Modelowanie i identyfikacja

Wykład 7: Metody podprzestrzeni

Gniewomir Sarbicki

Dowolną rzeczywistą macierz B rozmiaru $n \times m$ możemy zapisać jako (SVD):

$$B = U \Lambda V^T, \tag{1}$$

gdzie $U \in O(n), V \in O(m)$, a $\Lambda_{ij} = \lambda_i \delta_{ij}, \lambda_i \geqslant 0$.

$$B = \sum_i \lambda_i u_i v_i^T. \tag{2}$$

Działanie dowolnej macierzy B na wektor x można opisać następująco:

- Znajdź składowe wektora x w bazie $\{v_i\}$.
- Pomnóż uzyskane składowe przez liczby λ_i
- Przez tak zmodyfikowane składowe pomnóż elementy bazy $\{u_i\}$

Zbiór $\{v_i : i < \min\{n, m\} \wedge \lambda_i \neq 0\}$ jest bazą ortonormalną przestrzeni wierszowej B,
Zbiór $\{u_i : i < \min\{n, m\} \wedge \lambda_i \neq 0\}$ jest bazą ortonormalną przestrzeni kolumnowej B.

Rząd B jest równy liczbie niezerowych elementów w macierzy Λ (niezerowych λ_i)

Przypadki szczególne, gdy $\Lambda_{ij} \in \{0, 1\}$:

- $n = m, \mathrm{rank}\Lambda = m \Rightarrow$ obrót
- $n > m, \mathrm{rank}\Lambda = m \Rightarrow$ zanurzenie izometryczne
- $\mathrm{rank}\Lambda < m \Rightarrow$ zanurzenie izometryczne rzutu

Zdefiniujemy pseudoodwrotność macierzy B jako:

$$B^{-1} = V\Lambda.^{-1}U^T, \tag{3}$$

gdzie $.^{-1}$ oznacza odwrotność po niezerowych elementach i transpozycję.

W przypadku macierzy odwracalnych (kwadratowych pełnego rzędu) otrzymujemy macierz odwrotną, dlatego nie wprowadzamy nowego oznaczenia.

Pseudoodwrotność macierzy B o kolumnach liniowo niezależnych możemy obliczyć za pomocą zwykłej odwrotności jako $B(B^T B)^{-1}$, a dla macierzy B o wierszach liniowo niezależnych $(BB^T)^{-1}B$.

Jeżeli rząd macierzy nie jest równy m (pełen rząd kolumnowy) ani n (pełen rząd wierszowy), to mówimy o macierzy pseudoodwrotnej Moora-Penrose'a (polecenie `pinv` w Matlabie, `scipy.linalg.pinv` w Pythonie).

Obliczamy ją wtedy za pomocą SVD (nieefektywne, przybliżone) lub wykorzystując rozkład Cholewskiego.

Mamy:

$$B^{-1}B = VV^T \stackrel{ozn}{=} \Pi_B = \text{rzut na przestrzeń wierszową } B \tag{4}$$

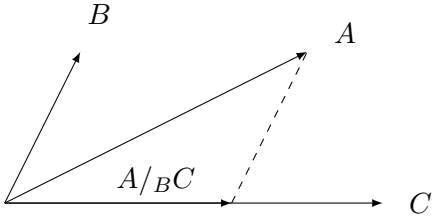
$$BB^{-1} = UU^T \stackrel{ozn}{=} \Pi_{B^T} = \text{rzut na przestrzeń kolumnową } B \tag{5}$$

Dla macierzy kwadratowych pełnego rzędu oba te rzuty to identyczności.

Oznaczamy: $A/B = A\Pi_B$. Jest to rzut ortogonalny przestrzeni wierszowej A na przestrzeń wierszową B.

Podobnie, jako $A/B^\perp \stackrel{ozn}{=} A\Pi_B^\perp$ oznaczamy rzut ortogonalny przestrzeni wierszowej A na dopełnienie ortogonalne przestrzeni wierszowej B.

Zachodzi: $A = A/B + A/B^\perp$.



Rzut ukośny wektora A na wektor C wzdłuż wektora B

W przypadku macierzy, mamy rozkład:

$$A = A/_B C + A/_C B + A/\begin{bmatrix} C \\ B \end{bmatrix}^{\perp} = A\begin{bmatrix} C \\ B \end{bmatrix}^{-1}\begin{bmatrix} C \\ 0 \end{bmatrix} + A\begin{bmatrix} C \\ B \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ B \end{bmatrix} + A/\begin{bmatrix} C \\ B \end{bmatrix}^{\perp} \tag{6}$$

Jeżeli macierze B i C mają ortogonalne przestrzenie wierszowe, to rzut ukośny redukuje się do rzutu ortogonalnego.

Wzór:

$$A/_B C = A \begin{bmatrix} C \\ B \end{bmatrix}^{-1} \begin{bmatrix} C \\ 0 \end{bmatrix}$$

Możemy przepisać jako:

$$A/_B C = (A/B^{\perp})(C/B^{\perp})^{-1} C \tag{7}$$

Poszukujemy minimalnego kąta między wektorami z przestrzeni wierszowych macierzy A i B, o rozkładach singularnych $A = U_A \Lambda_A V_A^T$ i $B = U_B \Lambda_B V_B^T$.

Dowolny wektor jednostkowy z przestrzeni wierszowej A możemy zapisać jako $x^T U_A V_A^T$, dla pewnego x o normie 1. Podobnie jako $y^T U_B V_B^T$ sparametryzujemy wektor jednostkowy z przestrzeni wierszowej B.

Poszukujemy $\max_{x,y:||x||=||y||=1} x^T U_A V_A^T V_B U_B^T y$. Jeżeli rozkładem singularnym $U_A V_A^T V_B U_B^T$ jest $\lambda_1 u_1 v_1^T + \ldots$ (w kolejności malejących λ_i), to maksimum tym będzie λ_1 i będzie ono realizowane dla wektorów $x = u_1$, $y = v_1$.

Jeżeli wyłączymy u_1 i v_1 z przestrzeni do których należą, to odpowiedzią o maksymalny cosinus kąta pomiędzy "resztami" jest λ_2 i jest on realizowany pomiędzy wektorami u_2 i v_2.

Pary wektorów (u_i, v_i) określają kierunki główne dla pary podprzestrzeni. Wartości singularne λ_i to cosinusy kątów głównych.

Kluczową macierz $U_A V_A^T V_B U_B^T$ obliczamy następująco:

$$(AA^T)^{-\frac{1}{2}} AB^T (BB^T)^{-\frac{1}{2}} = U_A \Lambda_A^{-1} U_A^T \cdot U_A \Lambda_A V_A^T V_B \Lambda_B U_B^T \cdot U_B \Lambda_B^{-1} U_B^T = U_A V_A^T V_B U_B^T \tag{8}$$

W przypadku macierzy o jednym wierszu: $A = \vec{a}^T, B = \vec{b}^T$:

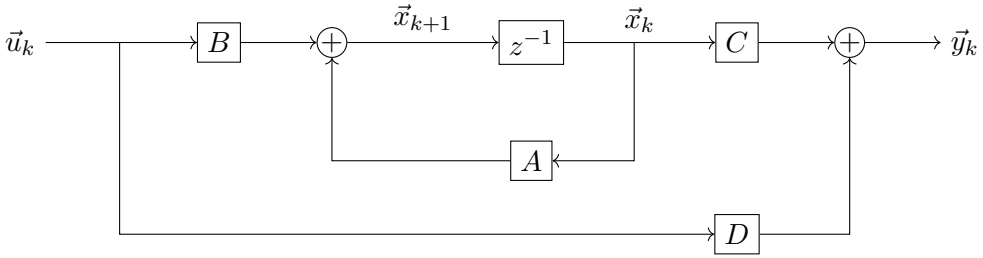
$$(\vec{a} \cdot \vec{a})^{-\frac{1}{2}} (\vec{a} \cdot \vec{b})(\vec{b} \cdot \vec{b})^{-\frac{1}{2}} = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \cdot ||\vec{b}||} = \cos \sphericalangle(\vec{a}, \vec{b}), \tag{9}$$

a kierunkami głównymi są oczywiście wektory $\vec{a}$ i $\vec{b}$.

Rozważmy układ dyskretny przedstawiony w przestrzeni stanów:

$$\vec{x}_{k+1} = A\vec{x}_k + B\vec{u}_k \tag{10}$$

$$\vec{y}_k = C\vec{x}_k + D\vec{u}_k \tag{11}$$



Wymiary obiektów w równaniach: $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^l$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{l\times n}$, $D \in \mathbb{R}^{l\times m}$,

Zakładamy, że układ jest sterowalny i obserwowalny, wprowadźmy macierze obserwowalności Γ_i i (odwróconą) sterowalności Δ_i:

$$\Gamma_i = \begin{bmatrix} C \\ CA \\ \vdots \\ C^{i-1}A \end{bmatrix} \qquad (12) \qquad \Delta_i = \begin{bmatrix} A^{i-1}B & \dots & AB & B \end{bmatrix}, \quad (13)$$

Mając dany ciąg wektorów $\vec{u}_i \in \mathbb{R}^m, i \in \{0, \ldots, s\}$ definiujemy macierz Hankela:

$$U_{0|2i-1} = \left[\begin{array}{ccccc} \vec{u}_0 & \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_{j-1} \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 & \ldots & \vec{u}_j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{u}_{i-1} & \vec{u}_i & \vec{u}_{i+1} & \ldots & \vec{u}_{j+i-2} \\ \hline \vec{u}_i & \vec{u}_{i+1} & \vec{u}_{i+2} & \ldots & \vec{u}_{j+i-1} \\ \vec{u}_{i+1} & \vec{u}_{i+2} & \vec{u}_{i+3} & \ldots & \vec{u}_{j+i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{u}_{2i-1} & \vec{u}_{2i} & \vec{u}_{2i+1} & \ldots & \vec{u}_{j+2i-2} \end{array}\right] \overset{ozn}{=} \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \end{bmatrix} \overset{ozn}{=} \begin{bmatrix} U_p \\ U_f \end{bmatrix} \quad (14)$$

inny rozkład:

$$U_{0|2i-1} = \left[\begin{array}{ccccc} \vec{u}_0 & \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_{j-1} \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 & \dots & \vec{u}_j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{u}_{i-1} & \vec{u}_i & \vec{u}_{i+1} & \dots & \vec{u}_{j+i-2} \\ \vec{u}_i & \vec{u}_{i+1} & \vec{u}_{i+2} & \dots & \vec{u}_{j+i-1} \\ \hline \vec{u}_{i+1} & \vec{u}_{i+2} & \vec{u}_{i+3} & \dots & \vec{u}_{j+i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{u}_{2i-1} & \vec{u}_{2i} & \vec{u}_{2i+1} & \dots & \vec{u}_{j+2i-2} \end{array}\right] \overset{ozn}{=} \begin{bmatrix} U_{0|i} \\ U_{i+1|2i-1} \end{bmatrix} \overset{ozn}{=} \begin{bmatrix} U_p^+ \\ U_f^- \end{bmatrix} \tag{15}$$

Indeks i powinien być "odpowiednio duży" - co najmniej kwadrat rzędu układu. Indeks j zaś taki, by wykorzystać wszystkie $\vec{u}_i$, czyli $j = s - 2i + 2$.

W podobny sposób z wektorów wyjść definiujemy macierz $Y_{0|2i-1}$ i jej bloki: Y_p, Y_f, Y_p^+, Y_f^-.

Tworzymy teraz dwie nowe macierze:

$$W_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix}, \qquad W_p^+ = \begin{bmatrix} U_p^+ \\ Y_p^+ \end{bmatrix} \tag{16}$$

Jako X_i będziemy oznaczali macierz zbudowaną z kolejnych j wektorów stanu zaczynając od wektora i-tego:

$$X_i = \begin{bmatrix} \vec{x}_i & \vec{x}_{i+1} & \dots & \vec{x}_{i+j-2} & \vec{x}_{i+j-1} \end{bmatrix} \tag{17}$$

Niech macierz H_i oznaczać będzie macierz blokową trójkątną dolną Toeplitza:

$$\begin{bmatrix} D & 0 & \dots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{i-2}B & \dots & CB & D \end{bmatrix} \tag{18}$$

Możemy przepisać równania układu (10-11) w kolejnych chwilach czasu jako równania macierzowe:

$$Y_p = \Gamma_i X_0 + H_i U_p \tag{19}$$

$$Y_f = \Gamma_i X_i + H_i U_f \tag{20}$$

$$X_i = A^i X_0 + \Delta_i U_p \tag{21}$$

Zakładamy, że:

- Sygnał wejściowy spełnia warunek trwałego pobudzania rzędu $2i$: $\operatorname{rank} \lim_{j\to\infty} U_{0|2i} U_{0|2i}^T = 2m$
  (Macierz korelacji wierszy macierzy $U_{0|2i}$ jest pełnego rzędu).
- Nie istnieją niezerowe wektory w przekroju podprzestrzeni wierszowych macierzy U_f (przyszłe wejścia) i X_0 (pierwsze j stanów).
- Wybrane przez użytkownika macierze wag: $W_1 \in \mathbb{R}^{li\times li}$ i $W_2 \in \mathbb{R}^{j\times j}$ są takie, że W_1 jest pełnego rzędu, a $\operatorname{rank}(W_p) = \operatorname{rank}(W_p W_2)$.

Definiujemy macierz $\mathcal{O}_i$ jako rzut ukośny: $\mathcal{O}_i = Y_f/_{U_f} W_p$. Jest ona równa $\mathcal{O} = \Gamma_i X_i$, a jej rząd jest równy rzędowi układu.
Używając rozkładu singularnego $W_1 \mathcal{O}_i W_2 = U_1 S_1 V_1^T$ otrzymamy:

$$\Gamma_i = W_1^{-1} U_1 S_1^{1/2} T \tag{22}$$

$$X_i = \Gamma_i^{-1} \mathcal{O}_i, \tag{23}$$

gdzie T jest dowolną macierzą odwracalną (transformacje podobieństwa w przestrzeni stanów).

Macierze stanu otrzymamy rozwiązując układ równań:

$$\begin{bmatrix} X_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X_i \\ U_{i|i} \end{bmatrix}, \tag{24}$$

gdzie:

$$X_{i+1} = \Gamma_{i-1}^{-1} \mathcal{O}_{i-1}, \tag{25}$$

$$\mathcal{O}_{i-1} = Y_f^- /_{U_f^-} W_p^+ \tag{26}$$

Dowód:

Z równania na Y_p (19) wyznaczamy X_0 i wstawiamy do równania na X_i (21):

$$X_i = \left[\ \Delta_i - A^i \Gamma_i^{-1} H_i \ \middle|\ A^i \Gamma_i^{-1}\ \right] \left[\frac{U_p}{Y_p}\right] \overset{ozn}{=} L_p W_p \tag{27}$$

Następnie wstawiamy X_i do równania na Y_f (20):

$$Y_f = \Gamma_i L_p W_p + H_i U_f \tag{28}$$

Pomnożmy te równanie z prawej strony przez $\Pi_{U_f^\perp}$:

$$Y_f \Pi_{U_f^\perp} = \Gamma_i L_p W_p \Pi_{U_f^\perp}, \tag{29}$$

$$Y_f / U_f^\perp = \Gamma_i L_p W_p / U_f^\perp, \tag{30}$$

$$\mathcal{O} = Y_f /_{U_f} W_p = Y_f / U_f^\perp \left(W_p / U_f^\perp\right)^{-1} W_p = \Gamma_i L_p W_p = \Gamma_i X_i \tag{31}$$

Dowód:

Macierz $W_1 \mathcal{O} W_2$, o rozkładzie singularnym

$$\begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ U_2^T \end{bmatrix} \tag{32}$$

jest iloczynem macierzy $W_1 \Gamma_1$ i macierzy $X_i W_2$:

$$W_1 \Gamma_i X_i W_2 = U_1 S_1^{1/2} S_1^{1/2} V_1^T, \tag{33}$$

stąd:

$$W_1 \Gamma_i = U_1 S_1^{1/2} T \tag{34}$$

$$X_i W_2 = T^{-1} S_1^{1/2} V_1^T, \tag{35}$$

Stąd obliczamy $\Gamma_i = W_1^{-1} U_1 S_1^{1/2} T$ i dalej $X_i = \Gamma_i^{-1} \mathcal{O}_i$

Przepisujemy równania na Y_f (19 - 21) dla zmienionych zakresów chwil czasowych:

$$Y_p^+ = \Gamma_{i+1} X_0 + H_{i+1} U_p^+ \tag{36}$$

$$Y_f^- = \Gamma_{i-1} X_{i+1} + H_{i-1} U_f^- \tag{37}$$

$$X_{i+1} = L_p^+ W_p^+ \tag{38}$$

Postępując podobnie obliczamy $X_{i+1} = \Gamma_{i-1}^{-1} Y_f^- /_{U_f^i} W_p^+ \ \square$

W szczególności, wybór $W_1 = I$, $W_2 = \Pi_{U_f^\perp}$ prowadzi do: $W_1 \mathcal{O}_i W_2 = Y_f / U_f^\perp$.

Rozważmy układ bez przenoszenia:

$$x_{k+1} = A_k x_k + B_k u_k + w_k \tag{39}$$

$$y_k = C_k x_k + v_k \tag{40}$$

zmienne losowe $w_k \sim \mathcal{N}(0, Q_k)$ i $v_k \sim \mathcal{N}(0, R_k)$ są w różnych chwilach czasu niezależne.

Macierz kowariancji błędu estymacji: $P_k = \mathbb{E}(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T$.

Estymata $\hat{x}_{k|k-1}$ wyznaczana ze znajomości układu jest uaktualniana proporcjonalnie do uchybu estymacji (innowacji):

$$\hat{x}_k = \underbrace{\hat{x}_{k|k-1}}_{\text{człon predykcyjny}} + \underbrace{K_k(y_k - C_k\hat{x}_{k|k-1})}_{\text{człon korekcyjny}} = (I - K_k C_k)\hat{x}_{k|k-1} + K_k y_k \tag{41}$$

Macierz kowariancji uaktualni się wtedy następująco (założenie niezależności szumów):

$$P_k = (I - K_k C_k)P_{k|k-1}(I - K_k C_k)^T + K_k R_k K_k^T \tag{42}$$

Macierz K_k chcemy dobrać tak, by zminimalizować błąd średniokwadratowy, czyli $\mathrm{Tr} P_k$.

Różniczkując $\mathrm{Tr}P_k$ po wyrazach macierzy K_k (i przyrównując do 0) otrzymamy:

$$-2(C_k P_{k|k-1})^T + 2K_k(C_k P_{k|k-1} C_k^T + R_k) = 0 \tag{43}$$

i stąd:

$$K_k = P_{k|k-1} C_k^T (C_k P_{k|k-1} C_k^T + R_k)^{-1} = P_{k|k-1} C_k^T Y^{-1}, \tag{44}$$

gdzie Y to macierz kowariancji innowacji. Podstawiając to do wzoru na P_k otrzymamy:

$$P_k = (I - K_k C_k) P_{k|k-1} \tag{45}$$

$\hat{x}_{k+1|k}$ - nieobciążona estymata a priori wektora x_k jest równa: $A_k \hat{x}_k + B_k u_k$, a stąd:

$$P_{k+1|k} = A_k P_k A_k^T + Q_k \tag{46}$$

Otrzymujemy algorytm rekurencyjny:

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1} + B_{k-1}u_{k-1} \tag{47}$$

$$P_{k|k-1} = A_{k-1}P_{k-1}A_{k-1}^T + Q_{k-1} \tag{48}$$

$$K_k = P_{k|k-1}C_k^T(C_kP_{k|k-1}C_k^T + R_k)^{-1} \tag{49}$$

$$\hat{x}_k = (I - K_kC_k)\hat{x}_{k|k-1} + K_ky_k \tag{50}$$

$$P_k = (I - K_kC_k)P_{k|k-1} \tag{51}$$

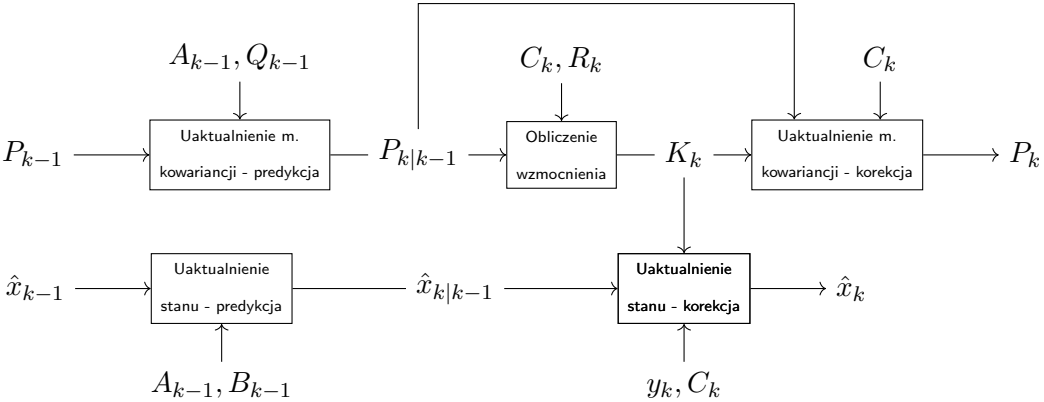
Opisany algorytm nazywamy fitrem Kálmána (Thiele, Swerling, Bucy, Stratonowicz). Łącząc równania (50) i (47) otrzymamy równanie:

$$\hat{x}_{k+1|k} = A_k(I - K_kC_k)\hat{x}_{k|k-1} + B_ku_k + A_kK_ky_k, \tag{52}$$

które wprowadzając zmienną $e_k = y_k - C_k\hat{x}_{k|k-1}$ (innowacja) możemy zapisać jako:

$$\hat{x}_{k+1|k} = A_k\hat{x}_{k|k-1} + B_ku_k + A_kK_ke_k \tag{53}$$

$$y_k = C_k\hat{x}_{k|k-1} + e_k \tag{54}$$


A_{k-1}, Q_{k-1}
C_k, R_k
C_k
P_{k-1}
Uaktualnienie m. kowariancji - predykcja
$P_{k|k-1}$
Obliczenie wzmocnienia
K_k
Uaktualnienie m. kowariancji - korekcja
P_k
$\hat{x}_{k-1}$
Uaktualnienie stanu - predykcja
$\hat{x}_{k|k-1}$
Uaktualnienie stanu - korekcja
$\hat{x}_k$
A_{k-1}, B_{k-1}
y_k, C_k

Asymptotycznie, dla stałych w czasie parametrów układu i szumu, filtr Kalmana osiąga swój punkt stacjonarny:

$$P = A(I - KC)PA^T + Q \tag{55}$$

$$K = PC^T(CPC^T + R)^{-1} \tag{56}$$

Wstawiając drugie równanie do pierwszego, otrzymamy równanie, które powinna spełniać macierz P (macierz kowariancji błędu estymacji a priori):

$$P = A(I - PC^T(CPC^T + R)^{-1}C)PA^T + Q \tag{57}$$

- dyskretne równanie Riccatiego.

Macierz wzmocnienia K otrzymamy wstawiając rozwiązanie równania (57) do (56)

W przypadku, gdy szumy są zależne: $\mathbb{E}(w_k v_k^T) = S_k \neq 0$, modyfikacji ulega człon predykcyjny filtru (47 - 48):

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1} + B_{k-1}u_{k-1} + S_{k-1}R_{k-1}^{-1}(y_{k-1} - C_{k-1}\hat{x}_{k-1}) \tag{58}$$

$$P_{k|k-1} = (A_{k-1} + S_{k-1}R_{k-1}^{-1}C_{k-1})P_{k-1}(A_{k-1} + S_{k-1}R_{k-1}^{-1}C_{k-1})^T + Q_{k-1} - S_{k-1}R_{k-1}^{-1}S_{k-1}^T \tag{59}$$

$$K_k = P_{k|k-1}C_k^T(C_kP_{k|k-1}C_k^T + R_k)^{-1} \tag{60}$$

$$\hat{x}_k = (I - K_kC_k)\hat{x}_{k|k-1} + K_ky_k \tag{61}$$

$$P_k = (I - K_kC_k)P_{k|k-1} \tag{62}$$

W przypadku stacjonarnym asymptotycznie $K = PC^T(CPC^T + R)^{-1}$ dla P spełniającego dyskretne równanie Riccatiego ze zmodyfikowanymi A i Q:

$$P = (A - SR^{-1}C)(I - PC^T(CPC^T + R)^{-1}C)P(A - SR^{-1}C)^T + Q - SR^{-1}S^T \tag{63}$$

Równania stanu estymaty $\hat{x}_k$ zmodyfikują się o składnik zależny od uchybu estymaty wyjścia w kroku poprzednim:

$$\hat{x}_{k+1|k} = A_k\hat{x}_{k|k-1} + B_k u_k + (A_k P_{k|k-1} C_k^T + S_k)(C_k P_{k|k-1} C_k^T + R_k)^{-1} e_k \tag{64}$$

$$y_k = C_k \hat{x}_k + e_k \tag{65}$$

Zauważmy, że możemy przepisać równanie Riccatiego (63) tak, by zawierało tak określoną macierz wzmocnienia:

$$P = APA^T - (APC^T + S)(CPC^T + R)^{-1}(APC^T + S)^T + Q \tag{66}$$

Będziemy rozważać układ stochastyczny bez sygnału wejściowego:

$$\vec{x}_{k+1} = A\vec{x}_k + \vec{w}_k \tag{67}$$

$$\vec{y}_k = C\vec{x}_k + \vec{v}_k \tag{68}$$

Przypomnijmy, że przyjmujemy, że szumy w różnych chwilach czasu są niezależne i mają macierz kowariancji:

$$\mathbb{E}\left(\begin{bmatrix} w_p \\ v_p \end{bmatrix}\begin{bmatrix} w_q^T & v_q^T \end{bmatrix}\right) = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}\delta_{pq} \tag{69}$$

W tej sytuacji, układ jest stabilny ($\sigma(A)$ w kole jednostkowym) $\Leftrightarrow$ x_k jest procesem stacjonarnym. Będziemy zakładać, że para $(A, Q^{1/2})$ jest sterowalna (szum wejściowy wzbudza wszystkie mody układu).

Definujemy macierze kowariancji:

$$\mathbb{E}\left(\begin{bmatrix} x_{k+i} \\ y_{k+i} \end{bmatrix} \begin{bmatrix} x_k^T & y_k^T \end{bmatrix}\right) \stackrel{df}{=} \begin{bmatrix} \Sigma_i & G_i \\ G_i^T & \Lambda_i \end{bmatrix} \tag{70}$$

Spełniają one równania:

$$\Sigma_0 = A\Sigma_0 A^T + Q \tag{71}$$

$$G_1 = A\Sigma_0 C^T + S \tag{72}$$

$$\Lambda_0 = C\Sigma_0 C^T + R \tag{73}$$

$$\Lambda_i = CA^{i-1}G_1 \tag{74}$$

$$\Lambda_{-i} = G_1^T (A^{i-1})^T C^T \tag{75}$$

Wprowadzamy korelacyjną rozszerzoną macierz sterowalności:

$$\Delta_i^c = \left[\begin{array}{c|c|c|c|c} A^{i-1}G_1 & A^{i-2}G_1 & \dots & AG_1 & G_1 \end{array} \right] \tag{76}$$

Oraz macierze blokowe Toeplitza:

$$C_i = \begin{bmatrix} \Lambda_i & \Lambda_{i-1} & \dots & \Lambda_1 \\ \Lambda_{i+1} & \Lambda_i & \dots & \Lambda_2 \\ \vdots & \ddots & \ddots & \vdots \\ \Lambda_{2i-1} & \Lambda_{2i-2} & \dots & \Lambda_i \end{bmatrix}, \qquad L_i = \begin{bmatrix} \Lambda_0 & \Lambda_{-1} & \dots & \Lambda_{1-i} \\ \Lambda_1 & \Lambda_0 & \dots & \Lambda_{2-i} \\ \vdots & \ddots & \ddots & \vdots \\ \Lambda_{i-1} & \Lambda_{i-2} & \dots & \Lambda_0 \end{bmatrix} \tag{77}$$

Są one blokami macierzy korelacji wektora danych wyjściowych:
$C_i = \mathbb{E}(Y_f Y_p^T) \overset{j\to\infty}{\longleftarrow} \frac{1}{j} Y_f Y_p^T$, $L_i = \mathbb{E}(Y_f Y_f^T) = \mathbb{E}(Y_p Y_p^T) \overset{j\to\infty}{\longleftarrow} \frac{1}{j} Y_f Y_f^T = \frac{1}{j} Y_p Y_p^T$.

Załóżmy, że estymujemy stan układu (o znanych A i C) za pomocą niestacjonarnego filtru Kalmana:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + K_k^f e_k \tag{78}$$

$$y_k = C\hat{x}_{k|k-1} + e_k \tag{79}$$

Estymata i jej uchyb są procesami niezależnymi: $\mathbb{E}(\hat{x}_{k|k-1}(\hat{x}_{k|k-1} - x_k)) = 0$, stąd:

$$\Sigma_0 = \mathbb{E}(x_k x_k^T) = \mathbb{E}(\hat{x}_{k|k-1}\hat{x}_{k|k-1}^T) + \mathbb{E}((\hat{x}_{k|k-1} - x_k)(\hat{x}_{k|k-1} - x_k)^T) \stackrel{df}{=} \widetilde{P}_{k-1} + P_{k|k-1} \tag{80}$$

Macierz korelacji innowacji jest równa: $Y_k = CP_{k|k-1}C^T + R = \Lambda_0 - C\widetilde{P}_{k-1}C^T$.

Macierze $\widetilde{P}_k$ i K_k^f spełniają zależności rekurencyjne:

$$K_k^f = (G_1 - A\widetilde{P}_{k-1}C^T)(\Lambda_0 - C\widetilde{P}_{k-1}C^T)^{-1} \tag{81}$$

$$\widetilde{P}_k = A\widetilde{P}_{k-1}A^T + (G_1 - A\widetilde{P}_{k-1}C^T)(\Lambda_0 - C\widetilde{P}_{k-1}C^T)^{-1}(G_1 - A\widetilde{P}_{k-1}C^T)^T \tag{82}$$

W granicy, rozwiązaniem stacjonarnym jest : $K^f = (G_1 - A\widetilde{P}C^T)(\Lambda_0 - C\widetilde{P}C^T)^{-1}$,

$$\widetilde{P} = A\widetilde{P}A^T + (G_1 - A\widetilde{P}C^T)(\Lambda_0 - C\widetilde{P}C^T)^{-1}(G_1 - A\widetilde{P}C^T)^T \tag{83}$$

Model (78,79) z tak określonymi macierzami nazywać będziemy modelem **innowacji w przód**.

Estymatę $\hat{x}_k$ i macierz kowariancji P_k można wyrazić jako:

$$\hat{x}_k = \Delta_k^c L_k^{-1} \begin{bmatrix} y_0 \\ \vdots \\ y_{k-1} \end{bmatrix} \tag{84}$$

$$P_k = \Delta_k^c L_k^{-1} (\Delta_k^c)^T \tag{85}$$

Równanie estymaty w przód $\hat{x}_k$ możemy zapisać za pomocą macierzy $\hat{X}_i = [\hat{x}_i, \hat{x}_{i+1}, \ldots, x_{i+j-1}]$ kolejnych j estymat w przód:

$$\hat{X}_i = \Delta_i^c L_i^{-1} Y_p \tag{86}$$

Zauważmy, że:

$$\Gamma_i \hat{X}_i = \Gamma_i \Delta_i^c L_i^{-1} Y_p = C_i L_i^{-1} Y_p = Y_f Y_p^T (Y_p Y_p^T)^{-1} Y_p = Y_f Y_p Y_p^{-1} = Y_f / Y_p \stackrel{df}{=} \mathcal{O}_i \tag{87}$$

Postępując jak poprzednio (teoria identyfikacji układów deterministycznych), mamy swobodę wyboru macierzy W_1 (pełnego rzędu) i W_2 (nie zmniejszającą rzędu wierszowego Y_p) i dalej:

$$W_1 \mathcal{O}_i W_2 = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 S_1 V_1^T \tag{88}$$

$$\Gamma_i = W_1^{-1} U_1 S_1^{1/2} T \tag{89}$$

$$\Delta_i^c = \Gamma_i^{-1} C_i \tag{90}$$

$$\hat{X}_i = \Gamma_i^{-1} \mathcal{O}_i \tag{91}$$

gdzie T to dowolna macierz nieosobliwa (swoboda wyboru bazy przestrzeni stanów).

Rozważamy teraz układ stochastyczny z sygnałem wejściowym:

$$\vec{x}_{k+1} = A\vec{x}_k + B\vec{u}_k + \vec{w}_k \tag{92}$$

$$\vec{y}_k = C\vec{x}_k + D\vec{u}_k + \vec{v}_k \tag{93}$$

Zakładamy, że para $[A, B|Q^{1/2}]$ jest sterowalna (wszystkie mody są osiągalne przez wejścia lub przez szum).

Rozbijamy układ na część deterministyczną i stochastyczną: $\vec{x}_k = \vec{x}_k^{(d)} + \vec{x}_k^{(s)}$, $\vec{y}_k = \vec{y}_k^{(d)} + \vec{y}_k^{(s)}$. Obie części podlegają równaniom stanu:

$$\begin{cases} \vec{x}_{k+1}^{(d)} = A\vec{x}_k^{(d)} + B\vec{u}_k^d \\ \vec{y}_k^{(d)} = C\vec{x}_k^{(d)} + D\vec{u}_k^d \end{cases}$$

$$\begin{cases} \vec{x}_{k+1}^{(s)} = A\vec{x}_k^{(s)} + \vec{w}_k \\ \vec{y}_k^{(s)} = C\vec{x}_k^{(s)} + \vec{v}_k^d \end{cases}$$

O procesach $\{\vec{u}_k\}$, $\{\vec{x}_k^{(d)}\}$, $\{\vec{x}_k^{(s)}\}$, $\{\vec{y}_k\}$ zakładamy, że są quasi-stacjonarne.

Dla wyjść obu części definiujemy oddzielne macierze Hankela $Y^{(d)}_{0|2i-1}$ i $Y^{(s)}_{0|2i-1}$.
Podobnie definiujemy oddzielne macierze $X^{(d)}_0$, $X^{(d)}_i$, $X^{(s)}_0$, $X^{(s)}_i$.

Definiujemy macierze kowariancji:

$$R^{uu} = \mathbb{E}(U_{0|2i-1}U^T_{0_2i-1}) \overset{j\to\infty}{\longleftarrow} \frac{1}{j}U_{0|2i-1}U^T_{0_2i-1} \tag{94}$$

$$S^{xu} = \mathbb{E}(X_0U^T_{0_2i-1}) \overset{j\to\infty}{\longleftarrow} \frac{1}{j}X_0U^T_{0_2i-1} \tag{95}$$

$$\Sigma^d = \mathbb{E}(X_0X^T_0) \overset{j\to\infty}{\longleftarrow} \frac{1}{j}X_0X^T_0 \tag{96}$$

O macierzy R^{uu} zakładamy, że jest pełnego rzędu (warunek trwałego wzbudzenia).

Będziemy również używać macierzy C_i, L_i oraz Δ^c_i.

Równania stanu możemy zapisać w wersji macierzowej:

$$Y_p = \Gamma_i X_0^d + H_i^d U_p + Y_p^s \tag{97}$$

$$Y_f = \Gamma_i X_i^d + H_i^d U_i + Y_f^s \tag{98}$$

$$X_f^d = A^i X_p^d + \Delta_i^d U_p \tag{99}$$

Niech początkową estymatą będzie x_0, a jej macierz kowariancji: P_0. Równania estymaty możemy wyrazić następująco:

$$\hat{x}_k = A\hat{x}_{k-1} + Bu_k + K_{k-1}(y_{k-1} - C\hat{x}_{k-1} - DU_{k-1}) \tag{100}$$

$$K_{k-1} = (G - AP_{k-1}C^T)(\Lambda_0 - CP_{k-1}C^T)^{-1} \tag{101}$$

$$P_k = AP_{k-1}A^T + (G - AP_{k-1}C^T)(\Lambda_0 - CP_{k-1}C^T)^{-1}(G - AP_{k-1}C^T)^T \tag{102}$$

możemy zapisać je następująco:

$$\hat{x}_k = [A - \Omega_k\Gamma_k|\Delta_k^c - \Omega_k H_k^d|\Omega_k][\hat{x}_0|u_0 \dots u_{k-1}|y_0 \dots y_{k-1}] \tag{103}$$

$$P_k = A^k P_0 (A^T)^k + (\Delta_k^c - A^k P_0 \Gamma_k^T)(L_k - \Gamma_k P_0 \Gamma_k^T)^{-1}(\Delta_k^c - A^k P_0 \Gamma_k^T)^T \tag{104}$$

gdzie

$$\Omega_k = (\Delta_k^c - A^k P_0 \Gamma_k^T)(L_k - \Gamma_k P_0 \Gamma_k^T)^{-1} \tag{105}$$

Możemy znów zebrać równania na estymaty w jedno równanie macierzowe:

$$\hat{X}_i = [A^i - \Omega_i \Gamma_i | \Delta_i^d - \Omega_i H_i^d | \Omega_i] \begin{bmatrix} \hat{X}_0 \\ \hline U_p \\ \hline Y_p \end{bmatrix} \tag{106}$$

Zauważmy, że wzór zawiera $\hat{X}_0$ i P_0. Nadamy im następujące wartości:

$$\hat{X}_0 = S^{xy}(R^{uu})^{-1} \begin{bmatrix} U_p \\ U_f \end{bmatrix} \tag{107}$$

$$P_0 = -(\Sigma^d - S^{xy}(R^{uu})^{-1}(S^{xy})^T) \tag{108}$$

Można je zapisać jako:

$$\hat{X}_0 = X_p^d / \begin{bmatrix} U_p \\ U_f \end{bmatrix}, \qquad P_0 = \mathbb{E}\hat{X}_0 \hat{X}_0^T$$

Zdefiniujmy macierz:

$$\mathcal{Z}_i \stackrel{df}{=} Y_f / \begin{bmatrix} W_p \\ U_f \end{bmatrix} \tag{109}$$

Przy takim wyborze $\hat{X}_0$ i P_0 zachodzi: $\mathcal{Z}_i = \Gamma_i \hat{X}_i + H_i^d U_f$

Przez $\widetilde{X}_i$ oznaczymy tęsamą macierz dla innego wyboru wielkości początkowych:

$$\hat{X}_0 = X_p^d /_{U_f} U_p \tag{110}$$

$$P_0 = -(\Sigma^d - S^{xy}(R^{uu})^{-1}(S^{xy})^T) \tag{111}$$

Spełnia ona równanie:

$$\mathcal{O}_i \stackrel{df}{=} Y_f /_{U_f} W_p = \Gamma_i \widetilde{X}_i \tag{112}$$