu_t + w_t \\ y_t &= Hx_t + v_t\end{aligned}\tag{1}$$

x_t - wektor zmiennych stanu

u_t - wektor sterowania - znany

y_t - wektor wyjścia - obserwowany

A, B, H - znane macierze

$w_t \sim WN(0, W)$

$v_t \sim WN(0, V)$

W i V - diagonalne macierze kowariancji.

Faza predykcji

Faza predykcji opisana jest następującymi wzorami:

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + Bu_t \quad (2)$$

$$P_{t|t-1} = AP_{t-1|t-1}A' + W \quad (3)$$

We wzorach pojawia się nowa macierz $P_{t|t-1}$ - jest to macierz kowariancji dla prognozy pod warunkiem chwili $t - 1$, definiowana jako:

$$P_{t|t-1} = E[(x_t - \hat{x}_{t|t-1})(x_t - \hat{x}_{t|t-1})'] \quad (4)$$

natomiast $P_{t|t}$ to macierz kowariancji dla x_t w chwili t definiowana jako:

$$P_{t|t} = E[(x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})'] \quad (5)$$

W chwili uaktualniania $P_{t-1|t-1}$ jest znana - obliczana jest ona bowiem w chwili $t - 1$.

Faza uaktualnienia

W momencie dokonania pomiaru następuje uaktualnienie wartości zmiennych stanu, na podstawie następujących wzorów:

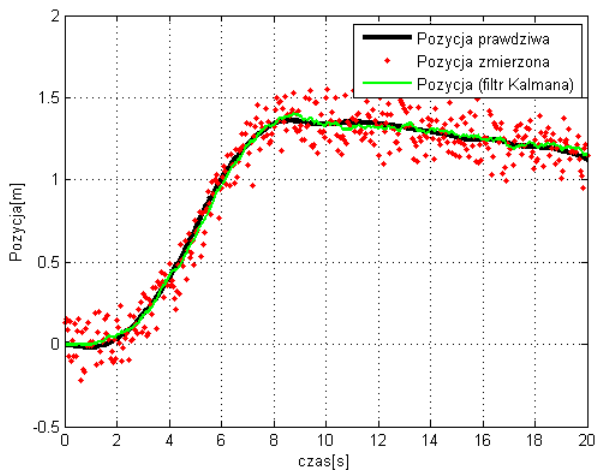
$$K_t = (P_{t|t-1}H')(HP_{t|t-1}H' + V)^{-1} \quad (6)$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - H\hat{x}_{t|t-1}) \quad (7)$$

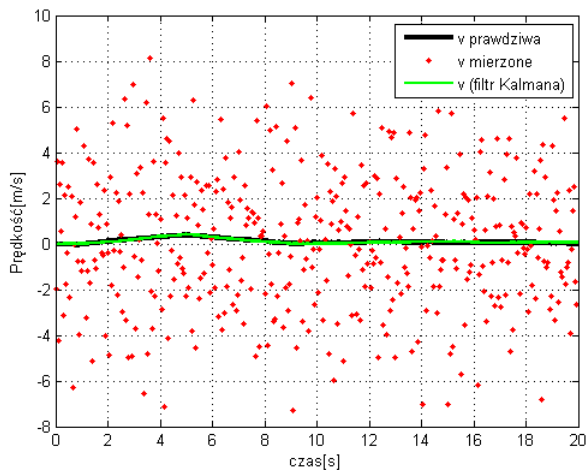
$$P_{t|t} = P_{t|t-1} - K_tHP_{t|t-1} \quad (8)$$

Najbardziej kluczową wartością jest K_t nazywane wzmocnieniem Kalmana - określa wpływ zaobserwowanego odchylenia wartości obserwowanej od predykcji ($y_t - H\hat{x}_{t|t-1}$) na aktualizację wartości zmiennych stanu - im K_t większe, tym większa waga będzie przykładana obserwacji.

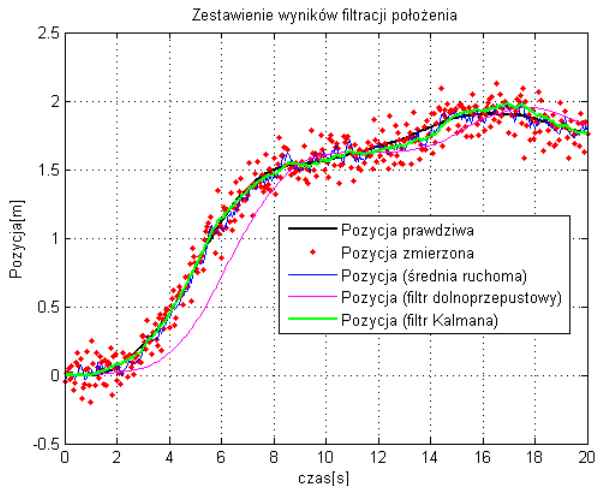
Filtr Kalmana - charakterystyka czasowa - położenie



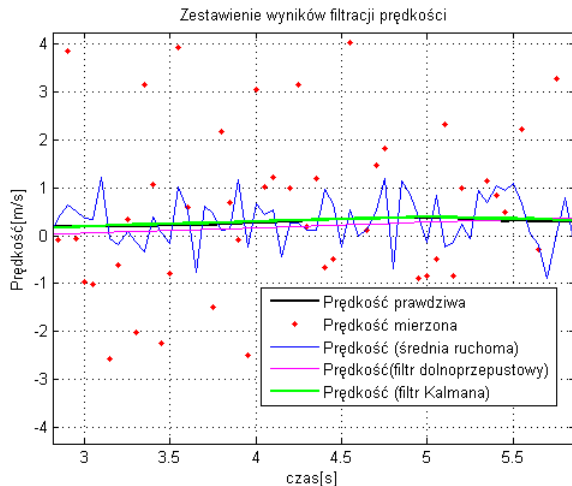
Filtr Kalmana - charakterystyka czasowa - prędkość



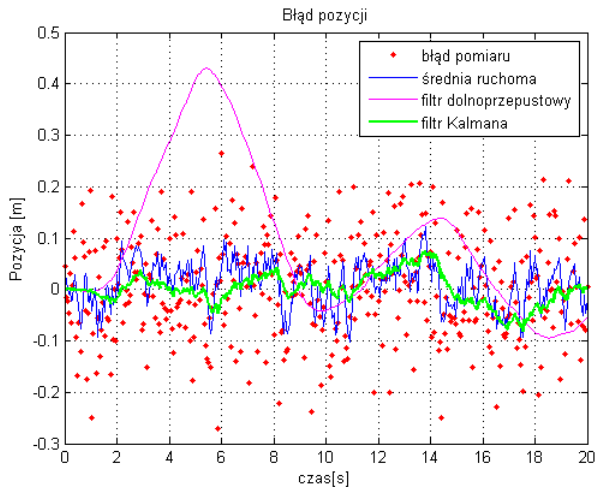
Zestawienie filtrów- charakterystyka czasowa - położenie



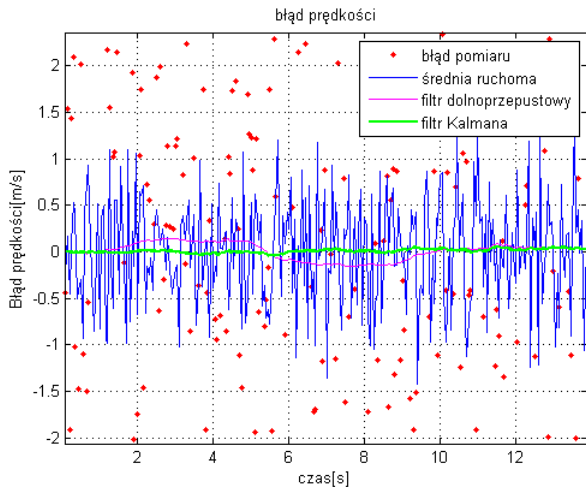
Zestawienie filtrów- charakterystyka czasowa - prędkość



Zestawienie filtrów- błąd położenia



Zestawienie filtrów- błąd prędkości



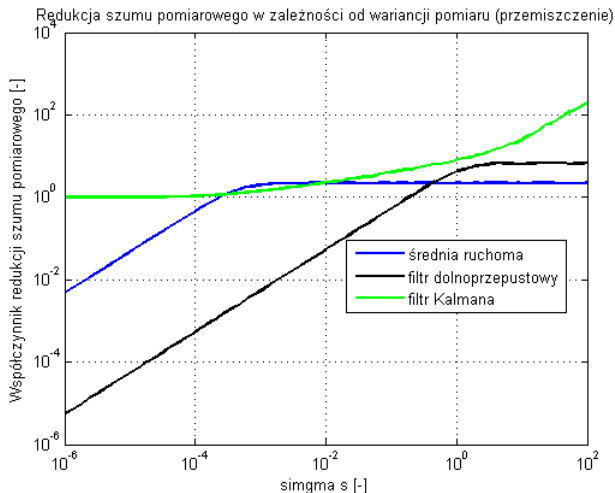
Filtr Kalmana - współczynnik tłumienia szumów

$$WTS = \frac{RMS(x_R - x_P)}{RMS(x_R - x_F)} = \frac{RMS(e_P)}{RMS(e_F)}$$

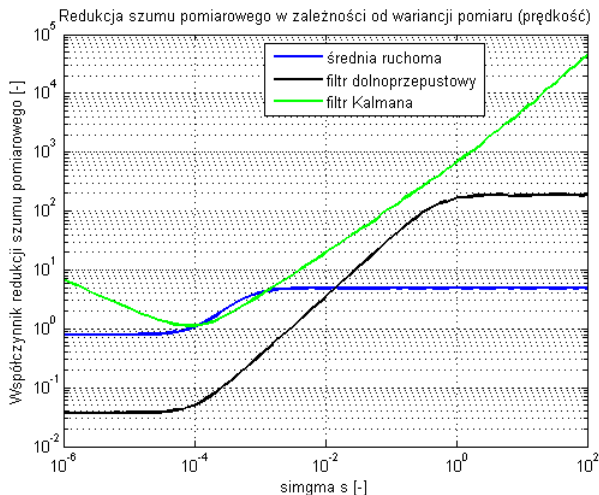
gdzie x_R - pozycja rzeczywista, x_P - pozycja zmierzona obarczona błędem, x_F - pozycja po filtracji.

W związku z czym im większy współczynnik WTS tym lepsze tłumienie szumów pomiarowych i lepsza estymacja zmiennych stanu.

Zestawienie filtrów - współczynnik tłumienia szumów - położenie



Zestawienie filtrów - współczynnik tłumienia szumów - prędkość



Prosty przykład ekonomiczny - trend liniowy

Bardzo prostym i intuicyjnym przykładem jest model trendu liniowego - który możemy zapisać w przestrzeni stanów w postaci:

$$\begin{aligned} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \\ y(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + v(t) \end{aligned} \tag{9}$$

Parametry α i β są nieobserwowanymi zmiennymi stanu, które w procesie filtracji chcemy estymować.

Filtr - główna idea. Trend liniowy - równania.

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$
$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + v_t$$

$$t = 1$$

$$y_1 = \alpha_1 + v_1$$

$$\alpha_1 = \alpha_0 + \beta_0$$

$$\beta_1 = \beta_0$$

$$t = 2$$

$$y_2 = \alpha_2 + v_2 = \alpha_0 + 2\beta_0$$

$$\alpha_2 = \alpha_1 + \beta_1 = \alpha_0 + \beta_0 + \beta_1 = \alpha_0 + 2\beta_0$$

$$\beta_2 = \beta_1 = \beta_0$$

Jaki więc zysk ma ekonomista z wykorzystania przestrzeni stanów?



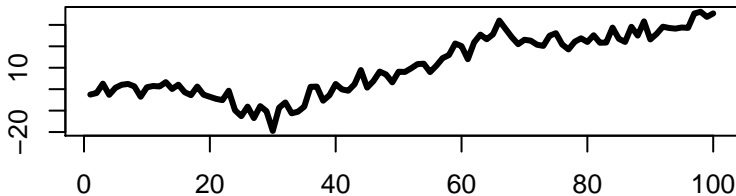
Model trendu liniowego ze zmieniającymi się parametrami

Co więcej przedstawienie modelu w przestrzeni stanów pozwala nam rozszerzyć klasę omawianych modeli. Zaprezentowany przebieg trendu możemy rozszerzyć na przypadek zmieniających się w czasie parametrów np. α i β podlegają błędzeniu losowemu - wtedy równania stanu będą wyglądały następująco:

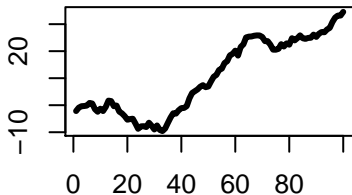
$$\begin{aligned} \begin{bmatrix} \alpha_{t+1} \\ \beta_{t+1} \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} + w_t \\ y_t &= [1 \quad 0] \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} + v_t \end{aligned} \tag{10}$$

Model trendu liniowego ze zmieniającymi się parametrami

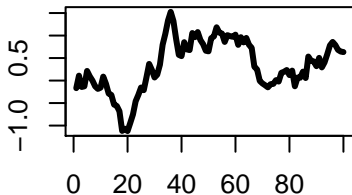
Proces



Alpha

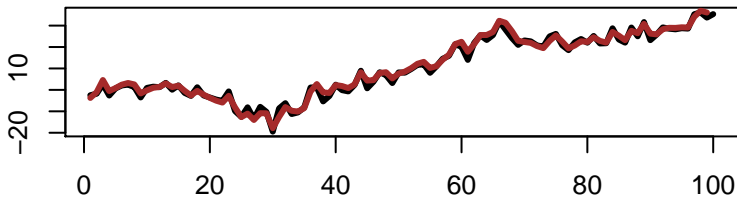


Beta

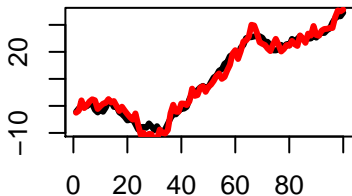


Model trendu liniowego ze zmieniającymi się parametrami - realizacje i wartości estymowane

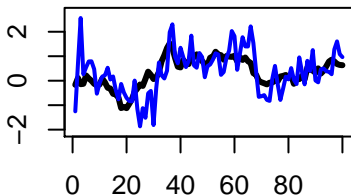
Proces



Alpha

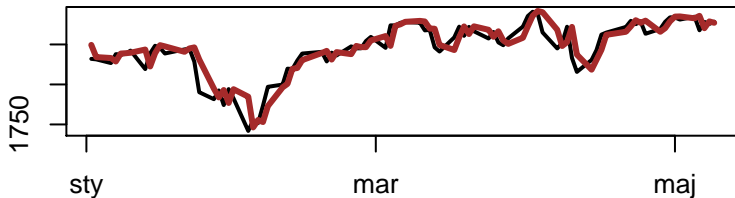


Beta

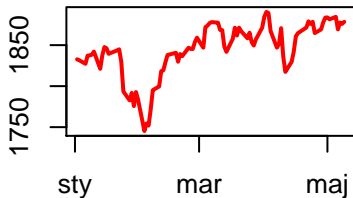


Model trendu liniowego - przykład empiryczny

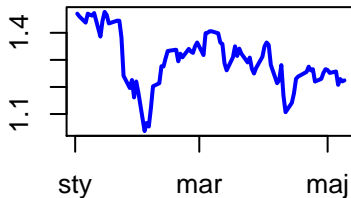
SP500



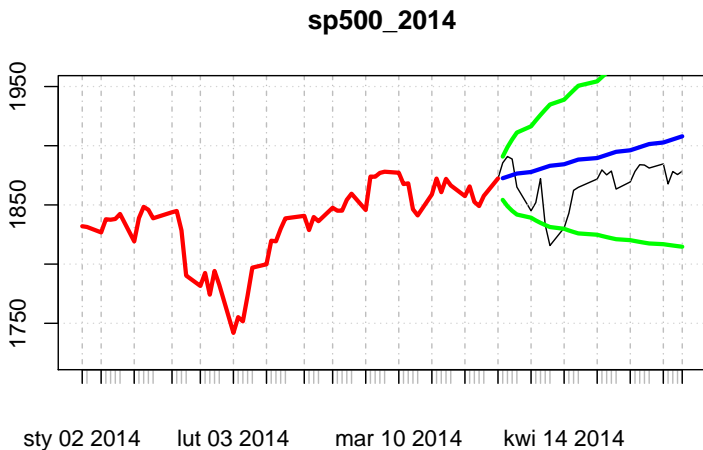
Alpha



Beta



Model trendu liniowego - przykład empiryczny - predykcja



Dylematy?

- ▶ W modelach znanych z teorii sterowania większość parametrów przestrzeni stanu jest znana - są to własności wynikające z fizyki.
- ▶ W przypadku modeli ekonomicznych w zasadzie wszystkie parametry podlegają estymacji, w tym rozmiar przestrzeni stanu (np. klasyczny problem związany z doбором zmiennych).

Bibliografia - podstawy



J.J.F. Commandeur and S.J. Koopman. *An Introduction to State Space Time Series Analysis*. PRACTICAL ECONOMETRICS SERIES. OUP Oxford, 2007. ISBN: 9780191607806.



Ramsey Faragher. "Understanding the Basis of the Kalman Filter Via a Simple and Intuitive Derivation [Lecture Notes]." In: *IEEE Signal Process. Mag.* 29.5 (2012), pp. 128–132.



Kędzierski J. "Filtr Kalmana - zastosowania w prostych układach sensorycznych". In: *Koło Naukowe Robotyków KoNaR* (2007).

Bibliografia - układy dynamiczne






K. Ogata. *State space analysis of control systems*. Prentice-Hall international series in electrical engineering. Prentice-Hall, 1967.
URL: http://books.google.pl/books?id=Z%5C_JQAAAAMAAJ.



Stanisław Osowski. *Modelowanie i symulacja układów i procesów dynamicznych*. Oficyna Wydawnicza Politechniki Warszawskiej, 2007.

Bibliografia - Filtr Kalmana w R

-  Fernando Tusell. “Kalman Filtering in R”. In: *Journal of Statistical Software* 39.2 (Mar. 1, 2011), pp. 1–27. ISSN: 1548-7660. URL: <http://www.jstatsoft.org/v39/i02>.
-  Petrone Petris and Campagnoli. *Dynamic Linear Models with R*. Springer, 2009.
-  Giovanni Petris. “An R Package for Dynamic Linear Models”. In: *Journal of Statistical Software* 36.12 (Oct. 13, 2010), pp. 1–16. ISSN: 1548-7660.