



MI

Metody Identyfikacji

wykład #7

1. *Identyfikacja modeli parametrycznych cd.*
 - a) *Metody dwuetapowe*
 - b) *Metody rekurencyjne*
 - c) *Rozszerzenia metody najmniejszych kwadratów*

Identyfikacja dwuetapowa z pośrednim modelem nieparametrycznym

- Jeśli struktura jest nieznana (a tak jest najczęściej) to można najpierw znaleźć model nieparametryczny (nie ma potrzeby na wiedzę *a priori*) a potem dopiero znajdujemy strukturę i parametry.
- Dwa podejścia:
 1. Odpowiedź na pobudzenie nieokresowe a potem MNK
 2. Metoda korelacyjna a potem MNK



Odpowiedź na pobudzenie nieokresowe a potem MNK

- Kilka M odpowiedzi na to samo pobudzenie $u_j(k)$ a potem uśrednienie wyjść $y_j(k)$ w celu eliminacji wpływu addytywnego zakłócenia stochastycznego $n_j(k)$

$$\bar{y}(k) = \frac{1}{M} \sum_{j=1}^M y_j(k)$$

- Otrzymujemy wartość oczekiwaną

$$E\{ \bar{y}(k) \} = y_u(k) + E\{ \bar{n}(k) \}$$

- Jeśli wartość oczekiwana zakłócenia jest zerowa, to wtedy obie wartości oczekiwane są identyczne.
- Model ARX

$$y(k) = -a_1 y(k-1) - a_2 y(k-2) - \dots - a_m y(k-m) \\ + b_1 u(k-d-1) + b_2 u(k-d-2) + \dots + b_m u(k-d-m)$$

Odpowiedź na pobudzenie nieokresowe a potem MNK

- Przepisujemy model do postaci wektorowej dla $1 \leq k \leq l$

$$\begin{pmatrix} y(1) \\ y(2) \\ y(3) \\ \vdots \\ y(l) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 & u(-d) & \dots & 0 \\ -y(1) & 0 & \dots & 0 & u(1-d) & \dots & 0 \\ -y(2) & -y(1) & \dots & 0 & u(2-d) & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -y(l-1) & -y(l-2) & \dots & -y(l-m) & u(l-d-1) & \dots & u(l-d-m) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \\ b_1 \\ \vdots \\ b_m \end{pmatrix}$$

- Co zapisujemy jako $y = R\Theta$ i otrzymujemy:

$$e = y - R\Theta$$

$$V = e^T e$$

$$\hat{\Theta} = (R^T R)^{-1} R^T y$$

Odpowiedź na pobudzenie nieokresowe a potem MNK

- Estymata jest spójna, tzn. dla liczby eksperymentów $M \rightarrow \infty$ otrzymujemy:
 - Wartość oczekiwaną błędu równą zero
 - Wartość oczekiwaną parametrów równą rzeczywistym parametrom
 - Wartość oczekiwaną wariancji parametrów równą zero
- Zalety metody:
 - Unikanie zakłóceń i szumów poprzez uśrednianie
- Wady:
 - Konieczność wielokrotnego przeprowadzenia eksperymentu

Metoda korelacyjna a potem MNK

- Jeśli na wejściu podajemy sygnał pseudolosowy to funkcja autokorelacji ma postać

$$R_{uu}(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{k=0}^N u(k-\tau)u(k)$$

- A funkcja korelacji wzajemnej

$$R_{uy}(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{k=0}^N u(k-\tau)y(k)$$

- Funkcje korelacji mogą być wyznaczone rekurencyjnie

$$\hat{R}_{uy}(\tau, k) = \hat{R}_{uy}(\tau, k-1) + \frac{1}{k+1} (u(k-\tau)y(k) - \hat{R}_{uy}(\tau, k-1))$$

- Model ARX jak poprzednio

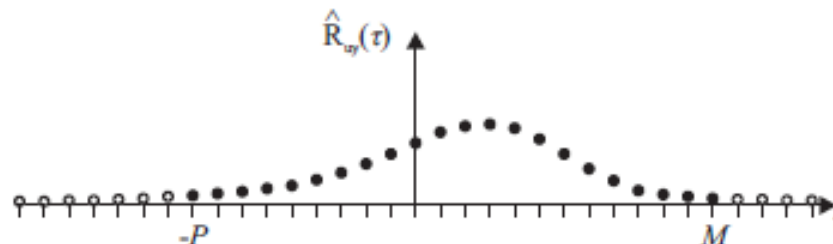
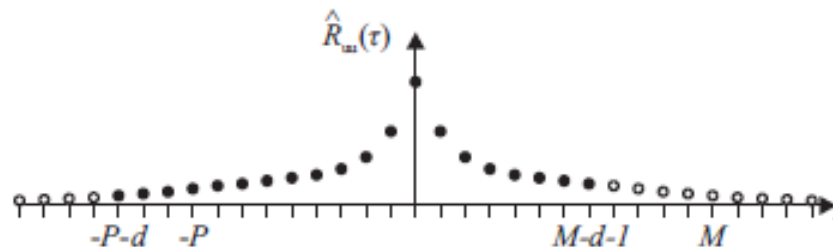
$$y(k) = -a_1 y(k-1) - a_2 y(k-2) - \dots - a_m y(k-m) \\ + b_1 u(k-d-1) + b_2 u(k-d-2) + \dots + b_m u(k-d-m)$$

Metoda korelacyjna a potem MNK

- Tym samym dla funkcji korelacji wzajemnej otrzymujemy

$$R_{uy}(\tau) = -a_1 R_{uy}(\tau - 1) - a_2 R_{uy}(\tau - 2) - \dots - a_m R_{uy}(\tau - m) \\ + b_1 R_{uu}(\tau - d - 1) + b_2 R_{uy}(\tau - d - 2) + \dots + b_m R_{uu}(\tau - d - m)$$

- Równanie to stanowi podstawę metody identyfikacji
- Przyjmujemy, że wartości funkcji korelacji wzajemnej wykorzystywane do estymaty parametrów modelu są różne od zera dla $-P \leq \tau \leq M$ i równe \approx zeru dla $\tau < -P$ i $\tau > M$



Metoda korelacyjna a potem MNK

- Ostatecznie otrzymujemy układ równań

$$\begin{pmatrix} R_{uy}(-P+m) \\ \vdots \\ R_{uy}(-1) \\ R_{uy}(0) \\ R_{uy}(1) \\ \vdots \\ R_{uy}(M) \end{pmatrix} = \begin{pmatrix} -R_{uy}(-P+m+1) & \dots & -R_{uy}(-P) & R_{uu}(-P+m-d-1) & \dots \\ \vdots & & \vdots & \vdots & \\ -R_{uy}(-2) & \dots & -R_{uy}(-1-m) & R_{uu}(-2-d) & \dots \\ -R_{uy}(-1) & \dots & -R_{uy}(-m) & R_{uu}(-d-1) & \dots \\ -R_{uy}(0) & \dots & -R_{uy}(1-m) & R_{uu}(-d) & \dots \\ \vdots & & \vdots & \vdots & \\ -R_{uy}(M-1) & \dots & -R_{uy}(M-m) & R_{uu}(M-d-1) & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_m \\ b_1 \\ \vdots \end{pmatrix}$$

- Co zapisujemy jako $R_{uy} = S\Theta$ a zastosowanie MNK prowadzi do estymaty parametrów

$$\hat{\Theta} = (S^T S)^{-1} S^T R_{uy}$$

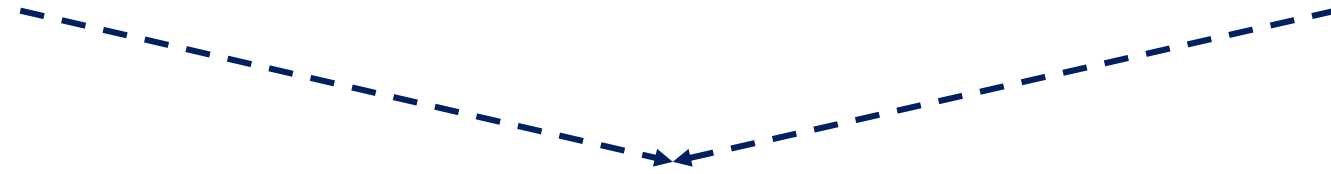
Metoda korelacyjna a potem MNK

- Wersja nierekurencyjna

1. Wczytanie wektorów wejść $u(k)$ i wyjść $y(k)$
2. Wyznaczenie autokorelacji wejścia $R_{uu}(\tau)$ oraz korelacji wzajemnej wejście-wyjście $R_{uy}(\tau)$
3. Wyznaczenie estymaty

- Wersja rekurencyjna

1. W k -tym kroku wyznaczamy autokorelację wejścia $R_{uu}(\tau, k)$ oraz korelację wzajemną wejście-wyjście $R_{uy}(\tau, k)$
2. Wyznaczamy estymatę parametrów w każdym kroku, lub też w dłuższych okresach.


$$\hat{\Theta} = (S^T S)^{-1} S^T R_{uy}$$

Rekurencyjna MNK (RLS – Recursive Least Squares)

- Pierwszy opis metody: *Gauss* (1809)
- Pierwsze zastosowanie: *Lee* (1964)
- Metoda nierekurencyjna (dla kroku k):

$$\hat{\theta}(k) = P(k)\Psi^T(k)y(k)$$

$$P(k) = (\Psi^T(k)\Psi(k))^{-1}$$

$$y(k) = \begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(k) \end{pmatrix}$$

$$\Psi(k) = \begin{pmatrix} \psi^T(1) \\ \psi^T(2) \\ \vdots \\ \psi^T(k) \end{pmatrix}$$

$$\psi^T = (-y(k-1) -y(k-2) \dots -y(k-m) | u(k-d-1) \dots u(k-d-m))$$

Rekurencyjna MNK (RLS – Recursive Least Squares)

- Dla kroku $k+1$ otrzymujemy:

$$\hat{\theta}(k+1) = P(k+1)\Psi^T(k+1)y(k+1)$$

- Równanie to można podzielić na:

$$\begin{aligned}\hat{\theta}(k+1) &= P(k+1) \begin{pmatrix} \Psi(k) \\ \psi^T(k+1) \end{pmatrix}^T \begin{pmatrix} y(k) \\ y(k+1) \end{pmatrix} \\ &= P(k+1)(\Psi(k)y(k) + \psi^T(k+1)y(k+1))\end{aligned}$$

- Podstawiając estymatę z kroku k w postaci $\Psi(k)y(k) = P^{-1}(k)\hat{\Theta}(k)$ otrzymamy

$$\hat{\theta}(k+1) = \hat{\theta}(k) + (P(k+1)P^{-1}(k) - I)\hat{\theta}(k) + P(k+1)\psi(k+1)y(k+1)$$

$$P^{-1}(k) = P^{-1}(k+1) - \psi(k+1)\psi^T(k+1)$$

$$\begin{aligned}P(k+1) &= \left(\begin{pmatrix} \Psi(k) \\ \psi^T(k+1) \end{pmatrix}^T \begin{pmatrix} \Psi(k) \\ \psi^T(k+1) \end{pmatrix} \right)^{-1} \\ &= (P^{-1}(k) + \psi(k+1)\psi^T(k+1))^{-1}\end{aligned}$$

Rekurencyjna MNK

- Otrzymujemy:

$$\hat{\Theta}(k+1) = \hat{\Theta}(k) + P(k+1)\psi(k+1)$$

Nowa estymata parametru Stara estymata parametru Wektor korekcji

$$\left(y(k+1) \quad \psi^T(k+1)\hat{\Theta}(k) \right)$$

Nowy pomiar Predykcja pomiaru na podstawie ostatniej dostępnej estymaty $\hat{y}(k+1|k)$

- A ostatecznie

$$\hat{\Theta}(k+1) = \hat{\Theta}(k) + P(k+1)\psi(k+1)e(k+1)$$

Rekurencyjna MNK

- Można zrezygnować z odwracania macierzy $P(k+1)$:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma(k) = P(k+1)\psi(k+1) = \frac{1}{\psi^T(k+1)P(k)\psi(k+1) + 1}P(k)\psi(k+1)$$

- Macierz $P(k+1)$ jest przeskalowaną estymatą macierzy kowariancji błędu estymacji

$$P(k+1) = (I - \gamma(k)\psi^T(k+1))P(k)$$

$$E\{P(k+1)\} = \frac{1}{\sigma_e^2} \text{cov } \Delta\theta(k+1)$$

Rekurencyjna MNK

- Inicjalizacja metody:

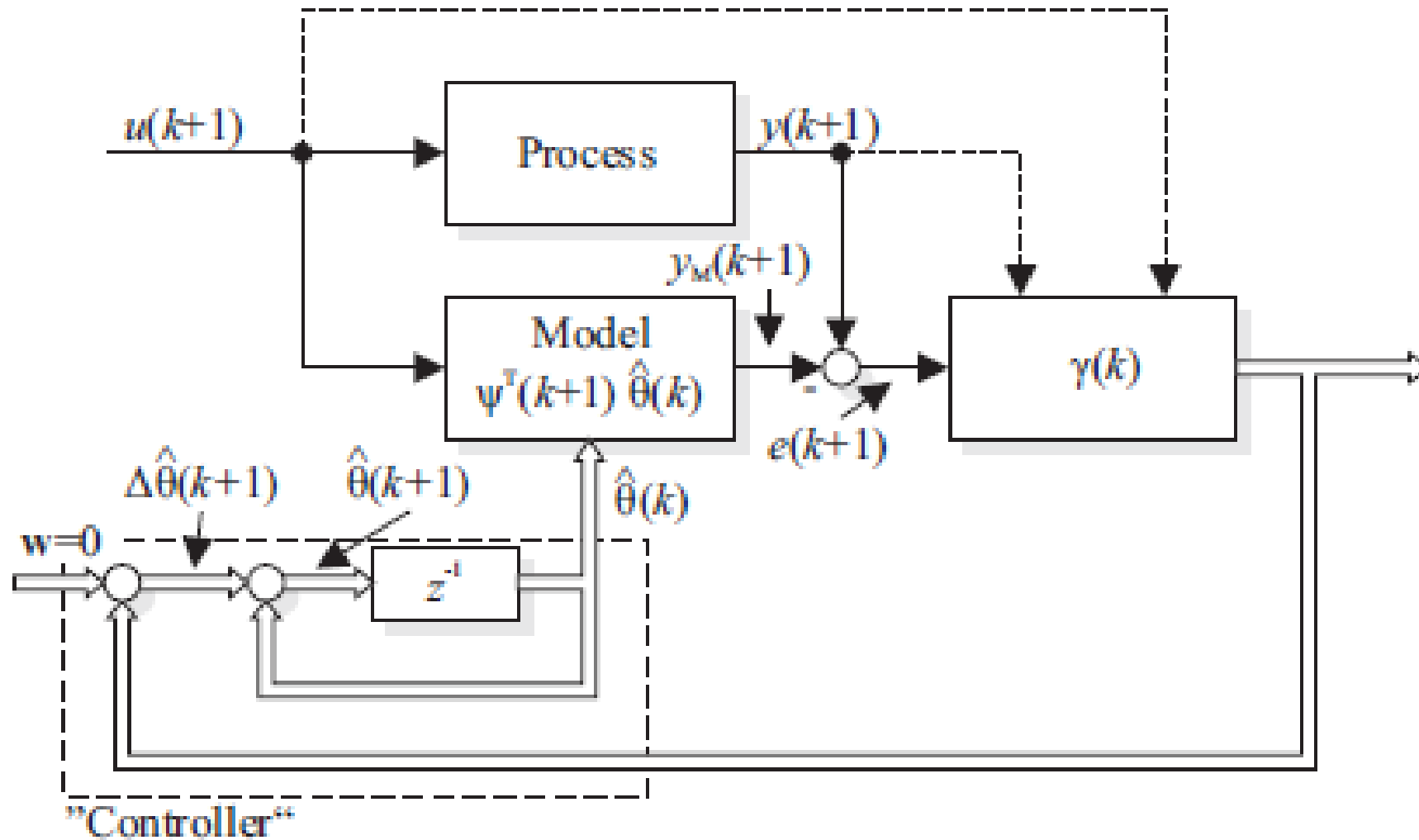
- Rozpocząć od metody nierekurencyjnej dla przynajmniej $2m$ rekordów i potem wystartować rekurencję
- Wykorzystać estymatę a priori parametrów, ich kowariancji oraz kowariancji błędu
- wystartowanie z warunkiem początkowym $P(0) = P_0$, $\theta(0) = \theta_0$, gdzie P_0 jest macierzą dodatnio określoną i dobrana tak aby

$$P^{-1}(k) = P^{-1}(0) + \Psi^T(k)\Psi(k)$$

- np. $P_0 = \alpha I$, gdzie α jest dużą liczbą.

$$\alpha \gg \frac{1}{u^2(0)}$$

Schemat blokowy rekurencyjnej MNK



Metoda ważona MNK (WLS – Weighted Least Squares)

- Jak do tej pory każdy element $e(k)$ miał takie same znaczenie
- Ogólnie można wprowadzić wagi różne dla poszczególnych chwil czasowych
- Wprowadzamy funkcję celu:

$$V = w(m+d)e^2(m+d) + w(m+d+1)e^2(m+d+1) + \dots \\ + w(m+d+N)e^2(m+d+N)$$

- A w ogólnej postaci

$$V = e^T W e$$

- Gdzie W jest symetryczną dodatnio określoną macierzą
 - Tylko część symetryczna ma wpływ
 - Dodatnio określona dla istnienia jednoznacznego wyniku

$$W = \begin{pmatrix} w(m+d) & 0 & \dots & 0 \\ 0 & w(m+d+1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w(m+d+N) \end{pmatrix}$$

Metoda ważona MNK

- W prosty sposób otrzymujemy

$$\hat{\Theta} = (\Psi^T W \Psi)^{-1} \Psi^T W y$$

- Analogicznie otrzymujemy zależności dla ważonej metody najmniejszych kwadratów

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma(k) = \frac{1}{\psi^T(k+1)P_W(k)\psi(k+1) + \frac{1}{w(k+1)}} P_W(k)\psi(k+1)$$

$$P_W(k+1) = (I - \gamma(k)\psi^T(k+1))P_W(k).$$

- Porównując z metoda podstawową zmienia się parametr korekcyjny $\gamma(k)$ a tym samym $P_w(k+1)$

Metoda rekurencyjna MNK z zapominaniem eksponentialnym

- Wprowadzenie parametrów zapominania zależnych od czasu, np.

$$w(k) = \lambda^{(m+d+N)-k} = \lambda^{N'-k}, \quad 0 < \lambda < 1$$

- Gdzie λ jest nazywana współczynnikiem zapominania.
- Poniżej pokazana są przykładowe parametry dla $N=50$.

k	1	10	20	30	40	47	48	49	50
$\lambda = 0.99$	0.61	0.67	0.73	0.82	0.90	0.97	0.98	0.99	1
$\lambda = 0.95$	0.08	0.13	0.21	0.35	0.60	0.85	0.90	0.95	1

$$W(m+d+n) = \begin{pmatrix} \lambda^N & & & & \\ & \lambda^{N-1} & & & \\ & & \ddots & & \\ & & & \lambda^2 & \\ & & & & \lambda \\ & & & & & 1 \end{pmatrix}$$

Metoda rekurencyjna MNK z zapominaniem eksponentyjalnym

- Po nadejściu kolejnego pomiaru macierz jest uaktualniana

$$W(k+1) = \begin{pmatrix} \lambda W(k) & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

- Ostatecznie metoda sprowadzona jest do postaci:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma_w(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma_w(k) = \frac{1}{\psi^T(k+1)P_w(k)\psi(k+1) + \lambda} P_w(k)\psi(k+1)$$

$$P_w(k+1) = (I - \gamma_w(k)\psi^T(k+1))P_w(k)\frac{1}{\lambda}$$

- Dobór współczynnika zapominania jest kompromisem pomiędzy tłumieniem zakłóceń $\gamma \rightarrow 1$ a lepszym śledzeniem procesów niestacjonarnym $\gamma < 1$
- W praktyce stosuje się wielkości około $0.9 < \gamma < 0.995$

Ogólniejsze metody

- Model Box-Jenkins umożliwia bardziej szerokie uwzględnianie zakłóceń

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} u(k) + \frac{D(z^{-1})}{C(z^{-1})} v(k)$$

$$n(k) = \frac{D(z^{-1})}{C(z^{-1})} v(k)$$

- Metoda działa jedynie wtedy, gdy błąd estymaty $e(k)$ jest nieskorelowany. Dzieje się tak jedynie wtedy, gdy zakłócenie $n(k)$ jest generowane poprzez koloryzowanie białego szumu $v(k)$ – rzadko spełnione.
- Poniżej będą przedstawione metody posiadające szerszy zakres stosowalności

Ogólna MNK (GLS – Generalized Least Squares)

- Zakładamy, że nieskorelowany sygnał błędu

$$A(z^{-1})y(z) - B(z^{-1})z^{-d}u(z) = e(z)$$

- Jest zastąpiony przez sygnał skorelowany, szum kolorowy $\xi(k)$, generowany jak poniżej:

$$\xi(z) = \frac{1}{F(z^{-1})} e'(z), \text{ gdzie } e'(z) \text{ jest nieskorelowany.}$$

- Wielomian $F(z^{-1})$ jest nieznany, zatem zaproponowany jest algorytm iteracyjny.



Ogólna MNK (GLS – Generalized Least Squares)

Krok 1

- Stosujemy metodę MNK do pomiarów na przedziale

$$m + d \leq k \leq m + d + N$$

- dla modelu

$$A(z^{-1})y(z) - B(z^{-1})z^{-d}u(z) = \xi(z)$$

- Estymaty $\hat{\theta}_1$ są obciążone, a $\xi(k)$ skorelowany.

Ogólna MNK (GLS – Generalized Least Squares)

Krok 2

- Sygnał $\xi(k)$ jest wyznaczany dla parametrów $\hat{\theta}_1$. Wykorzystanie modelu AR

$$\xi(k) = \psi_{\xi}^T(k)f + e'(k)$$

- Prowadzi do

$$\psi_{\xi}^T(k) = [-\xi(k-1), -\xi(k-2), \dots, -\xi(k-v)]$$

$$f^T = [f_1, f_2, \dots, f_v]$$

- Rząd wielomianu powinien być wyznaczony dokładnie, tj. $v = m$.
- Parametry otrzymujemy z metody najmniejszych kwadratów:

$$\hat{f} = [\Xi^T \Xi]^{-1} \Xi^T \xi,$$

gdzie Ξ jest utworzony z wierszy $\psi_{\xi}^T(k)$.

Ogólna MNK (GLS – Generalized Least Squares)

Krok 3

- Wejście $u(k)$ i wyjście $y(k)$ filtrujemy poprzez

$$G_F(z^{-1}) = \hat{F}(z^{-1})$$

- zatem

$$\tilde{u}(z) = G_F(z^{-1})u(z) \text{ oraz } \tilde{y}(z) = G_F(z^{-1})y(z)$$



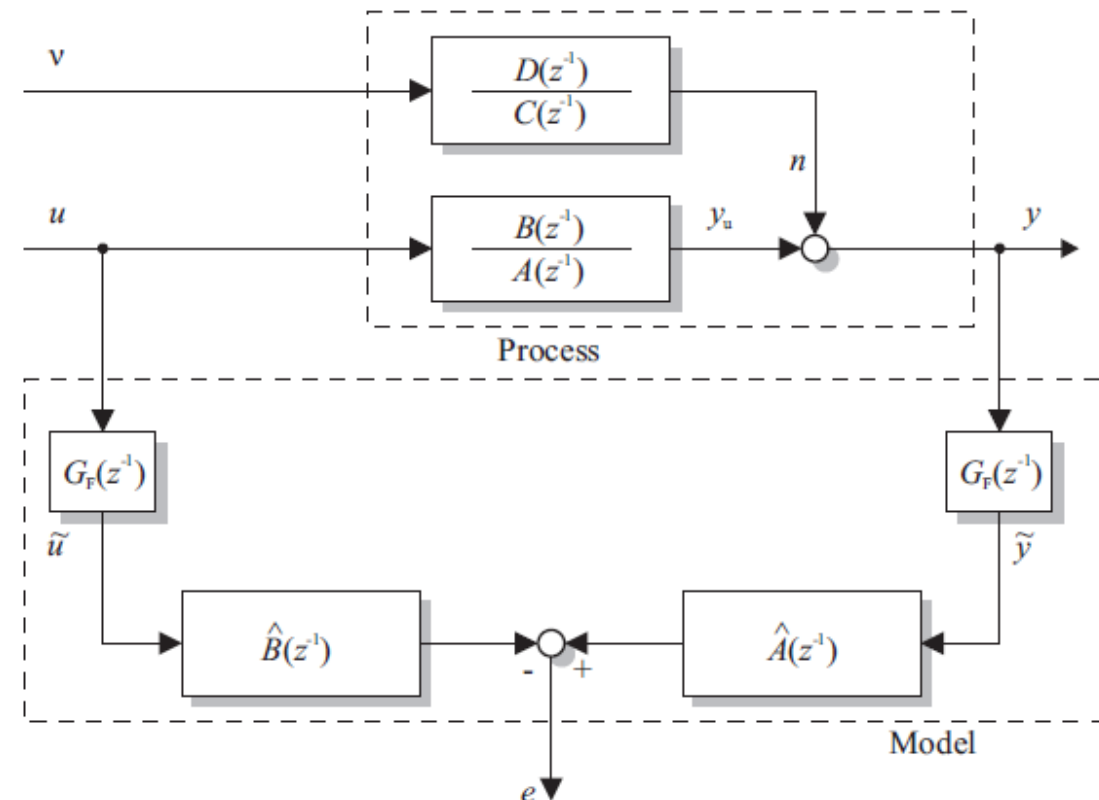
Ogólna MNK (GLS – Generalized Least Squares)

Krok 4

- Metoda najmniejszych kwadratów jest zastosowana do sygnałów filtrowanych

$$A(z^{-1})\tilde{y}(z) - B(z^{-1})z^{-d}\tilde{u}(z) = \xi(z)$$

- i otrzymujemy wektor parametrów $\hat{\theta}_2$.



Ogólna MNK (GLS – Generalized Least Squares)

Krok 5

- Kroki od 2 do 4 powtarzamy aż $\hat{\theta}_j$ nie zmieniają się w sposób znaczący pomiędzy iteracjami.
- Metoda ogólna MNK prowadzi do nieobciążonej estymaty szumu generowany poprzez:

$$G_v(z) = \frac{n(z)}{v(z)} = \frac{D(z^{-1})}{C(z^{-1})} = \frac{1}{A(z^{-1})F(z^{-1})}$$

Ogólna Rekurencyjna MNK (RGLS – Recursive Method of Generalized Least Squares)

- Z metody ogólnej MNK można wyprowadzić wersję rekursywną

$$\hat{\theta}(k+1) = \hat{\theta}(k) + (\psi^T(k+1)\tilde{P}(k)\tilde{\psi}(k+1) + 1)^{-1} \\ \tilde{P}(k)\tilde{\psi}(k+1)(\tilde{y}(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\tilde{P}(k+1) = \tilde{P}(k) \left(I - \tilde{\psi}^T(k+1)\tilde{\psi}(k+1)\tilde{P}(k) \right. \\ \left. (\tilde{\psi}^T(k+1)\tilde{P}(k)\tilde{\psi}(k+1) + 1)^{-1} \right)$$

$$\tilde{f}(k+1) = \tilde{f}(k) + (\psi_{\xi}^T(k+1)Q(k)\psi_{\xi}(k+1) + 1)^{-1} \\ Q(k)\psi_{\xi}(k+1)(\xi(k+1) - \psi_{\xi}^T(k+1)\tilde{f}(k))$$

$$Q(k+1) = Q(k) \left(I - \psi_{\xi}^T(k+1)\psi_{\xi}(k+1)Q(k) \right. \\ \left. (\psi_{\xi}^T(k+1)Q(k)\psi_{\xi}(k+1) + 1)^{-1} \right) .$$

Rozszerzona MNK (ELS – Extended Least Squares)

- Zamiast MNK dla modelu ze skorelowanym sygnałem $\varepsilon(z)$

$$A(z^{-1})\tilde{y}(z) - B(z^{-1})z^{-d}\tilde{u}(z) = \varepsilon(z)$$

- Stosujemy model ARMAX ze skorelowanym sygnałem $\varepsilon(z) = D(z^{-1})e'(z^{-1})$.
- W takim przypadku możemy połączyć rekurencyjne metody dla obiektu dynamicznego oraz procesu stochastycznego.
- Dla modelu

$$y(k) = \psi^T(k)\hat{\theta}(k-1) + e(k)$$

Wprowadzamy rozszerzony wektor

$$\begin{aligned} \psi^T(k) &= (-y(k-1) \dots -y(k-m) | u(k-d-1) \dots \\ &\quad u(k-d-m) | \hat{v}(k-1) \dots \hat{v}(k-p)) \\ \hat{\theta}^T &= (\hat{a}_1 \dots \hat{a}_m | \hat{b}_1 \dots \hat{b}_m | \hat{d}_1 \dots \hat{d}_p) . \end{aligned}$$

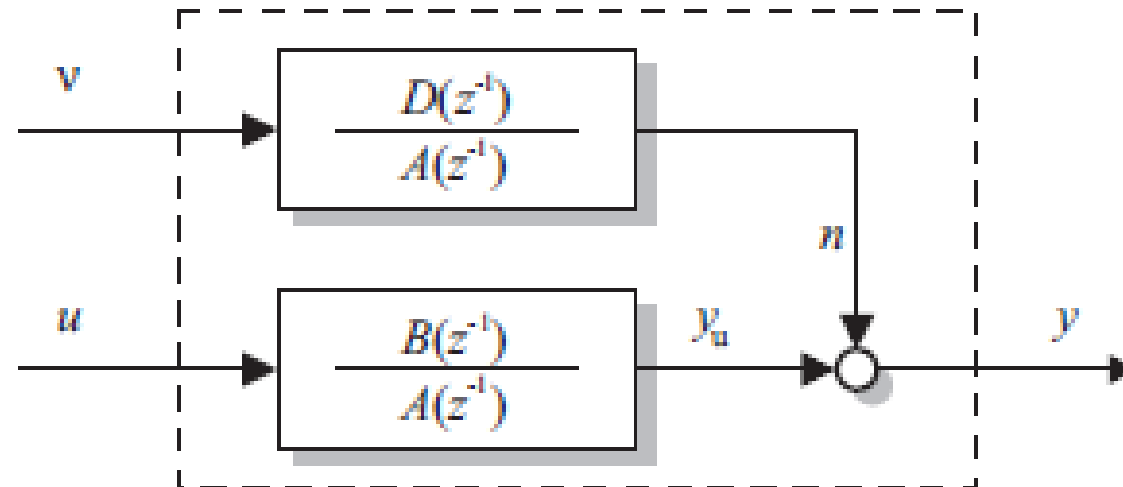
Rozszerzona MNK (ELS – Extended Least Squares)

- I sprowadzamy rozwiązanie do metody rekurencyjnej

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma(k) = P(k+1)\psi(k+1) = \frac{1}{\psi^T(k+1)P(k)\psi(k+1) + 1}P(k)\psi(k+1)$$

$$P(k+1) = (I - \gamma(k)\psi^T(k+1))P(k)$$



Method of Bias Correction (CLS)

- Dotychczasowe metody starały się uniknąć obciążenia metody zakładając odpowiednie założenia na tworzenie sygnału zakłócającego.
- Alternatywne podejście próbuje wyznaczyć obciążenie metody (*bias*) a potem jego wykorzystanie do korekcji.
- Niemniej założenia też są bardzo ostre, gdyż sygnał $v(t)$ musi być białym szumem.

Method of Bias Correction (CLS)

- Zakładamy model w postaci

$$y(z) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} u(z) + n(z)$$

gdzie $n(z)$ jest białym szumem.

- Bias opisany jest przez

$$E\{b(N+1)\} = -E\{R^{-1}(N+1)\} \underbrace{\left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right)}_S \theta_0 \sigma_n^2$$

- Jest on wykorzystany do korekcji estymaty wektora parametrów

$$\begin{aligned} \hat{\theta}_{\text{CLS}}(N+1) &= \hat{\theta}_{\text{LS}}(N+1) - b(N+1) \\ &= \hat{R}^{-1}(N+1) \frac{1}{N+1} \Psi^T(N+1) y(N+1) \\ &\quad + \hat{R}^{-1}(N+1) S \hat{\theta}_{\text{CLS}}(N+1) \sigma_n^2 \end{aligned}$$

Method of Bias Correction (CLS)

- Otrzymujemy postać algorytmu

$$\hat{\theta}_{\text{CLS}}(N + 1) = (R(N + 1) - S\sigma_n^2)^{-1} \frac{1}{N + 1} \psi^T(N + 1) y(N + 1)$$

- Gdzie wariancja jest opisana poprzez i może być liczona korelacyjnie

$$\sigma_n^2(N + 1) = E\{n^2(k)\} = \frac{1}{N + 1 - 2m} n^T(N + 1) n(N + 1)$$

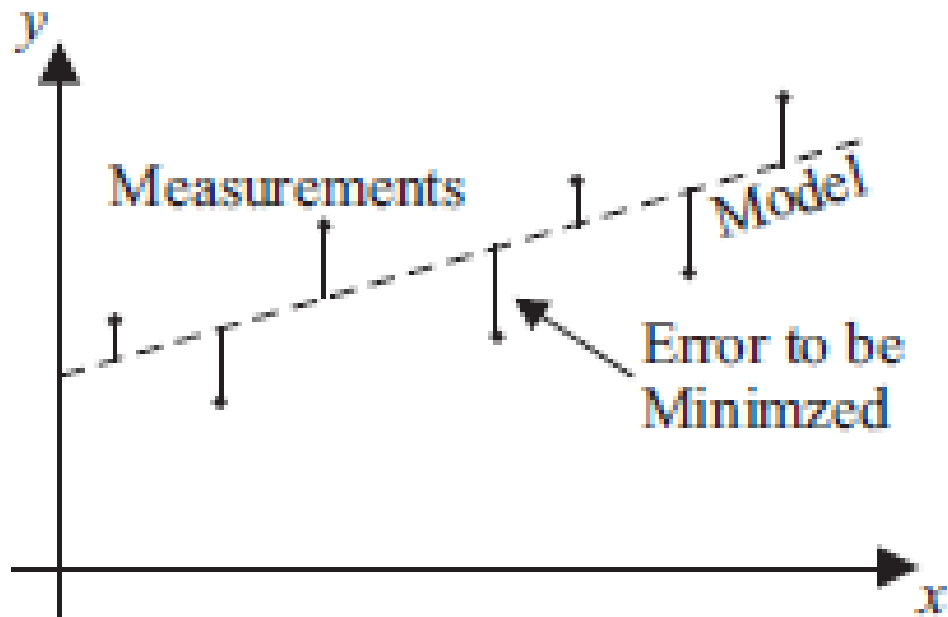


Total Least Squares (TLS)

- Klasyczny model zakłada, że tylko wyjście jest zakłócone, tzn.

$$y - e = \Psi \hat{\Theta}$$

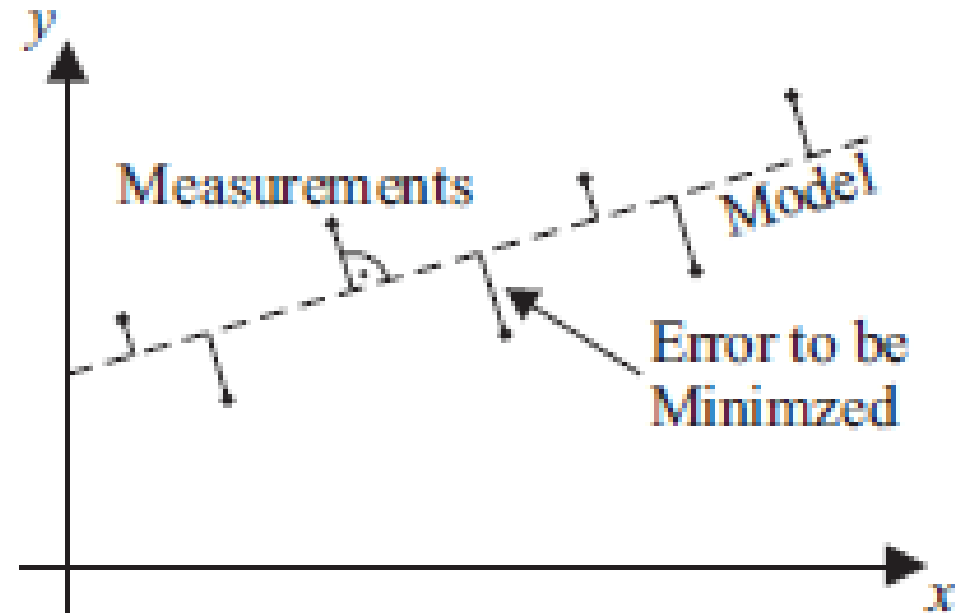
$$\hat{\Theta} = \arg \min \|e\|_2^2$$



- Obecnie zakładamy model w postaci

$$y - e = (\Psi + F) \hat{\Theta}$$

- Czyli błąd dotyczy **nie tylko wyjścia** ale również pomiarów w wektorze regresji Ψ



Total Least Squares (TLS)

- Model możemy przepisać do postaci

$$\underbrace{\begin{pmatrix} \Psi & y \end{pmatrix}}_C + \underbrace{\begin{pmatrix} F & e \end{pmatrix}}_A \begin{pmatrix} \hat{\theta} \\ -1 \end{pmatrix} = 0$$

- $A = (F, e)$ rozszerzona macierz błędów
- Rozmiar macierzy C będzie wynosił $N \times (m + 1)$
- Norma funkcji minimalizującej (*norma Frobeniusa*)

$$\|A\|_F^2 = \sum_{i=1}^M \sum_{j=1}^N A_{ij}^2$$

$$\hat{\theta} = \arg \min \|A\|_F^2$$

Total Least Squares (TLS)

- Macierz C może być wyznaczona jako

$$C = U\Sigma V^T$$

gdzie $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{n+1})$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n+1}$

co jest rozkładem SVD macierzy C

- Macierz \tilde{C} można zapisać jako

$$\tilde{C} = U\Sigma V^T \Leftrightarrow U^{-1}\tilde{C}(V^T)^{-1} = U^T\tilde{C}V = S$$

- Tym samym funkcja celu będzie w postaci

$$\|C - \tilde{C}\|_F^2 = \|U\Sigma V^T - USV^T\|_F^2 = \|\Sigma - S\|_F^2$$

- Macierz S jest diagonalna $S = \text{diag}(s_1, s_2, \dots, s_{n+1})$ i funkcja celu ma postać:

$$V = \|\Sigma - S\|_F^2 = \sum_{i=1}^{m+1} (\sigma_i - s_i)^2 \qquad V = \|\Sigma - S\|_F^2 = \sigma_{n+1}^2$$

Total Least Squares (TLS)

- Ponieważ

$$V := \begin{matrix} & \begin{matrix} n & d \end{matrix} \\ \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} & \begin{matrix} n \\ d \end{matrix} \end{matrix}$$

- Otrzymujemy rozwiązanie

$$\hat{\theta} = -V_{22}^{-1}V_{12}$$

- Metoda TLS dobrze pasuje do identyfikacji obiektów dynamicznych
- TLS jest blisko związana z PCA (Principal Component Analysis) wykorzystywanej w statystyce do wyznaczania korelacji danych i do zmniejszania rozmiaru tych danych.

Metoda Zmiennej Instrumentalnej (IV – Instrumental Variable)

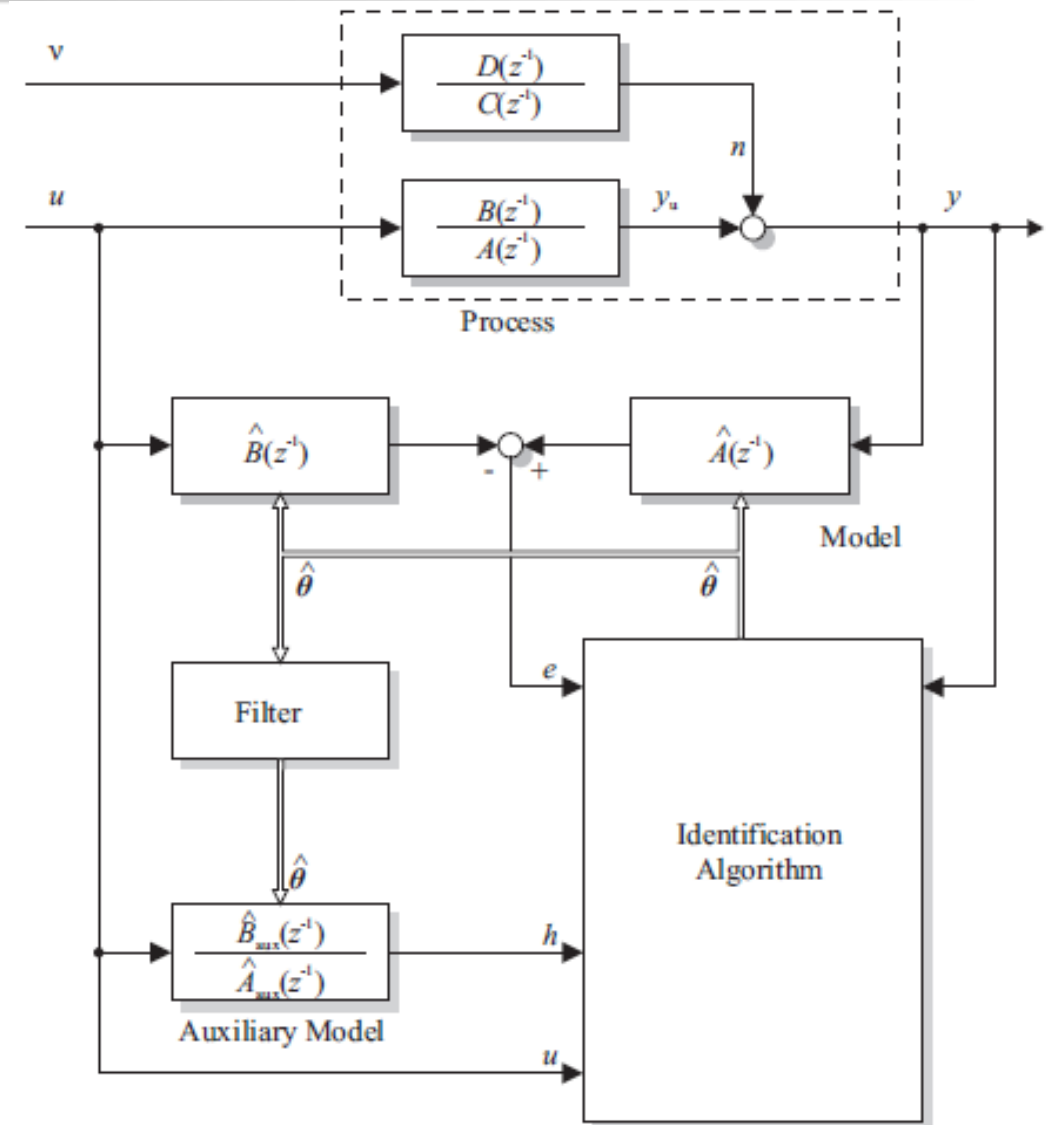
- Metoda bezpośrednia w celu unikania obciążenia estymaty
- Model w postaci $e = y - \Psi\theta$ mnożymy obustronnie przez macierz W – macierz zmiennych instrumentalnych

$$W^T e = W^T y - W^T \Psi \theta$$

- I otrzymujemy rozwiązanie w postaci

$$\hat{\theta} = (W^T \Psi)^{-1} W^T y$$

- Problem jak znaleźć zmienną instrumentalną:
 - Nieskorelowana z $n(k)$
 - Skorelowana z $u(k)$ i $y_u(k)$



Metoda Zmiennej Instrumentalnej (IV – Instrumental Variable)

- Różne metody wyboru zmiennej instrumentalnej:

- Sygnał wejściowy

$$w^T = (u(k-1-\delta) \dots u(k-m-\delta) | u(k-d-1) \dots u(k-d-m))$$

- Sygnał w oparciu o estymatę niezakłóconego wyjścia $h(k) = \hat{y}_u(k)$

$$w^T = (-h(k-1) \dots -h(k-m) | u(k-d-1) \dots u(k-d-m))$$

$$h(k) = \hat{y}_u(k) = \psi^T(k) \hat{\theta}(k)$$

- Podejście iteracyjne

1. W pierwszej iteracji stosujemy zmienną instrumentalną w postaci sygnału wejściowego
2. W kolejnej iteracji na podstawie pierwszej iteracji wyniku oraz zmiennej instrumentalnej wygenerowanej w oparciu o estymatę niezakłóconego wyjścia wyliczamy kolejną estymatę parametrów
3. Powtarzamy krok 2 aż parametry już się nie poprawiają

Metody Aproksymacji Stochastycznej (STA – Stochastic Approximation)

- Metody rekurencyjne
- Mniej wymagające obliczeniowo niż RLS
- Wykorzystujemy gradientowe metody optymalizacji do znalezienia wektora parametrów
 - Robbins-Monro
 - Kiefer-Wolfowitz
- Rzadko stosowane w praktyce, jako że trudno zagwarantować zbieżność metody



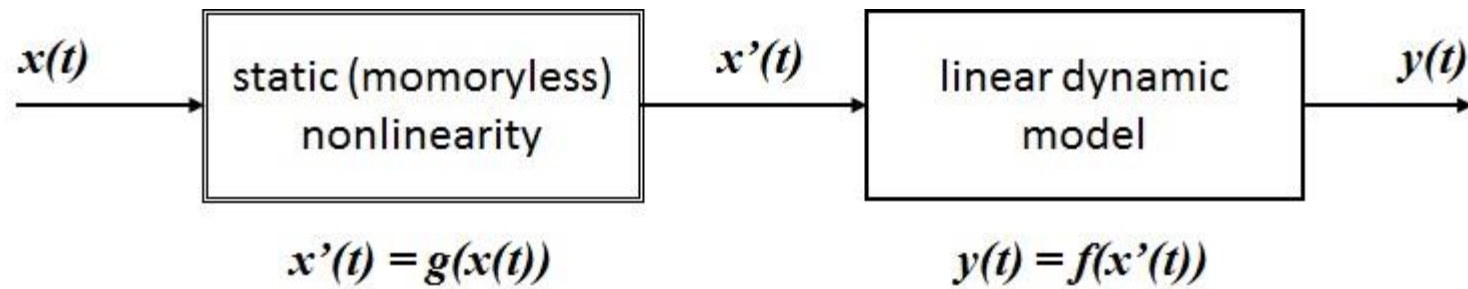
Modele nieliniowe

- Model Hammersteina
- Model Wienera
- Szeregi nieliniowe różnych funkcji bazowych:
 - Volter
 - Laguerre
- Modele NARMAX

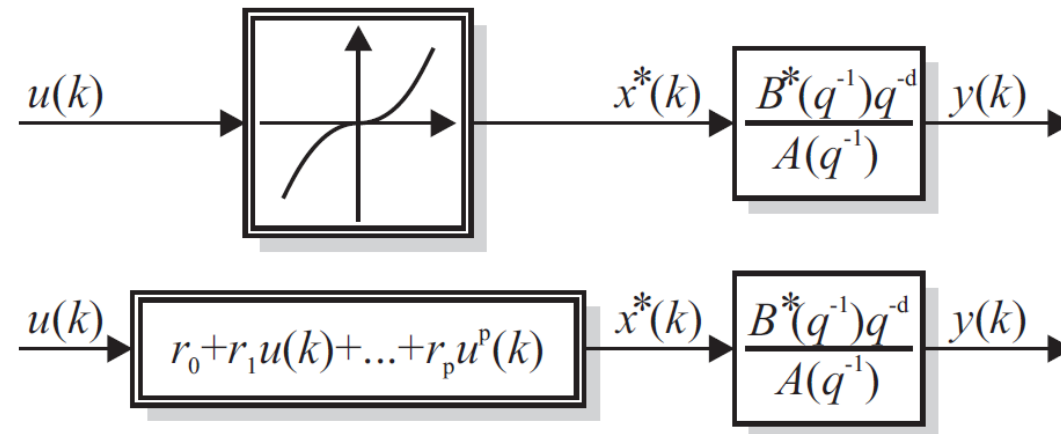


Model Hammersteina

Wersja ogólna



Wersja szczegółowa - podstawowa



Model Hammersteina podstawowy

- Statyczna nieliniowość w postaci wielomianu

$$x^*(k) = r_0 + r_1 u(k) + r_2 u^2(k) + \dots + r_p u^p(k)$$

- Dynamika liniowa – model ARX

$$A(q^{-1})y(k) = B^*(q^{-1})q^{-d}x^*(k)$$

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_m q^{-m}$$

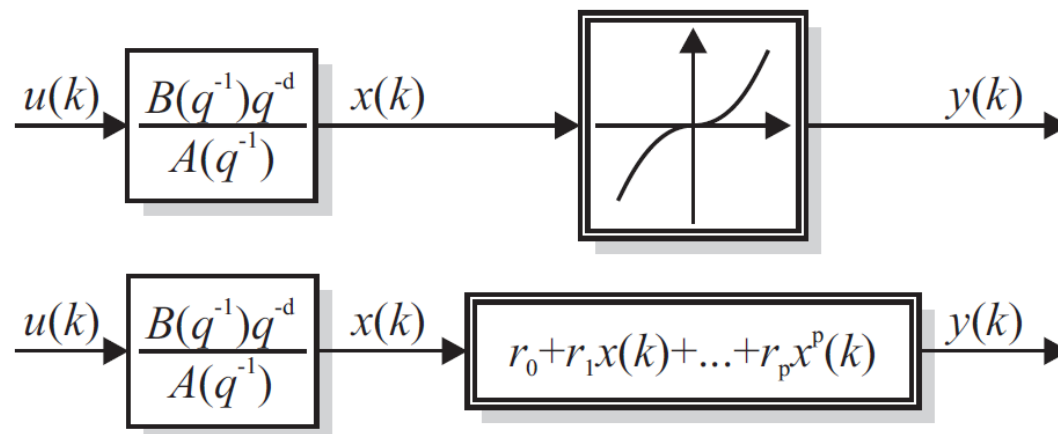
$$B^*(q^{-1}) = b_1^* q^{-1} + \dots + b_m^* q^{-m} ,$$

Model Hammersteina podstawowy

- Uczenie – proste rozszerzenie metody najmniejszych kwadratów
 - Rozszerzony wektor regresji

$$\psi^T(k) = (-y(k-1), \dots, u(k-1), \dots, u^2(k-1), \dots, u^3(k-1), \dots)$$

Model Wienera



Model Wienera

- Ogólna postać:

$$A_1(q^{-1})y(k) + A_2(q^{-1})y^2(k) + \dots + A_l(q^{-1})y^l(k) = c_{00} + B(q^{-1})u(k-d)$$

- Transmitancja liniowa dynamiczna

$$A(q^{-1})x(k) = B(q^{-1})q^{-d}u(k)$$

- Statyczna nieliniowość wielomianowa

$$y(k) = r_0 + r_1x(k) + r_2x^2(k) + \dots + r_lx^l(k)$$

- Otrzymujemy wersję podstawową

$$y(k) = r_0 + r_1 \frac{B(q^{-1})q^{-d}}{A(q^{-1})}u(k) + r_2 \left(\frac{B(q^{-1})q^{-d}}{A(q^{-1})} \right)^2 u^2(k) + \dots$$



Modele Hammersteina-Wienera

- Optymalizacja nieliniowa
- Uczenie iteracyjne

