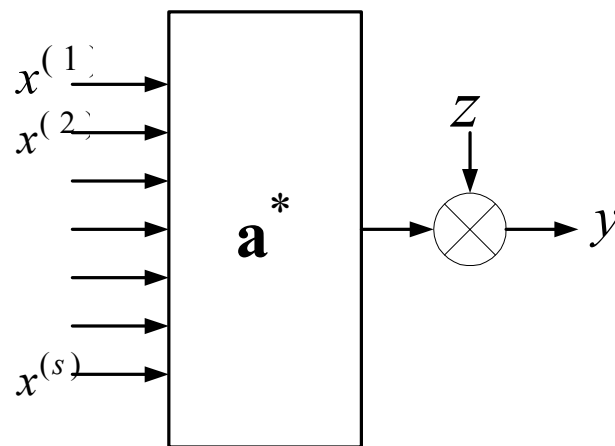


6.

**Identyfikacja wielowymiarowych systemów statycznych
metodą najmniejszych kwadratów**

1. Przedmiot identyfikacji

System



$$y = F^*(\mathbf{x}, z) = F(\mathbf{x}, z, \mathbf{a}^*), \text{ gdzie } \mathbf{x} = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(s)} \end{bmatrix} - \text{mierzone, } \mathbf{a}^* = \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_s^* \end{bmatrix} - \text{zestaw współczynników konkretyzujących } F()$$

$F()$ – znane, \mathbf{a}^* – nieznane

informacja aprioryczna

$$F^* \in \{F(\mathbf{x}, z, \mathbf{a}); \mathbf{a} \in R^s\}$$

istnieje dokładnie jedno \mathbf{a}^* , takie że $F^*(\mathbf{x}, z) = F(\mathbf{x}, z, \mathbf{a}^*)$ dla każdej pary (\mathbf{x}, z) (**identyfikowalność**)

informacja pomiarowa

$$\{\mathbf{x}_k, y_k\}_{k=1}^N$$

cel

znaleźć przepis $\Psi()$ (algorytm identyfikacji)

$$\Psi(\text{inf. aprioryczna, inf. pomiarowa}) = \mathbf{a}_N \sim \mathbf{a}^*$$

założenia o zakłóceniach

$$\begin{aligned} \text{jest addytywne, tj. } y &= F(\mathbf{x}, \mathbf{a}^*) + z \\ \{z_k\}_{k=1}^N &\text{ - ciąg i.i.d.} \\ \mathbf{E}z &= 0 \\ \text{var}z &< \infty \\ z, \mathbf{x} &\text{ - niezależne} \end{aligned}$$

natura procesów

z – losowe, \mathbf{x} – deterministyczne lub losowe, y – losowe, \mathbf{a}_N – losowe

własności \mathbf{a}_N

- praktyczne: wartości $\mathbf{E}\mathbf{a}_N[i]$ i $\text{var}\mathbf{a}_N[i]$
- asymptotyczne: fakt i typ zbieżności $\mathbf{a}_N \rightarrow \mathbf{a}^*$

2. Metoda najmniejszych kwadratów (Gauss 1809)

wskaźnik teoretyczny oceniający jakość dopasowania

$$Q_J(\mathbf{a}) = \mathbf{E}\{y - F(\mathbf{x}, \mathbf{a})\}^2 \rightarrow \min_{\mathbf{a}}$$

$$\begin{aligned} Q_J(\mathbf{a}) &= \mathbf{E}\{F(\mathbf{x}, \mathbf{a}^*) + z - F(\mathbf{x}, \mathbf{a})\}^2 = \mathbf{E}\{[F(\mathbf{x}, \mathbf{a}^*) - F(\mathbf{x}, \mathbf{a})]^2 + 2[F(\mathbf{x}, \mathbf{a}^*) - F(\mathbf{x}, \mathbf{a})]z + z^2\} = \\ &= \mathbf{E}\{F(\mathbf{x}, \mathbf{a}^*) - F(\mathbf{x}, \mathbf{a})\}^2 + \text{var}z \end{aligned}$$

konkluzja

$$\mathbf{a}^* = \arg \min_{\mathbf{a}} Q_J(\mathbf{a})$$

w praktyce nie jest możliwe obliczenie $Q_J(\mathbf{a})$ (brak znajomości odpowiednich rozkładów), proponuje się estymację wartości oczekiwanej za pomocą wartości średniej

$$\widehat{Q}_{J,N}(\mathbf{a}) = \frac{1}{N} \sum_{k=1}^N \{y_k - F(\mathbf{x}_k, \mathbf{a})\}^2 \rightarrow \min_{\mathbf{a}}$$

uśredniane składniki $\{y_k - F(\mathbf{x}_k, \mathbf{a})\}^2$ tworzą ciąg i.i.d., zatem na podstawie *MPWL*

$$\widehat{Q}_{J,N}(\mathbf{a}) \xrightarrow{p.1} Q_J(\mathbf{a}), \text{ gdy } N \rightarrow \infty$$

pytanie: czy prawdziwe jest następujące wnioskowanie

$$\widehat{Q}_{J,N}(\mathbf{a}) \xrightarrow{p.1} Q_J(\mathbf{a}) \implies \arg \min_{\mathbf{a}} \widehat{Q}_{J,N}(\mathbf{a}) \xrightarrow{p.1} \arg \min_{\mathbf{a}} Q_J(\mathbf{a})$$

odpowiedź: tylko wtedy, gdy

$$\sup_{\mathbf{a}} (\widehat{Q}_{J,N}(\mathbf{a}) - Q_J(\mathbf{a})) \xrightarrow{p.1} 0$$

3. Liniowy system statyczny *MIMO*

$$\begin{aligned} y &= \mathbf{a}^{*T} \mathbf{x} + z = \mathbf{x}^T \mathbf{a}^* + z \\ y_k &= \mathbf{x}_k^T \mathbf{a}^* + z_k, \quad k = 1, 2, \dots, N \end{aligned}$$

opis macierzowo–wektorowy

$$X_N = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix} = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(s)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(s)} \\ \dots & \dots & \dots & \dots \\ x_N^{(1)} & x_N^{(2)} & \dots & x_N^{(s)} \end{bmatrix}, \quad Y_N = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad Z_N = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}$$

równanie pomiarów

$$Y_N = X_N \mathbf{a}^* + Z_N$$

wektor wyjść modelu

$$\bar{Y}_N = X_N \mathbf{a}$$

empiryczne kryterium jakości

$$\begin{aligned}\bar{Q}_{J,N}(\mathbf{a}) &= \sum_{k=1}^N \{y_k - \mathbf{x}_k^T \mathbf{a}\}^2 = (Y_N - X_N \mathbf{a})^T (Y_N - X_N \mathbf{a}) = \|Y_N - X_N \mathbf{a}\|_2^2 = \dots - \text{kwadrat normy euklidesowej} \\ \dots &= (Y_N^T - \mathbf{a}^T X_N^T)(Y_N - X_N \mathbf{a}) = Y_N^T Y_N - \mathbf{a}^T X_N^T Y_N - Y_N^T X_N \mathbf{a} + \mathbf{a}^T X_N^T X_N \mathbf{a} \rightarrow \min_{\mathbf{a}}\end{aligned}$$

gradient

$$\nabla_{\mathbf{a}} \bar{Q}_{J,N}(\mathbf{a}) = -2X_N^T Y_N + 2X_N^T X_N \mathbf{a} = 0$$

równanie normalne

$$X_N^T X_N \mathbf{a} = X_N^T Y_N$$

$$\dim X_N^T X_N = s \times s, \dim X_N^T Y_N = s \times 1, \dim \mathbf{a} = s \times 1$$

- kwestia istnienia rozwiązania

$$\text{lincol} X_N^T X_N = \text{lincol} X_N^T Y_N$$

- kwestia jednoznaczności rozwiązania

$$\begin{aligned}\det X_N^T X_N &> 0 \quad (\neq 0) \\ \text{inaczej } \text{rank} X_N^T X_N &= s \text{ (nieosobliwa)}\end{aligned}$$

Fakt: $\text{rank} X_N^T X_N = \text{rank} X_N^T = \text{rank} X_N$.

Warunek jednoznaczności: $\text{rank} X_N = s$

1) $N \geq s$ – dostatecznie dużo pomiarów (warunek konieczny, ale nie wystarczający)

2) wśród N pomiarów musi się znaleźć s niezależnych liniowo wektorów

Gdy $\text{rank} X_N < s$, wtedy rozwiązanie równania normalnego nie jest jednoznaczne.

- kwestia nazwy równania (normalność-prostopadłość)

$$X_N^T(Y_N - X_N \mathbf{a}) = 0$$

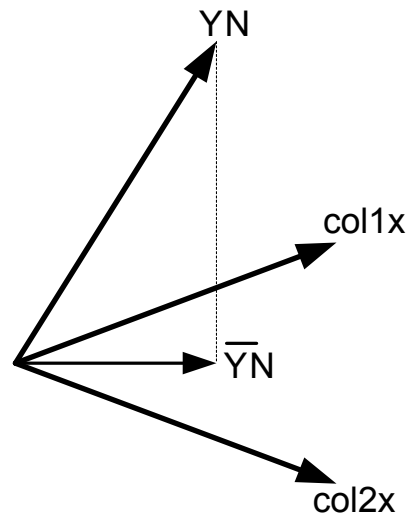
$$X_N^T(Y_N - \bar{Y}_N) = 0, \text{ gdzie } \bar{Y}_N - \text{wyjście modelu o parametrach } \mathbf{a}$$

$$X_N^T = [col_1x, col_2x, \dots, col_sx]$$

$$\bar{Y}_N = a_1 col_1x + a_2 col_2x + \dots + a_s col_sx - \text{liniowa kombinacja wektorów kolumnowych}$$

$$\text{wniosek : } \bar{Y}_N \in \text{lincol} X_N^T$$

$$\text{natomiast w ogólności : } Y_N \notin \text{lincol} X_N^T$$



$$X_N^T(Y_N - \bar{Y}_N) = 0 \Rightarrow (col_{xi})^T(Y_N - \bar{Y}_N) = 0 \text{ dla każdego } i = 1, 2, \dots, s$$

wniosek: wektor różnicy $Y_N - \bar{Y}_N$ jest prostopadły do każdej z kolumn col_{xi} (rzut prostopadły jest "najkrótszy")

4. Własności asymptotyczne estymatora NK – wejście losowe

$$\mathbf{a}_N = (X_N^T X_N)^{-1} X_N^T Y_N$$

$\{x_k^{(i)}\}, \{z_k\}$ – dwa niezależne od siebie ciągi losowe typu i.i.d., $\mathbf{E}z_k = 0$

$$y = \mathbf{x}^T \mathbf{a}^* + z$$

$$\mathbf{x}y = \mathbf{x}\mathbf{x}^T \mathbf{a}^* + \mathbf{x}z$$

$$\mathbf{E}\mathbf{x}y = (\mathbf{E}\mathbf{x}\mathbf{x}^T) \mathbf{a}^* + \mathbf{E}\mathbf{x}z$$

prawdziwe $\mathbf{a}^* = (\mathbf{E}\mathbf{x}\mathbf{x}^T)^{-1} \mathbf{E}\mathbf{x}y$, macierz $\mathbf{E}\mathbf{x}\mathbf{x}^T$ musi być nieosobliwa (wymaganie dot. rozkładów)

korzystając z alternatywnych postaci

$$X_N^T X_N = \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^T, \quad X_N^T Y_N = \sum_{k=1}^N \mathbf{x}_k y_k \quad (\text{wyprowadzić!})$$

estymator NK można zapisać następująco

$$\mathbf{a}_N = \left(\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \left(\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k y_k \right)$$

elementy $\{\mathbf{x}_k \mathbf{x}_k^T\}$ są ciągami typu i.i.d. (stac.)

$\mathbf{E}\mathbf{x}_k \mathbf{x}_k^T = \mathbf{E}\mathbf{x}\mathbf{x}^T$ – dla każdego $k = 1, 2, \dots, N$

wniosek

$$\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^T \xrightarrow{p.1} \mathbf{E}\mathbf{x}\mathbf{x}^T, \text{ gdy } N \rightarrow \infty$$

a ponieważ odwracanie macierzy jest operacją ciągłą, to

$$\left(\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \xrightarrow{p.1} (\mathbf{E}\mathbf{x}\mathbf{x}^T)^{-1}, \text{ gdy } N \rightarrow \infty \quad (1)$$

elementy $\{\mathbf{x}_k y_k\}$ są ciągami typu i.i.d. (stac.)

$\mathbf{E}\mathbf{x}_k y_k = \mathbf{E}\mathbf{x}y$ – dla każdego $k = 1, 2, \dots, N$

wniosek

$$\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k y_k \xrightarrow{p.1} \mathbf{E}\mathbf{x}y, \text{ gdy } N \rightarrow \infty \quad (2)$$

zatem na podstawie (1) i (2)

$$\mathbf{a}_N \xrightarrow{p.1} \mathbf{a}^*, \text{ gdy } N \rightarrow \infty$$

Przy losowym pobudzeniu i powyższych założeniach, \mathbf{a}_N jest estymatorem **mocno zgodnym** parametrów \mathbf{a}^*

5. Własności asymptotyczne estymatora NK – wejście deterministyczne

$$\begin{aligned} \mathbf{a}_N &= (X_N^T X_N)^{-1} X_N^T Y_N = (X_N^T X_N)^{-1} X_N^T (X_N \mathbf{a}^* + Z_N) = \mathbf{a}^* + \left(\frac{1}{N} X_N^T X_N \right)^{-1} \left(\frac{1}{N} X_N^T Z_N \right) = \\ &= \mathbf{a}^* + \left(\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^T \right)^{-1} \left(\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k z_k \right) \end{aligned}$$

- 1) czynnik $\left(\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^T \right)^{-1}$ jest czysto deterministyczny
- 2) $\{\mathbf{x}_k z_k\}$ – niezależne (nie muszą być typu i.i.d.)
- 3) $\mathbf{E} \mathbf{x}_k z_k = \mathbf{x}_k \mathbf{E} z_k = 0$, dla każdego k
- 4) dodatkowy warunek dotyczący deterministycznego wejścia (patrz II wersja MPWL)

$$\text{var} \mathbf{x}_k^{(i)} z_k = [\mathbf{x}_k^{(i)}]^2 \text{var} z_k = [\mathbf{x}_k^{(i)}]^2 \sigma_z^2 < \infty$$

Jeżeli dodatkowo

$$\sigma_z^2 \sum_{k=1}^N \frac{[\mathbf{x}_k^{(i)}]^2}{k^2} < \infty, \text{ dla każdego } i = 1, 2, \dots, s$$

to

$$\frac{1}{N} \sum_{k=1}^N \mathbf{x}_k z_k \xrightarrow{p.1} 0, \text{ gdy } N \rightarrow \infty$$

Przykład ograniczonego wejścia deterministycznego

$$\left| \mathbf{x}^{(i)} \right| \leq d_i$$

tj.

$\mathbf{x} \in [-d_1, d_1] \times [-d_2, d_2] \times \dots \times [-d_s, d_s]$ – wielowymiarowy przedział

6. Własności estymatora NK dla $N < \infty$

analiza obciążenie

oznaczając $L_N = (X_N^T X_N)^{-1} X_N^T$ otrzymujemy $\mathbf{a}_N = L_N Y_N$ (estymator **liniowy**)

$$\mathbf{a}_N = L_N Y_N = L_N (X_N \mathbf{a}^* + Z_N) = \mathbf{a}^* + L_N Z_N$$

$$\mathbf{E} \mathbf{a}_N = \mathbf{a}^* + \mathbf{E} \{L_N\} \mathbf{E} \{Z_N\} = \mathbf{a}^*$$

wniosek: \mathbf{a}_N jest nieobciążony zarówno dla wejścia losowego, jak i deterministycznego

analiza macierzy kowariancji

$$\text{cov}(\mathbf{a}_N) = \mathbf{E} \{(\mathbf{a}_N - \mathbf{a}^*)(\mathbf{a}_N - \mathbf{a}^*)^T\}$$

$$\mathbf{a}_N - \mathbf{a}^* = L_N Z_N$$

$$\text{cov}(\mathbf{a}_N) = \mathbf{E} \{L_N Z_N Z_N^T L_N^T\} = L_N (\mathbf{E} Z_N Z_N^T) L_N^T$$

struktura macierzy kowariancji zakłóceń $\mathbf{E} Z_N Z_N^T$ jest następująca

$$\mathbf{E} Z_N Z_N^T = \begin{bmatrix} \mathbf{E} z_1^2 & \mathbf{E} z_1 z_2 & \dots & \mathbf{E} z_1 z_N \\ \mathbf{E} z_2 z_1 & \mathbf{E} z_2^2 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{E} z_N z_1 & \dots & \dots & \mathbf{E} z_N^2 \end{bmatrix} = \sigma_z^2 I$$

zatem

$$\text{cov}(\mathbf{a}_N) = \sigma_z^2 L_N L_N^T$$

postać macierzy $L_N L_N^T$

$$L_N L_N^T = (X_N^T X_N)^{-1} X_N^T X_N (X_N^T X_N)^{-1} = (X_N^T X_N)^{-1}$$

wniosek

$$\text{cov}(\mathbf{a}_N) = \sigma_z^2 (X_N^T X_N)^{-1}$$

planowanie ortogonalne

gdy $X_N^T X_N$ jest macierzą diagonalną (eksperyment czynny – wejście deterministyczne)

6.1. Optimalność \mathbf{a}_N w klasie estymatorów LUE

klasa LUE

$$\{\alpha_N : \alpha_N = M_N Y_N, \quad \mathbf{E} \alpha_N = \mathbf{a}^*, \quad M_N - \text{deterministyczna}\}$$

wartość oczekiwana

$$\mathbf{E} \alpha_N = \mathbf{E} M_N Y_N = \mathbf{E} M_N X_N \mathbf{a}^* + \mathbf{E} M_N Z_N \quad \text{wniosek: } M_N X_N = I$$

macierz kowariancji

$$\begin{aligned} \text{cov}(\alpha_N) &= \mathbf{E} \{ M_N Z_N Z_N^T M_N^T \} = M_N (\mathbf{E} Z_N Z_N^T) M_N^T \\ \text{cov}(\alpha_N) &= \sigma_z^2 M_N M_N^T \end{aligned}$$

pokażemy, że

$$\text{cov}(\mathbf{a}_N) \leq \text{cov}(\alpha_N)$$

co oznacza znak \leq w odniesieniu do macierzy?

Definicja 1 Macierz $A \in R^{n,n}$ jest nieujemnie określona (zapisujemy $A \geq 0$), jeśli dla każdego wektora $w \in R^n$ zachodzi

$$w^T A w \geq 0$$

Twierdzenie 1 Macierz postaci AA^T jest zawsze nieujemnie określona.

podstawmy $A = M_N - L_N$

$$(M_N - L_N)(M_N - L_N)^T = (M_N - L_N)(M_N^T - L_N^T) = M_N M_N^T - M_N L_N^T - L_N M_N^T + L_N L_N^T = (***)$$

sposobrzeżenia

$$\begin{aligned} M_N L_N^T &= M_N X_N (X_N^T X_N)^{-1} = I (X_N^T X_N)^{-1} = (X_N^T X_N)^{-1} = L_N L_N^T \\ L_N M_N^T &= (M_N L_N^T)^T = L_N L_N^T \end{aligned}$$

zatem

$$(***) = M_N M_N^T - L_N L_N^T$$

zaś na podstawie Twierdzenia

$$\begin{aligned} M_N M_N^T - L_N L_N^T &\geq 0 & / \cdot \sigma_z^2 \\ \text{cov}(\alpha_N) - \text{cov}(\mathbf{a}_N) &\geq 0 \end{aligned}$$