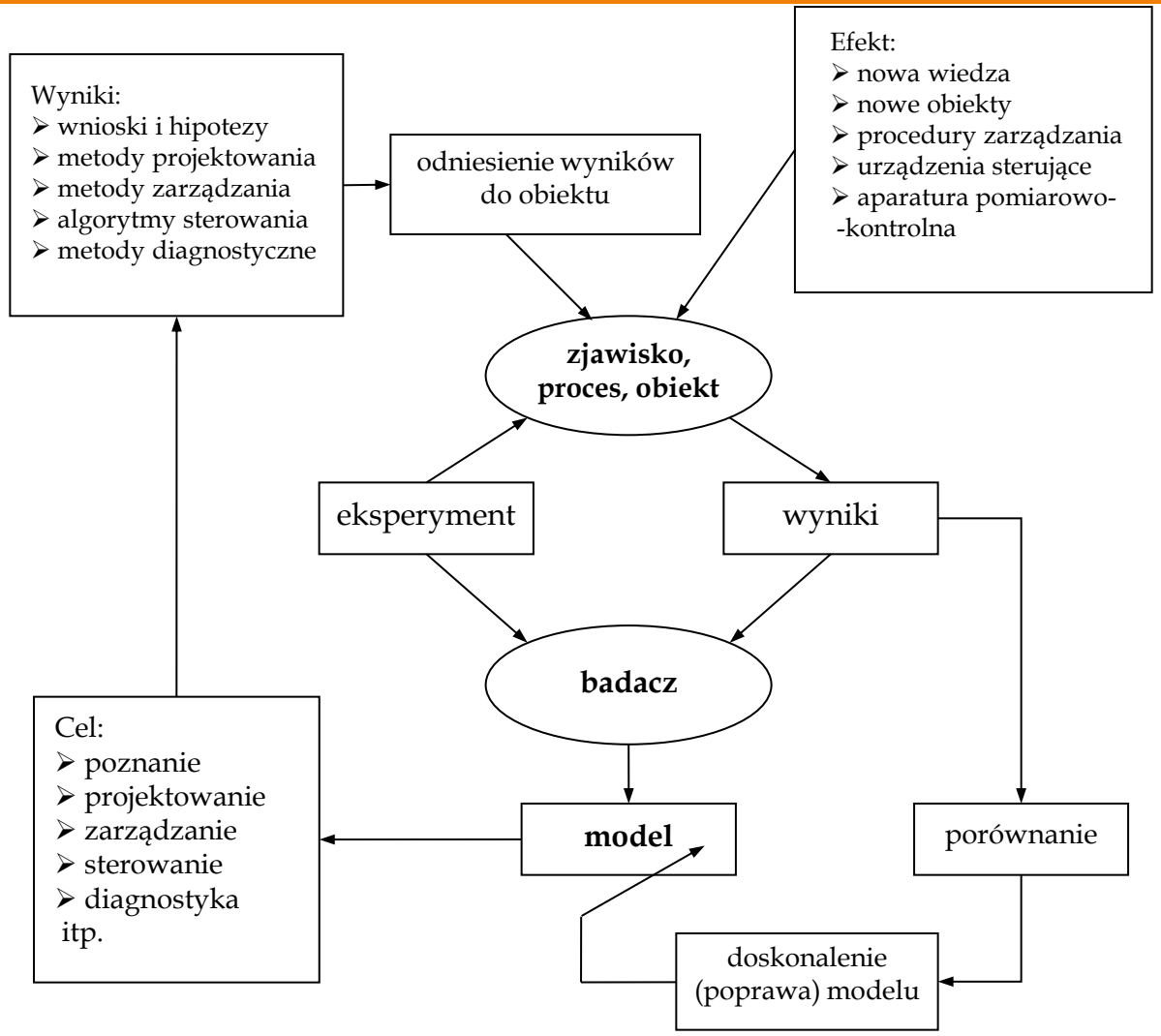
Optymalizacja systemów

Wykład 14. Uczenie maszynowe. Podstawowe zadania identyfikacji.

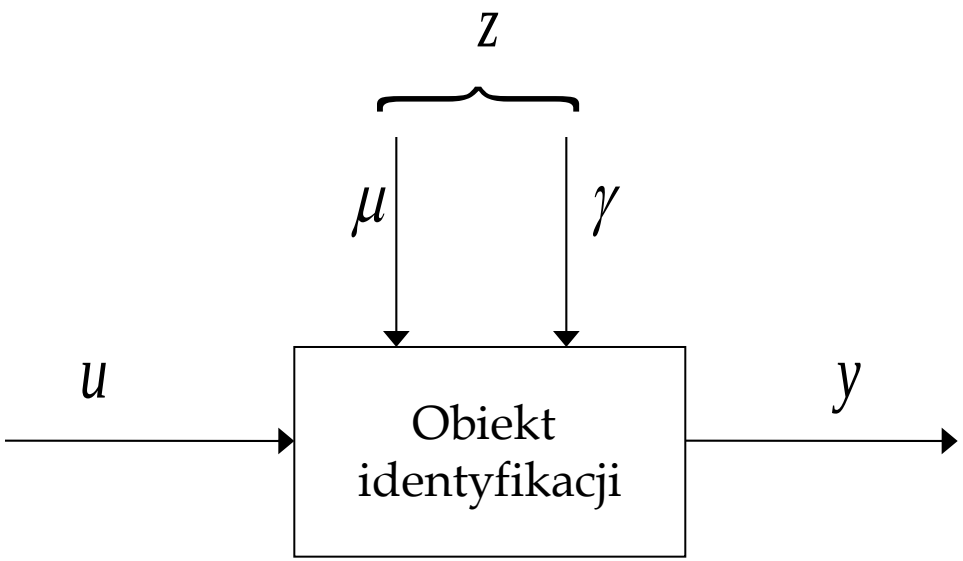
Model w badaniach systemowych



Zadanie identyfikacji

1. Określenie obiektu identyfikacji
2. Określenie klasy modeli
3. Organizacja eksperymentu
4. Wyznaczenie algorytmu identyfikacji
5. Wykonanie identyfikatora

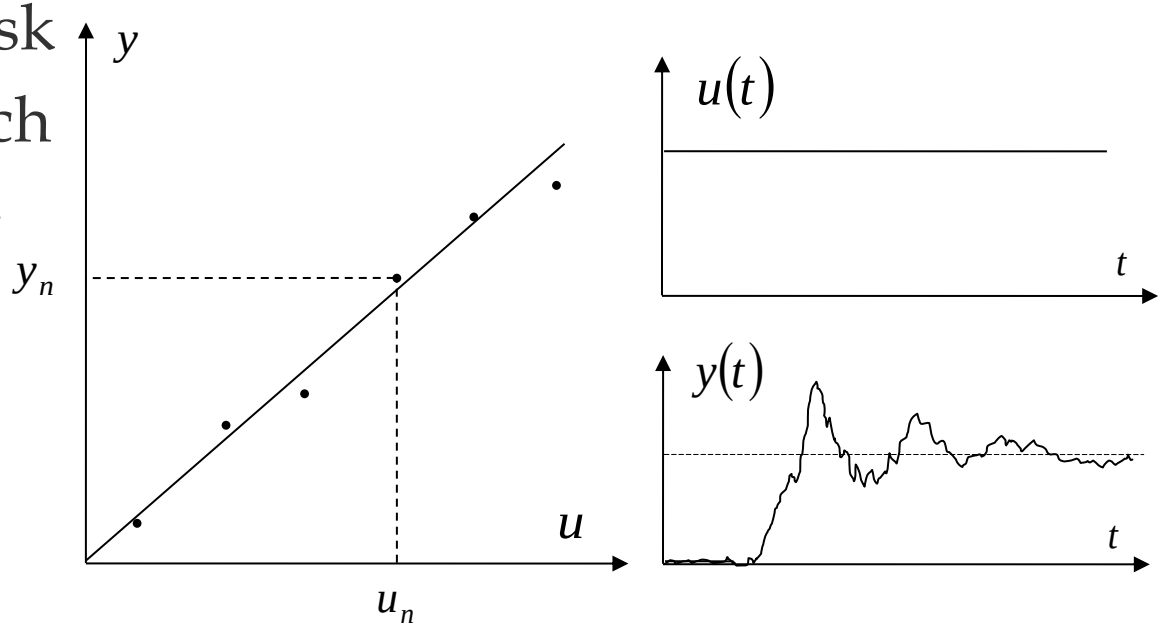
Ad.1. Określenie obiektu identyfikacji



u – wejście
y – wyjście
μ – zakłócenia mierzalne
γ – zakłócenia niemierzalne

Ad.2. Określenie klasy modeli

- Analiza zjawisk
- Analiza danych pomiarowych

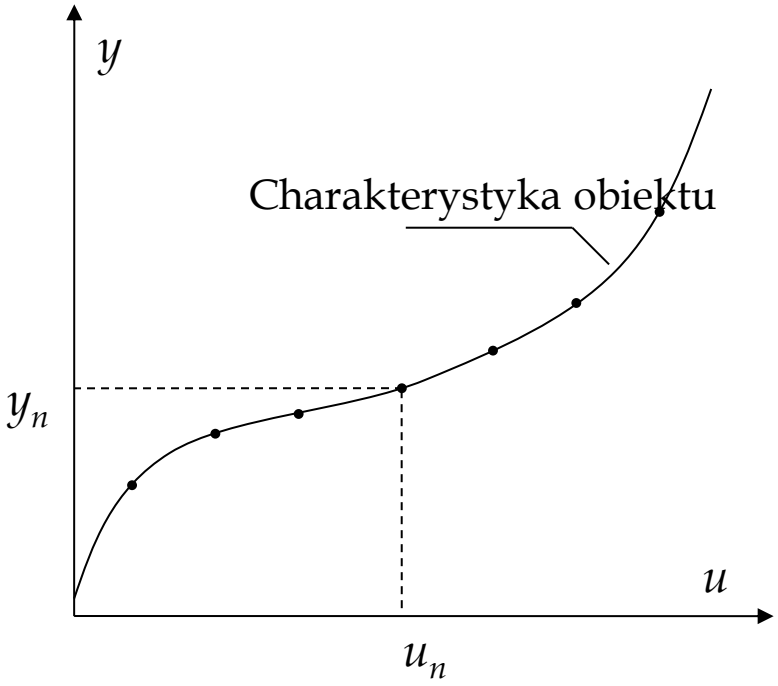


- Model arbitralny
- Model oparty na wiedzy eksperta

Ad.2. Określenie klasy modeli

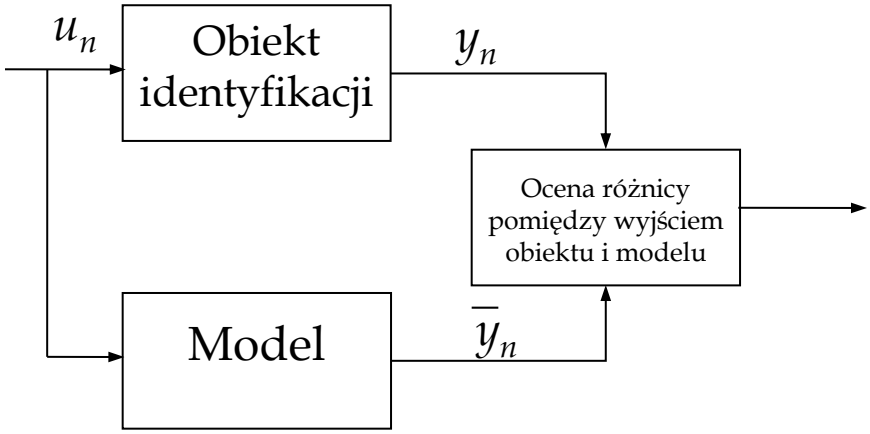
- Obiekt w klasie modeli

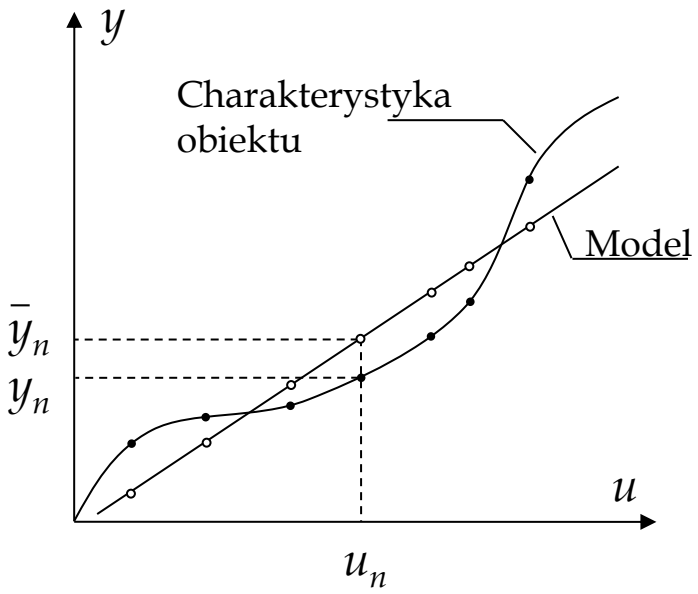




Ad.2. Określenie klasy modeli

- Wybór optymalnego modelu

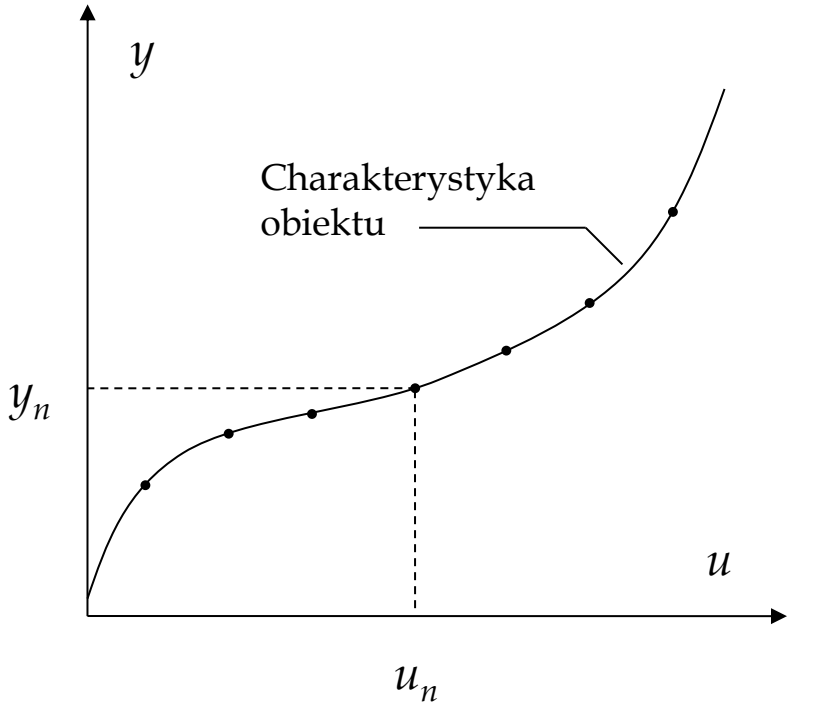




Ad.3. Organizacja eksperymentu

Obiekt statyczny

$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$



Ad.3. Organizacja eksperymentu

Obiekt dynamiczny

$$U_T = \{u(t)\}_{t=t_0}^{T}, \quad Y_T = \{y(t)\}_{t=t_0}^{T},$$

Obszerwacje dyskretne

$$t_1, t_2, \cdots, t_N, t_n \in [t_0\ T], n = 1,2, \ldots, N$$

$$U_N = \{u(t_n)\}_{n=1}^{N}, \quad Y_N = \{y(t_n)\}_{n=1}^{N}.$$

Dynaamiczny obiekt dyskretny

$$U_N = \{u_n\}_{n=1}^{N}, \quad Y_N = \{y_n\}_{n=1}^{N}.$$

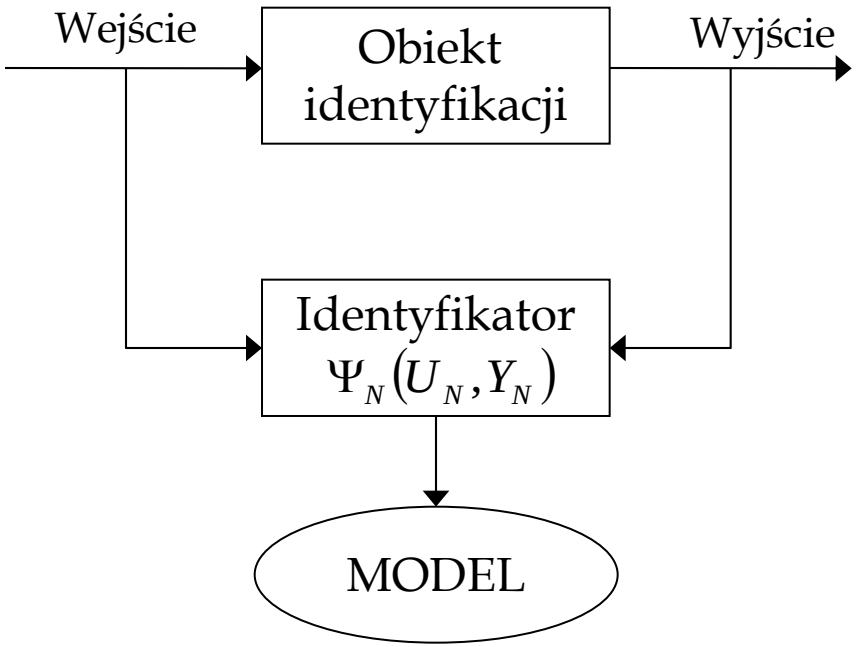
U_N

Ad.3. Organizacja eksperymentu

- Eksperyment pasywny
  - Sygnały wejściowe : U_N, $\{u(t)\}_{t_0}^{T}$, $\{u(t_n)\}_{n=0}^{N}$, $\{u_n\}_{n=0}^{N}$ są tylko obserwowane

- Eksperyment aktywny
  - Sygnały wejściowe : U_N , $\{u(t)\}_{t_0}^{T}$, $\{u(t_n)\}_{n=0}^{N}$, $\{u_n\}_{n=0}^{N}$ mogą być zaprojektowane (zaplanowane)

Ad.4. Wyznaczenie algorytmu identyfikacji



U_N – Seria pomiarowa wejść

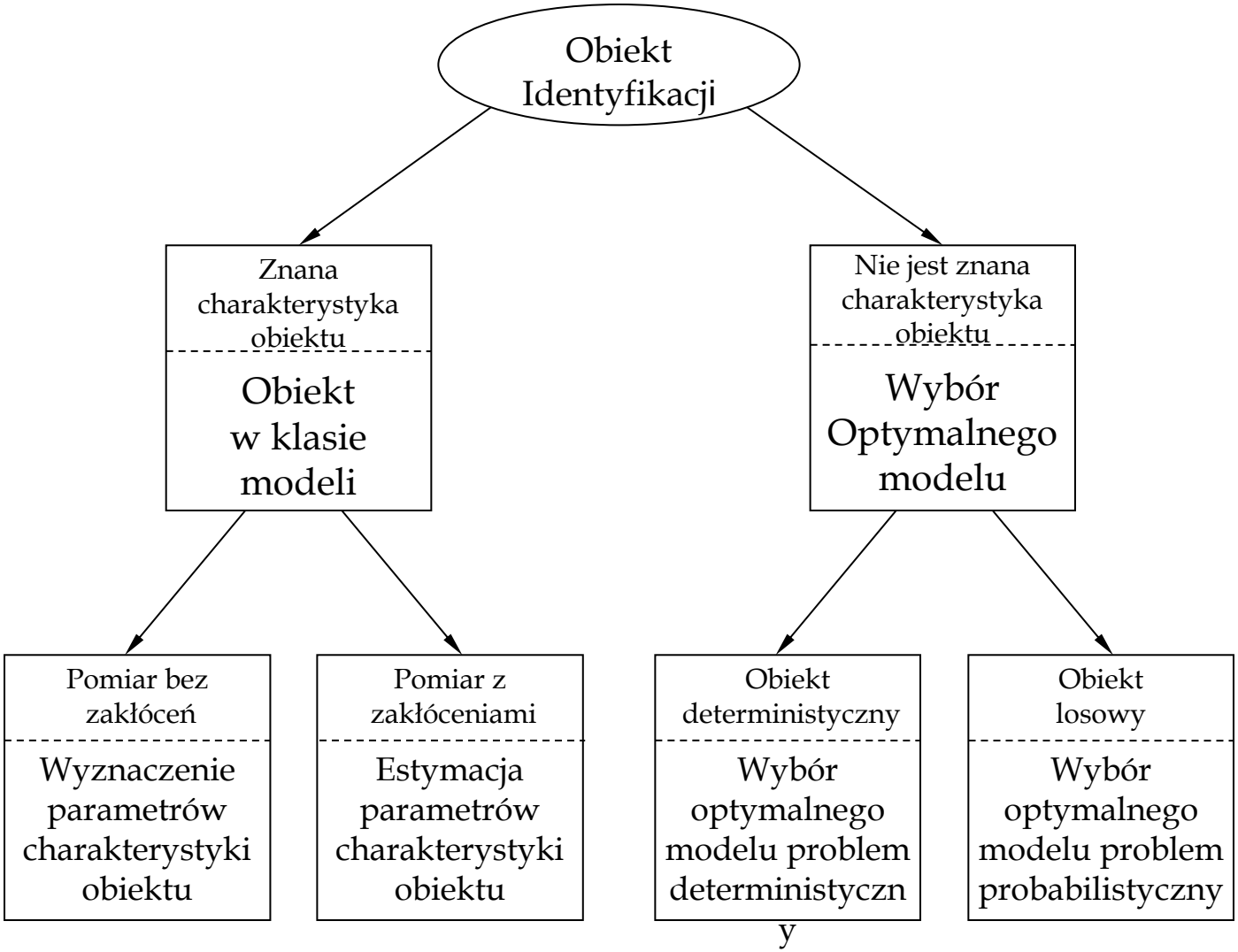
Y_N – Wyniki pomiarów wyjść

Ψ_N – Algorytm identyfikacji

Ad. 5. Wykonanie identyfikatora

- Algorytm identyfikacji
  - Program komputerowy
  - Realizacja sprzętowa

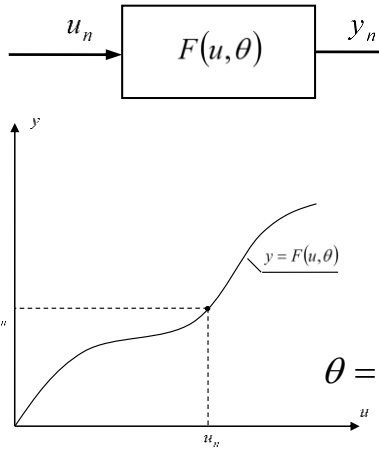
Typowe zadania identyfikacji



Podstawowe zadania identyfikacji podsumowanie

Deterministyczny

Obiekt w klasie modeli



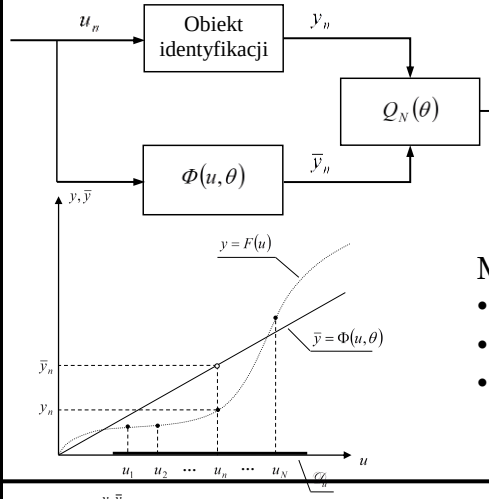
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

$$y_n = F(u_n, \theta),$$
$$n = 1, 2, \ldots, N$$

$$Y_N = \bar{F}(U_N, \theta)$$

$$\theta = \bar{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

Wybór optymalnego modelu

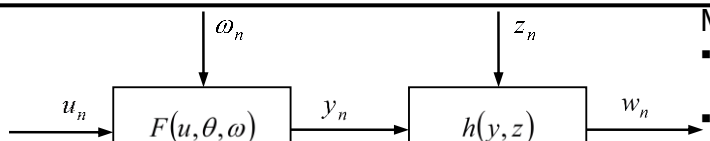


$$\bar{y} = \Phi(u, \theta_N^*)$$
$$\theta_N^* \rightarrow Q_N(\theta_N^*) = \min_{\theta} Q_N(\theta)$$
$$\theta_N^* = \Psi_N(U_N, Y_N)$$
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

Model jest optymalny:

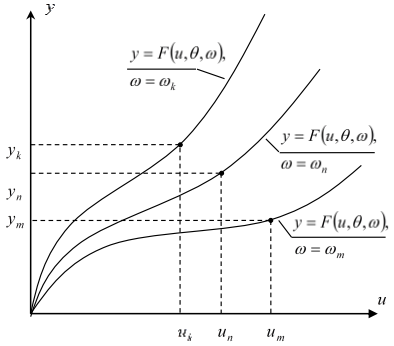
- dla zadanej serii pomiarowej
- przyjętego modelu
- przyjętego wskaźnika jakości identyfikacji

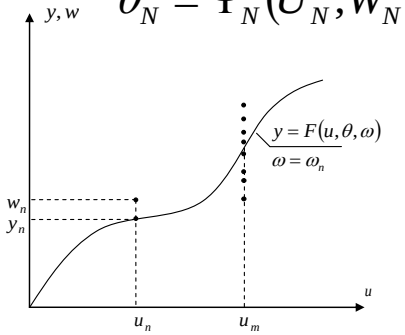
Losowy



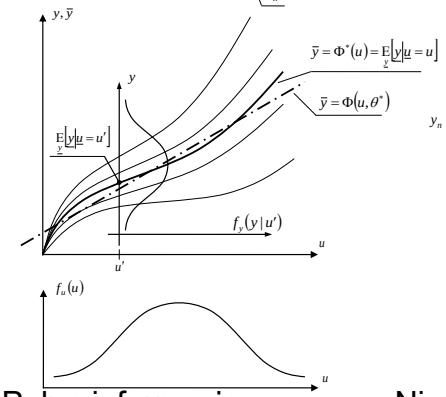
Metody:

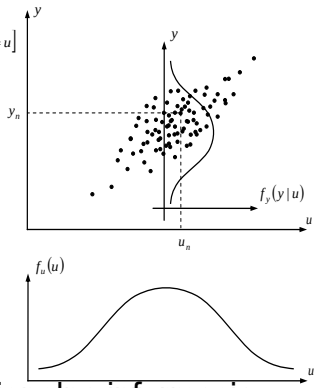
- Najmniejszych kwadratów
- Maksymalnej wiarogodności
- Bayes'a





$$\theta_N = \Psi_N(U_N, W_N)$$





Pełna informacja

- Regresja I rodzaju
- Regresja II rodzaju

Niepełna informacja

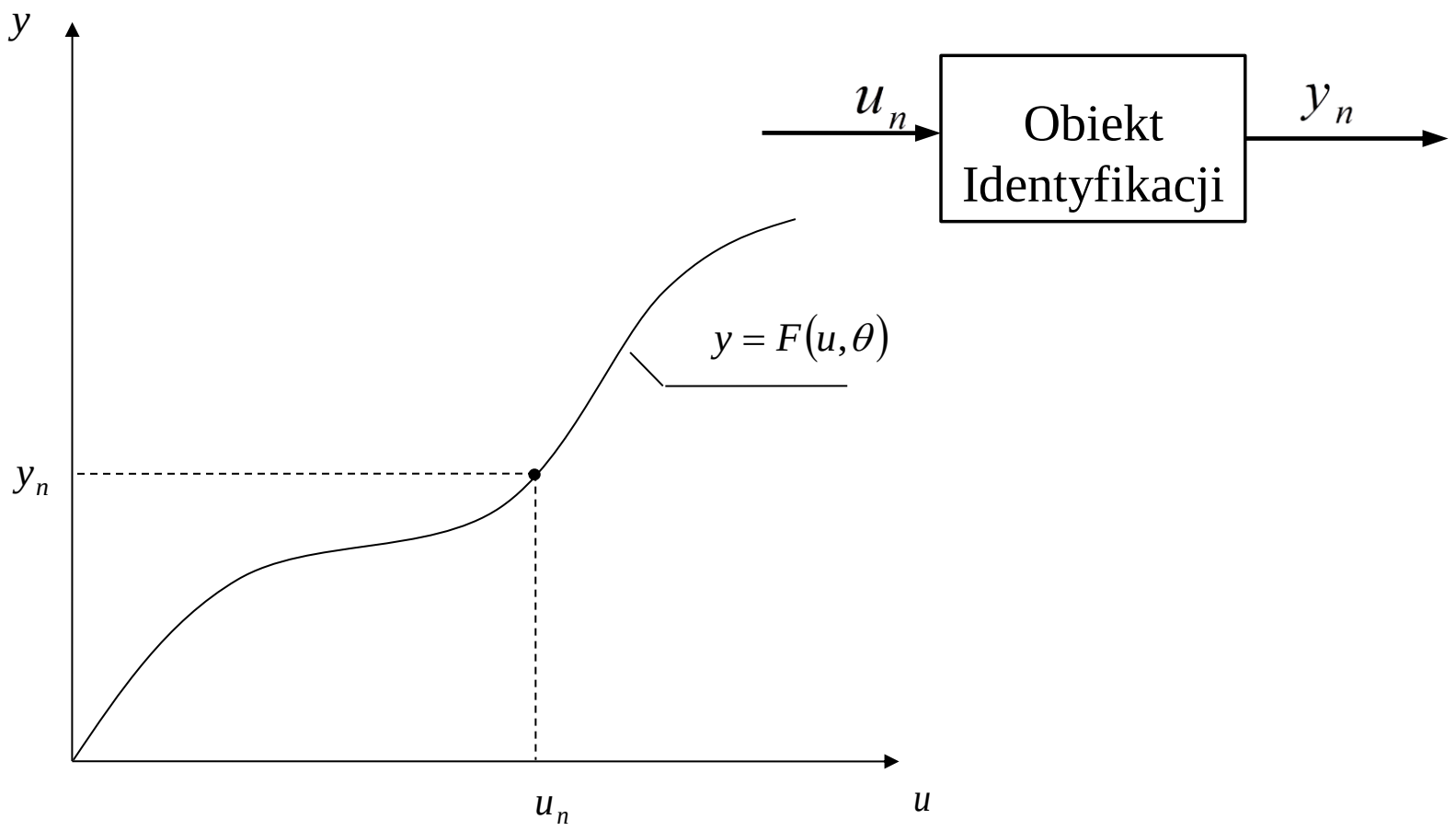
- Estymacja wskaźnika jakości
- Estymacja parametrów rozkładu
- Estymacja rozkładu

Obiekt w klasie modeli Przypadek Deterministyczny

15

Wyznaczeenie parametrów modelu

Wyznaczenie parametrów charakterystyki obiektu



Wyznaczenie parametrów charakterystyki obiektu

- Sformułowanie problemu

Charakterystyka statyczna obiektu: $y = F(u, \theta)$

F – znana funkcja

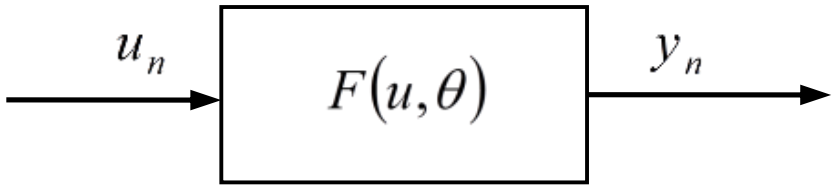
u – wektor wejść $u \in \mathscr{U} \subseteq \mathscr{R}^S$ $\mathscr{U}$ – przestrzeń wejść

y – wektor wyjść $y \in \mathscr{Y} \subseteq \mathscr{R}^L$ $\mathscr{Y}$ – przestrzeń w yjść

θ – nieznany parametr charakterystyki obiektu $\theta \in \Theta \subseteq \mathscr{R}^R$

Θ – przestrzeń parametrów

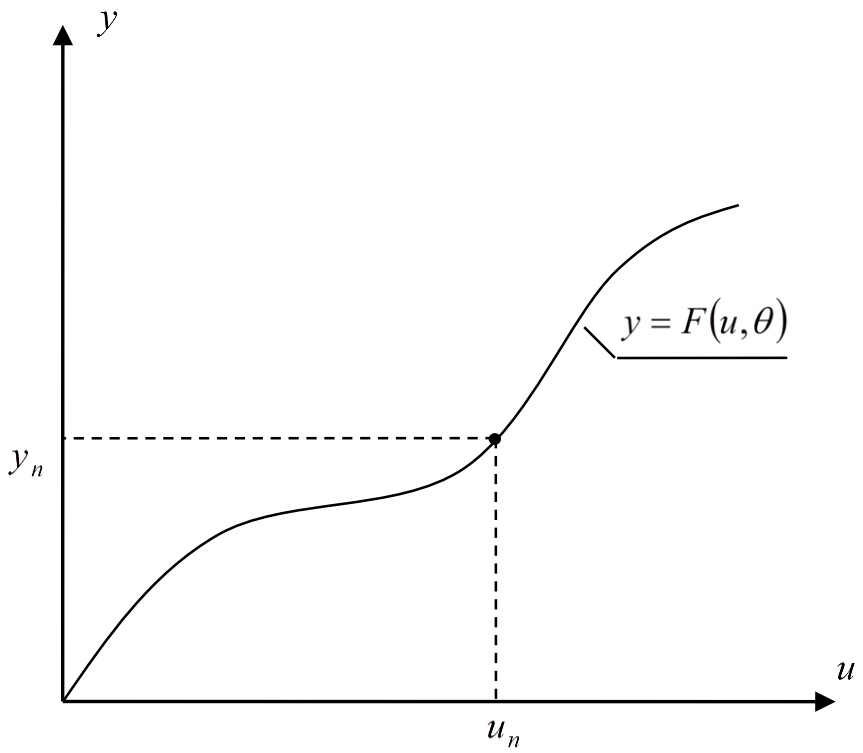
Wyznaczenie parametrów charakterystyki obiektu



Pomiary:

$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$

$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$



Wyznaczenie parametrów charakterystyki obiektu

Punkty pomiarowe pokrywają się z punktami z charakterystyki obiektu, zatem dla każdego pomiaru prawdziwe jest równanie:

$$y_n = F(u_n, \theta), \quad n = 1, 2, \ldots, N$$

W zawartym zapisie ma ono postać:

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix} = \begin{bmatrix} F(u_1,\theta) & F(u_2,\theta) & \cdots & F(u_N,\theta) \end{bmatrix}$$

Oznaczając przez:

$$\begin{bmatrix} F(u_1,\theta) & F(u_2,\theta) & \cdots & F(u_N,\theta) \end{bmatrix} \overset{df}{=} \overline{F}(U_N,\theta)$$

Układ równań przyjmuje postać:

$$Y_N = \overline{F}(U_N,\theta)$$

Wyznaczenie parametrów charakterystyki obiektu

Rozwiązanie (rozwikłanie) równania $Y_N = \bar{F}(U_N, \theta)$ względem parametrów θ daje algorytm identyfikacji tj.:

$$\theta = \bar{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

gdzie:

$\bar{F}^{-1}$ - Funkcja odwrotna względem θ

Ψ_N - Algorytm identyfikacji

Liczba koniecznych pomiarów N – spełnia warunek: $N \times L \geq R$

System jest identyfikowalny jeżeli istnieje taka seria pomiarowa wejść U_N (seria identyfikująca), która wraz z odpowiednią serią pomiarów wyjść Y_N jednoznacznie określa wektor parametrów θ, czyli istnieje jednoznaczne rozwiązanie względem θ równania: $Y_N = \bar{F}(U_N, \theta)$

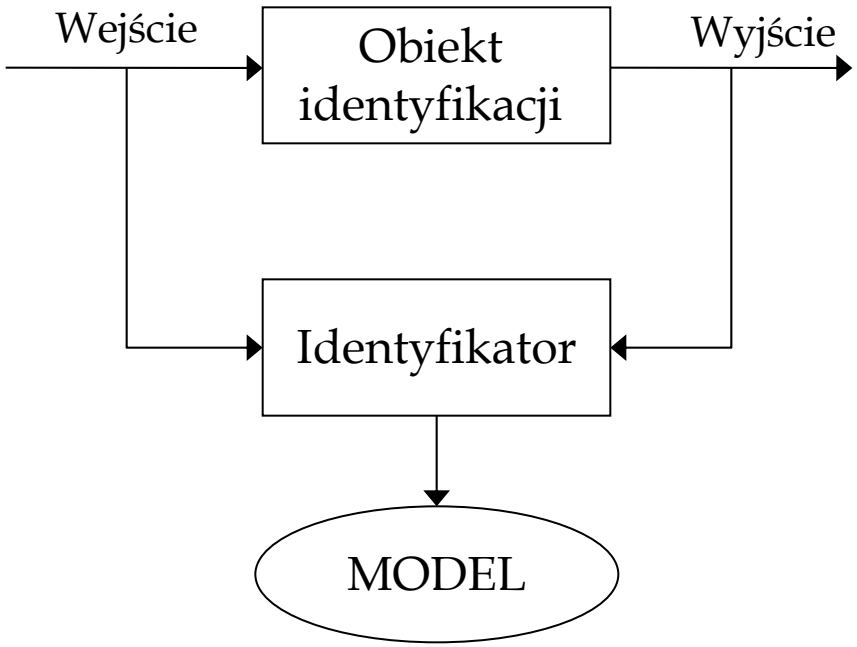
Obiekt w klasie modeli Przypadek Probabilistyczny

21

Estymacja parametrów modelu

Zadanie identyfikacji uczenie maszynowe

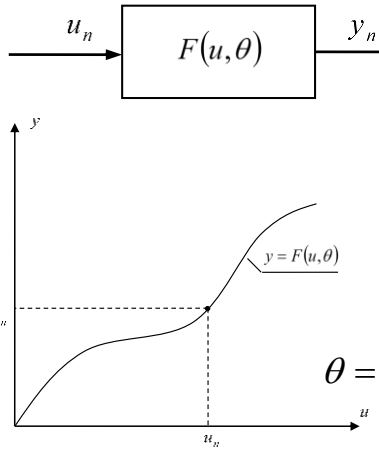
Zadanie identyfikacji (czenie maszynowe) – proces tworzenia modelu matematycznego obiektu na podstawie obserwacji obiektu



Podstawowe zadania identyfikacji podsumowanie

Deterministyczny

Obiekt w klasie modeli



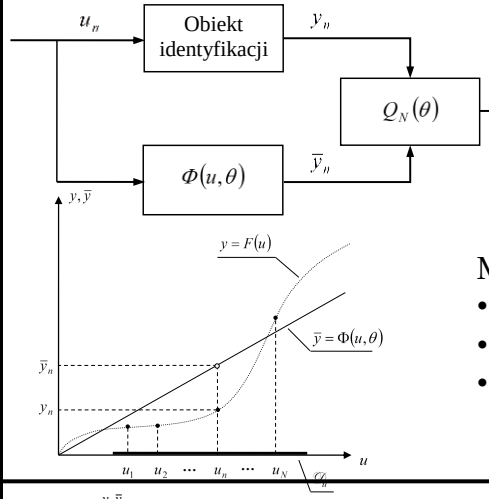
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

$$y_n = F(u_n, \theta),$$
$$n = 1, 2, \ldots, N$$

$$Y_N = \bar{F}(U_N, \theta)$$

$$\theta = \bar{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

Wybór optymalnego modelu

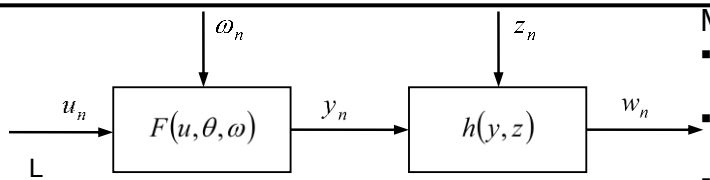


$$\bar{y} = \Phi(u, \theta_N^*)$$
$$\theta_N^* \rightarrow Q_N(\theta_N^*) = \min_{\theta} Q_N(\theta)$$
$$\theta_N^* = \Psi_N(U_N, Y_N)$$
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

Model jest optymalny:

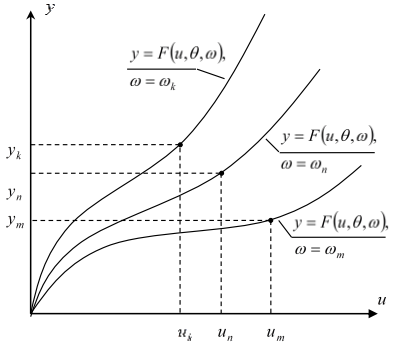
- dla zadanej serii pomiarowej
- przyjętego modelu
- przyjętego wskaźnika jakości identyfikacji

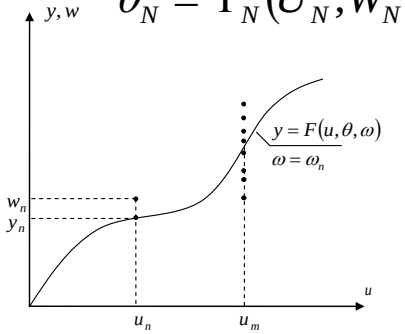
Losowy



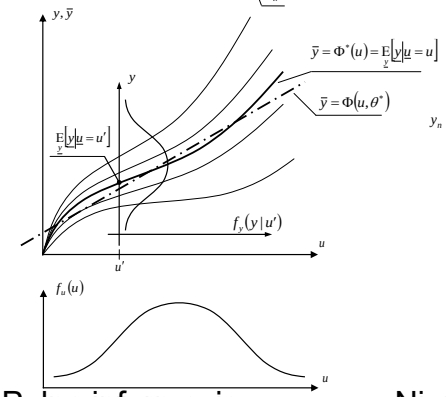
Metody:

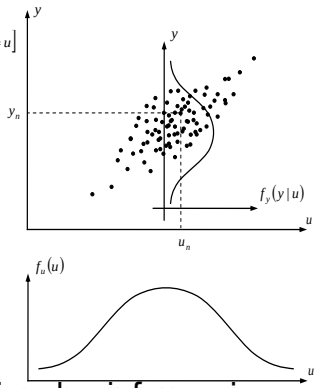
- Najmniejszych kwadratów
- Maksymalnej wiarogodności
- Bayes'a





$$\theta_N = \Psi_N(U_N, W_N)$$





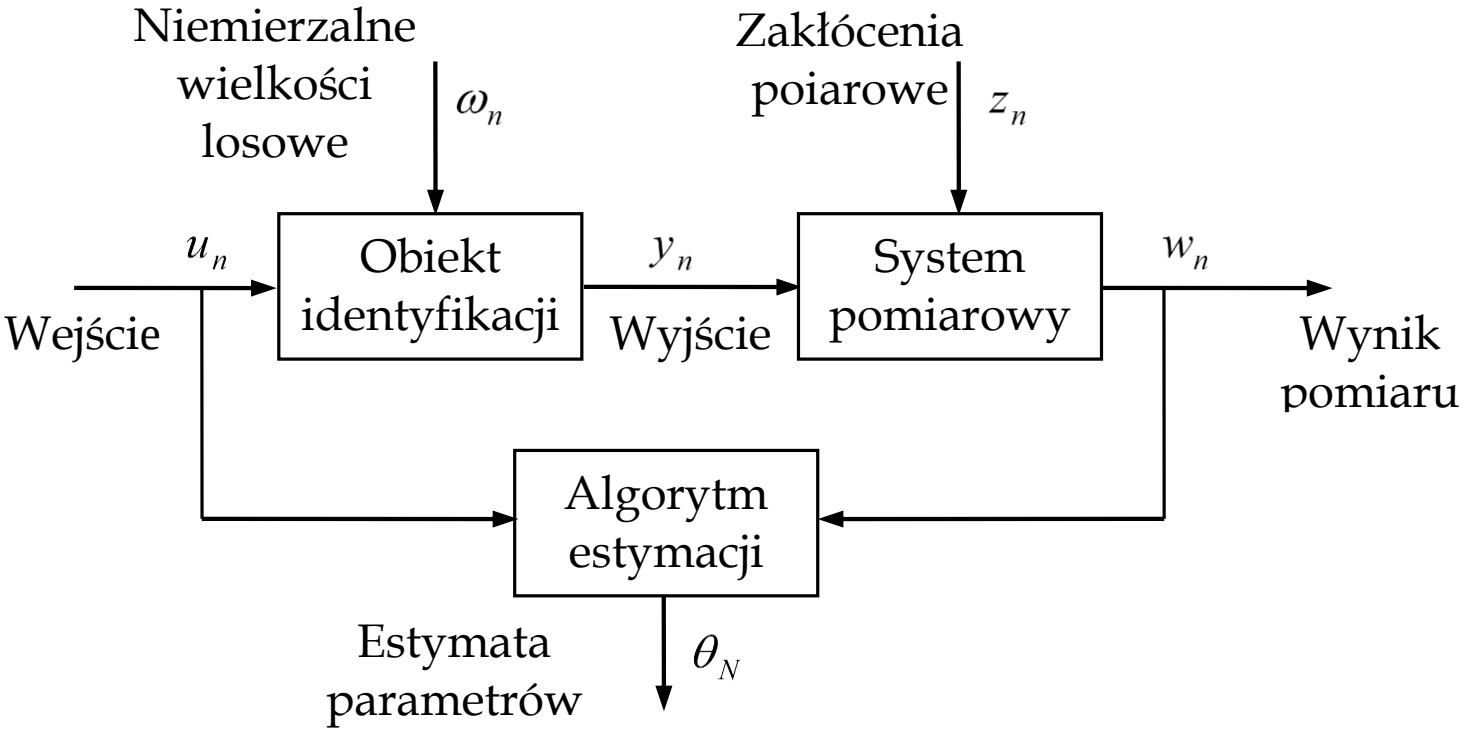
Pełna informacja

- Regresja I rodzaju
- Regresja II rodzaju

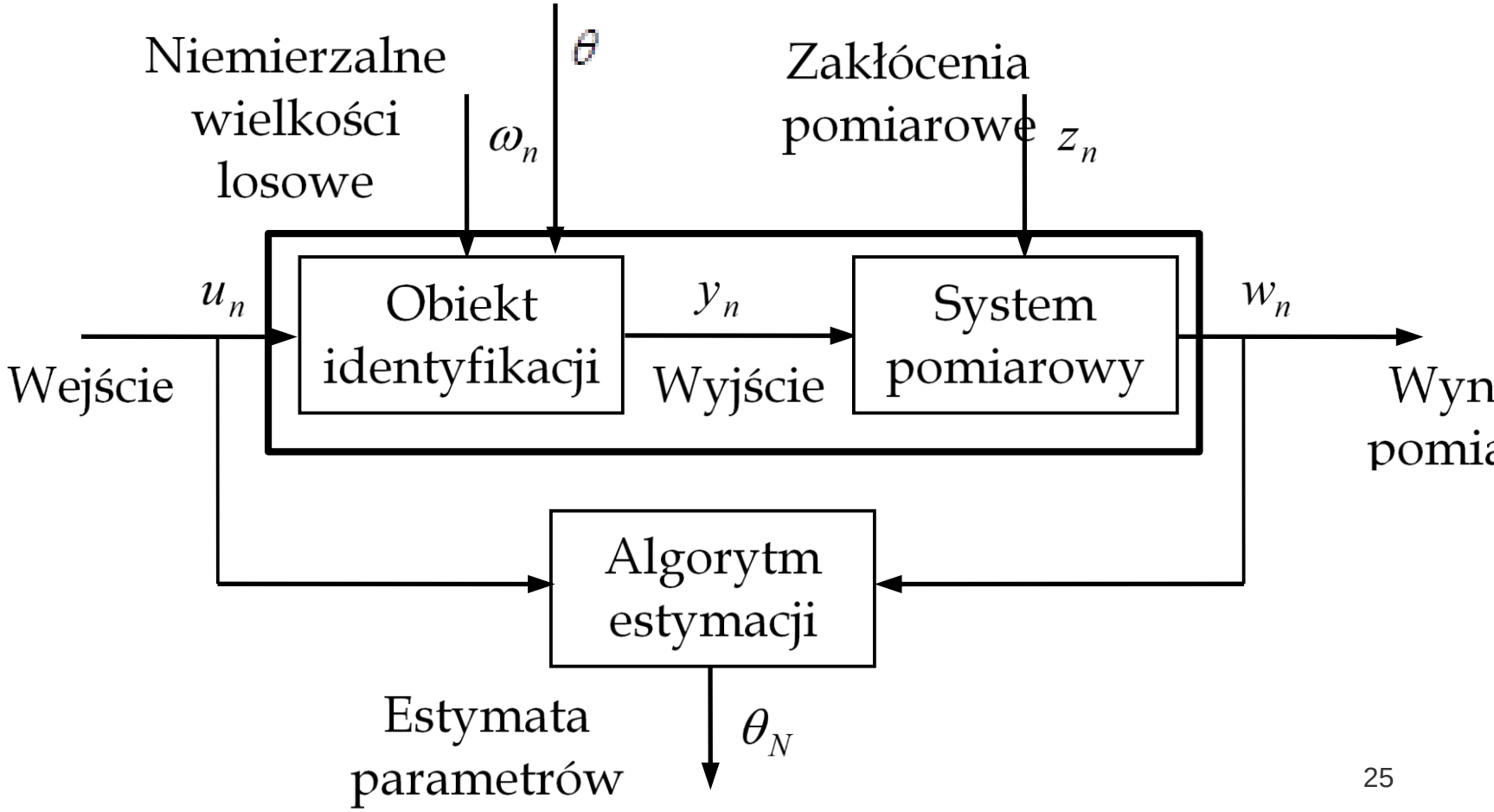
Niepełna informacja

- Estymacja wskaźnika jakości
- Estymacja parametrów rozkładu
- Estymacja rozkładu

Estymacja parametrów charakterystyki statycznej obiektu

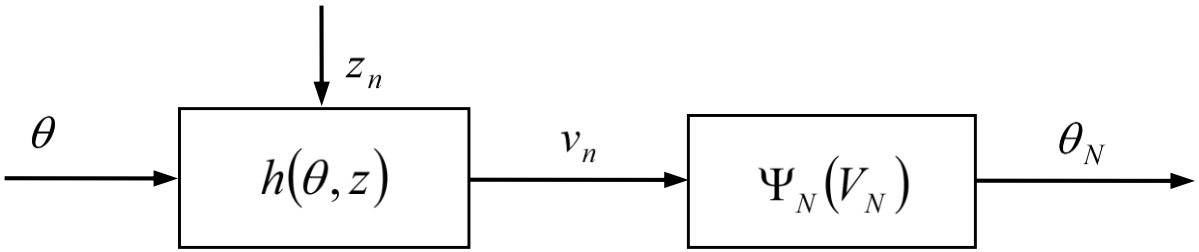


Estymacja parametrów charakterystyki statycznej obiektu



Estymacja parametrów charakterystyki statycznej obiektu

- Zakłócony pomiar wielkości fizycznej



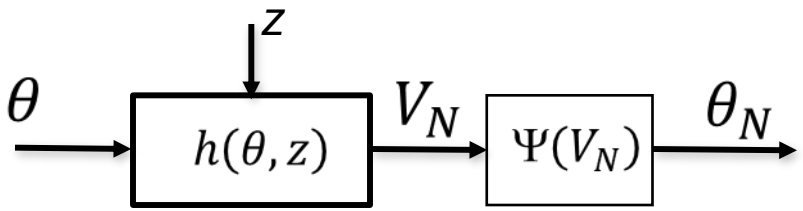
Zakłócony pomiar wielkości fizycznej

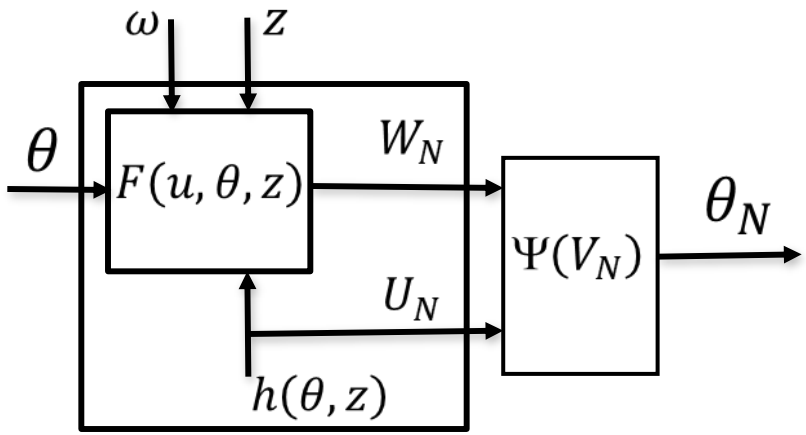
Poszukujemy algorytmu estymacji:

$$\theta_N = \Psi_N(V_N)$$

- Możliwe rozwiązania:
  - Metoda najmniejszych kwadratów
  - Metoda maksymalnej wiarogodności
  - Metody Bayes'a

ω_n

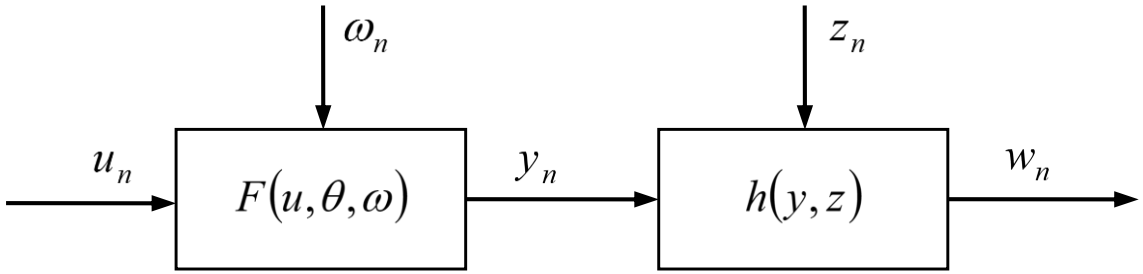


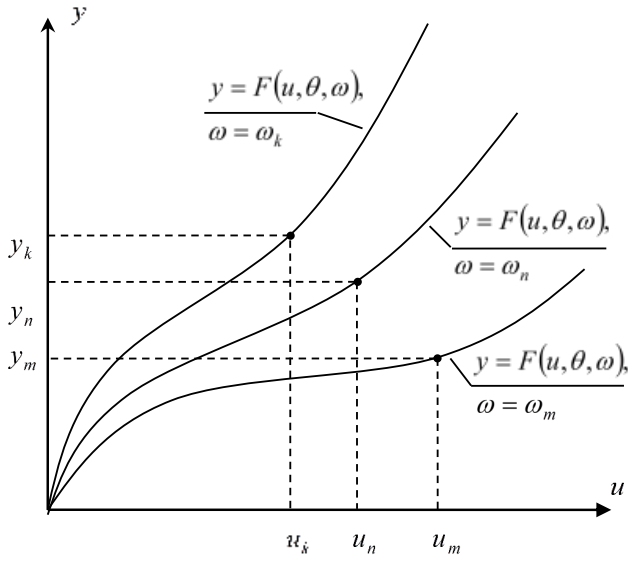


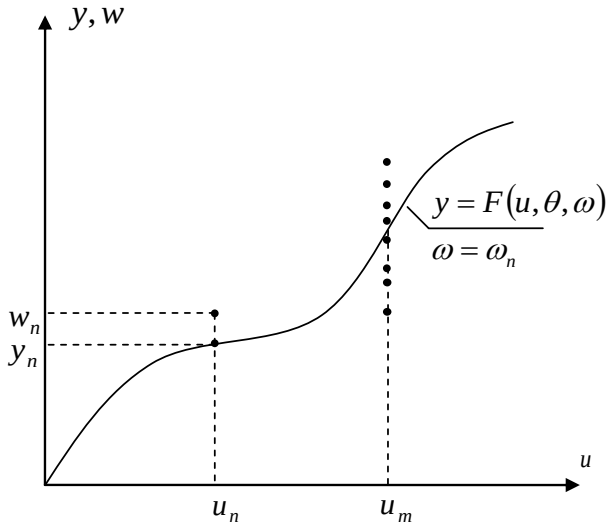
$$V_N \begin{bmatrix} W_N \\ U_N \end{bmatrix}$$

Pomiar wyjścia z zakłóceniami oraz losowo zmienny parametr obiektu

Pomiar wyjścia z zakłóceniami oraz losowo zmienny parametr obiektu

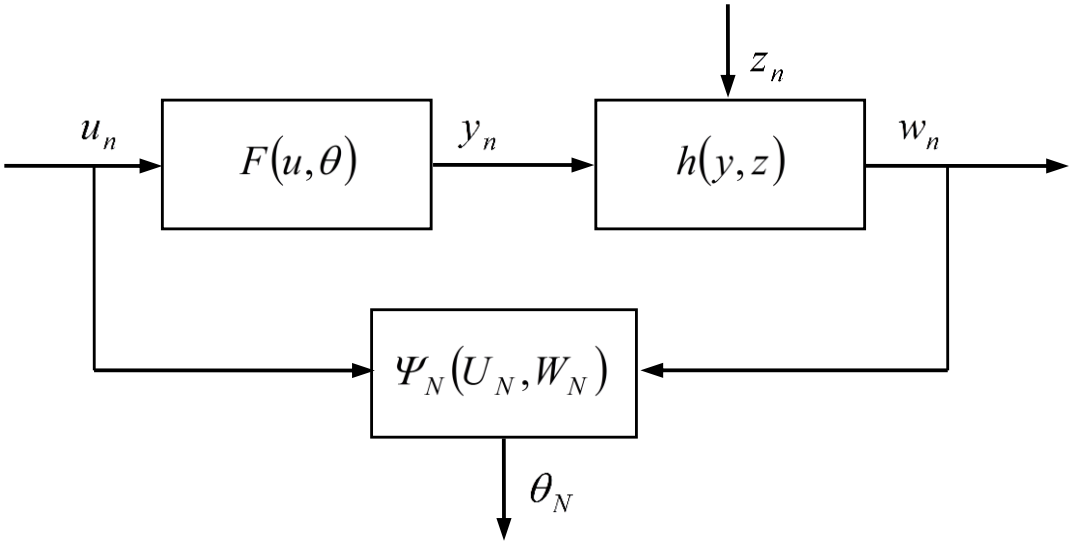






Estymacja parametrów charakterystyki statycznej obiektu

- Obiekt deterministyczny, zakłócony pomiar wyjścia



gdzie:

$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N]$

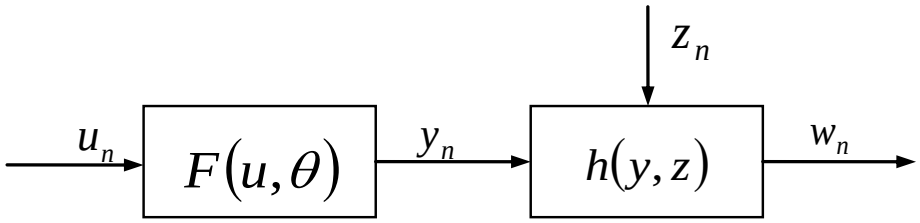
$W_N = [w_1 \quad w_2 \quad \cdots \quad w_N]$

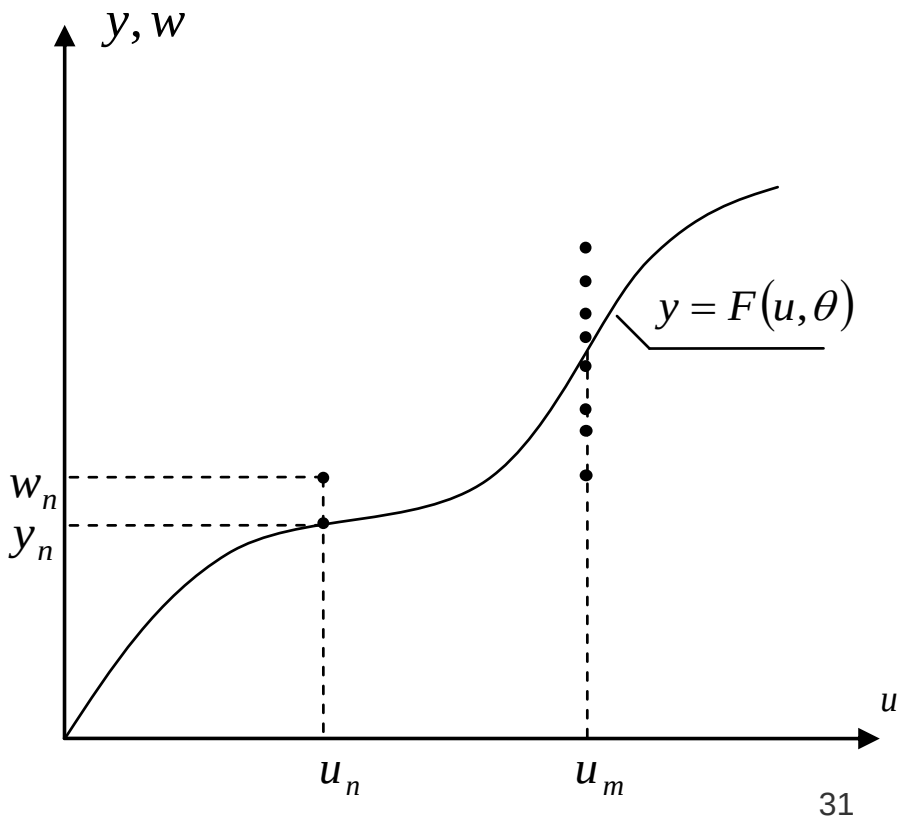
Ψ_N – algorytm estymacji

θ_N – estymata parametru θ

Estymacja parametrów charakterystyki statycznej obiektu

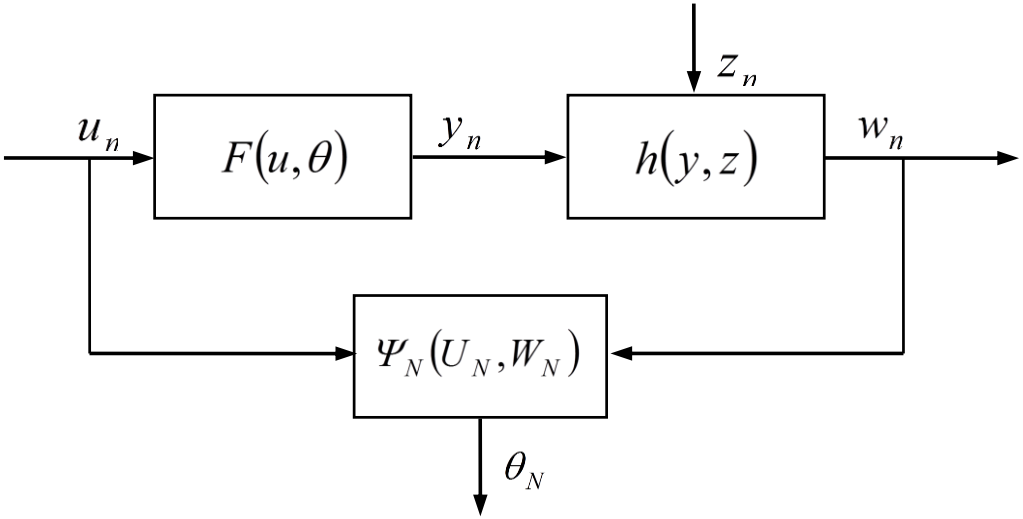
- Pomiar wyjścia obiektu z zakłóceniami pomiarowymi





Estymacja parametrów charakterystyki statycznej obiektu

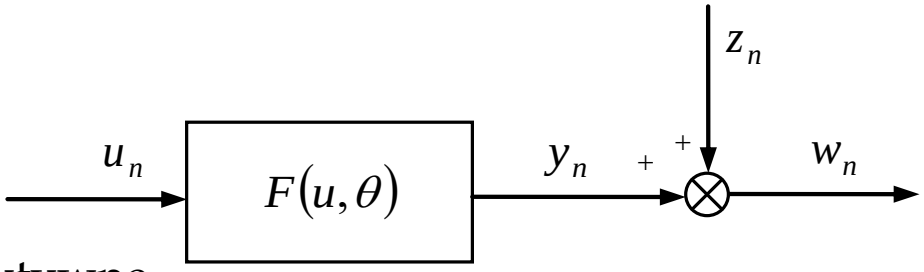
Poszukujemy algorytmu estymacji: $\theta_N = \Psi_N(U_N, W_N)$



- Możliwe podejścia:
  - Metoda najmniejszych kwadratów
  - Metoda maksymalnej wiarogodności
  - Metody Bayes'a

Metoda najmniejszych kwadratów

Przypadek jednowymiarowym wyjściem $L=1$



Założenia:

$w=h(y,z)=y+z$ – zakłócenia addytywne

$E_{\underline{z}}[\underline{z}]=0$ – wartość oczekiwana zakłóceń równa zero

$Var_{\underline{z}}[\underline{z}]<\infty$ – skończona wariancja zakłóceń

Estymator metodą najmniejszych kwadratów otrzymujemy minimalizując empiryczną wariancję zakłóceń :

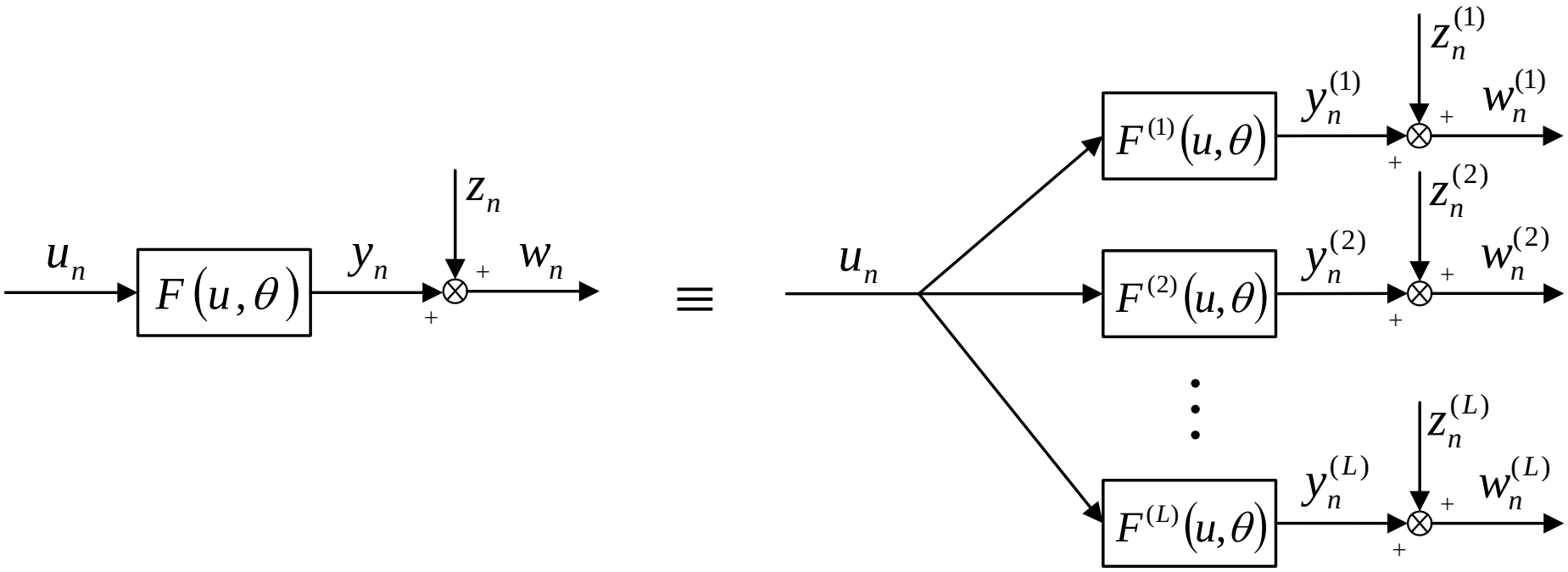
$$Var_{zN}(U_N,W_N,\theta)=\frac{1}{N}\sum_{n=1}^{N}(w_n-y_n)^2=\frac{1}{N}\sum_{n=1}^{N}(w_n-F(u_n,\theta))^2$$

Algorytm estymacji otrzymujemy w wyniku rozwiązania zadania:

$$\theta_N=\Psi_N(U_N,W_N)\quad\rightarrow\quad Var_{zN}(U_N,W_N,\theta_N)=\min_{\theta\in\Theta}Var_{zN}(U_N,W_N,\theta)$$

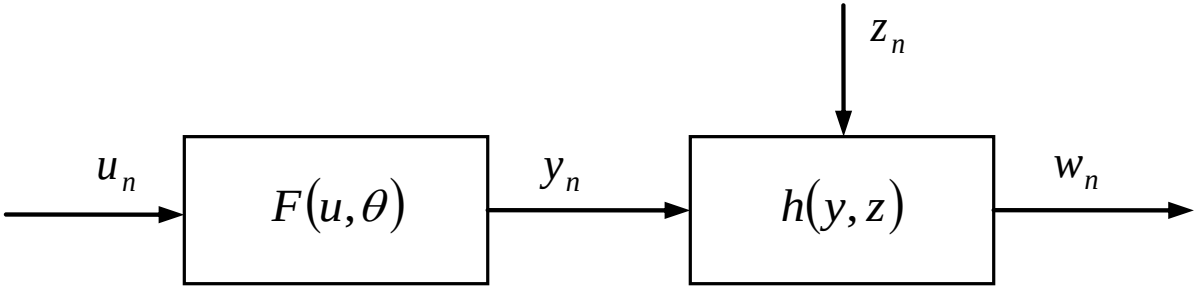
Metoda najmniejszych kwadratów

Przypadek z wielowymiarowym wyjściem L dekomponujemy na L zadań z jednowymiarowym wyjściem:



Metoda maksymalnej wiarogodności

Założenia



Charakterystyka statyczna obiektu: $y = F(u,\theta)$ - znana funkcja

z_n – zakłócenia pomiarowe są niezależną wartością zmiennej losowej Z, dla której znany jest gęstość rozkładu prawdopodobieństwa $f_z(z)$

$\underline{w} = h(y,\underline{z})$ – sposób nakładania się zakłóceń na mierzone wyjście opisani jest funkcji *h,* która jest wzajemnie jednoznaczna względem zakłóceń

Dla zadanej serii wejść $u_n, n = 1, 2, \ldots, N$ otrzymujemy wyniki zakłóconego pomiaru wyjść $w_n, n = 1, 2, \ldots, N$

Metoda maksymalnej wiarogodności

Zakłócone wyniki pomiarów wyjścia w_n są wartościami zmiennej losowej $\underline{w}$, której funkcję gęstości rozkładu prawdopodobieństwa należy powiązać z systemem pomiarowym oraz obiektem identyfikacji. Zmienna losowa $\underline{w}$ jest wynikiem przekształcenia zmiennej losowej $\underline{z}$ w systemie pomiarowym z uwzględnieniem charakterystyki obiekty, tj.: $\underline{w} = h(\bar{F}(u,\theta), \underline{z})$, zatem funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{w}$ – $f_w(w,\theta;u)$ wyznaczamy:

$$f_w(w,\theta;u) = f_z\left(h_z^{-1}(F(u,\theta),w)\right)\cdot|J_h|$$

gdzie: J_h - macierz Jacobiego $J_h = \dfrac{\partial h_z^{-1}(y,w)}{\partial w}$

Funkcja wiarogodności ma postać:

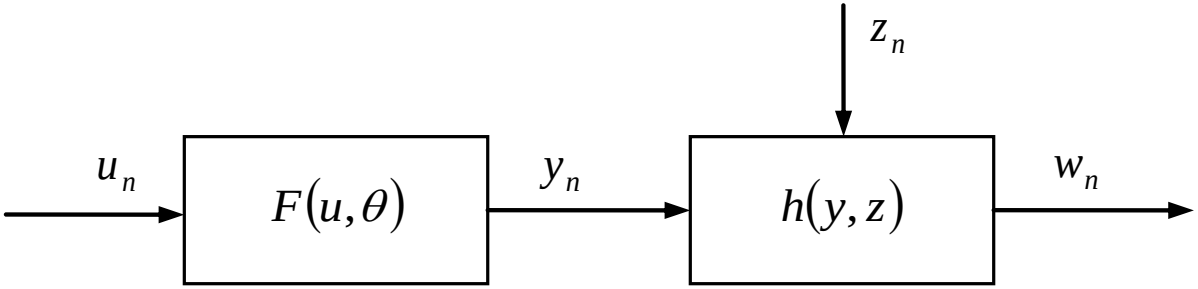
$$L_N(W_N,\theta;U_N) = \prod_{n=1}^{N} f_w(w_n,\theta;u_n) = \prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n,\theta),w_n)\right)|J_h|$$

W wyniku maksymalizacji funkcji wiarogodności otrzymujemy algorytm estymacji:

$$\theta_N = \Psi_N(U_N,W_N) \quad \rightarrow \quad L_N(W_N,\theta_N;U_N) = \max_{\theta\in\Theta} L_N(W_N,\theta;U_N)$$

Metody Bayes' a

Założenia



Charakterystyka statyczna obiektu: $y = F(u,\theta)$ - znana funkcja

z_n – zakłócenia pomiarowe są niezależną wartością zmiennej losowej $\underline{z}$, dla której znany jest gęstość rozkładu prawdopodobieństwa $f_z(z)$

$\underline{w} = h(y,\underline{z})$ – sposób nakładania się zakłóceń na mierzone wyjście opisani jest funkcji *h,* która jest wzajemnie jednoznaczna względem zakłóceń

θ - wektor parametrów jest wartością zmiennej losowej $\underline{\theta}$, $\theta \in \Theta \subseteq \mathscr{R}^R$, dla której funkcja gęstości rozkładu prawdopodobieństwa - $f_\theta(\theta)$ jest znana, $f_\theta(\theta)$ - gęstość rozkładu prawdopodobieństwa a' priori

$L(\theta,\bar{\theta})$ - zadana jest funkcja strat

Dla zadanej serii wejść $u_n, n = 1, 2, \ldots, N$ otrzymujemy wyniki zakłóconego pomiaru wyjść $w_n, n = 1, 2, \ldots, N$

Metody Bayes' a

Rryzyko podjęcia decyzji:

$$R\left(\overline{\Psi}\right) \overset{\text{df}}{=} \underset{\underline{\theta},\underline{W_N}}{E}\left[L\left(\underline{\theta},\underline{\bar{\theta}} = \overline{\Psi}\left(U_N,\underline{W_N}\right)\right)\right] = \int_{\mathcal{W}_N}\int_{\Theta} L\left(\theta,\overline{\Psi}\left(U_N,W_N\right)\right) f\left(\theta,W_N;U_N\right) d\theta\, dW_N$$

gdzie $f\left(\theta,W_N;U_N\right)$ jest łączną gęstością rozkładu prawdopodobieństwa wektora parametrów oraz wyników pomiarów, dla zadanej serii pomiarowej wejść U_N

Należy wyznaczyć algorytm estymacji, który minimalizuje ryzyko, czyli:

$$\Psi_N \rightarrow \; R\left(\Psi_N\right) = \min_{\overline{\Psi}} R\left(\overline{\Psi}\right)$$

Zauważmy, że $f\left(\theta,W_N;U_N\right) = f'\left(\theta \middle| W_N;U_N\right) f''\left(W_N;U_N\right)$

Metody Bayes' a

Ryzyko podjęcia decyzji przyjmie postać

$$R(\overline{\Psi}) = \int_{\mathscr{W}_N} \int_{\Theta} L\left(\theta, \overline{\Psi}(U_N, W_N)\right) f'\left(\theta \middle| W_N; U_N\right) d\theta \; f''(W_N; U_N)\, dW_N$$

Ryzyko warunkowe:

$$r\left(\overline{\theta}, W_N; U_N\right) \overset{\text{df}}{=} \underset{\underline{\theta}}{E}\left[L\left(\underline{\theta}, \overline{\theta}\right) \middle| W_N; U_N\right] = \int_{\Theta} L\left(\theta, \overline{\theta} = \overline{\Psi}(U_N, W_N)\right) f'\left(\theta \middle| W_N; U_N\right) d\theta$$

W wyniku minimalizacji powyższego otrzymujemy algorytm estymacji:

$$\theta_N = \Psi_N(U_N, W_N) \quad \rightarrow \quad r(\theta_N, W_N; U_N) = \min_{\overline{\theta} \in \Theta} r\left(\overline{\theta}, W_N; U_N\right)$$

Metody Bayes' a

Pozostaje wyznaczyć warunkową gęstość rozkładu prawdopodobieństwa (rozkład prawdopodobieństwa a' posteriori):

zatem:
$$f'(\theta|W_N;U_N)=\frac{f(\theta,W_N;U_N)}{\int_\Theta f(\theta,W_N;U_N)d\theta}=\frac{f(\theta,W_N;U_N)}{f''(W_N;U_N)}$$

gdzie:
$$f'(\theta|W_N;U_N)=\frac{f_\theta(\theta)\,f_{WN}(W_N|\theta;U_N)}{f''(W_N;U_N)}=\frac{f_\theta(\theta)\,f_{WN}(W_N|\theta;U_N)}{\int_\Theta f_\theta(\theta)\,f_{WN}(W_N|\theta;U_N)d\theta}$$

$$f_{WN}(W_N|\theta;U_N)=\prod_{n=1}^{N} f_w(w_n|\theta;u_n)$$

Biorąc pod uwagę charakterystykę obiektu oraz opis systemu pomiarowego obserwowana zmienna losowa ma postać: $\underline{w}=h(F(u,\theta),\underline{z})$

Warunkowy rozkład obserwowanej zmiennej losowej wyraża się wzorem:

$$f_w(w|\theta;u)=f_z\left(h_z^{-1}(F(u,\theta),w)\right)|J_h|$$

Ostatecznie:

$$f'(\theta|W_N;U_N)=\frac{f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n,\theta),w_n)\right)|J_h|}{\int_\Theta f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n,\theta),w_n)\right)|J_h|d\theta}\quad\alpha\quad f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n,\theta),w_n)\right)|J_h|$$

Metody Bayes' a

- Metoda średniej a' posteriori

Specjalna postać funkcji strat: $L(\theta,\bar{\theta})=[\theta-\bar{\theta}]^T[\theta-\bar{\theta}]$

Ryzyko warunkowe: $r(\bar{\theta},W_N;U_N)\overset{df}{=}\underset{\underline{\theta}}{E}\left[[\underline{\theta}-\bar{\theta}]^T[\underline{\theta}-\bar{\theta}]|W_N;U_N\right]=$

$$=\underset{\underline{\theta}}{E}\left[\underline{\theta}^T\underline{\theta}|W_N;U_N\right]-2\bar{\theta}\underset{\underline{\theta}}{E}\left[\underline{\theta}^T|W_N;U_N\right]+\bar{\theta}^T\bar{\theta}$$

W wyniku minimalizacji powyższego, tj.:

$$\underset{\bar{\theta}}{\text{grad}}\, r(\bar{\theta},W_N;U_N)\Big|_{\bar{\theta}=\theta_N}=-2\underset{\underline{\theta}}{E}[\underline{\theta}|W_N;U_N]+2\theta_N=0_N$$

otrzymujemy algorytm estymacji:

$$\theta_N=\Psi_N(U_N,W_N)=\underset{\underline{\theta}}{E}[\underline{\theta}|W_N;U_n]=\int_{\Theta}\theta f'(\theta|W_N;U_N)d\theta$$

Metody Bayes' a

- Metoda maksymalnego prawdopodobieństwa a' posteriori

Specjalna postać funkcji strat: $L(\theta,\bar{\theta}) = -\delta(\theta - \bar{\theta})$

Ryzyko warunkowe: $r(\bar{\theta}, W_N; U_N) = -\int_{\Theta} \delta(\theta - \bar{\theta}) f'(\theta | W_N; U_N)\, d\theta = -f'(\bar{\theta} | W_N; U_N)$

Algorytm estymacji otrzymujemy w wyniku maksymalizacji rozkładu prawdopodobieństwa a' posteriori, tj.:

$$\theta_N = \Psi_N(U_N, W_N) \quad \rightarrow \quad f'(\theta_N | W_N; U_N) = \max_{\bar{\theta}\in\Theta} f'(\bar{\theta} | W_N; U_N)$$

$$f'(\theta_N | W_N; U_N) \quad \alpha \quad f_\theta(\theta_N) \prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n, \theta_N), w_n)\right) |J_h|$$

Ostatecznie $\theta_N = \Psi_N(U_N, W_N)$ takie, że:

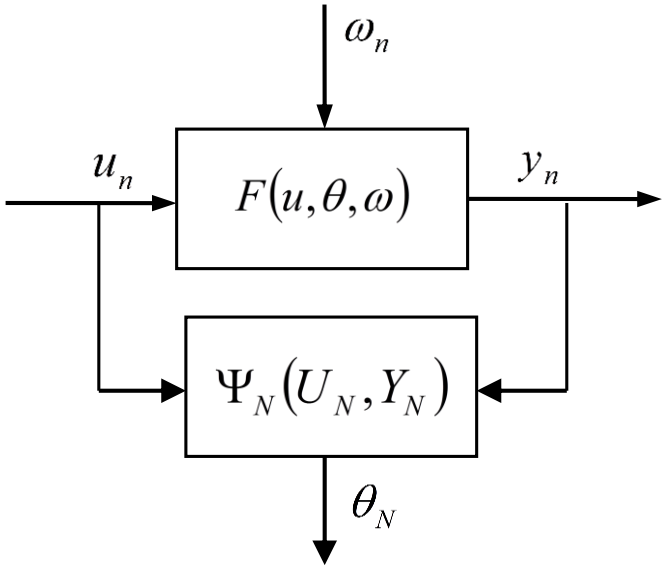
$$f_\theta(\theta_N) \prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n, \theta_N), w_n)\right) |J_h| = \max_{\bar{\theta}\in\Theta} f_\theta(\bar{\theta}) \prod_{n=1}^{N} f_z\left(h_z^{-1}(F(u_n, \bar{\theta}), w_n)\right) |J_h|$$

Estymacja parametów obiektu z loswą wielkością w obiekcie

43

Estymacja parametrów charakterystyki statycznej obiektu

- Niemierzalny losowy parametr w obiekcie



gdzie:

$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N]$$

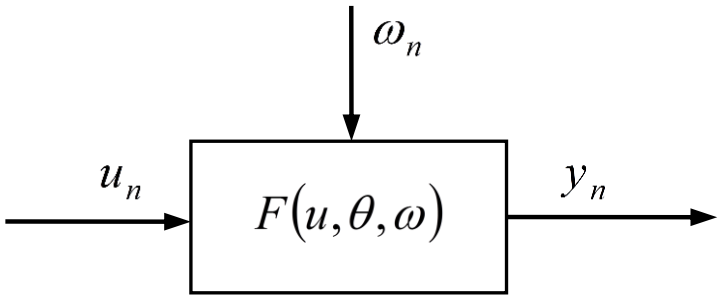
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

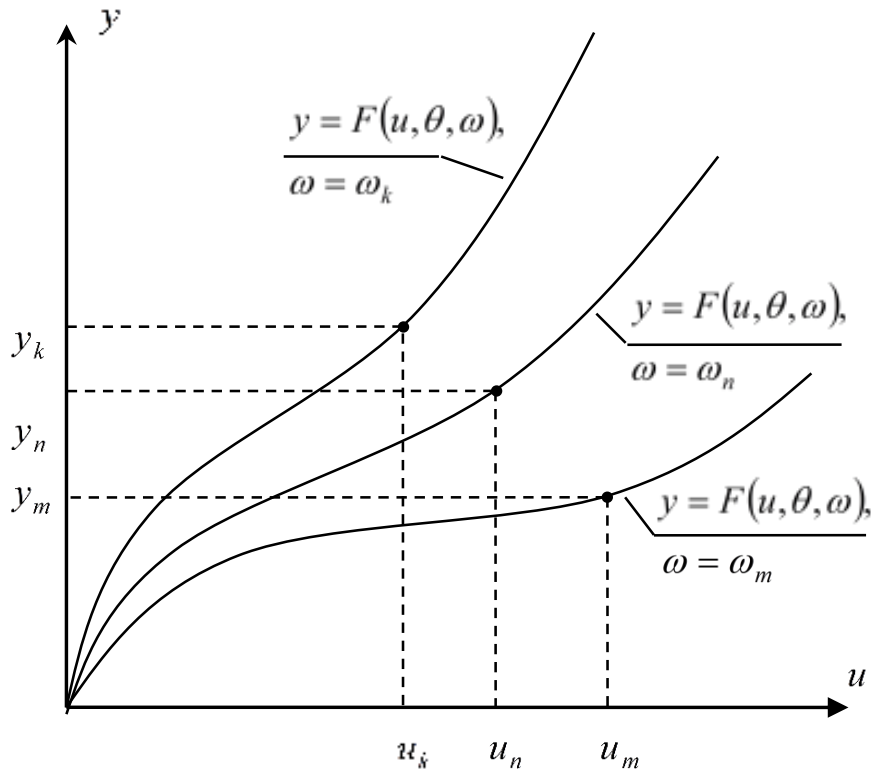
Ψ_N – algorytm estymacji

θ_N – estymata parametru θ

Niemierzalny losowy parametr obiektu

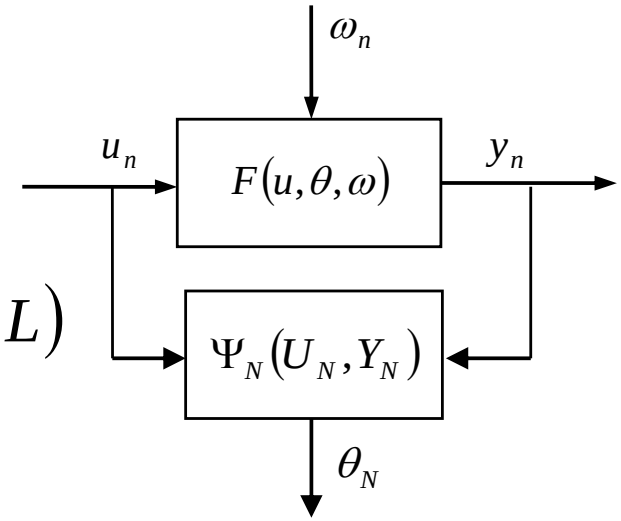
- Pomiary z obiektu z losowo zmiennym parametrem





Niemierzalny losowy parametr obiektu

- Sformułowanie problemu



Charakterystyka obiektu: $y = F(u,\theta,\omega)$

Losowy parametr obiektu:

$$\omega \in \Omega \subseteq \mathcal{R}^L, \quad (\dim y = \dim \omega = L)$$

F – wzajemnie jednoznaczna funkcja względem ω

$$\omega = F_{\omega}^{-1}(u,\theta,y)$$

ω_n – wartość zmiennej losowej $\underline{\omega}$ z przestrzeni $\Omega \subseteq \mathcal{R}^L$

Znana funkcja gęstości rozkładu prawdopodobieństwa $f_{\omega}(\omega)$

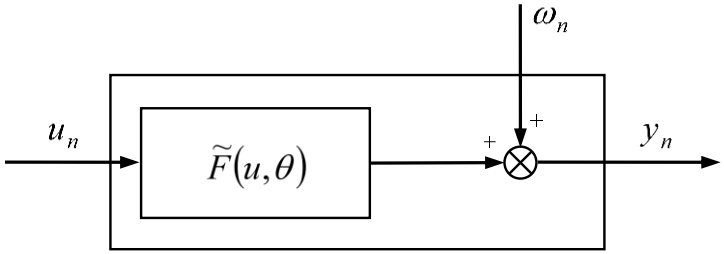
Pomiary: $U_N = [u_1 \quad u_2 \quad \cdots \quad u_N]$, $Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$

Algorytm estymacji: $\theta_N = \Psi_N(U_N, Y_N)$

Metoda najmniejszych kwadratów

Założenie:

Charakterystyka obiektu:



$$y = F(u,\theta,\omega) = \widetilde{F}(u,\theta) + \omega$$

$$\underset{\omega}{E}[\omega] = 0 \qquad \underset{\omega}{Var}[\omega] < \infty$$

Metoda najmniejszych kwadratów minimalizuje empiryczną wariancje

$$Var_{\omega N}(U_N, Y_N, \theta) = \frac{1}{N}\sum_{n=1}^{N}\left(y_n - \widetilde{F}(u_n,\theta)\right)^2$$

Algorytm estymacji otrzymujemy rozwiązując zadanie:

$$\theta_N = \Psi_N(U_N, Y_N) \quad \rightarrow \quad Var_{\omega N}(U_N, Y_N, \theta_N) = \min_{\theta\in\Theta} Var_{\omega N}(U_N, Y_N, \theta)$$

Metoda maksymalnej wiarogodności

Założenia:

ω_n wartość zmiennej losowej $\underline{\omega}$, z funkcją gęstości rozkładu prawdopodobieństwa $f_\omega(\omega)$

Dla zadanego wejścia u_n, wyjście $y_n, n = 1, 2, \ldots, N$ jest mierzone

Sekwencja $y_n, n = 1, 2, \ldots, N$ jest ciągiem wartości zmiennej losowej $\underline{y}$

$$\underline{y} = F(u, \theta, \underline{\omega})$$

Funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{y}$

$$f_y(y, \theta; u) = f_\omega\left(F_\omega^{-1}(u, \theta, y)\right) \cdot |J_F|$$

gdzie J_F jest macierzą Jakobiego

$$J_F = \frac{\partial F_\omega^{-1}(u, \theta, y)}{\partial y}$$

Funkcja wiarogodności:

$$L_N(Y_N, \theta; U_N) = \prod_{n=1}^{N} f_y(y_n, \theta; u_n) = \prod_{n=1}^{N} f_\omega\left(F_\omega^{-1}(u_n, \theta, y_n)\right) |J_F|$$

Estimation algorithm:

$$\theta_N = \Psi_N(U_N, Y_N) \quad \rightarrow \quad L_N(Y_N, \theta_N; U_N) = \max_{\theta \in \Theta} L_N(Y_N, \theta; U_N)$$

Metody Bayes'a

Założenia:

Dodatkowo zakładamy, że θ jest wartością zmiennej losowej $\underline{\theta}$,
o funkcji gęstości rozkładu prawdopodobieństwa $f_{\theta}(\theta)$

$$r\left(\bar{\theta}, Y_N; U_N\right) \overset{df}{=} \underset{\underline{\theta}}{E}\left[L\left(\underline{\theta}, \bar{\theta}\right) \middle| Y_N; U_N\right] = \int_{\Theta} L\left(\theta, \bar{\theta} = \bar{\Psi}\left(U_N, Y_N\right)\right) f'\left(\theta \middle| Y_N; U_N\right) d\theta$$

$$f_y\left(y \middle| \theta; u\right) = f_{\omega}\left(F_{\omega}^{-1}\left(u, \theta, y\right)\right) \left| J_F \right|$$

Gęstość rozkładu prawdopodobieństwa a posteriori :

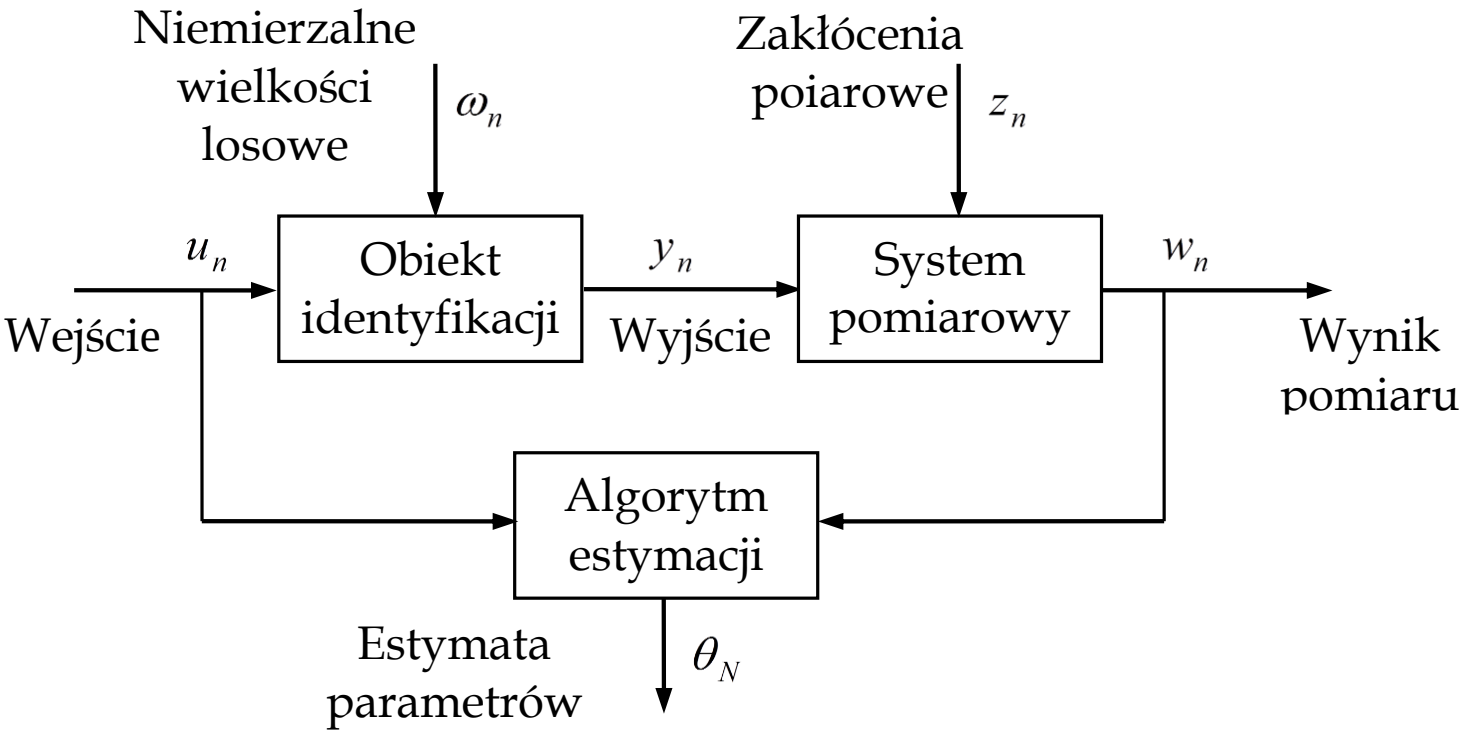
$$f'\left(\theta \middle| Y_N; U_N\right) = \frac{f_{\theta}(\theta) \cdot \prod_{n=1}^{N} f_{\omega}\left(F_{\omega}^{-1}\left(u_n, \theta, y_n\right)\right) \left| J_F \right|}{\int_{\Theta} f_{\theta}(\theta) \cdot \prod_{n=1}^{N} f_{\omega}\left(F_{\omega}^{-1}\left(u_n, \theta, y_n\right)\right) \left| J_F \right| d\theta}$$

Algorytm estymacji:

$$\theta_N = \Psi_N\left(U_N, Y_N\right) \quad \rightarrow \quad r\left(\theta_N, Y_N; U_N\right) = \min_{\bar{\theta} \in \Theta} r\left(\bar{\theta}, Y_N; U_N\right)$$

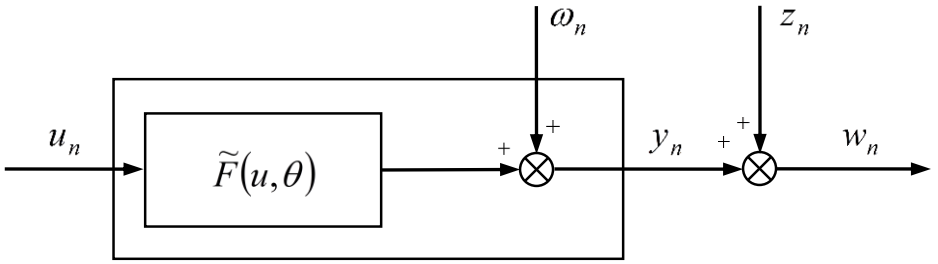
Estymacja parametrów charakterystyki statycznej obiektu

Niemierzalny losowy parametr w obiekcie oraz zakłócony pomiar wyjścia



Metoda najmniejszych kwadratów

Założenia:



$\underline{z}$, $\underline{\omega}$ – zmienne losowe, takie że $\underset{\underline{\omega}}{E}[\underline{\omega}]=0,\quad \underset{\underline{\omega}}{Var}[\underline{\omega}]<\infty,\quad \underset{\underline{z}}{E}[\underline{z}]=0,\quad \underset{\underline{z}}{Var}[\underline{z}]<\infty$

Wyniki pomiarów są wartościami zmiennej losowej: $\underline{w}=\widetilde{F}(u,\theta)+\underline{\omega}+\underline{z}$

$$\underset{\underline{\omega}+\underline{z}}{E}[\underline{\omega}+\underline{z}]=0,\qquad \underset{\underline{\omega}+\underline{z}}{Var}[\underline{\omega}+\underline{z}]<\infty$$

Empiryczna wariancja:

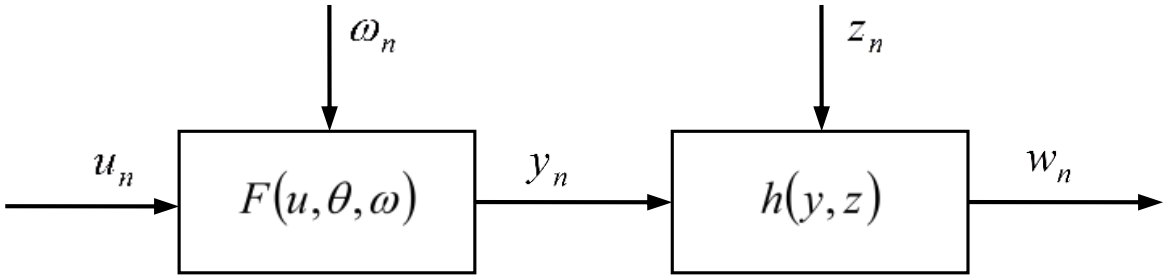
$$Var_{(\omega+z)N}(U_N,W_N,\theta)=\frac{1}{N}\sum_{n=1}^{N}\left(w_n-\widetilde{F}(u_n,\theta)\right)^2$$

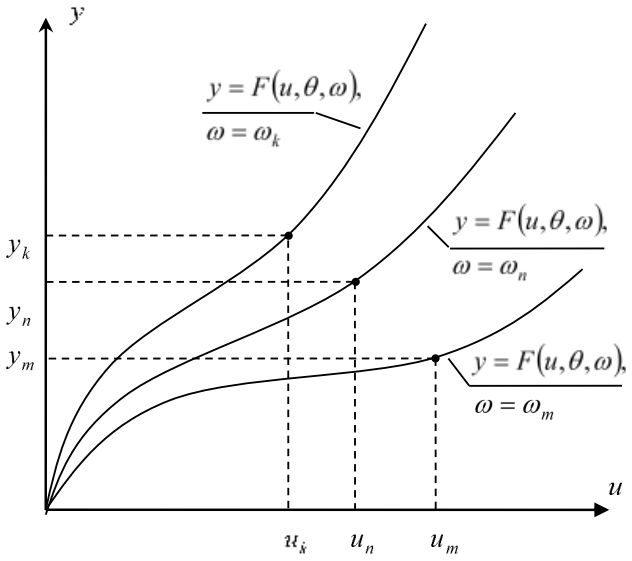
Algorytm estymacji otrzymujemy w wyniku rozwiązania zadania:

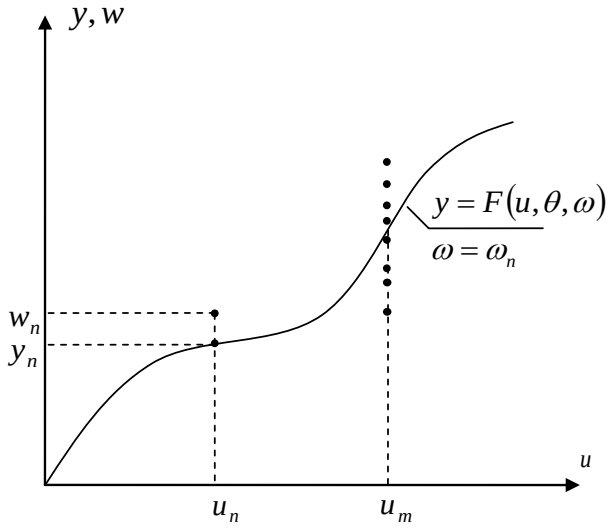
$$\theta_N=\Psi_N(U_N,Y_N)\quad\rightarrow\quad Var_{(\omega+z)N}(U_N,W_N,\theta_N)=\min_{\theta\in\Theta}Var_{(\omega+z)N}(U_N,W_N,\theta)$$

Pomiar wyjścia z zakłóceniami oraz losowo zmienny parametr obiektu

Pomiar wyjścia z zakłóceniami oraz losowo zmienny parametr obiektu







Metoda maksymalnej wiarogodności

Założenia:

ω_n jest wartością zmiennej losowej $\underline{\omega}$, o funkcji gęstości rozkładu prawdopodobieństwa $f_{\omega}(\omega)$

z_n jest wartością zmiennej losowej $\underline{z}$, o funkcji gęstości rozkładu prawdopodobieństwa $f_z(z)$

$w_n, n = 1, 2, \ldots, N$ są wartościami zmiennej losowej $\underline{w} = h(\underline{y}, \underline{z})$

Dla zadanego u_n, $y_n, n = 1, 2, \ldots, N$ są wartościami zmiennej losowej $\underline{y} = F(u, \theta, \underline{\omega})$

Funkcja h nie jest wzajemnie jednoznaczna względem $(\underline{y}, \underline{z})$, dlatego dodajemy $\underline{y} = \underline{y}$ i w konsekwencji otrzymujemy:

$$\underline{w} = h(\underline{y}, \underline{z})$$

$$\underline{y} = \underline{y}$$

Przekształcenie odwrotne:

$$\underline{z} = h_z^{-1}(\underline{y}, \underline{w})$$

$$\underline{y} = \underline{y}$$

Metoda maksymalnej wiarogodności

Łączna gęstość rozkładu prawdopodobieństwa pary zmiennych *(w,y)*

$$f_{wy}(w,y,\theta;u) = f_z\left(h_z^{-1}(y,w)\right)\left|J_h\right| f_y(y,\theta;u) = f_z\left(h_z^{-1}(y,w)\right)\left|J_h\right| f_\omega\left(F_\omega^{-1}(u,\theta,y)\right)\left|J_F\right|$$

Funkcja gęstości prawdopodobieństwa obserwowanej zmiennej w

$$f_w(w,\theta;u) = \int_{\mathscr{Y}} f_{wy}(w,y,\theta;u)dy = \int_{\mathscr{Y}} f_z\left(h_z^{-1}(y,w)\right) f_\omega\left(F_\omega^{-1}(u,\theta,y)\right)\left|J_F\right|\left|J_h\right|dy$$

Funkcja wiarogodności:

$$L_N(W_N,\theta;U_N) = \prod_{n=1}^{N} f_w(w_n,\theta;u_n) =$$

$$= \prod_{n=1}^{N} \int_{\mathscr{Y}} f_z\left(h_z^{-1}(y,w_n)\right) f_\omega\left(F_\omega^{-1}(u_n,\theta,y)\right)\left|J_F\right|\left|J_h\right|dy$$

Algorytm estymacji:

$$\theta_N = \Psi_N(U_N,Y_N) \quad \rightarrow \quad L_N(W_N,\theta_N;U_N) = \max_{\theta\in\Theta} L_N(W_N,\theta;U_N)$$

Metody Bayes'a

Założenia:

Dodatkowo zakładamy, że θ Jest wartością zmiennej losowej $\underline{\theta}$, o rozkładzie $f_\theta(\theta)$

Dla ustalonego parametru θ oraz wejścia u_n wartości wyjść $y_n, n=1, 2, \ldots, N$ są wartościami zmiennej losowej $\underline{y}$ pod warunkiem, że $\underline{\theta}=\theta$ oraz wejście jest równe u.

$w_n, n=1, 2, \ldots, N$ są wartościami zmiennej losowej $\underline{w}$ pod warunkiem, że $\underline{\theta}=\theta$

z_n jest wartością zmiennej losowej $\underline{z}$, o funkcji gęstości prawdopodobieństwa $f_z(z)$

W konsekwencji

$$\underline{w}=h(\underline{y},\underline{z})$$

$$f_w(w|\theta;u)=\int_{\mathscr{Y}} f_z\left(h_z^{-1}(y,w)\right)f_y(y|\theta;u)|J|dy=\int_{\mathscr{Y}} f_z\left(h_z^{-1}(y,w)\right)f_\omega\left(F_\omega^{-1}(u,\theta,y)\right)|J_F||J_h|dy$$

Metody Bayes'a

Funkcja gęstości rozkładu prawdopodobieństwa a posteriori:

$$f'(\theta|W_N;U_N)=\frac{f_\theta(\theta)\prod_{n=1}^{N}\int_{\mathcal{Y}} f_z\left(h_z^{-1}(y,w_n)\right)f_\omega\left(F_\omega^{-1}(u_n,\theta,y)\right)|J_F\|J_h|dy}{\int_\Theta f_\theta(\theta)\prod_{n=1}^{N}\int_{\mathcal{Y}} f_z\left(h_z^{-1}(y,w_n)\right)f_\omega\left(F_\omega^{-1}(u_n,\theta,y)\right)|J_F\|J_h|dy\,d\theta}$$

Ryzyko warunkowe:

$$r(\bar{\theta},W_N;U_N)\overset{\text{df}}{=}\underset{\underline{\theta}}{E}\left[L(\underline{\theta},\bar{\theta})|W_N;U_N\right]=\int_\Theta L(\theta,\bar{\theta}=\bar{\Psi}(U_N,W_N))\,f'(\theta|W_N;U_N)\,d\theta$$

Algorytm estymacji otrzymujemy rozwiązując zadanie:

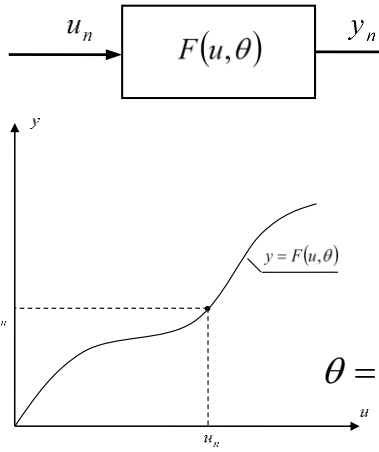
$$\theta_N=\Psi_N(U_N,W_N)\quad\rightarrow\quad r(\theta_N,W_N;U_N)=\min_{\bar{\theta}\in\Theta} r(\bar{\theta},W_N;U_N)$$

Wybór Optymalnego Modelu Przypadek Deterministyczny

Podstawowe zadania identyfikacji podsumowanie

Deterministyczny

Obiekt w klasie modeli



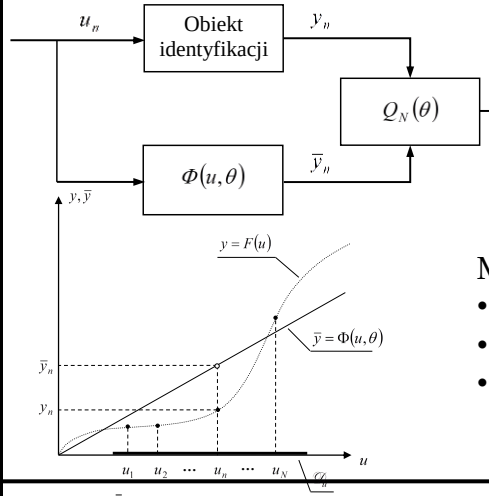
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

$$y_n = F(u_n, \theta),$$
$$n = 1, 2, \ldots, N$$

$$Y_N = \bar{F}(U_N, \theta)$$

$$\theta = \bar{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

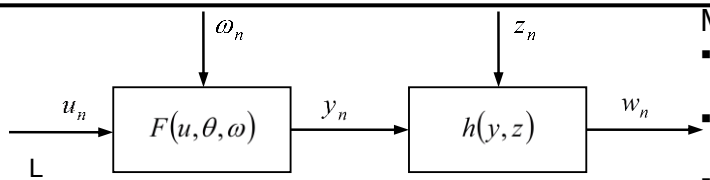
Wybór optymalnego modelu



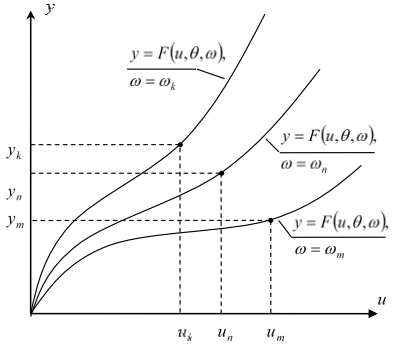
$$\bar{y} = \Phi(u, \theta_N^*)$$
$$\theta_N^* \rightarrow Q_N(\theta_N^*) = \min_{\theta} Q_N(\theta)$$
$$\theta_N^* = \Psi_N(U_N, Y_N)$$
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

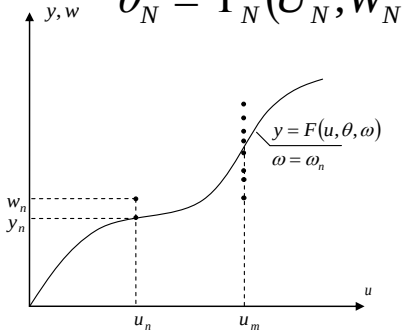
Model jest optymalny:
- dla zadanej serii pomiarowej
- przyjętego modelu
- przyjętego wskaźnika jakości identyfikacji

Losowy

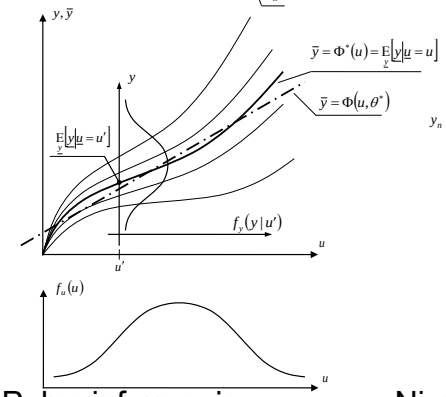


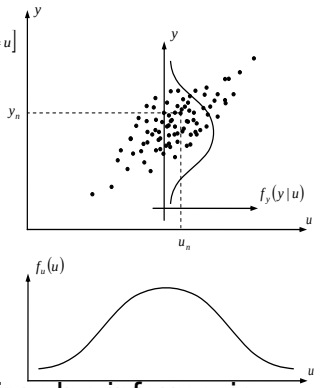
Metody:
- Najmniejszych kwadratów
- Maksymalnej wiarogodności
- Bayes'a





$$\theta_N = \Psi_N(U_N, W_N)$$

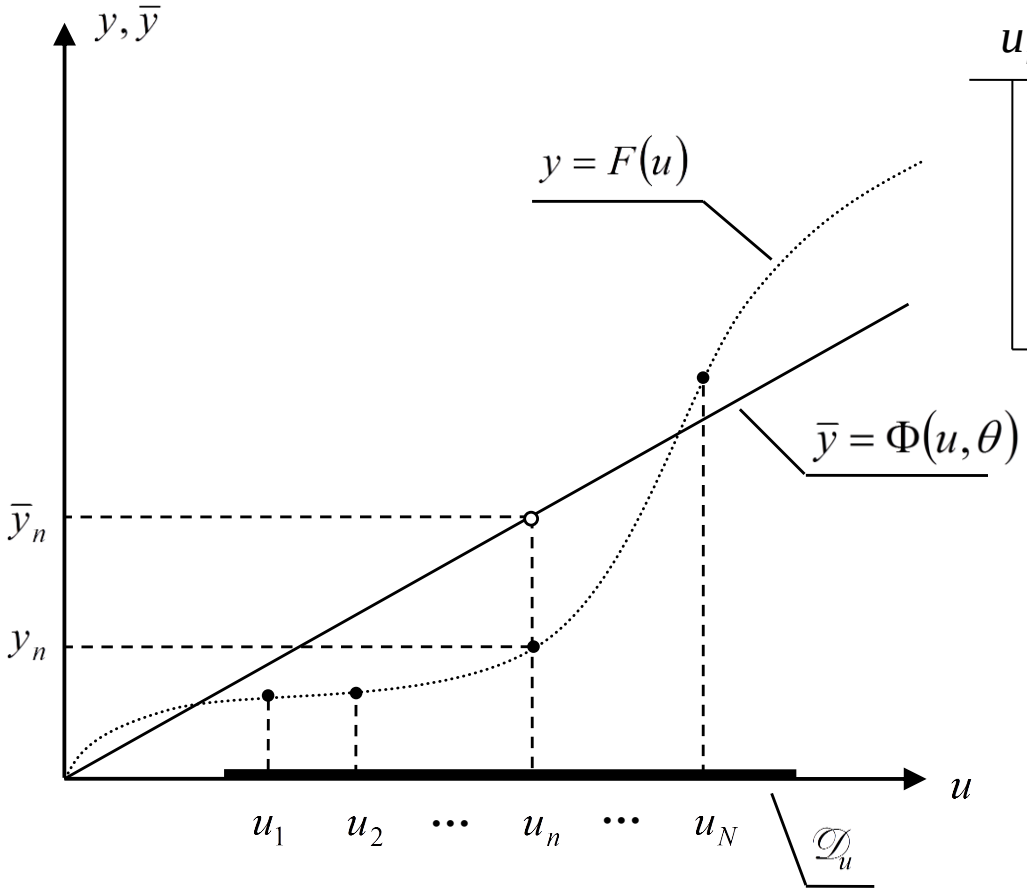


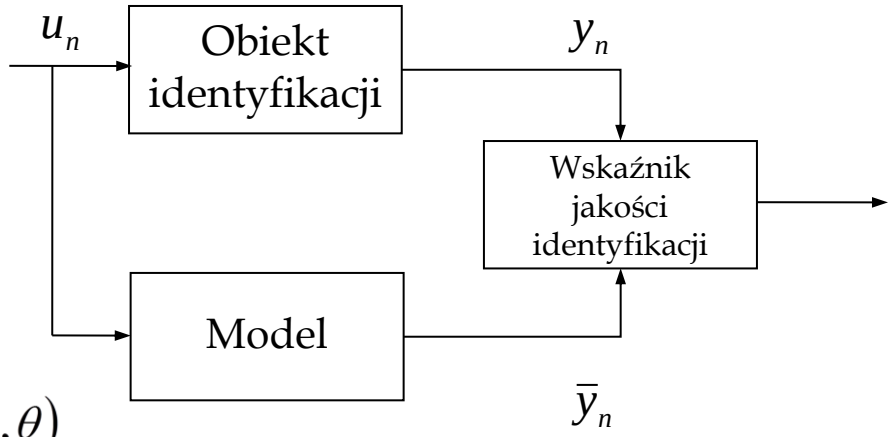


Pełna informacja
- Regresja I rodzaju
- Regresja II rodzaju

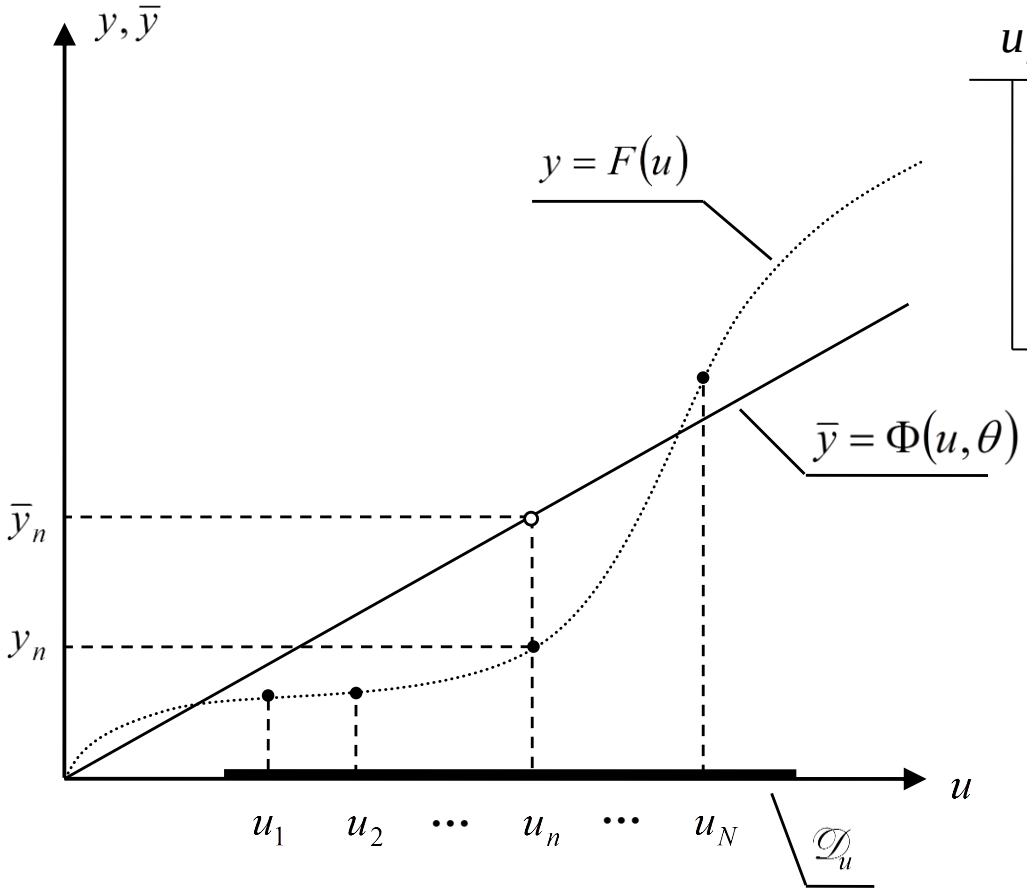
Niepełna informacja
- Estymacja wskaźnika jakości
- Estymacja parametrów rozkładu
- Estymacja rozkładu

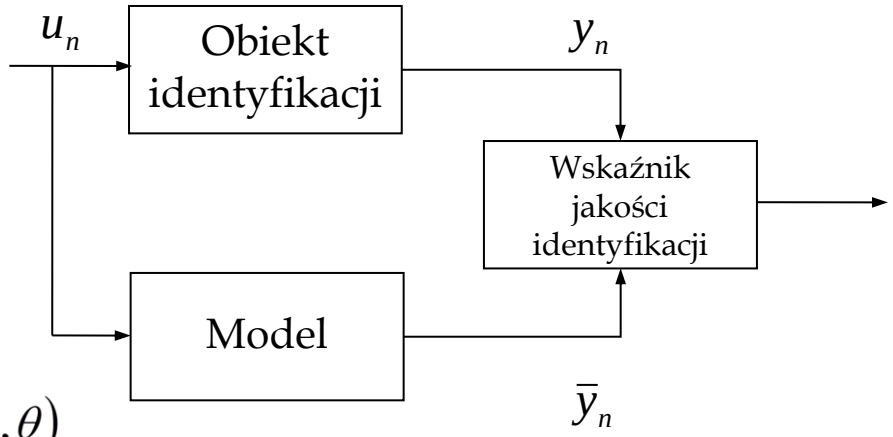
Wybór optymalnego modelu - Przypadek deterministyczny





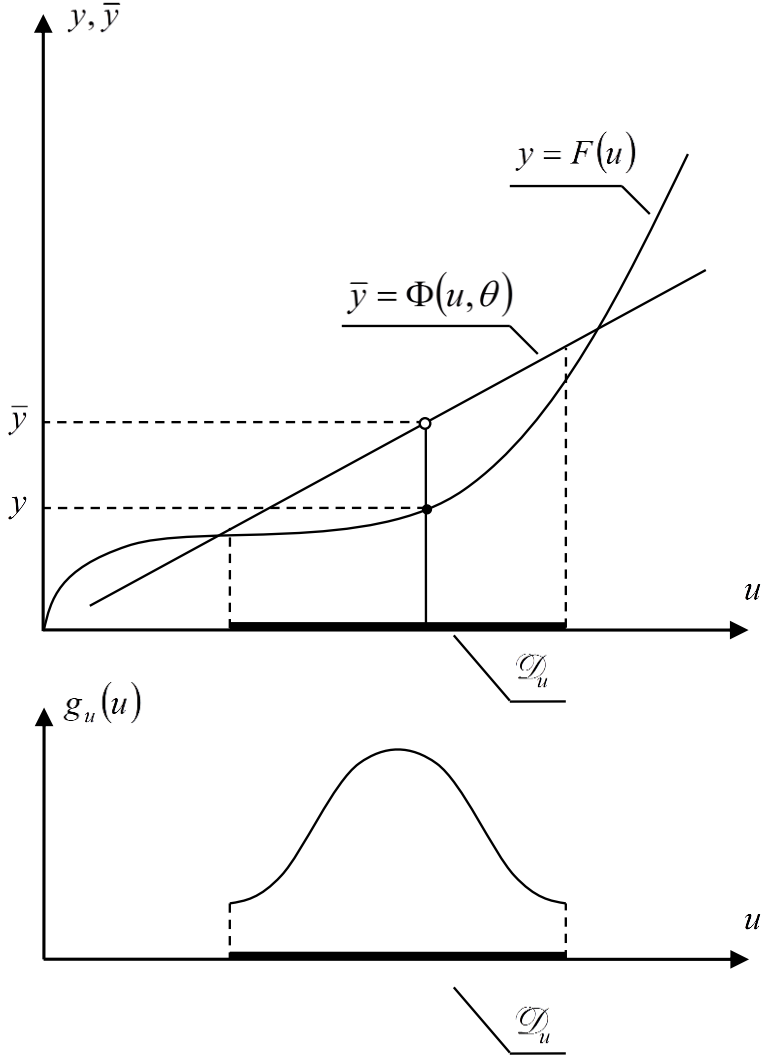
Wybór optymalnego modelu - Przypadek deterministyczny





Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu



Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu

Carakterystyka statczna obiektu: $y = F(u)$

F – znana funkcja

u – wektor wejść $u \in \mathcal{U} \subseteq \mathcal{R}^S$ $\mathcal{U}$ – przestrzeń wejść

y – wektor wyjść $y \in \mathcal{Y} \subseteq \mathcal{R}^L$ $\mathcal{Y}$ – przestrzeń wyjść

Funkcja aproksymująca (model): $\bar{y} = \Phi(u, \theta)$

Φ – arbitralnie zadana funkcja

$\bar{y}$ – wektor wyjść modelu $\bar{y} \in \overline{\mathcal{Y}} \subseteq \mathcal{R}^L$

θ – wektor parametrów modelu $\theta \in \Theta \subseteq \mathcal{R}^R$

$\mathcal{D}_u \subseteq \mathcal{U}$ – podzbiór przestrzeni wejść $g_u(u)$ - funkcja wagi

Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu

Miara różnicy pomiędzy wartością funkcji aproksymowanej i aproksymującej w punkcie u: $\forall u \in \mathscr{D}_u \quad q(y,\bar{y}) = q(F(u), \Phi(u,\theta))$

gdzie np.: $q(y,\bar{y}) = \sum_{l=1}^{L} \left(y^{(l)} - \bar{y}^{(l)}\right)^2$ $q(y,\bar{y}) = \sum_{l=1}^{L} \left|y^{(l)} - \bar{y}^{(l)}\right|$ L – wymiarowe wyjście

$q(y,\bar{y}) = (y-\bar{y})^2 = (F(u) - \Phi(u,\theta))^2$

$q(y,\bar{y}) = |y-\bar{y}| = |F(u) - \Phi(u,\theta)|$

jednowymiarowe wyjście $(L=1)$

Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu

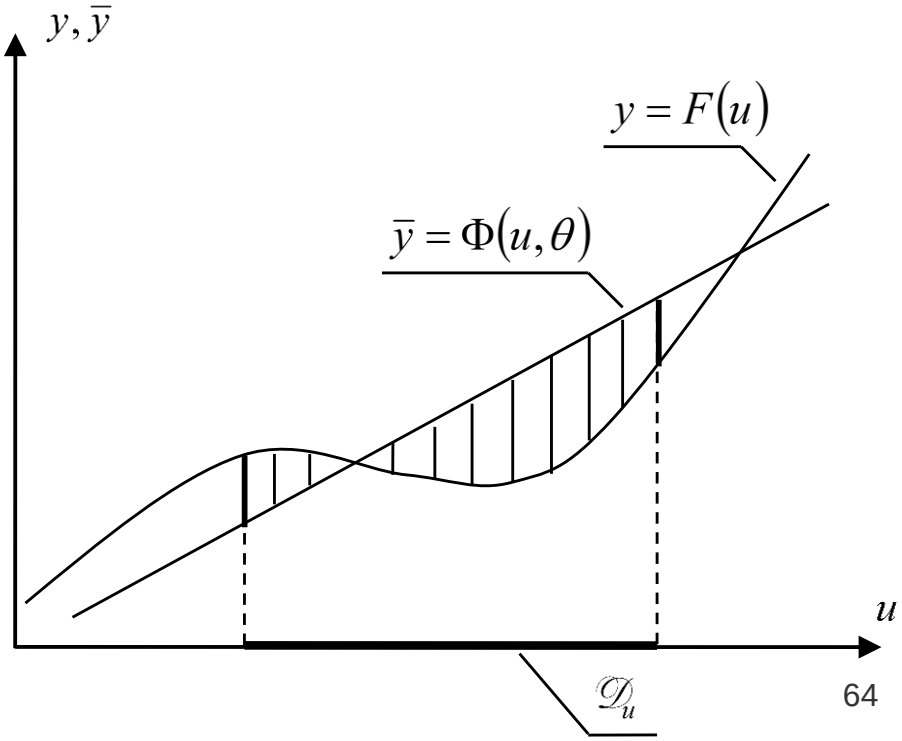
Wskaźnik jakości aproksymacji – miara różnicy pomiędzy funkcją Aproksymowaną $F(u)$ i aproksymującą $\Phi(u,\theta)$ w obszarze $\mathcal{D}_u$

$$Q(\theta) = \left\| F(u) - \Phi(u,\theta) \right\|_{\mathcal{D}_u}$$

$$Q(\theta) = \int_{\mathcal{D}_u} q\left(F(u), \Phi(u,\theta)\right) g_u(u)\, du$$

np. :

$$Q(\theta) = \int_{\mathcal{D}_u} \left| F(u) - \Phi(u,\theta) \right| g_u(u)\, du$$



Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu

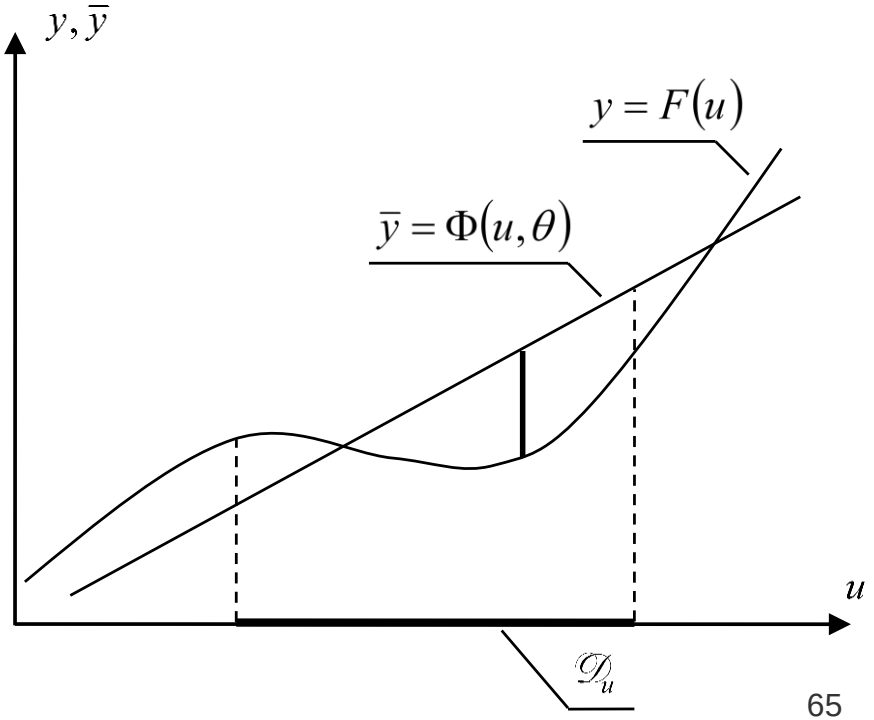
Wskaźnik jakości aproksymacji – miara różnicy pomiędzy funkcją Aproksymowaną $F(u)$ i aproksymującą $\Phi(u,\theta)$ w obszarze $\mathcal{D}_u$

$$Q(\theta) = \left\| F(u) - \Phi(u,\theta) \right\|_{\mathcal{D}_u}$$

$$Q(\theta) = \max_{u \in \mathcal{D}_u} \left\{ q\left(F(u), \Phi(u,\theta)\right) g_u(u) \right\}$$

np.:

$$Q(\theta) = \max_{u \in \mathcal{D}_u} \left\{ \left| F(u) - \Phi(u,\theta) \right| g_u(u) \right\}$$

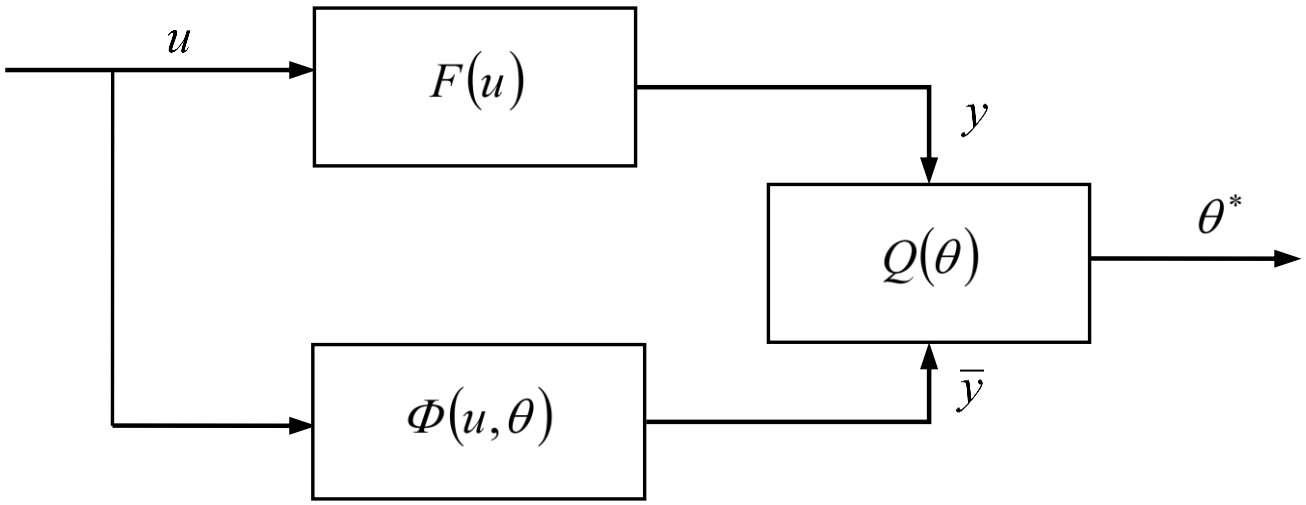


Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu

Optymalny model: $\bar{y} = \Phi(u, \theta^*)$

gdzie: θ^* – optymalny wektor parametrów modelu: $\theta^* \rightarrow \quad Q(\theta^*) = \min_{\theta \in \Theta} Q(\theta)$



Przypadek liniowo - kwadratowy

Model: $\bar{y} = \Phi(u,\theta) = \theta^T \varphi(u)$

gdzie $\theta \overset{df}{=} \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \\ \vdots \\ \theta^{(R)} \end{bmatrix}$ $\varphi(u) \overset{\text{df}}{=} \begin{bmatrix} \varphi_1(u) \\ \varphi_2(u) \\ \vdots \\ \varphi_R(u) \end{bmatrix}$ np. : $\varphi(u) = \begin{bmatrix} 1 \\ u \\ \vdots \\ u^{R-1} \end{bmatrix}$

Wskaźnik jakości aproksymacji:

$$Q(\theta) = \int_{\mathcal{D}_u} \left(F(u) - \Phi(u,\theta)\right)^2 g_u(u)\, du$$

czyli:

$$Q(\theta) = \int_{\mathcal{D}_u} \left(F(u) - \theta^T \varphi(u)\right)^2 g_u(u)\, du$$

Przypadek liniowo - kwadratowy

Warunek konieczny optymalności wektora θ :

$$\operatorname*{grad}_{\theta} Q(\theta) = -2\left[\int_{\mathcal{D}_u} \left(F(u) - \theta^{*T}\varphi(u)\right)\varphi(u) g_u(u)\, du\right] = 0_R, \qquad 0_R = \left.\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right\} R$$

Rozwiązanie powyższego równania jest następujące:

$$\theta^* = \left[\int_{\mathcal{D}_u} \varphi(u)\,\varphi^T(u)\, g_u(u)\, du\right]^{-1} \int_{\mathcal{D}_u} F(u)\,\varphi(u) g_u(u)\, du$$

Przypadek liniowo - kwadratowy

Zauważmy że w rozwiązaniu występuje macierz:

$$\int_{\mathcal{D}_u} \varphi(u)\,\varphi^T(u)\,g_u(u)\,du =$$

$$\begin{bmatrix} \int_{\mathcal{D}_u} \varphi_1(u)\,\varphi_1(u) g_u(u)\,du & \int_{\mathcal{D}_u} \varphi_1(u)\,\varphi_2(u) g_u(u)\,du & \cdots & \int_{\mathcal{D}_u} \varphi_1(u)\,\varphi_R(u) g_u(u)\,du \\ \int_{\mathcal{D}_u} \varphi_2(u)\,\varphi_1(u) g_u(u)\,du & \int_{\mathcal{D}_u} \varphi_2(u)\,\varphi_2(u) g_u(u)\,du & \cdots & \int_{\mathcal{D}_u} \varphi_2(u)\,\varphi_R(u) g_u(u)\,du \\ \vdots & \vdots & \ddots & \vdots \\ \int_{\mathcal{D}_u} \varphi_R(u)\,\varphi_1(u) g_u(u)\,du & \int_{\mathcal{D}_u} \varphi_R(u)\,\varphi_2(u) g_u(u)\,du & \cdots & \int_{\mathcal{D}_u} \varphi_R(u)\,\varphi_R(u) g_u(u)\,du \end{bmatrix}$$

Przypadek liniowo - kwadratowy

Jeżeli wybierzemy funkcje $\varphi_i(u)$ ortonormalne czyli takie że:

$$\int_{\mathcal{D}_u} \varphi_i(u)\,\varphi_j(u)\,g_u(u)\,du = \begin{cases} 1 & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j \end{cases}, \quad i, j = 1, 2, \ldots, R$$

Wówczas macierz : $\int_{\mathcal{D}_u} \varphi(u)\,\varphi^T(u)\,g_u(u)\,du = \mathbf{I}$ jest macierzą jednostkową

i algorytm aproksymacji upraszcza się do postaci:

$$\theta^* = \int_{\mathcal{D}_u} F(u)\,\varphi(u) g_u(u)\,du$$

Przypadek liniowo – kwadratowy
Przykład

Charakterystyka statyczna obiektu (funkcja aproksymowana):

$$y = F(u) = u^2$$

Model (funkcja aproksymująca):

$$\bar{y} = \Phi(u,\theta) = \theta^T \varphi(u) = \theta^{(1)} + \theta^{(2)} u \qquad \text{gdzie:} \quad \varphi(u) = \begin{bmatrix} 1 \\ u \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \end{bmatrix}$$

Podzbiór przestrzeni wejść (odcinek):

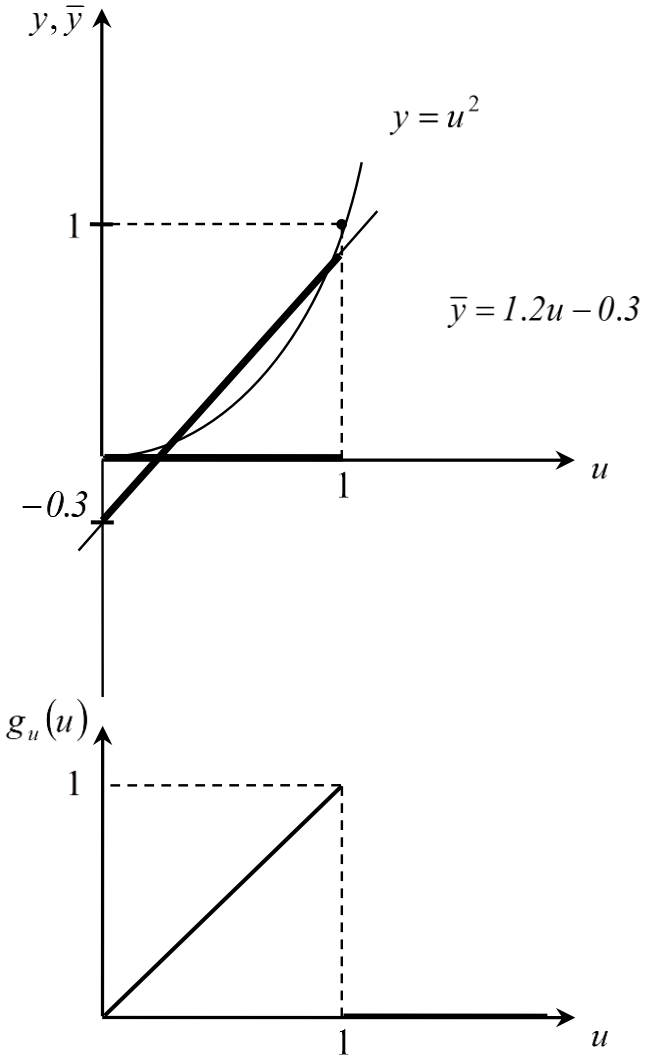
$$\mathcal{D}_u = \{u \in \mathcal{R} : 0 \le u \le 1\}$$

Funkcja wagi: $$g_u(u) = \begin{cases} u & \text{for} \quad u \in [0,1] \\ 0 & \text{for} \quad u \notin [0,1] \end{cases}$$

Wskaźnik jakości aproksymacji: $$Q(\theta) = \int_{\mathcal{D}_u} \left(F(u) - \theta^T \varphi(u)\right)^2 g_u(u)\, du$$

Przypadek liniowo – kwadratowy
Przykład

Interpretacja graficzna:



Przypadek liniowo – kwadratowy
Przykład

Wskaźnik jakości aproksymacji przyjmuje postać:

$$Q(\theta)=\int_{\mathcal{D}_u}\left(F(u)-\theta^T\varphi(u)\right)^2\, g_u(u)\,du=\int_0^1\left(u^2-\left[\theta^{(1)}\ \theta^{(2)}\right]\begin{bmatrix}1\\u\end{bmatrix}\right)^2 u\ du$$

Warunek konieczny optymalności jest następujący:

$$\operatorname*{grad}_{\theta} Q(\theta)=-2\left[\int_0^1\left(u^2-\left[\theta^{(1)*}\ \theta^{(2)*}\right]\begin{bmatrix}1\\u\end{bmatrix}\right)\begin{bmatrix}1\\u\end{bmatrix}u\ du\right]=0_2$$

a rozwiązanie ma postać:

$$\theta^*=\left[\int_0^1\begin{bmatrix}1\\u\end{bmatrix}\begin{bmatrix}1&u\end{bmatrix}u\ du\right]^{-1}\int_0^1 u^2\begin{bmatrix}1\\u\end{bmatrix}u\ du=\begin{bmatrix}\int_0^1 u\ du & \int_0^1 u^2\ du\\ \int_0^1 u^2\ du & \int_0^1 u^3\ du\end{bmatrix}^{-1}\begin{bmatrix}\int_0^1 u^3\ du\\ \int_0^1 u^4\ du\end{bmatrix}$$

Przypadek liniowo – kwadratowy
Przykład

Po prostych przekształceniach otrzymujemy:

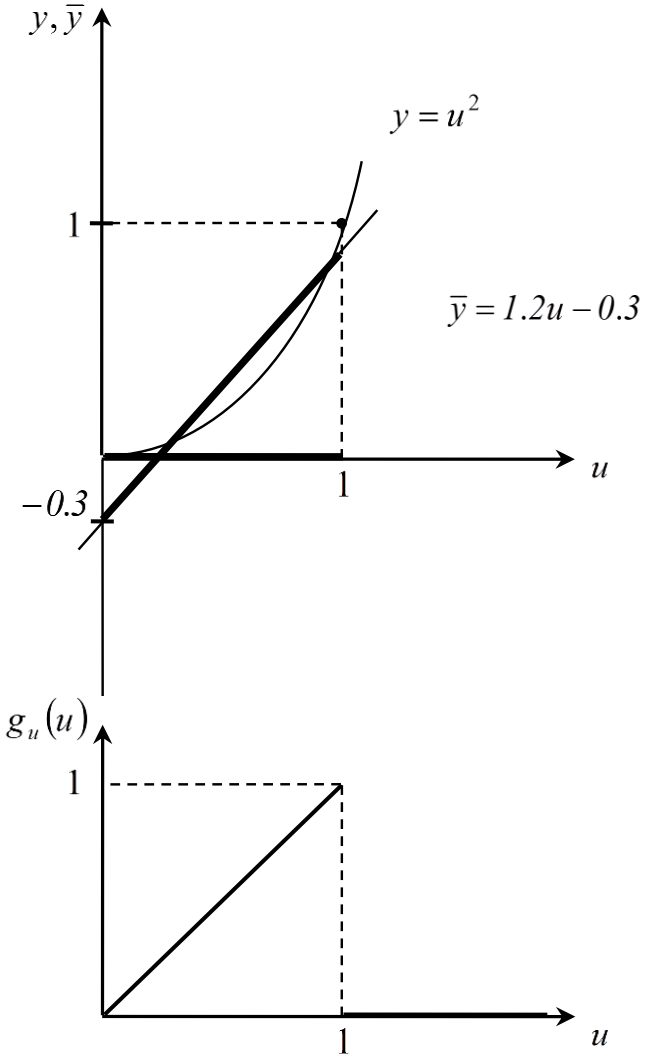
$$\theta^* = \begin{bmatrix} \theta^{(1)*} \\ \theta^{(2)*} \end{bmatrix} = \begin{bmatrix} \dfrac{\int_0^1 u^3\,du \times \int_0^1 u^3\,du - \int_0^1 u^4\,du \times \int_0^1 u^2\,du}{\int_0^1 u\,du \times \int_0^1 u^3\,du - \left(\int_0^1 u^2 du\right)^2} \\ \dfrac{\int_0^1 u^4\,du \times \int_0^1 u\,du - \int_0^1 u^3\,du \times \int_0^1 u^2\,du}{\int_0^1 u\,du \times \int_0^1 u^3\,du - \left(\int_0^1 u^2 du\right)^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\frac{1}{4}u^4\Big|_0^1 \times \frac{1}{4}u^4\Big|_0^1 - \frac{1}{5}u^5\Big|_0^1 \times \frac{1}{3}u^3\Big|_0^1}{\frac{1}{2}u^2\Big|_0^1 \times \frac{1}{4}u^4\Big|_0^1 - \left(\frac{1}{3}u^3\Big|_0^1\right)^2} \\ \dfrac{\frac{1}{5}u^5\Big|_0^1 \times \frac{1}{2}u^2\Big|_0^1 - \frac{1}{4}u^4\Big|_0^1 \times \frac{1}{3}u^3\Big|_0^1}{\frac{1}{2}u^2\Big|_0^1 \times \frac{1}{4}u^4\Big|_0^1 - \left(\frac{1}{3}u^3\Big|_0^1\right)^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\frac{1}{4}\times\frac{1}{4} - \frac{1}{5}\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{4} - \left(\frac{1}{3}\right)^2} \\ \dfrac{\frac{1}{5}\times\frac{1}{2} - \frac{1}{4}\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{4} - \left(\frac{1}{3}\right)^2} \end{bmatrix} = \begin{bmatrix} -0.3 \\ 1.2 \end{bmatrix}$$

Przypadek liniowo – kwadratowy
Przykład

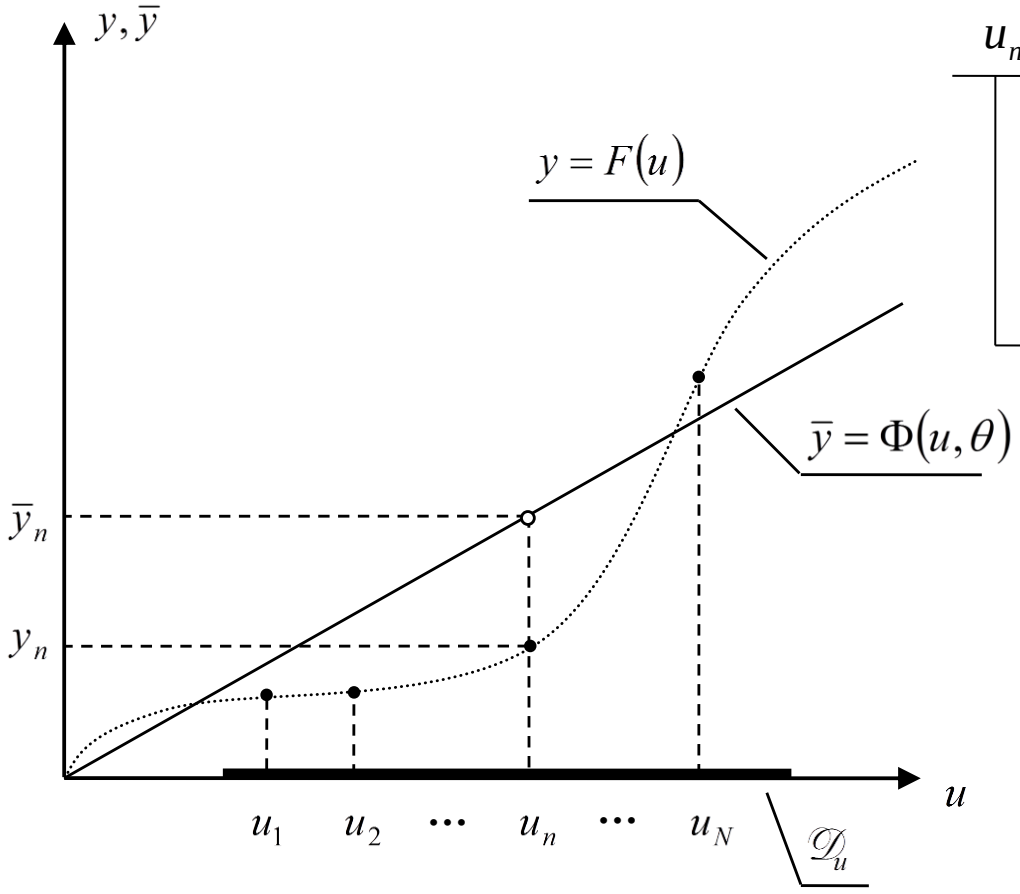
Model optymalny:

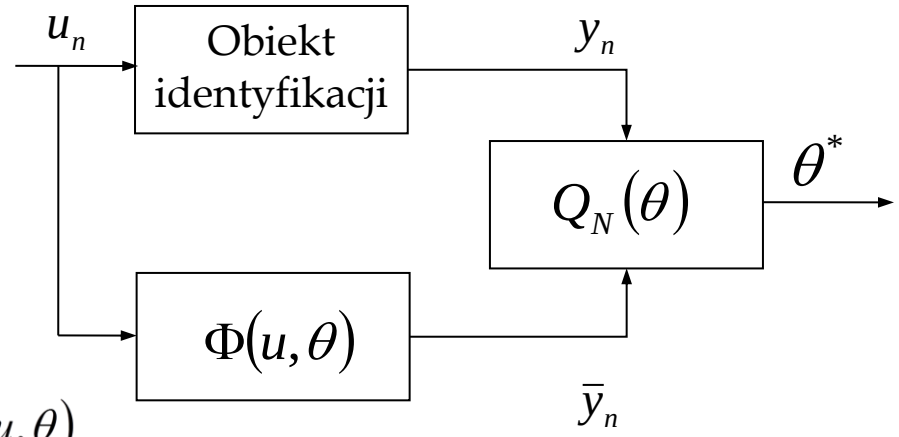
$$\bar{y} = 1.2\,u - 0.3$$

Interpretacja graficzna:



Wybór optymalnego modelu - Przypadek deterministyczny





Wybór optymalnego modelu - Przypadek deterministyczny

- Sformułowanie problemu

Eksperyment: $U_N = [u_1 \quad u_2 \quad \cdots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$

Funkcja aproksymująca (model): $\bar{y} = \Phi(u, \theta)$

Miara różnicy pomiędzy wartością zmierzoną wyjścia obiektu i wartością wyznaczoną z modelu dla tej samej wartości wejścia, tj. dla każdego pomiaru:

$$\forall n = 1, 2, \ldots, N \qquad q(y_n, \bar{y}_n) = q(y_n, \Phi(u_n, \theta))$$

Wybór optymalnego modelu - Przypadek deterministyczny

- Sformułowanie problemu

Wskaźnik jakości identyfikacji $Q_N(\theta) = \left\| Y_N - \bar{Y}_N(\theta) \right\|_{U_N}$

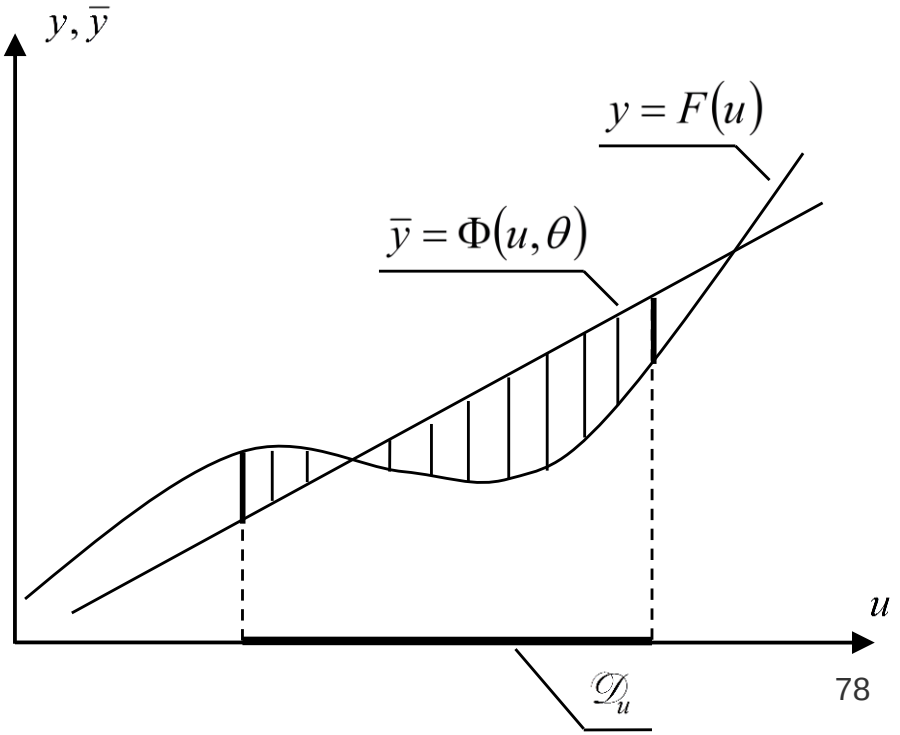
gdzie: $\bar{Y}_N(\theta) \overset{df}{=} \left[\Phi(u_1, \theta) \quad \Phi(u_2, \theta) \quad \cdots \quad \Phi(u_N, \theta) \right]$

$$Q_N(\theta) = \sum_{n=1}^{N} \alpha_n q(y_n, \bar{y}_n) = \sum_{n=1}^{N} \alpha_n q(y_n, \Phi(u_n, \theta))$$

np. :

$$Q_N(\theta) = \sum_{n=1}^{N} |y_n - \bar{y}_n| = \sum_{n=1}^{N} |y_n - \Phi(u_n, \theta)|$$

$$Q_N(\theta) = \sum_{n=1}^{N} (y_n - \bar{y}_n)^2 = \sum_{n=1}^{N} (y_n - \Phi(u_n, \theta))^2$$



Wybór optymalnego modelu - Przypadek deterministyczny

- Sformułowanie problemu

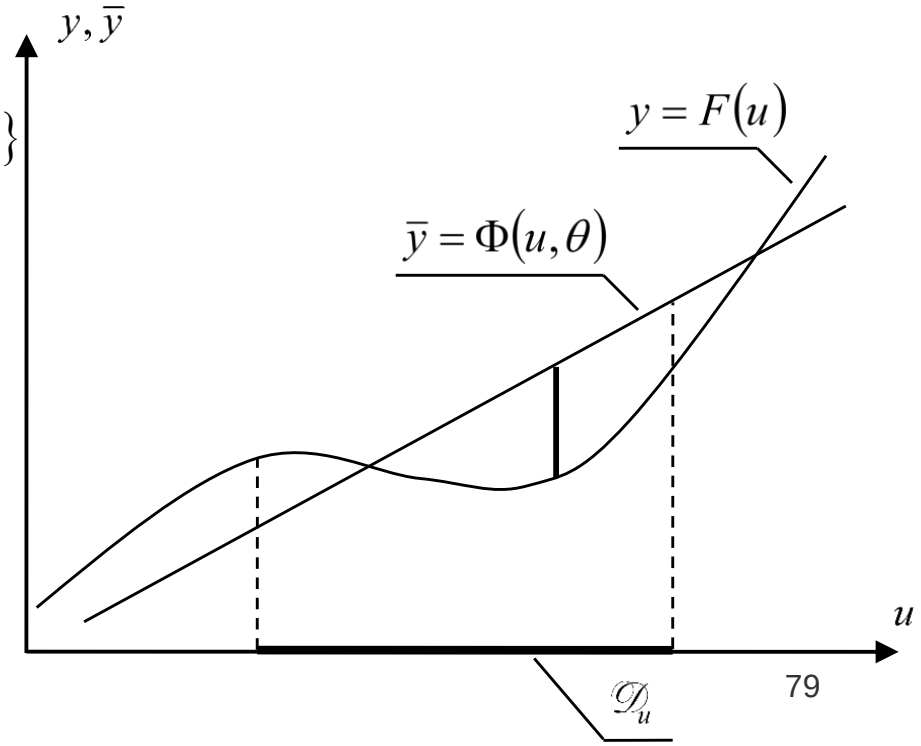
Wskaźnik jakości identyfikacji $Q_N(\theta) = \left\| Y_N - \bar{Y}_N(\theta) \right\|_{U_N}$

gdzie: $\bar{Y}_N(\theta) \overset{df}{=} \left[\Phi(u_1, \theta) \quad \Phi(u_2, \theta) \quad \cdots \quad \Phi(u_N, \theta) \right]$

$$Q_N(\theta) = \max_{1 \le n \le N} \{ \alpha_n q(y_n, \bar{y}_n) \} = \max_{1 \le n \le N} \{ \alpha_n q(y_n, \Phi(u_n, \theta)) \}$$

np. :

$$Q_N(\theta) = \max_{1 \le n \le N} \{ |y_n - \bar{y}_n| \} = \max_{1 \le n \le N} \{ |y_n - \Phi(u_n, \theta)| \}$$

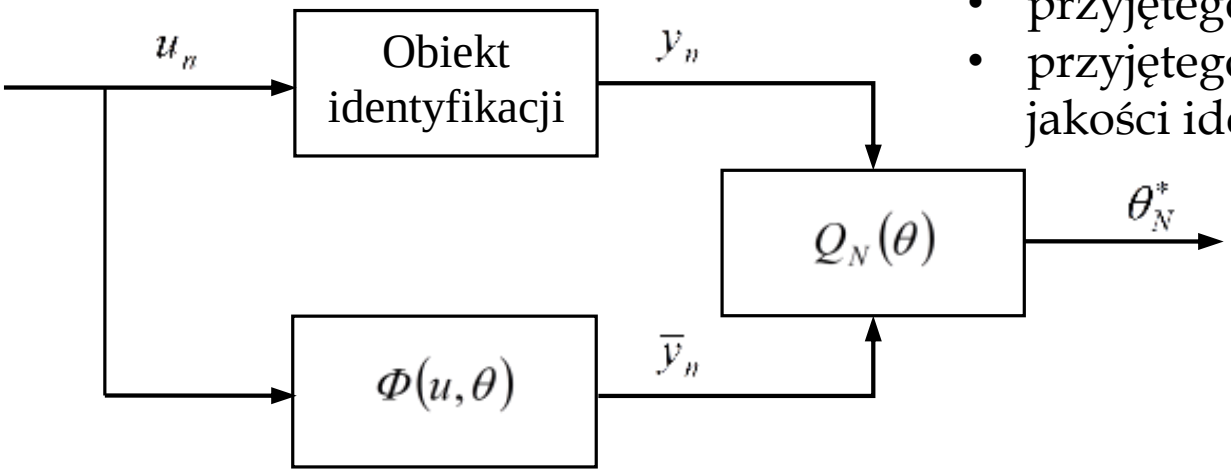


Wybór optymalnego modelu - Przypadek deterministyczny

- Sformułowanie problemu

Model optymalny: $\bar{y} = \Phi(u, \theta_N^*)$ $\theta_N^* \to Q_N(\theta_N^*) = \min_\theta Q_N(\theta)$

Model jest optymalny:
- dla zadanej serii pomiarowej
- przyjętego modelu
- przyjętego wskaźnika jakości identyfikacji



Zadanie aproksymacji

$$y = F(u) \qquad \bar{y} = \Phi(u, \theta^*)$$

gdzie: θ^* – optymalny wektor parametrów: $\theta^* \rightarrow \quad Q(\theta^*) = \min_{\theta \in \Theta} Q(\theta)$

$$Q(\theta) = \int_{\mathcal{D}_u} q(F(u), \Phi(u, \theta))\, g_u(u)\, du$$

Zadanie identyfikacji

$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N] \qquad \bar{y} = \Phi(u, \theta_N^*)$$

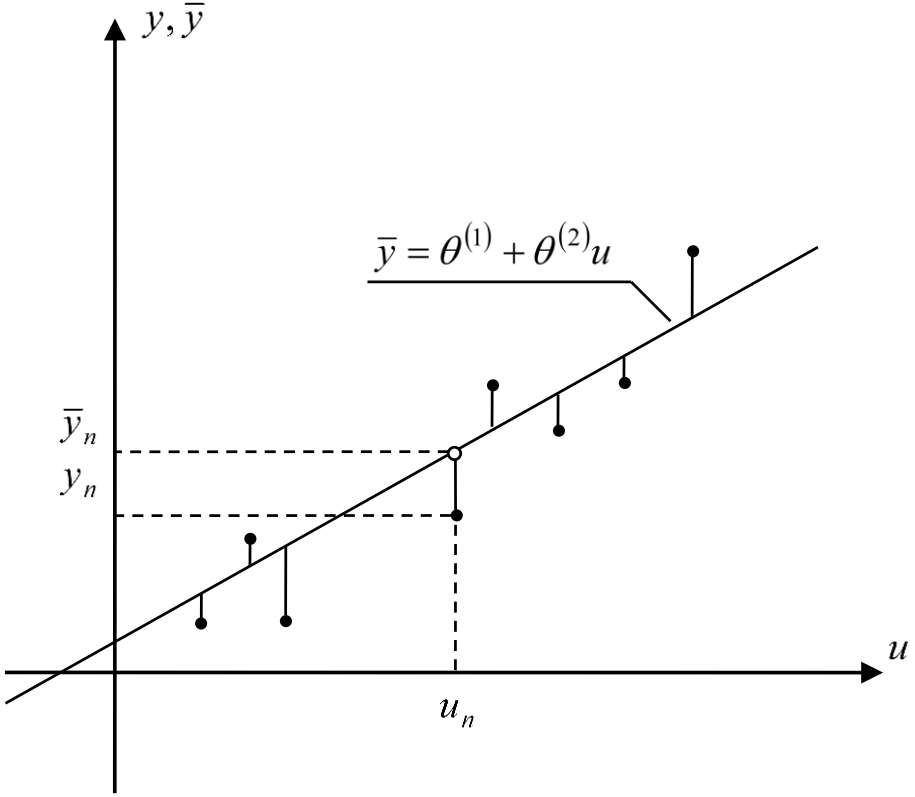
gdzie: θ_N^* – optymalny wektor parametrów: $\theta_N^* \rightarrow Q_N(\theta_N^*) = \min_{\theta} Q_N(\theta)$

$$Q_N(\theta) = \sum_{n=1}^{N} \alpha_n q(y_n, \bar{y}_n) = \sum_{n=1}^{N} \alpha_n q(y_n, \Phi(u_n, \theta))$$

$$\theta_N^* \sim \theta^* \qquad \theta_N^* \xrightarrow[U_N, \alpha_n \sim g_u(u)]{} \theta^*$$

Przypadek liniowo – kwadratowy
Przykład

Eksperyment: $U_N = [u_1 \quad u_2 \quad \cdots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$



Model: $\bar{y} = \Phi(u, \theta) = \theta^T \varphi(u) = \theta^{(1)} + \theta^{(2)} u$

gdzie: $\theta \overset{df}{=} \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \\ \vdots \\ \theta^{(R)} \end{bmatrix} \qquad \varphi(u) \overset{\text{df}}{=} \begin{bmatrix} \varphi_1(u) \\ \varphi_2(u) \\ \vdots \\ \varphi_R(u) \end{bmatrix}$

np..: $\theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \end{bmatrix} \qquad \varphi(u) = \begin{bmatrix} 1 \\ u \end{bmatrix},$

Model: $\bar{y} = \Phi(u, \theta) = \theta^T \varphi(u) = \theta^{(1)} + \theta^{(2)} u$

Wskaźnik jakości identyfikacji:

$$Q_N(\theta) = \sum_{n=1}^{N} (y_n - \bar{y}_n)^2 = \sum_{n=1}^{N} \left(y_n - \theta^T \varphi(u_n)\right)^2$$

Przypadek liniowo – kwadratowy
Przykład

Warunki konieczne optymalności mają postać

$$\operatorname*{grad}_{\theta} Q_N(\theta) = -2\sum_{n=1}^{N}\left(y_n - \theta_N^{*T}\varphi(u_n)\right)\varphi(u_n) = 0_2$$

Rozwiązanie powyższego równania przyjmuje postać:

$$\theta_N^* = \left[\sum_{n=1}^{N}\varphi(u_n)\varphi^T(u_n)\right]^{-1} \times \sum_{n=1}^{N} y_n\varphi(u_n)$$

a po podstawieniu i przekształceniach:

$$\theta_N^* = \begin{bmatrix}\theta_N^{*(1)} \\ \theta_N^{*(2)}\end{bmatrix} = \begin{bmatrix} \dfrac{\left(\sum_{n=1}^{N}u_n^2\right)\times\left(\sum_{n=1}^{N}y_n\right) - \left(\sum_{n=1}^{N}u_n\right)\times\left(\sum_{n=1}^{N}y_nu_n\right)}{N\sum_{n=1}^{N}u_n^2 - \left(\sum_{n=1}^{N}u_n\right)\times\left(\sum_{n=1}^{N}y_nu_n\right)} \\ \dfrac{N\sum_{n=1}^{N}y_nu_n - \left(\sum_{n=1}^{N}u_n\right)\times\left(\sum_{n=1}^{N}y_n\right)}{N\sum_{n=1}^{N}u_n^2 - \left(\sum_{n=1}^{N}u_n\right)\times\left(\sum_{n=1}^{N}y_nu_n\right)} \end{bmatrix}$$

Przypadek liniowo – kwadratowy Algorytm rekurencyjny

$$Q_N(\theta)=\sum_{n=1}^{N}(y_n-\bar{y}_n)^2=\sum_{n=1}^{N}\left(y_n-\theta^T\varphi(u_n)\right)^2 \quad U_N=[u_1 \quad u_2 \quad \cdots \quad u_N], \quad Y_N=[y_1 \quad y_2 \quad \cdots \quad y_N]$$

Wektor parametrów θ_N^* wyznaczamy na podstawie N pomiarów

$$\theta_N^*=\left[\sum_{n=1}^{N}\varphi(u_n)\varphi^T(u_n)\right]^{-1}\times\sum_{n=1}^{N}y_n\varphi(u_n)$$

Pojawia się nowy $N+1$ pomiar u_{N+1}, y_{N+1}

Jak wyznaczyć (poprawić) parametr korzystając z nowego pomiaru?

$$\theta_{N+1}^*=A\left(\theta_N^*, u_{N+1}, y_{N+1}\right)$$

$$Q_{N+1}(\theta)=\sum_{n=1}^{N+1}(y_n-\bar{y}_n)^2=\sum_{n=1}^{N+1}\left(y_n-\theta^T\varphi(u_n)\right)^2$$

$$\theta_{N+1}^*=\left[\sum_{n=1}^{N+1}\varphi(u_n)\varphi^T(u_n)\right]^{-1}\times\sum_{n=1}^{N+1`}y_n\varphi(u_n)=$$

$$\left[\sum_{n=1}^{N}\varphi(u_n)\varphi^T(u_n)+\varphi(u_{N+1})\varphi^T(u_{N+1})\right]^{-1}\times\left[\sum_{n=1}^{N`}y_n\varphi(u_n)+y_{N+1}\varphi(u_{N+1})\right]$$

Przypadek liniowo – kwadratowy Algorytm rekurencyjny

Przydatne przekształcenia algebraiczne:

$$\left(\mathbf{A}+\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}+\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{C}\mathbf{A}^{-1}$$

$\mathbf{B}=\varphi$ – wektor kolumnowy

$$\mathbf{D}^{-1}=1$$

$$\mathbf{C}=\varphi^{T}$$

$$\left(\mathbf{A}+\varphi\varphi^{T}\right)^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\varphi\left(1+\varphi^{T}\mathbf{A}^{-1}\varphi\right)^{-1}\varphi^{T}\mathbf{A}^{-1}$$

$$\mathbf{D}^{-1}=-1$$

$$\left(\mathbf{A}-\varphi\varphi^{T}\right)^{-1}=\mathbf{A}^{-1}+\mathbf{A}^{-1}\varphi\left(1-\varphi^{T}\mathbf{A}^{-1}\varphi\right)^{-1}\varphi^{T}\mathbf{A}^{-1}$$

Przypadek liniowo – kwadratowy Algorytm rekurencyjny

Oznaczmy przez: $\varphi_n = \varphi(u_n), n = 1, 2, \ldots, N+1$

$$\left(\sum_{n=1}^{N+1} \varphi_n \varphi_n^T\right)^{-1} = \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T + \varphi_{N+1}\varphi_{N+1}^T\right)^{-1} =$$

$$= \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T\right)^{-1} - \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T\right)^{-1} \frac{\varphi_{N+1}\varphi_{N+1}^T}{1+\varphi_{N+1}^T\left(\sum_{n=1}^{N} \varphi_n \varphi_n^T\right)^{-1}\varphi_{N+1}} \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T\right)^{-1}$$

Niech: $P_N = \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T\right)^{-1}$

Powyższe przyjmuje postać:

$$P_{N+1} = P_N - P_N \frac{\varphi_{N+1}\varphi_{N+1}^T}{1+\varphi_{N+1}^T P_N \varphi_{N+1}} P_N$$

Przypadek liniowo – kwadratowy
Algorytm rekurencyjny

$$\theta^*_{N+1} = \left(P_N - P_N \frac{\varphi_{N+1}\varphi^T_{N+1}}{1+\varphi^T_{N+1}P_N\varphi_{N+1}} P_N \right)\left(\sum_{n=1}^{N} \varphi_n y_n + \varphi_{N+1} y_{N+1} \right) =$$

$$= P_N \sum_{n=1}^{N} \varphi_n y_n + P_N \varphi_{N+1} y_{N+1} - \frac{P_N \varphi_{N+1}\varphi^T_{N+1} P_N}{1+\varphi^T_{N+1}P_N\varphi_{N+1}} \sum_{n=1}^{N} \varphi_n y_n - \frac{P_N \varphi_{N+1}\varphi^T_{N+1} P_N \varphi_{N+1} y_{N+1}}{1+\varphi^T_{N+1}P_N\varphi_{N+1}} =$$

$$= P_N \sum_{n=1}^{N} \varphi_n y_n + \frac{P_N \varphi_{N+1} y_{N+1} + P_N \varphi_{N+1} y_{N+1}\varphi^T_{N+1} P_N \varphi_{N+1} - P_N \varphi_{N+1} y_{N+1}\varphi^T_{N+1} P_N \varphi_{N+1} - P_N \varphi_{N+1}\varphi^T_{N+1} P_N \sum_{n=1}^{N} \varphi_n y_n}{1+\varphi^T_{N+1}P_N\kappa_{N+1}} =$$

$$= P_N \sum_{n=1}^{N} \varphi_n y_n + \frac{P_N \varphi_{N+1} y_{N+1} - P_N \varphi_{N+1}\varphi^T_{N+1} P_N \sum_{n=1}^{N} \varphi_n y_n}{1+\varphi^T_{N+1}P_N\varphi_{N+1}} y =$$

$$= P_N \sum_{n=1}^{N} \varphi_n y_n + \frac{P_N \varphi_{N+1} y_{N+1}}{1+\varphi^T_{N+1}P_N\varphi_{N+1}} \left(y_{N+1} + \varphi^T_{N+1} P_N \sum_{n=1}^{N} \varphi_n y_n \right)$$

$$= \theta^*_N + K_{N+1}\left(y_{N+1} - \varphi^T_{N+1}\theta_N \right)$$

gdzie: $K_{N+1} = \dfrac{P_N \varphi_{N+1}}{1+\varphi^T_{N+1}P_N\varphi_{N+1}}$

Przypadek liniowo – kwadratowy Algorytm rekurencyjny

Podsumowując, otrzymujemy następujący algorytm

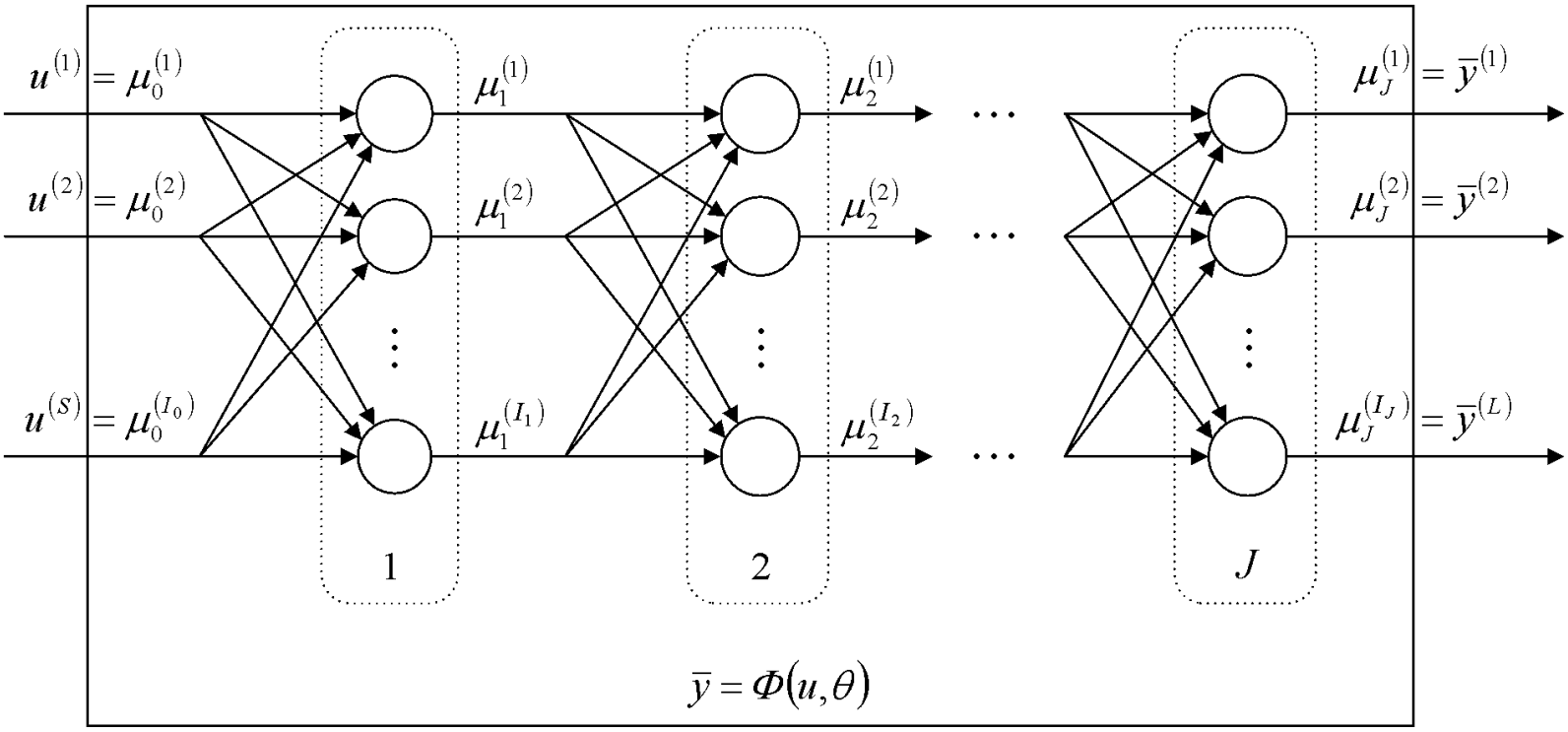
$$\theta_N^* = \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T \right)^{-1} \sum_{n=1}^{N} \varphi_n y_n = P_N \sum_{n=1}^{N} \varphi_n y_n \,, \text{ gdzie: } P_N = \left(\sum_{n=1}^{N} \varphi_n \varphi_n^T \right)^{-1}$$

$$\theta_{N+1}^* = A\left(\theta_N, u_{N+1}, y_{N+1}\right) = \theta_N^* + K_{N+1}\left[y_{N+1} - \varphi\left(u_{N+1}\right)^T \theta_N^* \right]$$

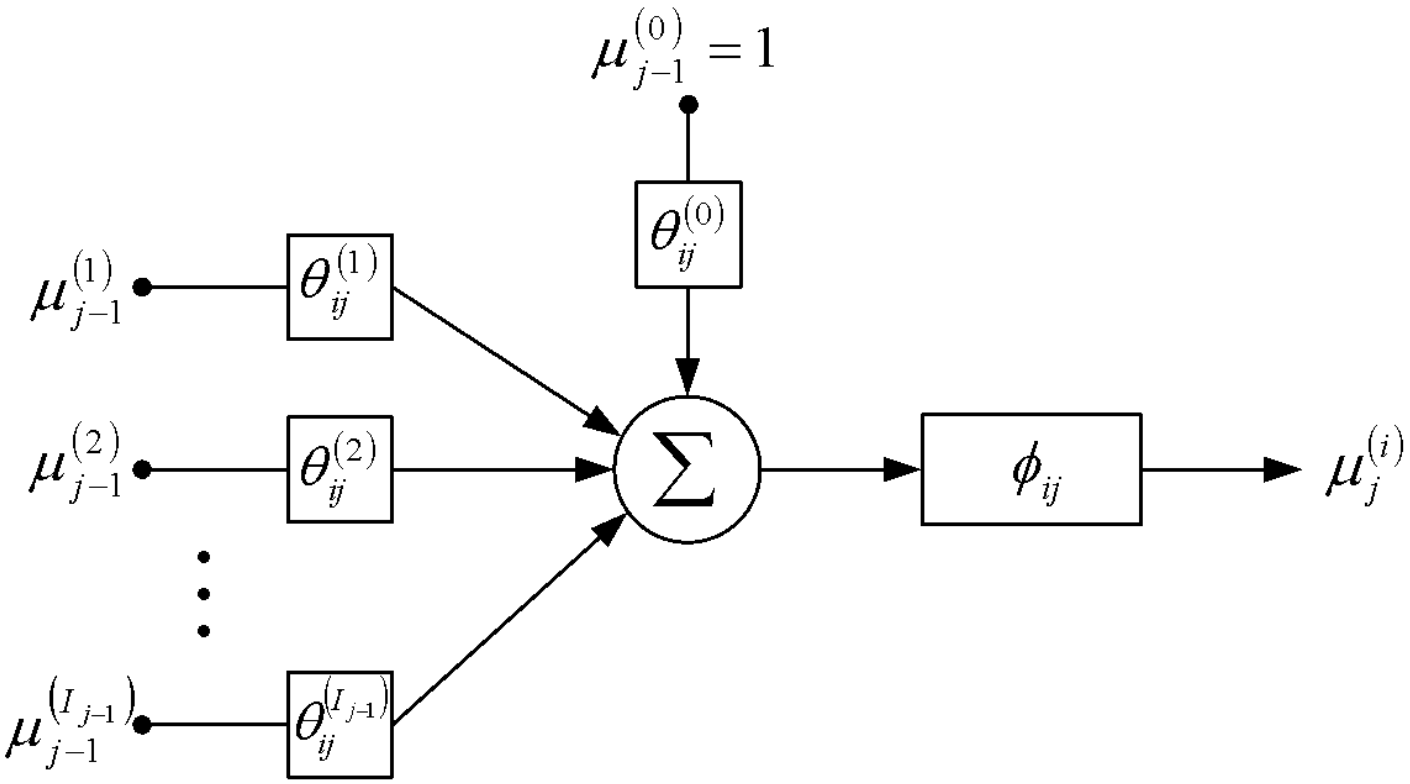
$$K_{N+1} = \frac{P_N \varphi\left(u_{N+1}\right)}{1 + \varphi\left(u_{N+1}\right)^T P_N \varphi\left(u_{N+1}\right)}$$

$$P_{N+1} = P_N - \frac{P_N \varphi\left(u_{N+1}\right) \varphi\left(u_{N+1}\right)^T P_N}{1 + \varphi\left(u_{N+1}\right)^T P_N \varphi\left(u_{N+1}\right)}$$

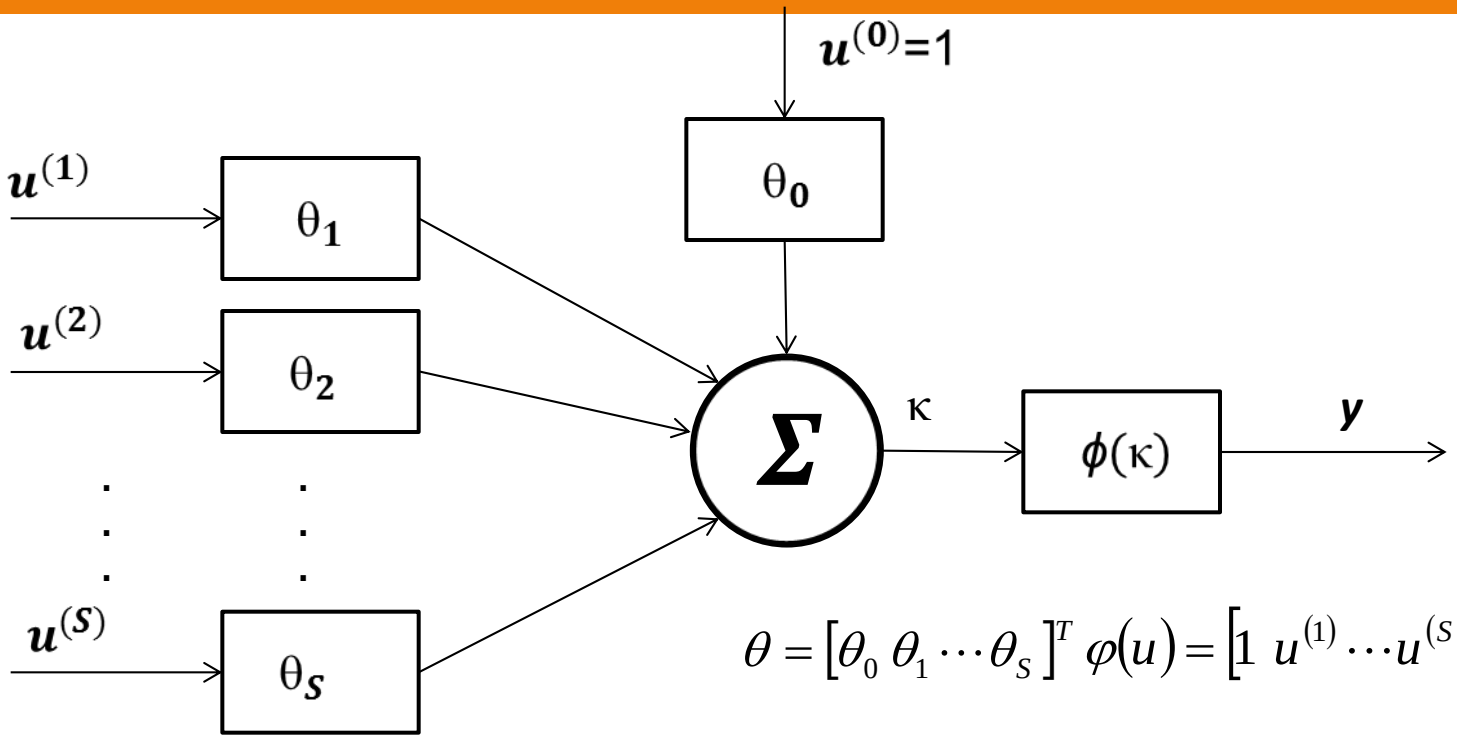
Wybór optymalnego modelu, a sieci neuronowe – szczególny przypadek



Model neuronu (wyrwany z sieci)



Model neuronu (uproszczenie)



$$\theta = \left[\theta_0\ \theta_1 \cdots \theta_S\right]^T \quad \varphi(u) = \left[1\ u^{(1)} \cdots u^{(S)}\right]^T$$

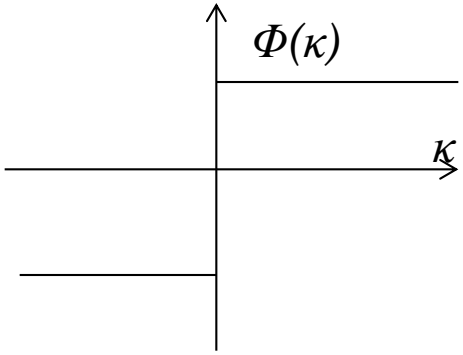
$$y = \phi\left(\sum_{s=1}^{S} \theta_s u^{(s)} + \theta_0\right) = \phi\left(\theta^T \varphi(u)\right)$$

Φ – funkcja aktywacji

Funkcja aktywacji

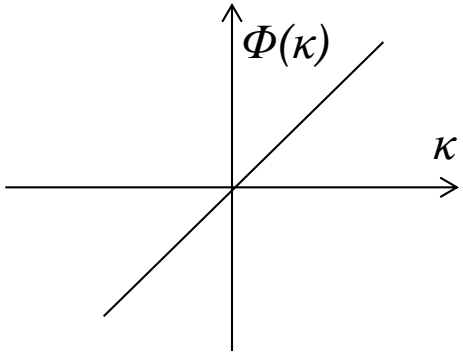
$$\phi(\kappa) = \begin{cases} 1 \; dla \; \kappa > 0 \\ -1 \; dla \; \kappa \leq 0 \end{cases}$$

Perceptron



$$\phi(\kappa) = \kappa$$

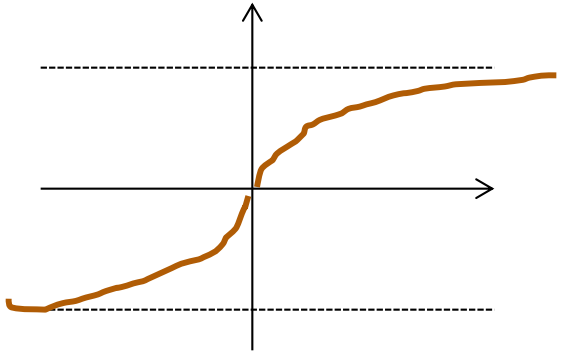
Adaline
Adaptive Linear Neuron



$$\phi(\kappa) = \frac{1}{1 + e^{-\beta\kappa}}$$

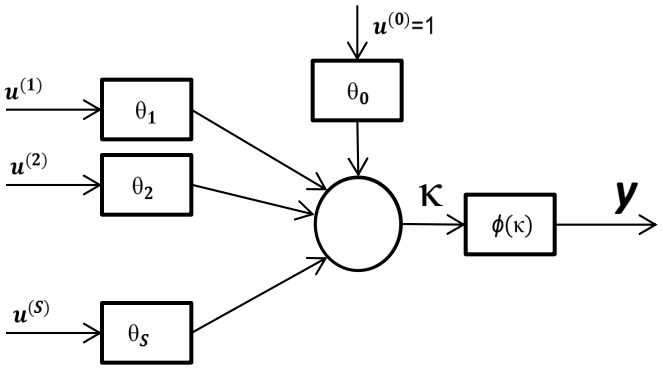
$$\phi(\kappa) = \tanh(\beta\kappa) = \frac{1 - e^{-\beta\kappa}}{1 + e^{-\beta\kappa}}$$

Neuron sigmoidalny



Uczenie neuronu

Neuron Adaline (Adaptive Linear Neuron) $\phi(\kappa) = \kappa$



$$y = \phi\left(\sum_{s=1}^{S} \theta_s u^{(s)} + \theta_0\right) = \phi\left(\theta^T \varphi(u)\right) = \theta^T \varphi(u)$$

$$\theta = \left[\theta_0\ \theta_1 \cdots \theta_S\right]^T \qquad \varphi(u) = \left[1\ u^{(1)} \cdots u^{(S)}\right]^T$$

Uczenie: $\left[u_1\ u_2 \cdots u_N\right] = U_N \qquad \left[y_1\ y_2 \cdots y_N\right] = Y_N$

$$Q_N(\theta) = \sum_{n=1}^{N} \left(y_n - \bar{y}_n\right)^2 = \sum_{n=1}^{N} \left(y_n - \theta^T \varphi(u_n)\right)^2$$

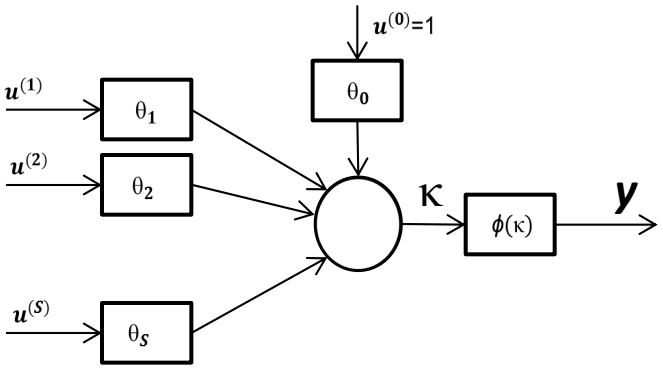
To już było

$$\theta_N^* = \left[\sum_{n=1}^{N} \varphi(u_n)\varphi^T(u_n)\right]^{-1} \times \sum_{n=1}^{N} y_n \varphi(u_n)$$

lub rekurencyjnie

Uczenie neuronu

Neuron Adaline (Adaptive Linear Neuron) $\phi(\kappa) = \kappa$



$$y = \phi\left(\sum_{s=1}^{S} \theta_s u^{(s)} + \theta_0\right) = \phi\left(\theta^T \varphi(u)\right) = \theta^T \varphi(u)$$

$$\theta = \left[\theta_0\ \theta_1 \cdots \theta_S\right]^T \qquad \varphi(u) = \left[1\ u^{(1)} \cdots u^{(S)}\right]^T$$

Uczenie: $\left[u_1\ u_2 \cdots u_N\right] = U_N \qquad \left[y_1\ y_2 \cdots y_N\right] = Y_N$

$$Q_N(\theta) = \sum_{n=1}^{N} \left(y_n - \bar{y}_n\right)^2 = \sum_{n=1}^{N} \left(y_n - \theta^T \varphi(u_n)\right)^2$$

To już było

$$\theta_N^* = \left[\sum_{n=1}^{N} \varphi(u_n)\varphi^T(u_n)\right]^{-1} \times \sum_{n=1}^{N} y_n \varphi(u_n)$$

lub rekurencyjnie

Algortm rekurencyjny

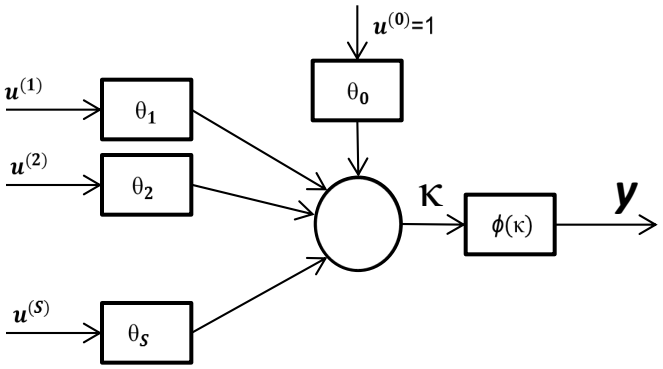
$$\theta^*_{N+1} = A\left(\theta^*_N, u_{N+1}, y_{N+1}\right) = \theta^*_N + K_{N+1}\left[y_{N+1} - \varphi\left(u_{N+1}\right)^T \theta^*_N\right]$$

$$K_{N+1} = \frac{P_N \varphi\left(u_{N+1}\right)}{1 + \varphi\left(u_{N+1}\right)^T P_N \varphi\left(u_{N+1}\right)}$$

$$P_{N+1} = P_N - \frac{P_N \varphi\left(u_{N+1}\right)\varphi\left(u_{N+1}\right)^T P_N}{1 + \varphi\left(u_{N+1}\right)^T P_N \varphi\left(u_{N+1}\right)}$$

Uczenie neuronu

Podobnie jak neuron Adaline (Adaptive Linear Neuron) $\phi(\kappa)$ - wzajemnie jednoznaczna



$$y = \phi\left(\sum_{s=1}^{S}\theta_s u^{(s)} + \theta_0\right) = \phi\left(\theta^T \varphi(u)\right)$$

$$\theta = \left[\theta_0\ \theta_1 \cdots \theta_S\right]^T \qquad \varphi(u) = \left[1\ u^{(1)} \cdots u^{(S)}\right]^T$$

Uczenie: $\left[u_1\ u_2 \cdots u_N\right] = U_N \qquad \left[y_1\ y_2 \cdots y_N\right] = Y_N$

$$Q_N(\theta) = \sum_{n=1}^{N}\left(\kappa_n - \bar{\kappa}_n\right)^2 = \sum_{n=1}^{N}\left(\phi^{-1}(y_n) - \theta^T\varphi(u_n)\right)^2$$ To już było

$$\theta_N^* = \left[\sum_{n=1}^{N}\varphi(u_n)\varphi^T(u_n)\right]^{-1} \times \sum_{n=1}^{N}\phi^{-1}(y_n)\varphi(u_n)$$

lub rekurencyjnie

Algorytm rekurencyjny

$$\theta^*_{N+1} = A\left(\theta^*_N, u_{N+1}, y_{N+1}\right) = \theta^*_N + K_{N+1}\left[\phi^{-1}\left(y_{N+1}\right) - \varphi\left(u_{N+1}\right)^T \theta^*_N\right]$$

$$K_{N+1} = \frac{P_N \varphi\left(u_{N+1}\right)}{1 + \varphi\left(u_{N+1}\right)^T P_N \varphi\left(u_{N+1}\right)}$$

$$P_{N+1} = P_N - \frac{P_N \varphi\left(u_{N+1}\right)\varphi\left(u_{N+1}\right)^T P_N}{1 + \varphi\left(u_{N+1}\right)^T P_N \varphi\left(u_{N+1}\right)}$$

Uczenie neuronu
reguła delta

$$Q(\theta) = \frac{1}{2}\left(y - \theta^T \varphi(u)\right)^2, \quad \theta_0$$ Neuron adaline

Numeryczna metoda optymalizacji (to będzie)

$$\theta^* \rightarrow Q(\theta^*) = \min_{\theta} Q(\theta)$$

$$\theta^*_{n+1} = \theta^*_n - \eta \nabla_\theta Q(\theta^*_n), \quad \theta^*_0$$ η – współczynnik uczenia

$$\nabla Q(\theta) = -\left(y - \theta^T \varphi(u)\right)\varphi(u), \quad \theta_0$$

$$\theta^*_{n+1} = \theta^*_n + \eta\left(y_{n+1} - \theta^{*T}_n \varphi(u_{n+1})\right)\varphi(u_{n+1}), \quad \theta^*_0$$

$$\Delta_n = \left(y_{n+1} - \theta^{*T}_n \varphi(u_{n+1})\right)$$

$$\theta^*_{n+1} = \theta^*_n + \eta \Delta_n \varphi(u_{n+1})$$

Uczenie neuronu
reguła delta

$$Q(\theta) = \frac{1}{2}\left(y - \phi\left(\theta^T \varphi(u)\right)\right)^2, \quad \theta_0$$

ϕ – funkcja aktywacji, różniczkowalna

Numeryczna metoda optymalizacji (to będzie)

$$\theta^* \rightarrow Q\left(\theta^*\right) = \min_{\theta} Q(\theta)$$

$$\theta^*_{n+1} = \theta^*_n - \eta \nabla_\theta Q\left(\theta^*_n\right), \quad \theta^*_0$$

η – współczynnik uczenia

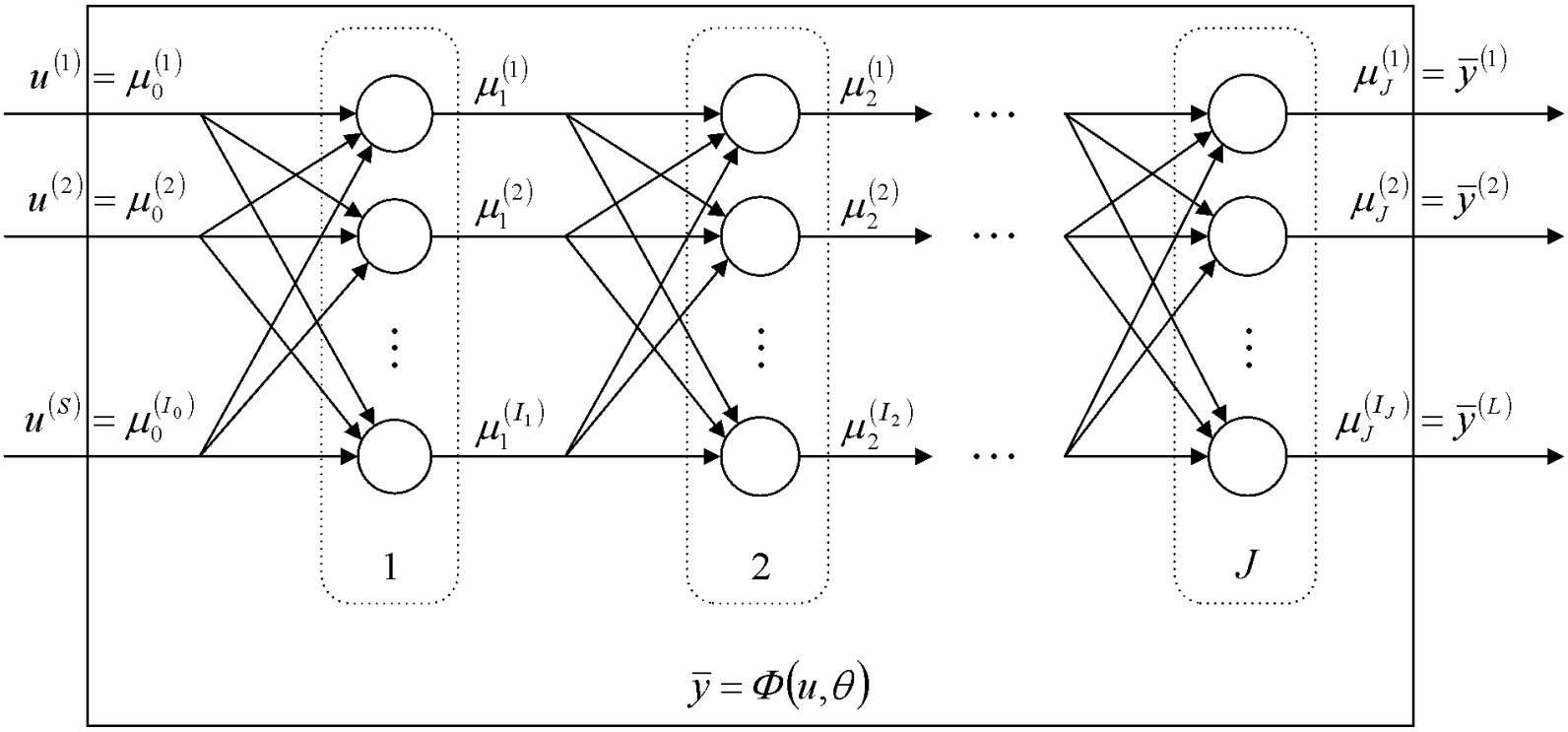
$$\nabla Q(\theta) = -\left(y - \phi\left(\theta^T \varphi(u)\right)\right)\frac{\partial \phi(\kappa)}{\partial(\kappa)}\varphi(u), \quad \theta_0$$

$$\theta^*_{n+1} = \theta^*_n + \eta\left(y_{n+1} - \phi\left(\theta^{*T}_n \varphi\left(u_{n+1}\right)\right)\right)\frac{\partial \phi(\kappa)}{\partial \kappa}\varphi\left(u_{n+1}\right), \quad \theta^*_0$$

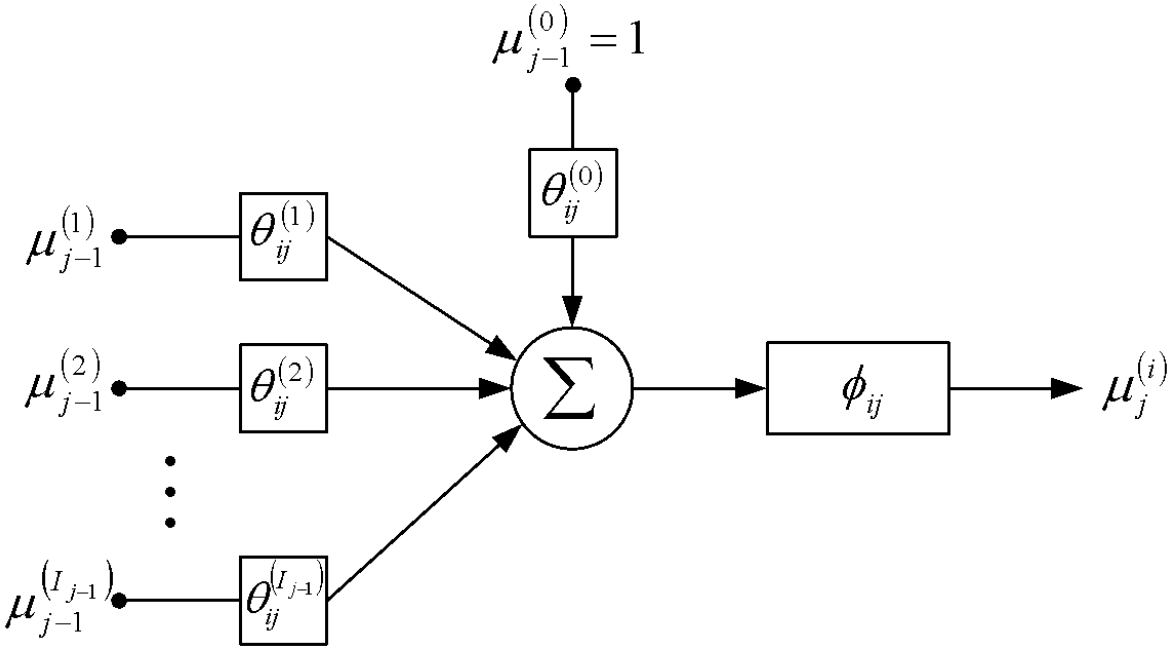
$$\Delta_n = \left(y_{n+1} - \phi\left(\theta^{*T}_n \varphi\left(u_{n+1}\right)\right)\right)\frac{\partial \phi(\kappa)}{\partial \kappa}$$

$$\theta^*_{n+1} = \theta^*_n + \eta \Delta_n \varphi\left(u_{n+1}\right)$$

Sieć wielowarstwowa



Model neuronu (wyrwany z sieci)

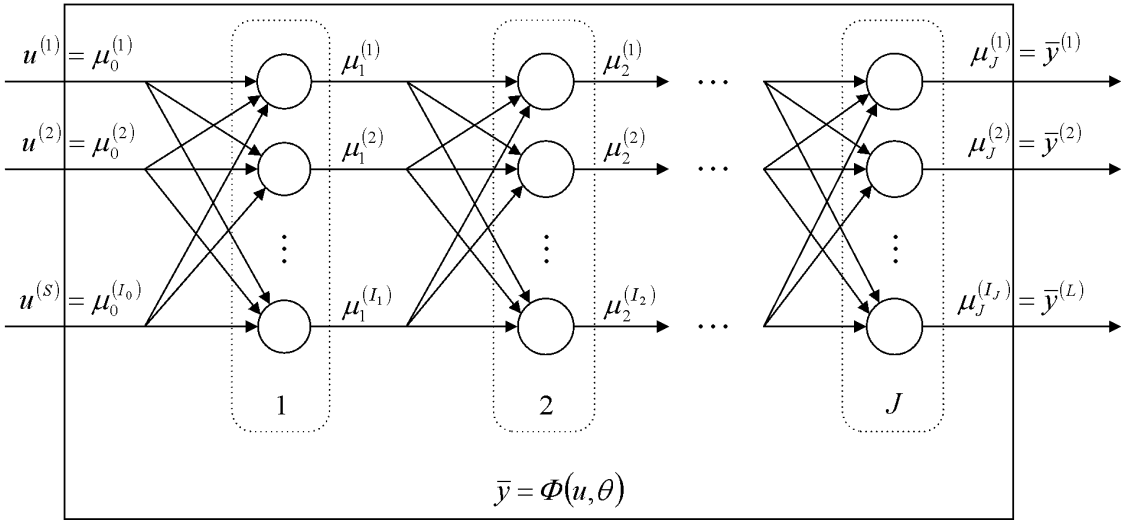


$$\mu_j^{(i)} = \phi_{ij}\left(\overline{\mu}_{j-1}^T \theta_{ij}\right).$$

$$\mu_j^{(i)} = \phi_{ij}\left(\sum_{s=1}^{I_{j-1}} \mu_{j-1}^{(s)}\theta_{ij}^{(s)} + \theta_{ij}^{(0)}\right), \qquad \mu_{j-1} = \begin{bmatrix} \mu_{j-1}^{(1)} \\ \mu_{j-1}^{(2)} \\ \vdots \\ \mu_{j-1}^{(I_{j-1})} \end{bmatrix},$$

$$\overline{\mu}_{j-1} \overset{\text{df}}{=} \begin{bmatrix} \mu_{j-1}^{(0)} \\ \mu_{j-1} \end{bmatrix} = \begin{bmatrix} \mu_{j-1}^{(0)} \\ \mu_{j-1}^{(1)} \\ \mu_{j-1}^{(2)} \\ \vdots \\ \mu_{j-1}^{(I_{j-1})} \end{bmatrix}, \qquad \theta_{ij} \overset{\text{df}}{=} \begin{bmatrix} \theta_{ij}^{(0)} \\ \theta_{ij}^{(1)} \\ \vdots \\ \theta_{ij}^{(I_{j-1})} \end{bmatrix},$$

Sieć wielowarstwowa



$$\mu_j = \begin{bmatrix} \mu_j^{(1)} \\ \mu_j^{(2)} \\ \vdots \\ \mu_j^{(I_j)} \end{bmatrix} = \begin{bmatrix} \phi_{1j}\left(\bar{\mu}_{j-1}^T \theta_{1j}\right) \\ \phi_{2j}\left(\bar{\mu}_{j-1}^T \theta_{2j}\right) \\ \vdots \\ \phi_{I_j j}\left(\bar{\mu}_{j-1}^T \theta_{I_j j}\right) \end{bmatrix} \overset{\text{df}}{=} \phi_j\left(\bar{\mu}_{j-1}, \theta_j\right),$$

$$\theta_j \overset{\text{df}}{=} \begin{bmatrix} \theta_{1j} & \theta_{2j} & \cdots & \theta_{I_j j} \end{bmatrix} = \begin{bmatrix} \theta_{1j}^{(0)} & \theta_{2j}^{(0)} & \cdots & \theta_{I_j j}^{(0)} \\ \theta_{1j}^{(1)} & \theta_{2j}^{(1)} & \cdots & \theta_{I_j j}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{1j}^{(I_{j-1})} & \theta_{2j}^{(I_{j-1})} & \cdots & \theta_{I_j j}^{(I_{j-1})} \end{bmatrix}.$$

$$\mu_j = \phi_j\left(\bar{\mu}_{j-1}, \theta_j\right), \; j = 1, 2, \; \ldots, J.$$

$$\mu_1 = \phi_1\left(\bar{\mu}_0, \theta_1\right) = \phi_1\left(u, \theta_1\right).$$

$$\mu_j = \phi_j\left(\bar{\mu}_{j-1}, \theta_j\right), \; j = 2, 3, \; \ldots, J-1.$$

$$\bar{y} = \mu_J = \phi_J\left(\bar{\mu}_{J-1}, \theta_J\right).$$

$$\bar{y} = \begin{bmatrix} \bar{y}^{(1)} \\ \bar{y}^{(2)} \\ \vdots \\ \bar{y}^{(L)} \end{bmatrix} = \begin{bmatrix} \mu_J^{(1)} \\ \mu_J^{(2)} \\ \vdots \\ \mu_J^{(I_J)} \end{bmatrix} = \phi_J\left(\phi_{J-1}\left(\ldots \phi_1\left(u, \theta_1\right), \ldots, \theta_{J-1}\right), \theta_J\right) \overset{\text{df}}{=} \begin{bmatrix} \Phi^{(1)}(u, \theta) \\ \Phi^{(2)}(u, \theta) \\ \vdots \\ \Phi^{(L)}(u, \theta) \end{bmatrix} \overset{\text{df}}{=} \Phi(u, \theta),$$

$$\theta \overset{\text{df}}{=} \{\theta_1, \theta_2, \ldots, \theta_J\}.$$

Uczenie sieci wielowarstwowej

$$Q_n(\theta) = [y_n - \Phi(u_n, \theta)]^T [y_n - \Phi(u_n, \theta)] = \sum_{l=1}^{L} \left(y_n^{(l)} - \Phi^{(l)}(u_n, \theta)\right)^2 = \sum_{l=1}^{L} \varepsilon_n^{(l)\,2} \qquad \varepsilon_n^{(l)} \stackrel{\text{df}}{=} y_n^{(l)} - \bar{y}_n^{(l)} = y_n^{(l)} - \Phi^{(l)}(u_n, \theta),\ l = 1, 2, \ldots, L$$

$$\left.\frac{\partial Q_n(\theta)}{\partial \theta_{ij}^{(s)}}\right|_{\theta_{ij}^{(s)} = \tilde{\tilde{\theta}}_{ij,n}^{(s)}} = -2\delta_{jn}^{(i)} \mu_{j-1,n}^{(s)},$$

$$\tilde{\tilde{\theta}}_{ij,n+1}^{(s)} = \tilde{\tilde{\theta}}_{ij,n}^{(s)} - \eta_n \left.\frac{\partial Q_n(\theta)}{\partial \theta_{ij}^{(s)}}\right|_{\theta_{ij}^{(s)} = \tilde{\tilde{\theta}}_{ij,n}^{(s)}}, \quad s = 0, 1, \ldots, I_{J-1},\ i = 1, 2, \ldots, I_J,\ j = 1, 2, \ldots, J,$$

$$\varepsilon_{jn}^{(i)} \stackrel{\text{df}}{=} \begin{cases} \varepsilon_n^{(i)} & \text{dla } j = J \\ \displaystyle\sum_{p=1}^{I_{j+1}} \delta_{j+1,n}^{(p)} \tilde{\tilde{\theta}}_{pj+1,n}^{(i)} & \text{dla } j = J-1, J-2, \ldots, 1 \end{cases}$$

$$\delta_{jn}^{(i)} \stackrel{\text{df}}{=} \varepsilon_{jn}^{(i)} \left.\frac{d\phi_{ij}(\kappa_j^{(i)})}{d\kappa_j^{(i)}}\right|_{\theta_{ij}^{(s)} = \tilde{\tilde{\theta}}_{ij,n}^{(s)}}$$

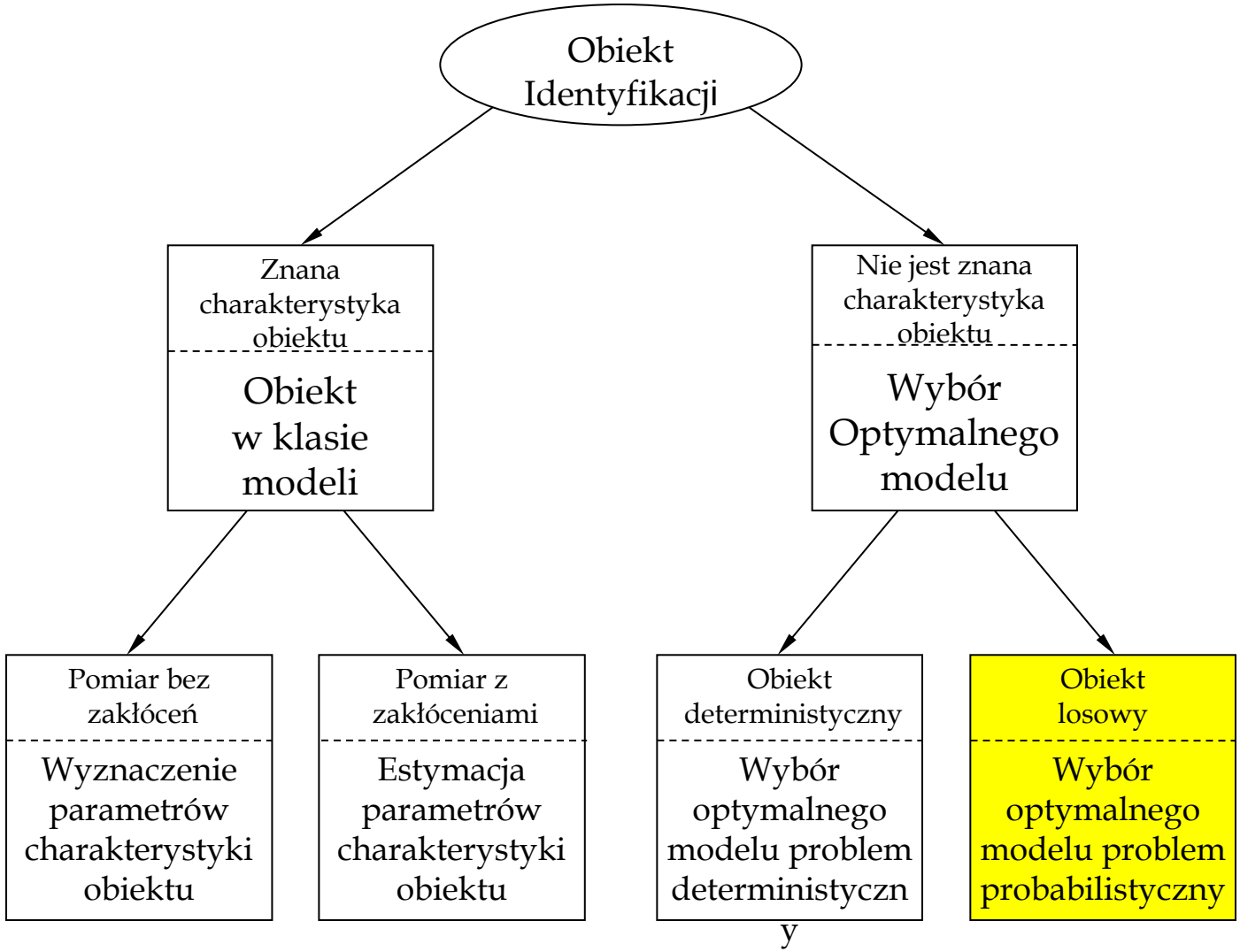
$$\kappa_j^{(i)} \stackrel{\text{df}}{=} \sum_{s=1}^{I_{j-1}} \mu_{j-1}^{(s)} \theta_{ij}^{(s)} + \theta_{ij}^{(0)}.$$

$$\tilde{\tilde{\theta}}_{ij,n+1}^{(s)} = \tilde{\tilde{\theta}}_{ij,n}^{(s)} + 2\eta_n \delta_{jn}^{(i)} \mu_{j-1,n}^{(s)},\ s = 0, 1, \ldots, I_{J-1},\ i = 1, 2, \ldots, I_J,\ j = 1, 2, \ldots, J.$$

Wybór Optymalnego Modelu Przypadek Probabilistyczny

∞ 104 ∞

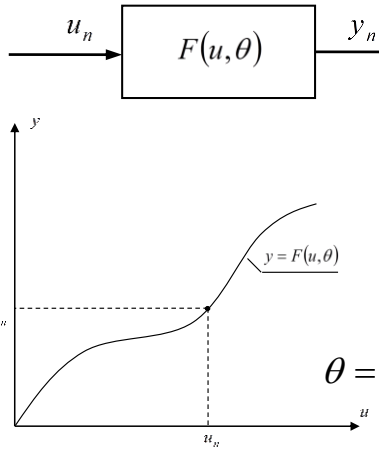
Typowe zadania identyfikacji



Podstawowe zadania identyfikacji podsumowanie

Deterministyczny

Obiekt w klasie modeli



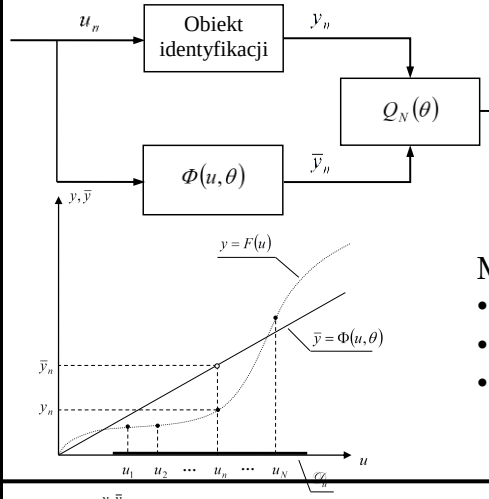
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

$$y_n = F(u_n, \theta),$$
$$n = 1, 2, \ldots, N$$

$$Y_N = \bar{F}(U_N, \theta)$$

$$\theta = \bar{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

Wybór optymalnego modelu

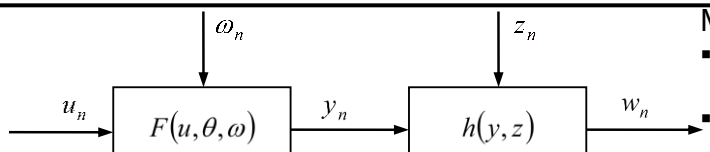


$$\bar{y} = \Phi(u, \theta_N^*)$$
$$\theta_N^* \rightarrow Q_N(\theta_N^*) = \min_{\theta} Q_N(\theta)$$
$$\theta_N^* = \Psi_N(U_N, Y_N)$$
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

Model jest optymalny:

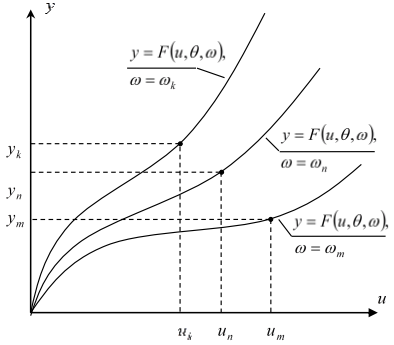
- dla zadanej serii pomiarowej
- przyjętego modelu
- przyjętego wskaźnika jakości identyfikacji

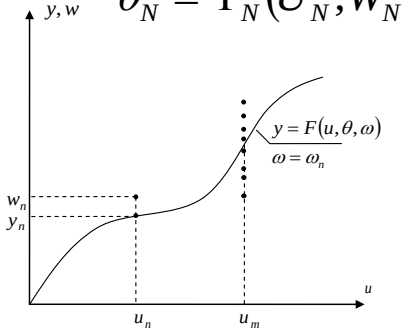
Losowy



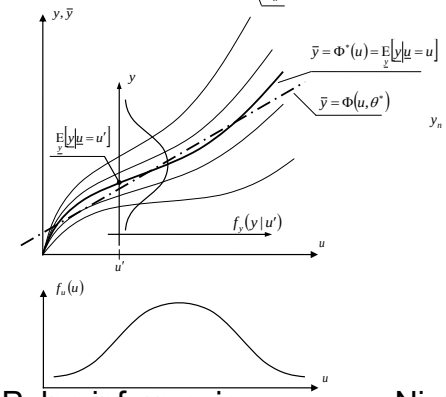
Metody:

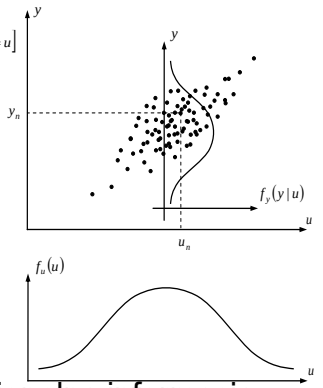
- Najmniejszych kwadratów
- Maksymalnej wiarogodności
- Bayes'a





$$\theta_N = \Psi_N(U_N, W_N)$$





Pełna informacja

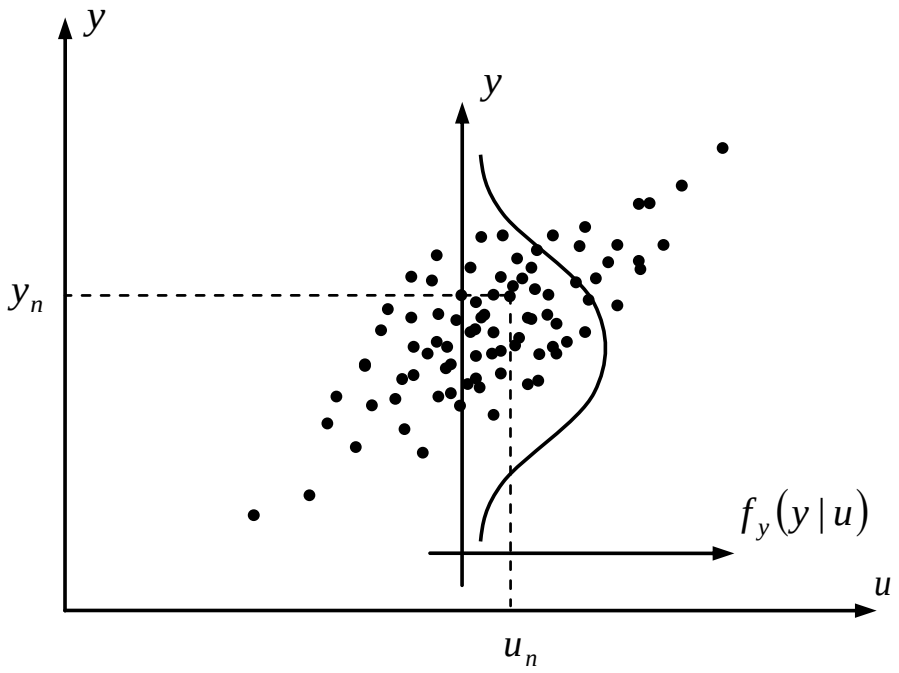
- Regresja I rodzaju
- Regresja II rodzaju

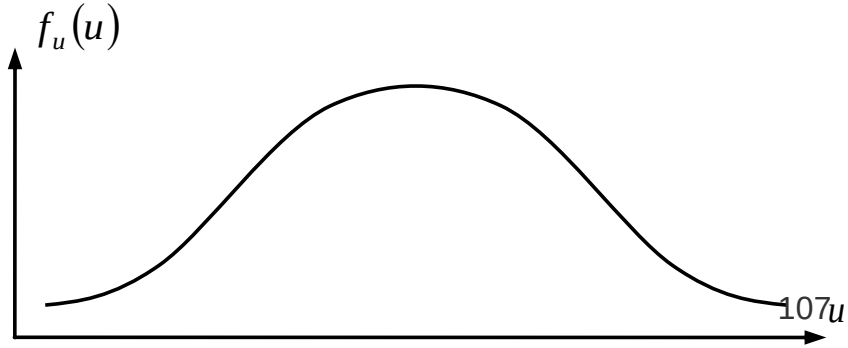
Niepełna informacja

- Estymacja wskaźnika jakości
- Estymacja parametrów rozkładu
- Estymacja rozkładu

$(u_n, y_n), n = 1, 2, \ldots, N$

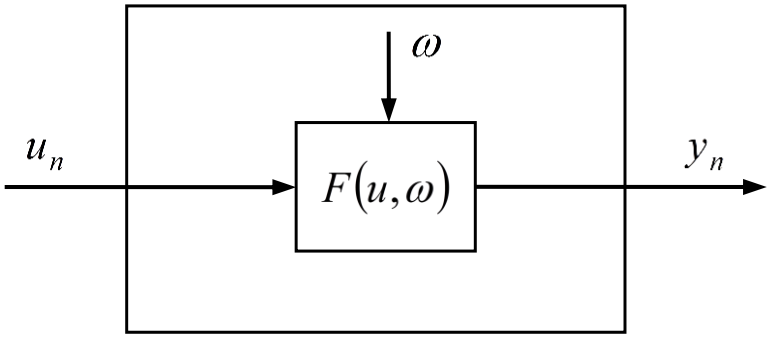
Są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$

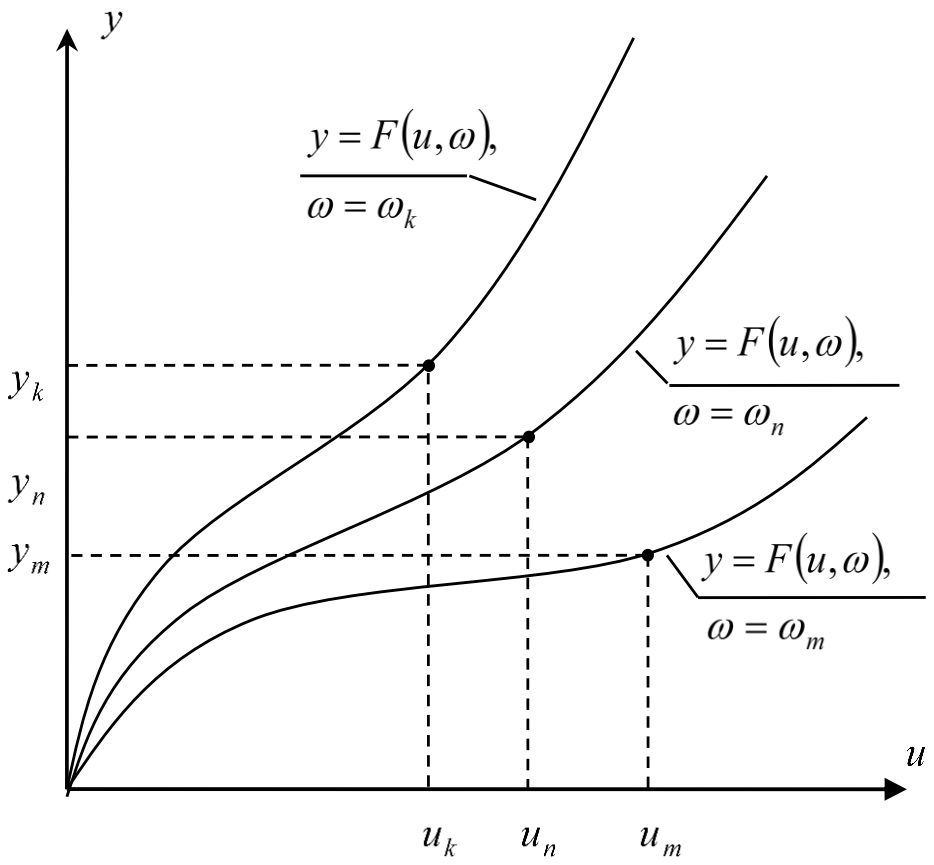




Wybór optymalnego modelu w warunkach losowych

- Obiekt identyfikacji z losowo zmienną wielkością ω





Wybór optymalnego modelu w warunkach losowych

- Sformułowanie problemu

Charakterystyka statyczna obiektu identyfikacji : $\underline{y} = F(u, \underline{\omega})$

ω_n – wartość zmiennej losowej $\underline{\omega}$ (losowo zmienna wielkość w obiekcie)

$f_\omega(\omega)$ - funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{\omega}$ $\left(\omega \in \Omega \subseteq \mathscr{R}^L\right)$,

y_n – wartość zmiennej losowej $\underline{y}$ – przekształcenie zmiennej losowej $\underline{\omega}$

F – funkcja wzajemnie jednoznaczna względem ω ,
a zatem istnieje funkcja odwrotna : $\underline{\omega} = F_\omega^{-1}(u, \underline{y})$

Wybór optymalnego modelu w warunkach losowych

Warunkowa gęstość rozkładu prawdopodobieństwa zmiennej losowej $\underline{y}$, pod warunkiem, że zmienna losowa $\underline{u} = u$:

$$f_y(y \mid u) = f_\omega\left(F_\omega^{-1}(u, y)\right) \cdot \left|J_F\right|$$

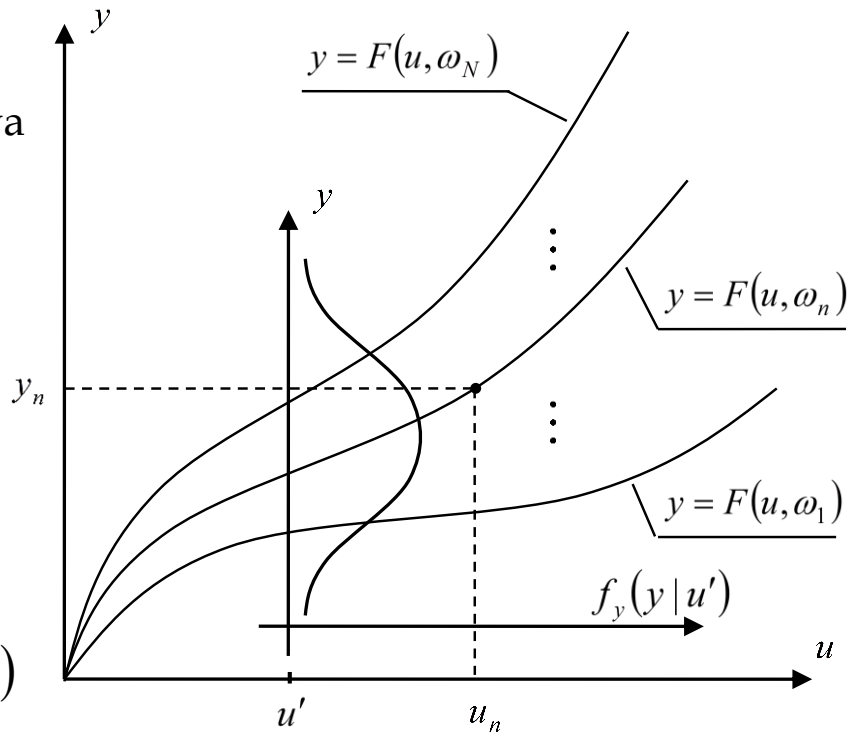
gdzie

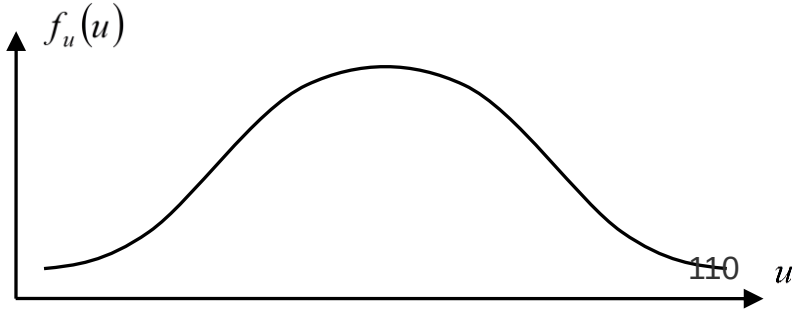
$$J_F = \frac{\partial F_\omega^{-1}(u, y)}{\partial \underline{y}}$$

u_n – wartość zmiennej losowej $\underline{u}$,
o gęstości rozkładu prawdopodobieństwa $f_u(u)$

Wejście u_n i wyjście y_n są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$

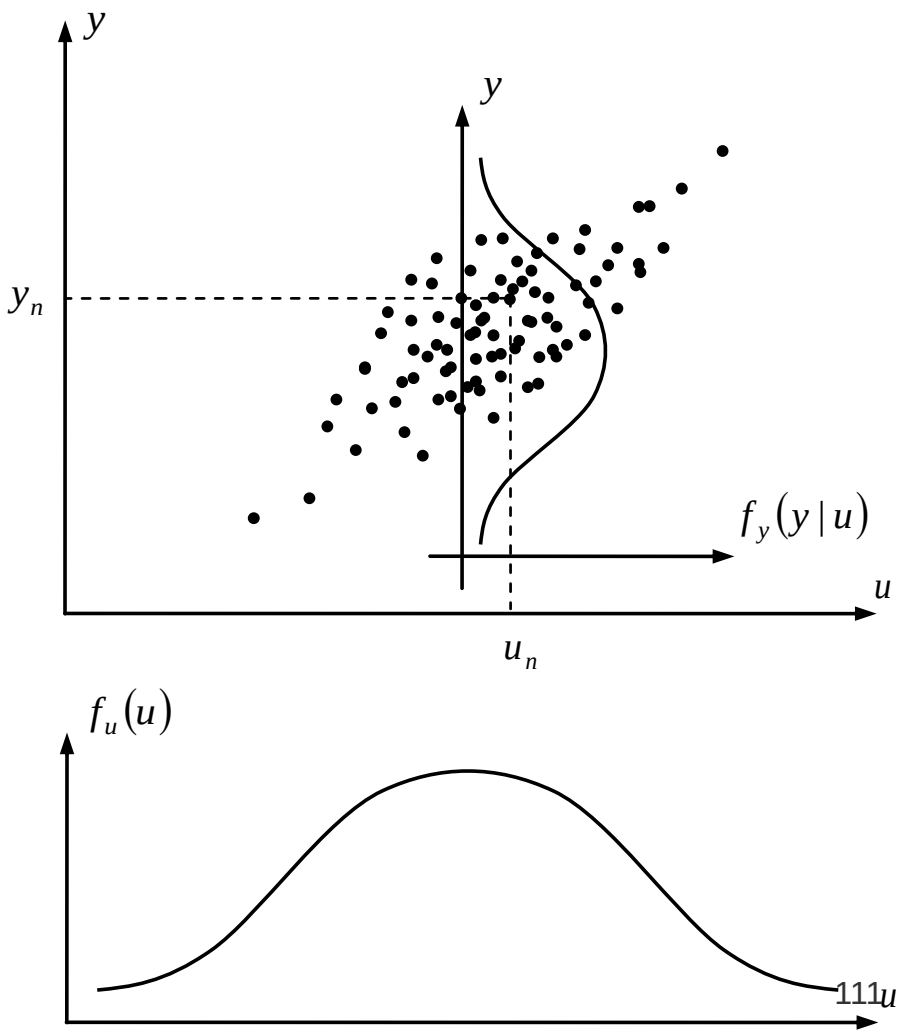
$$u \in \mathscr{U} \subseteq \mathscr{R}^S, \quad y \in \mathscr{Y} \subseteq \mathscr{R}^L$$





$(u_n, y_n), n = 1, 2, \ldots, N$

Są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$



Wybór optymalnego modelu w warunkach losowych

Dwa możliwe przypadki

Pełna informacja probabilistyczna

łączna gęstość rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ $f(u, y)$

lub
warunkowa gęstość rozkładu prawdopodobieństwa $f_y(y|u)$

i brzegowa $f_u(u)$

są znane

Niepełna informacja probabilistyczna

łączna gęstość rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ istnieje ale nie jest znana
W zamian dysponujemy pomiarami:

$$(u_n, y_n),\ n = 1, 2, \ldots, N$$

które są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$

Pełna informacja probabilistyczna

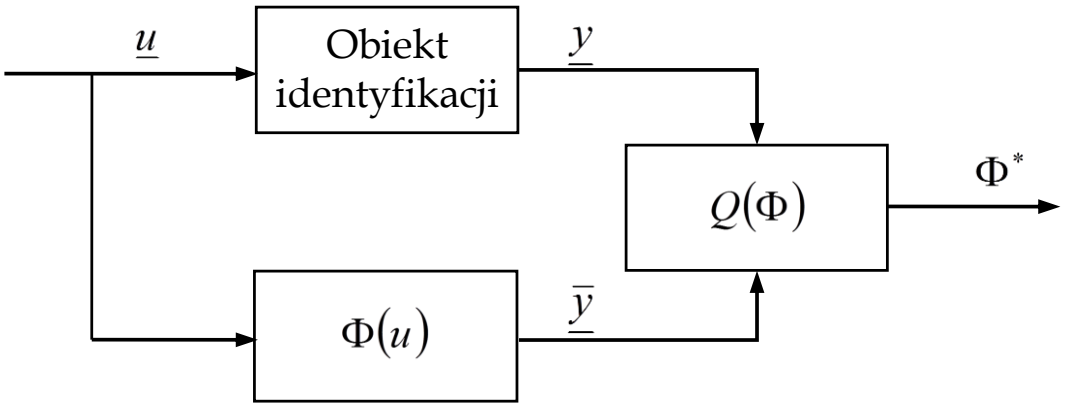
$$f(u,y) = f_y(y|u) \times f_u(u)$$ - jest znane

Model: $\bar{y} = \Phi(u)$ - szukamy funkcji Φ

Wskaźnik jakości: $$Q(\Phi) = \underset{\underline{u},\underline{y}}{E}\left[q\left(\underline{y}, \Phi(\underline{u})\right)\right] = \int_{\mathcal{U}}\int_{\mathcal{Y}} q(y, \Phi(u)) f(u,y)\, dy\, du$$

Model optymalny: $$\Phi^* \rightarrow \quad Q(\Phi^*) = \min_{\Phi} Q(\Phi)$$

$\bar{y} = \Phi^*(u)$



Pełna informacja probabilistyczna

$$Q(\Phi) = \underset{\underline{u},}{E}\left[\underset{\underline{y}}{E}\left[q\left(\underline{y}, \Phi(\underline{u})\right) | \underline{u} = u\right]\right] = \int_{\mathscr{U}}\left[\int_{\mathscr{Y}} q(y, \Phi(u)) f_y(y|u)\, dy \right] f_u(u)\, du$$

Dla ustalonego u warunkowa wartość oczekiwana ma postać:

$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[q\left(\underline{y}, \Phi(\underline{u})\right) | \underline{u} = u\right] = \int_{\mathscr{Y}} q(y, \Phi(u)) f_y(y|u)\, dy$$

Optymalny model otrzymujemy:

$$\Phi^*(u) \rightarrow \quad Q_u\left(\Phi^*(u)\right) = \min_{\Phi(u)} Q_u(\Phi(u))$$

Dla :

$$q(y, \bar{y}) = [y - \bar{y}]^T [y - \bar{y}]$$

Wskaźnik jakości ma postać:

$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[[y - \Phi(u)]^T [y - \Phi(u)] | \underline{u} = u\right] = \int_{\mathscr{Y}} [y - \Phi(u)]^T [y - \Phi(u)] f_y(y|u)\, dy$$

$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[\left[y-\Phi(u)\right]^T\left[y-\Phi(u)\right]|\underline{u}=u\right] = \int_{\mathcal{Y}}\left[y-\Phi(u)\right]^T\left[y-\Phi(u)\right]f_y(y|u)\,dy$$

$$Q(\Phi(u)) = E_{\underline{y}}\left[\underline{y}^T\underline{y}|u\right] - 2\Phi(u)E_{\underline{y}}\left[\underline{y}|u\right] + \Phi(u)^T\Phi(u)$$

$$\nabla_{\Phi(u)}Q(\Phi(u)) = -2E_{\underline{y}}\left[\underline{y}|u\right] + 2\Phi^*(u) = 0_L$$

$$\Phi^*(u) = E_{\underline{y}}\left[\underline{y}|u\right] = \int_{\mathcal{Y}} y f_y(y|u)dy$$

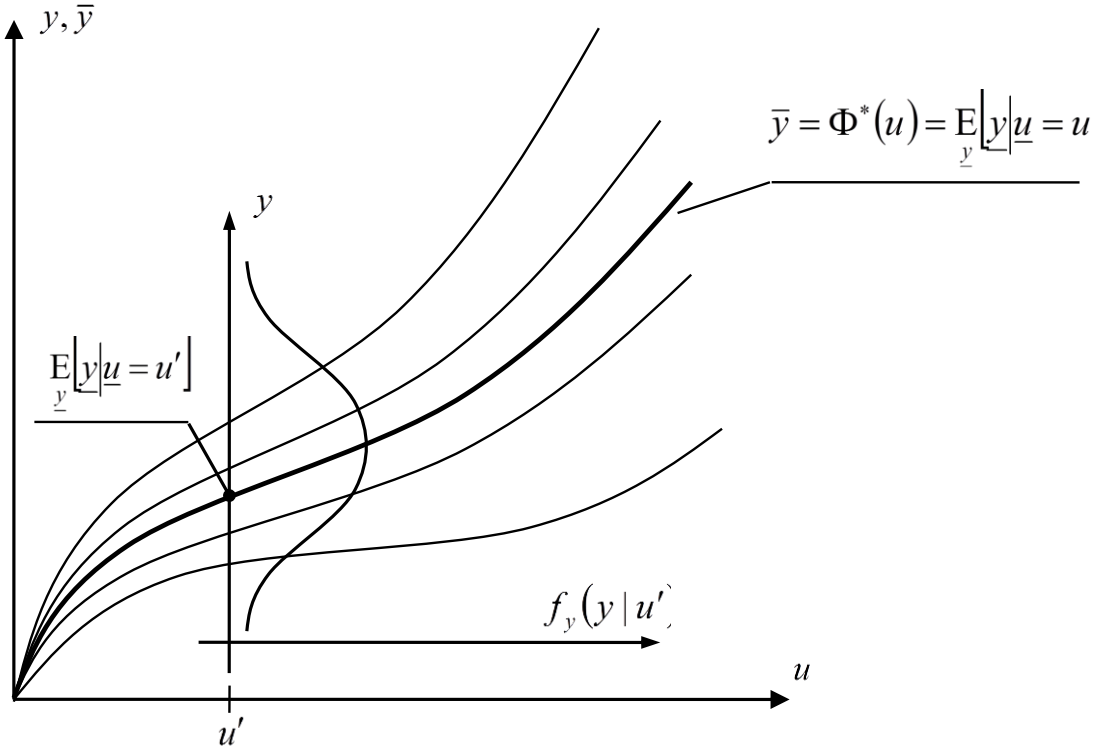
Pełna informacja probabilistyczna

- Regresja I rodzaju

$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[[y-\Phi(u)]^T[y-\Phi(u)]\middle|\underline{u}=u\right] = \int_{\mathscr{Y}}[y-\Phi(u)]^T[y-\Phi(u)]f_y(y|u)\,dy$$

W wyniku minimalizacji powyższego wskaźnika względem $\Phi(u)$ otrzymujemy optymalny model:

$$\bar{y} = \Phi^*(u) = \underset{\underline{y}}{E}\left[\underline{y}\middle|\underline{u}=u\right] = \int_{\mathscr{Y}} y\, f(y|u)dy$$



Pełna informacja probabilistyczna

- Przykład regresja I rodzaju dla rozkładu normalnego

Łączna gęstość rozkładu pary zmiennych losowych $(\underline{u}, \underline{y})$ ma postać:

$$f(u,y) = (2\pi)^{-\frac{L+S}{2}} \left|\Sigma^{-1}\right|^{\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\begin{bmatrix} u \\ y \end{bmatrix} - \begin{bmatrix} m_u \\ m_y \end{bmatrix}\right)^T \Sigma^{-1} \left(\begin{bmatrix} u \\ y \end{bmatrix} - \begin{bmatrix} m_u \\ m_y \end{bmatrix}\right)\right]$$

gdzie : Σ – macierz kowariancji, $\Sigma = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uy} \\ \Sigma_{yu} & \Sigma_{yy} \end{bmatrix}$

m_u, m_y - wartości oczekiwane zmiennych losowych odpowiednio : $\underline{u}$ i $\underline{y}$

Warunkowa funkcja gęstości prawdopodobieństwa ma postać:

$$f_y(y|u) = (2\pi)^{-\frac{L}{2}} \left|\Sigma_{y|u}^{-1}\right|^{\frac{1}{2}} \exp\left[-\frac{1}{2}\left(y - m_{y|u}\right)^T \Sigma_{y|u}^{-1}\left(y - m_{y|u}\right)\right]$$

Pełna informacja probabilistyczna

- Przykład regresja I rodzaju dla rozkładu normalnego

$$f(y|u) = (2\pi)^{-\frac{L}{2}} \left|\Sigma_{y|u}^{-1}\right|^{\frac{1}{2}} \exp\left[-\frac{1}{2}\left(y - m_{y|u}\right)^T \Sigma_{y|u}^{-1}\left(y - m_{y|u}\right)\right]$$

Macierz kowariancji: $\Sigma_{y|u} = \Sigma_{yy} - \Sigma_{yu}\Sigma_{uu}^{-1}\Sigma_{uy}$

Wartość oczekiwana: $m_{y|u} = m_y + \Sigma_{yu}\Sigma_{uu}^{-1}(u - m_u)$

Optymalny model: $\bar{y} = \Phi^*(u) = \underset{\underline{y}}{E}\left[\underline{y}|\underline{u} = u\right] = m_{y|u} = m_y + \Sigma_{yu}\Sigma_{uu}^{-1}(u - m_u)$

$$\bar{y} = \Phi^*(u) = Au + b$$

gdzie: $A = \Sigma_{yu}\Sigma_{uu}^{-1}$

$$b = m_y - \Sigma_{yu}\Sigma_{uu}^{-1}m_u$$

Pełna informacja probabilistyczna

Model: $\bar{y} = \Phi(u, \theta)$ **- szukamy wektora parametrów** θ

Φ zadana funkcja, θ – wektora parametrów modelu

$$u \in \mathscr{U} \subseteq \mathscr{R}^S, \quad \bar{y} \in \overline{\mathscr{Y}} \subseteq \mathscr{R}^L, \quad \theta \in \Theta \subseteq \mathscr{R}^R$$

Wskaźnik jakości: $$Q(\theta) = \underset{\underline{u},\underline{y}}{E}\left[q\left(\underline{y}, \Phi(\underline{u}, \theta)\right)\right] = \int_{\mathscr{U}}\int_{\mathscr{Y}} q(y, \Phi(u, \theta)) f(u, y)\, dy\, du$$

$$\theta^* \rightarrow \quad Q(\theta^*) = \min_{\theta \in \Theta} Q(\theta)$$

Optimalny model: $\bar{y} = \Phi(u, \theta^*)$

Pełna informacja probabilistyczna

- Regresja II rodzaju

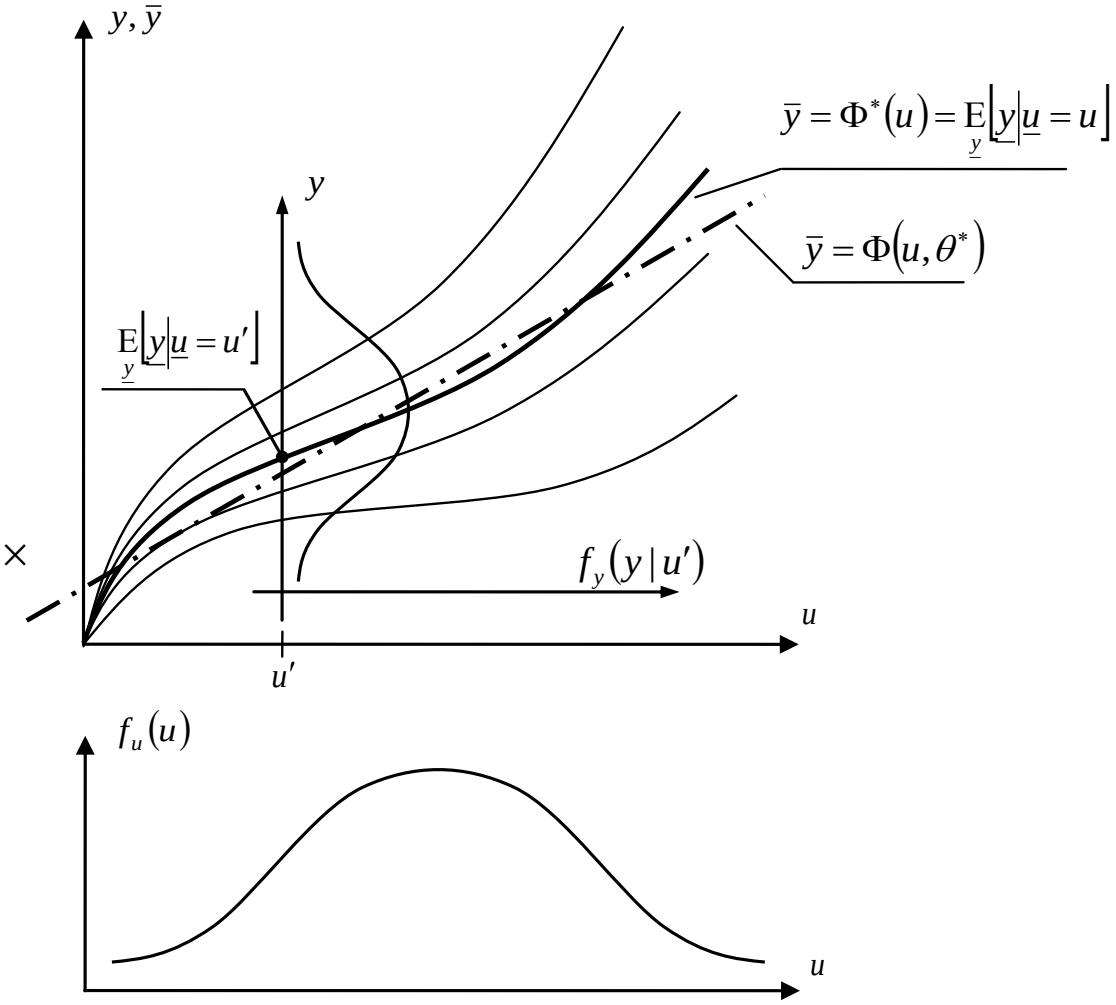
Dla:

$$q(y,\bar{y}) = [y-\bar{y}]^T [y-\bar{y}]$$

$$Q(\theta) = \int_{\mathcal{U}}\int_{\mathcal{Y}} [y-\Phi(u,\theta)]^T [y-\Phi(u,\theta)] \times f(u,y)\,dy\,du$$

$$\theta^* \rightarrow \quad Q(\theta^*) = \min_{\theta\in\Theta} Q(\theta)$$

$$\bar{y} = \Phi(u,\theta^*)$$



Pełna informacja probabilistyczna

Czy jest jakaś relacja pomiędzy regresją I rodzaju i II rodzaju ?

$$\Phi^*(u) \sim ? \sim \Phi(u, \theta^*)$$

Wskaźnik jakości:

$$Q(\theta) = \underset{\underline{u},\underline{y}}{E}\left[\left(\underline{y} - \Phi(\underline{u}, \theta)\right)^2\right] = \int_{\mathcal{U}}\int_{\mathcal{Y}} \left(y - \Phi(u, \theta)\right)^2 f(u, y)\, dy\, du$$

Przyjmując:

$$\left(\Phi^*(u) - \Phi^*(u) = 0\right) \text{ oraz } f(u, y) = f_y(y \mid u)\, f_u(u) \text{ otrzymujemy:}$$

$$Q(\theta) = \int_{\mathcal{U}}\int_{\mathcal{Y}} \left(\left(y - \Phi^*(u)\right) + \left(\Phi^*(u) - \Phi(u, \theta)\right)\right)^2 f_y(y|u) f_u(u)\, dy\, du$$

Pełna informacja probabilistyczna

$$Q(\theta) = \int_{\mathcal{U}} \int_{\mathcal{Y}} \left(y - \Phi^*(u)\right)^2 f_y(y|u) f_u(u) dy\, du +$$

$$+ 2 \int_{\mathcal{U}} \left[\int_{\mathcal{Y}} \left(y - \Phi^*(u)\right)\left(\Phi^*(u) - \Phi(u,\theta)\right) f_y(y|u) dy \right] f_u(u) du +$$

$$+ \int_{\mathcal{U}} \int_{\mathcal{Y}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 f_y(y|u) f_u(u) dy\, du =$$

$$= \int_{\mathcal{U}} Q_u\left(\Phi^*(u)\right) f_u(u) du +$$

$$+ 2 \int_{\mathcal{U}} \int_{\mathcal{Y}} \left(y - \Phi^*(u)\right) f_y(y|u) dy \left(\Phi^*(u) - \Phi(u,\theta)\right) f_u(u) du +$$

$$+ \int_{\mathcal{U}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 \int_{\mathcal{Y}} f_y(y|u) dy\, f_u(u)\, du$$

Pełna informacja probabilistyczna

$$Q(\theta) = \sigma_y^2 +$$

$$+ 2\int_{\mathcal{U}} \left[\int_{\mathcal{Y}} y f_y(y|u)dy - \Phi^*(u) \int_{\mathcal{Y}} f_y(y|u)dy \right] \left(\Phi^*(u) - \Phi(u,\theta)\right) f_u(u)du +$$

$$+ \int_{\mathcal{U}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 f_u(u)\, du =$$

$$= \sigma_y^2 + 2\int_{\mathcal{U}} \left[\Phi^*(u) - \Phi^*(u)\cdot 1\right]\left(\Phi^*(u) - \Phi(u,\theta)\right) f_u(u)du +$$

$$+ \int_{\mathcal{U}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 f_u(u)\, du =$$

$$= \sigma_y^2 + \int_{\mathcal{U}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 f_u(u)\, du$$

Pełna informacja probabilistyczna

$$Q(\theta) = \int_{\mathcal{U}} \int_{\mathcal{Y}} \left(y - \Phi^*(u)\right)^2 f(u,y)\,dy\,du + \int_{\mathcal{U}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 f_u(u)\,du$$

$$\theta^* \to \min_\theta Q(\theta) = \min_\theta \int_{\mathcal{U}} \left(\Phi^*(u) - \Phi(u,\theta)\right)^2 f_u(u)\,du$$

Regresja I rodzaju

Funkcja wagi

Regresja II rodzaju jest najlepszym przybliżeniem regresji I rodzaju z funkcją wagi będącą gęstością rozkładu prawdopodobieństwa wejścia

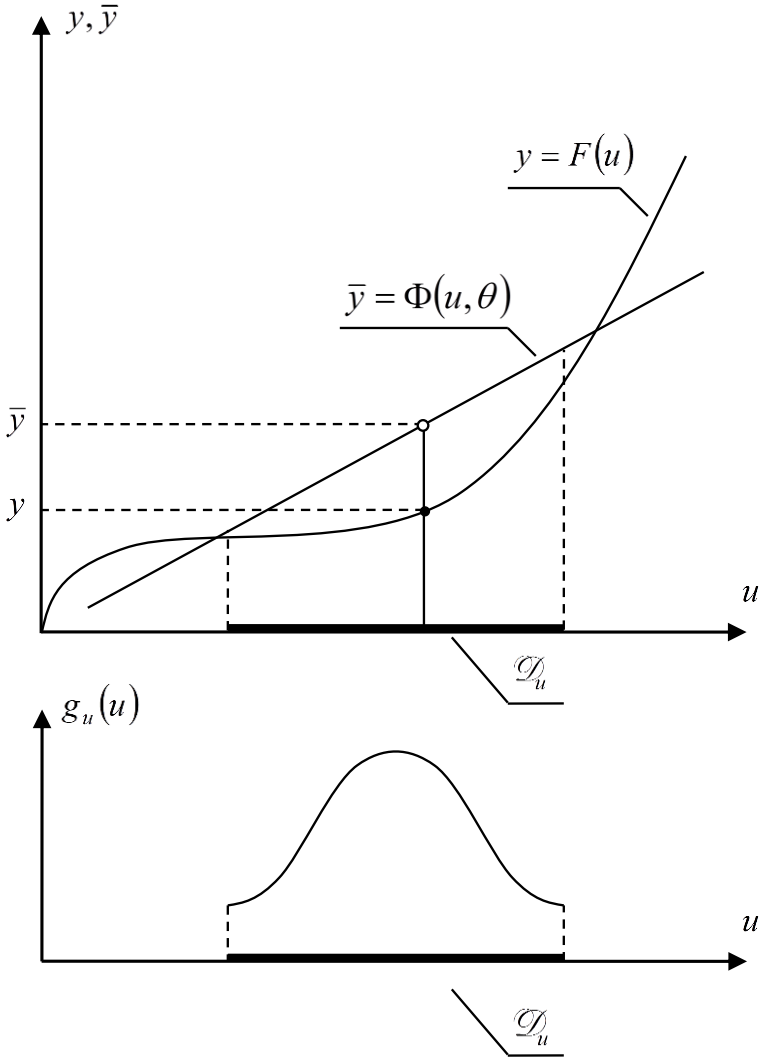
Zadanie aproksymacji charakterystyki statycznej obiektu

- Sformułowanie problemu

$$Q(\theta) = \int_{\mathscr{D}_u} q\left(F(u), \Phi(u,\theta)\right) g_u(u)\, du$$

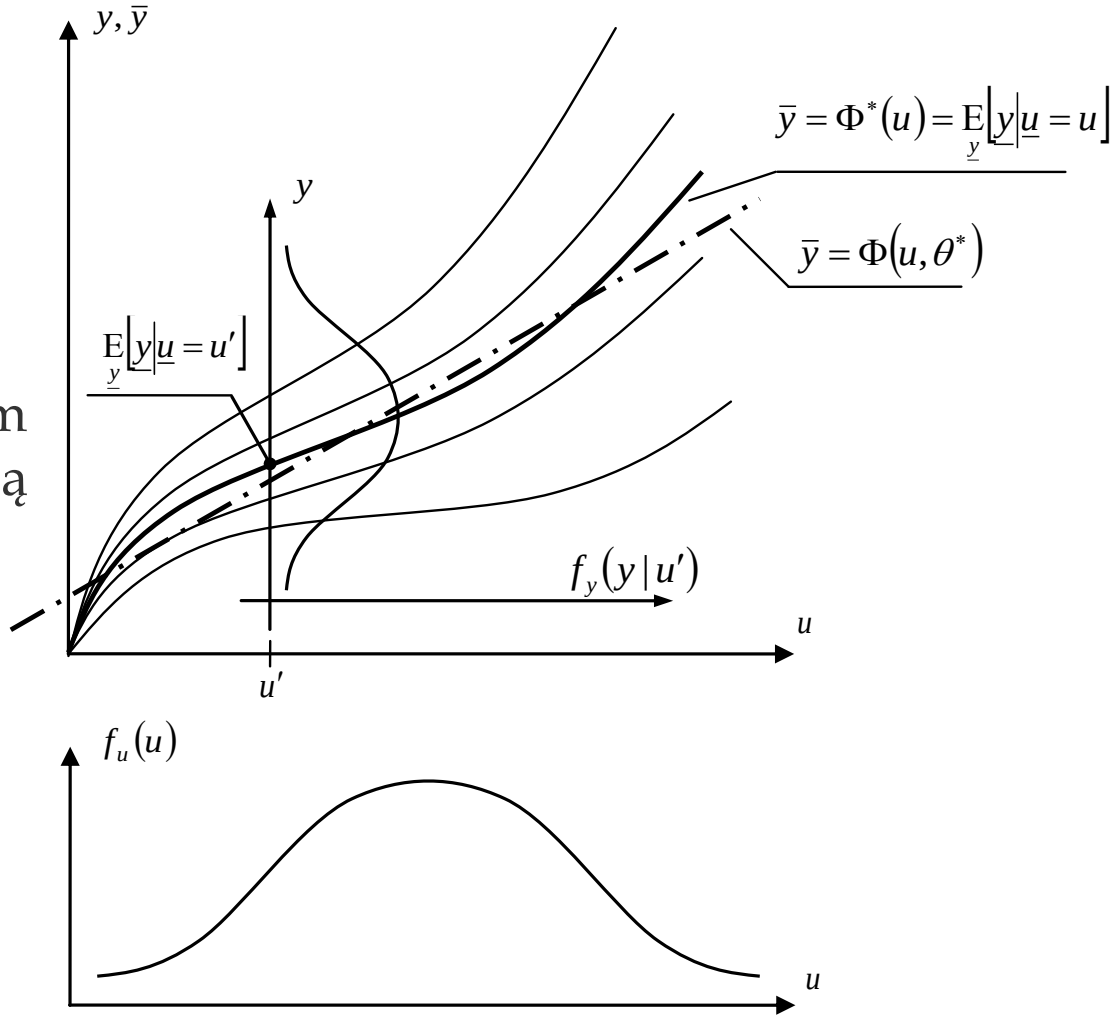
$$\theta_N^* \rightarrow Q_N\left(\theta_N^*\right) = \min_{\theta} Q_N(\theta)$$

$$\bar{y} = \Phi(u, \theta_N^*)$$



Pełna informacja probabilistyczna

- Relacja pomiędzy regresją I rodzaju a regresją II rodzaju
  - Regresja II rodzaju jest najlepszym przybliżeniem regresji I rodzaju z funkcją wagi będącą gęstością rozkładu prawdopodobieństwa wejścia



Wybór optymalnego modelu w warunkach losowych

Dwa możliwe przypadki

Pełna informacja probabilistyczna

łączna gęstość rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ $f(u, y)$

lub
warunkowa gęstość rozkładu prawdopodobieństwa $f_y(y|u)$

i brzegowa $f_u(u)$

są znane

Niepełna informacja probabilistyczna

łączna gęstość rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ istnieje ale nie jest znana
W zamian dysponujemy pomiarami:

$$(u_n, y_n),\ n = 1, 2, \ldots, N$$

które są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$

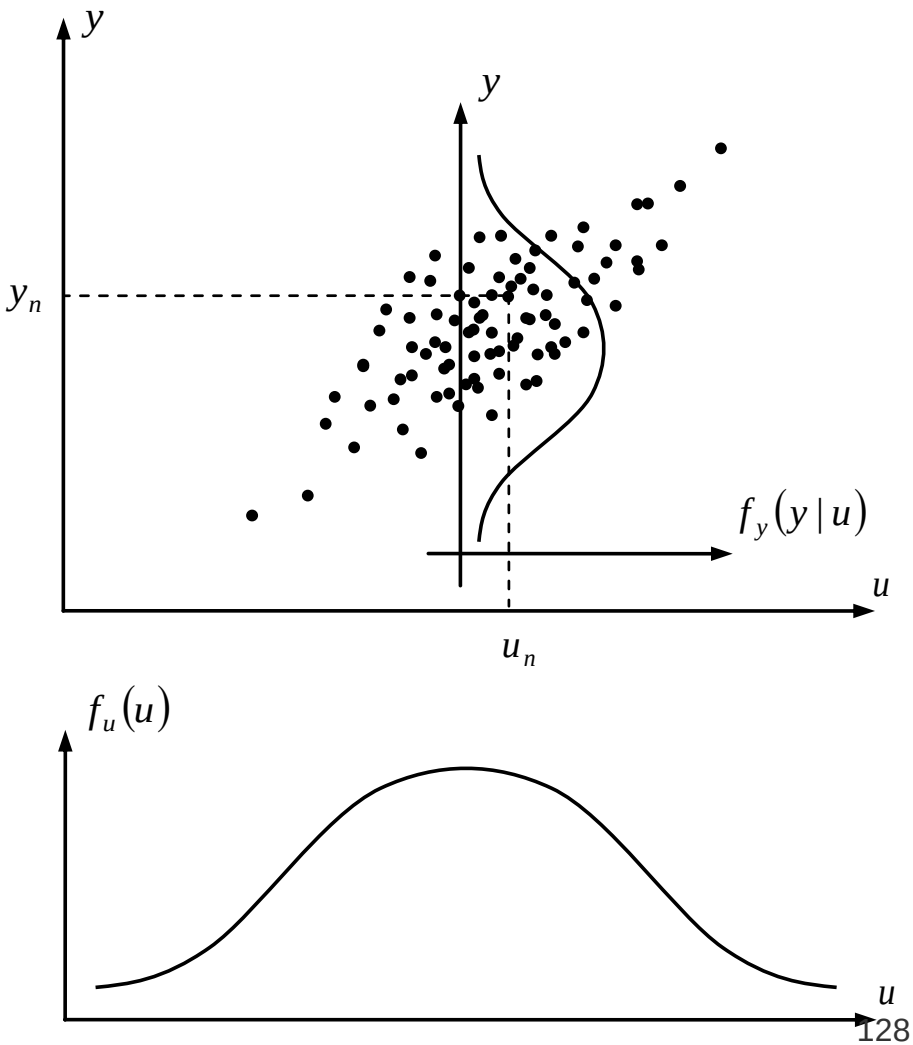
Niepełna informacja probabilistyczna

Łączna gęstość rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ istnieje ale nie jest znana

W zamian dysponujemy pomiarami:

$(u_n, y_n), n = 1, 2, \ldots, N$

które są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$



Typowe podejścia

- Estymacja wskaźników jakości
- Estymacja parametrów rozkładu prawdopodobieństwa
- Nieparametryczna estymacja gęstości rozkładów prawdopodobieństwa

Niepełna informacja probabilistyczna
Empiryczne oszacowanie wskaźników jakości

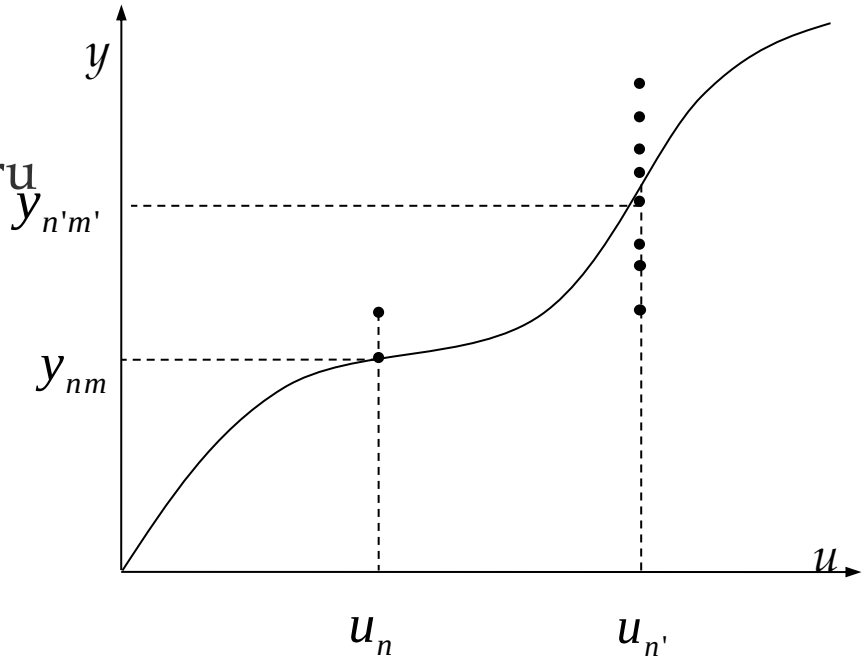
$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[q(\underline{y}, \Phi(\underline{u}))\,|\,\underline{u} = u\right] = \int_{\mathscr{Y}} q(y, \Phi(u)) f_y(y\,|\,u)\,dy$$

Eksperyment:

Dla ustalonego wejścia: u_n dokonujemy wielokrotnego pomiaru wyjścia y_{nm}.

W wyniku otrzymujemy:

$$(u_n, y_{nm}),\ m = 1, 2, \ldots, M_n,\ n = 1, 2, \ldots, \overline{N}$$



gdzie:

M_n – liczba pomiarów wyjścia dla ustalonego wejścia u_n.

$\overline{N}$ – liczba różnych wartości wejść

$N = \sum_{n=1}^{\overline{N}} M_n$ – liczba wszystkich pomiarów

Niepełna informacja probabilistyczna
Empiryczne oszacowanie wskaźników jakości

Oszacowanie wskaźnika jakości:

$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[q(\underline{y}, \Phi(\underline{u}))\,|\,\underline{u} = u\right] = \int_{\mathscr{Y}} q(y, \Phi(u)) f_y(y\,|\,u)\,dy$$

dla $u = u_n, \quad n = 1, 2, \ldots, \overline{N}$

$$Q_{uM_n}(\Phi(u_n)) = \frac{1}{M_n} \sum_{m=1}^{M_n} q(y_{nm}, \Phi(u_n))$$

$$\Phi^*_{M_n}(u_n) \rightarrow \quad Q_{uM_n}\left(\Phi^*_{M_n}(u_n)\right) = \min_{\Phi(u_n)} Q_{uM_n}(\Phi(u_n))$$

$$\Phi^*_{M_n}(u_n) \qquad n = 1, 2, \ldots, \overline{N} \qquad\qquad N = \sum_{n=1}^{\overline{N}} M_n$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie wskaźników jakości

Dla $$q(y,\bar{y}) = [y-\bar{y}]^T [y-\bar{y}]$$

Wskaźnik jakości:

$$Q_u(\Phi(u)) = \underset{\underline{y}}{E}\left[[y-\Phi(u)]^T [y-\Phi(u)] \middle| \underline{u}=u\right] = \int_{\mathcal{Y}} [y-\Phi(u)]^T [y-\Phi(u)] f_y(y|u)\, dy$$

dla ustalonego $u = u_n$

$$Q_{uM_n}(\Phi(u_n)) = \frac{1}{M_n}\sum_{m=1}^{M_n} [y_{nm}-\Phi(u_n)]^T [y_{nm}-\Phi(u_n)]$$

Estymacja funkcji regresji I rodzaju:

$$\bar{y} = \Phi^*(u) = \underset{\underline{y}}{E}\left[\underline{y} \middle| \underline{u}=u\right] = \int_{\mathcal{Y}} y\, f(y|u)\, dy$$

$$\bar{y}_n = \frac{1}{M_n}\sum_{m=1}^{M_n} y_{nm},\ n = 1, 2, \ldots \bar{N}$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie wskaźników jakości

Oszacowanie wskaźnika jakości:

$$Q(\theta) = \underset{\underline{u},\underline{y}}{E}\left[(y - \Phi(u,\theta))^2\right] = \iint_{\mathcal{UY}} (y - \Phi(u,\theta))^2 \, f(u,y)\,dy\,du$$

$$Q_N(\theta) = \frac{1}{N}\sum_{n=1}^{N} q(y_n, \Phi(u_n\theta))$$

lub

$$Q_N(\theta) = \frac{1}{N}\sum_{n=1}^{\bar{N}}\sum_{m=1}^{M_n} q(y_{nm}, \Phi(u_n, \theta))$$

$$\theta_N^* \to Q_N(\theta_N^*) = \min_{\theta\in\Theta} Q_N(\theta)$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Estymacja parametrów gęstości rozkładu prawdopodobieństwa

Parametry gęstości rozkładu prawdopodobieństwa nie są znane

$$f(u,y) = f_\alpha(u,y,\alpha)$$

f_α - known function, $\alpha \in \mathcal{A}$ - nieznany wektor parametrów

Funkcja wiarogodności

$$L(U_N, Y_N, \alpha) = \prod_{n=1}^{N} f_\alpha(u_n, y_n, \alpha)$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Oszacowanie:

$$\alpha_N = \Psi_N(U_N, Y_N) \quad \rightarrow \quad L(U_N, Y_N, \alpha_N) = \max_{\alpha \in \mathscr{A}} L(U_N, Y_N, \alpha)$$

Wskaźnik jakości:

$$Q_{u\alpha_N}(\Phi(u), \alpha_N) = \int_{\mathscr{Y}} q(y, \Phi(u)) \frac{f_\alpha(u, y, \alpha_N)}{\int_{\mathscr{Y}} f_\alpha(u, y, \alpha_N) dy} dy$$

$$\Phi^*_{\alpha_N}(u) \rightarrow Q_{u\alpha_N}\left(\Phi^*_{\alpha_N}(u), \alpha_N\right) = \min_{\Phi(u)} Q_{u\alpha_N}\left(\Phi(u), \alpha_N\right)$$

Regresja I rodzaju

$$\bar{y} = \Phi^*_{\alpha_N}(u, \alpha_N) = \int_{\mathscr{Y}} y \frac{f_\alpha(u, y, \alpha_N)}{\int_{\mathscr{Y}} f_\alpha(u, y, \alpha_N) dy} dy$$

Wskaźnik jakości:

$$Q_{\alpha_N}(\theta, \alpha_N) = \int_{\mathscr{U}} \int_{\mathscr{Y}} q(y, \Phi(u, \theta)) f_\alpha(u, y, \alpha_N) dy\, du$$

$$\theta^*_{\alpha_N} \rightarrow \quad Q_{\alpha_N}\left(\theta^*_{\alpha_N}, \alpha_N\right) = \min_{\theta \in \Theta} Q_{\alpha_N}(\theta, \alpha_N)$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

$\underline{x}$ – zmienna losowa, funkcja gęstości rozkładu prawdopodobieństwa $f_x(x)$ istnieje lecz nie jest znana. dim $x = P$

Obserwacje: $[x_1\ x_2\ \cdots\ x_N] = X_N$

Estymator Parzena :

$$f_x(x) \approx f_{xN}(x; X_N) = \frac{1}{Nh^P(N)} \sum_{n=1}^{N} K_x\left(\frac{x - x_n}{h(N)}\right)$$

gdzie:

$$h(N) > 0,\ \lim_{N\to\infty} h(N) = 0,\ \lim_{N\to\infty} h^P(N) = \infty$$

przykład $h(N) = cN^{-\alpha}$ $\quad c > 0, 0 < \alpha < \frac{1}{P}$

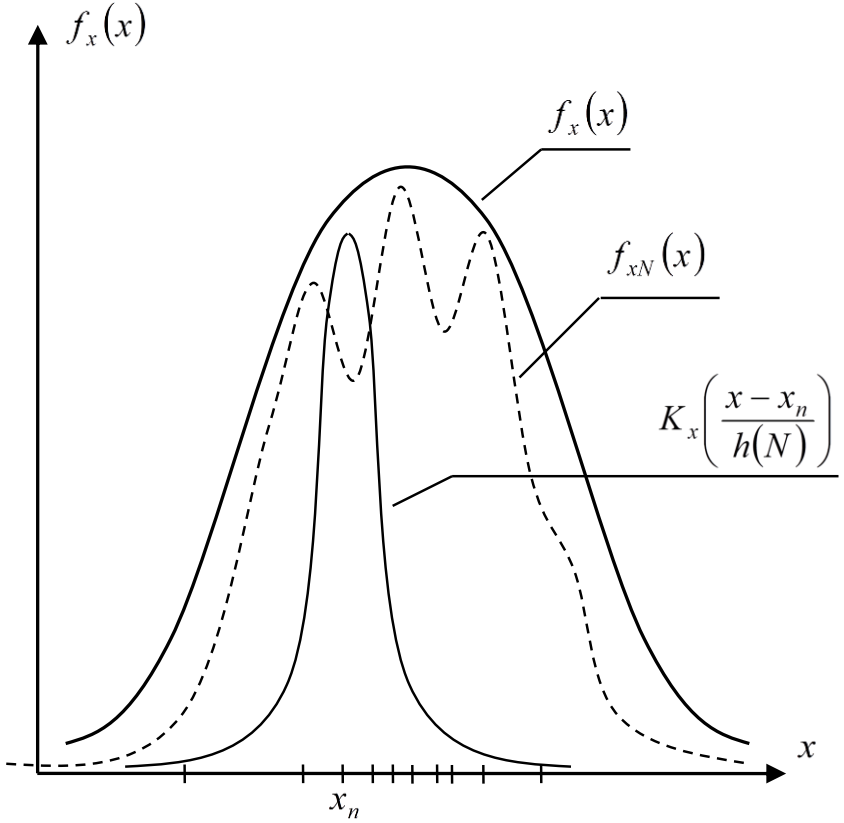
K_x – znana funkcja jądra, taka że:

$$\sup_{x\in\mathcal{X}} |K_x(x)| < \infty,\ \int_{\mathcal{X}} K_x(x)dx = 1,\ \int_{\mathcal{X}} |K_x(x)|dx < \infty,\ \lim_{\|x\|\to\infty} |K_x(x)| = 0$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Przykłady funkcji jądra K_x

$$K_x(x)=\begin{cases}1 \; dla \; \|x\| \le 1\\ 0 \; dla \; \|x\| > 1\end{cases}, \quad K_x(x)=(2\pi)^{-\frac{P}{2}} e^{-\frac{1}{2}\|x\|^2}, \quad K_x(x)=\pi^{-P}\prod_{p=1}^{P}\sqrt{1+\left|x^{(p)}\right|}$$



Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Estymator Parzena funkcji gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$:

$$f_N(u, y; U_N, Y_N) = \frac{1}{N\, h^{L+S}(N)} \sum_{n=1}^{N} K_u\left(\frac{u-u_n}{h(N)}\right) K_y\left(\frac{y-y_n}{h(N)}\right)$$

Estymator Parzena funkcji gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{u}$:

$$f_{uN}(u; U_N) = \frac{1}{N\, h^{S}(N)} \sum_{n=1}^{N} K_u\left(\frac{u-u_n}{h(N)}\right)$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Estymator Parzena funkcji gęstości rozkładu warunkowego prawdopodobieństwa $f_y(y|u)$:

$$f_{yN}(y|u;U_N,Y_N)=\frac{f_N(u,y;U_N,Y_N)}{f_{uN}(u;U_N)}=$$

$$=\frac{1}{h^L(N)}\frac{\sum_{n=1}^{N}K_u\left(\frac{u-u_n}{h(N)}\right)K_y\left(\frac{y-y_n}{h(N)}\right)}{\sum_{n=1}^{N}K_u\left(\frac{u-u_n}{h(N)}\right)}$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Konsekwentnie:

$$Q_{uN}(\Phi(u);U_N,Y_N)=\int_{\mathscr{Y}} q(y,\Phi(u))\,f_{yN}(y|u;U_N,Y_N)\,dy$$

$$\bar{y}=\Phi_N^*(u;U_N,Y_N)\rightarrow Q_{uN}(\Phi_N^*(u,U_N,Y_N),U_N,Y_N)=\min_{\Phi(u)} Q_{uN}(\Phi(u),U_N,Y_N)$$

Dla: $q(y,\bar{y})=[y-\bar{y}]^T[y-\bar{y}]$

Regresja I rodzaju

$$\bar{y}=\Phi_N^*(u;U_N,Y_N)=\int_{\mathscr{Y}} y\ f_{yN}(y|u;U_N,Y_N)\,dy=$$

$$=\frac{1}{h^L(N)}\frac{\sum_{n=1}^{N}K_u\left(\frac{u-u_n}{h(N)}\right)\int_{\mathscr{Y}} y\,K_y\left(\frac{y-y_n}{h(N)}\right)dy}{\sum_{n=1}^{N}K_u\left(\frac{u-u_n}{h(N)}\right)}$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Dla specjalnej postaci funkcji $K_y(y)$, takiej że $K_y(y)=K_y(-y)$ mamy:

$$\frac{1}{h^L(N)}\int_{\mathscr{Y}} y\,K_y\left(\frac{y-y_n}{h(N)}\right)dy=y_n$$

Wówczas regresja I rodzaju przyjmuje postać:

$$\bar{y}=\Phi_N^*(u;U_N,Y_N)=\frac{\sum_{n=1}^{N} y_n K_u\left(\frac{u-u_n}{h(N)}\right)}{\sum_{n=1}^{N} K_u\left(\frac{u-u_n}{h(N)}\right)}$$

Niepełna informacja probabilistyczna
Empiryczne oszacowanie rozkładów prawdopodobieństwa

Optimalny model:

$$Q_N(\theta; U_N, Y_N) = \int_{\mathcal{U}} \int_{\mathcal{Y}} q(y, \Phi(u, \theta)) f_N(u, y; U_N, Y_N)\, dy\, du$$

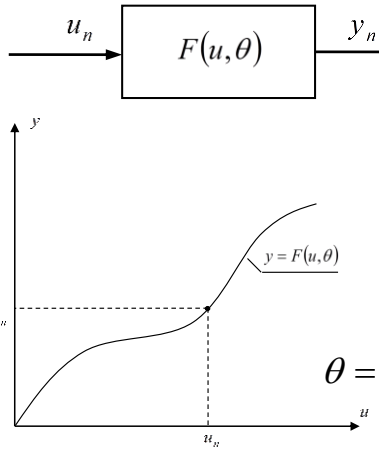
$$f_N(u, y; U_N, Y_N) = \frac{1}{N h^{L+S}(N)} \sum_{n=1}^{N} K_u\left(\frac{u - u_n}{h(N)}\right) K_y\left(\frac{y - y_n}{h(N)}\right)$$

$$\theta_N^*(U_N, Y_N) \rightarrow \quad Q_N(\theta_N^*; U_N, Y_N) = \min_{\theta \in \Theta} Q_N(\theta; U_N, Y_N)$$

Podstawowe zadania identyfikacji podsumowanie

Deterministyczny

Obiekt w klasie modeli



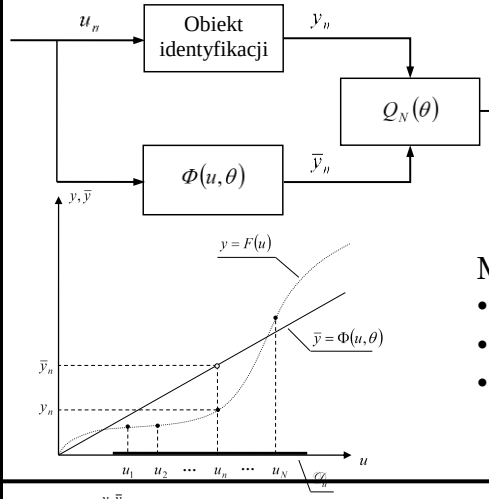
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

$$y_n = F(u_n, \theta),$$
$$n = 1, 2, \ldots, N$$

$$Y_N = \bar{F}(U_N, \theta)$$

$$\theta = \bar{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

Wybór optymalnego modelu

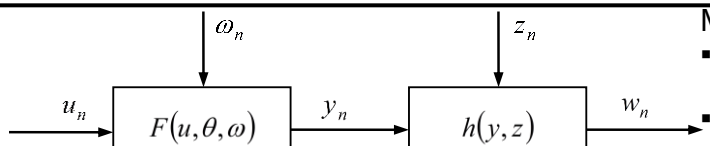


$$\bar{y} = \Phi(u, \theta_N^*)$$
$$\theta_N^* \to Q_N(\theta_N^*) = \min_\theta Q_N(\theta)$$
$$\theta_N^* = \Psi_N(U_N, Y_N)$$
$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N],$$
$$Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]$$

Model jest optymalny:

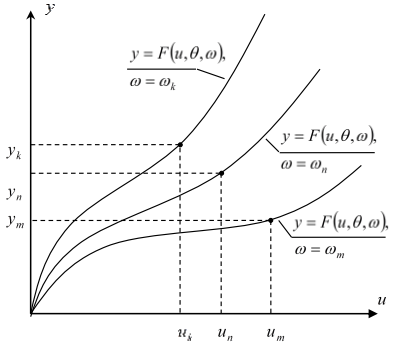
- dla zadanej serii pomiarowej
- przyjętego modelu
- przyjętego wskaźnika jakości identyfikacji

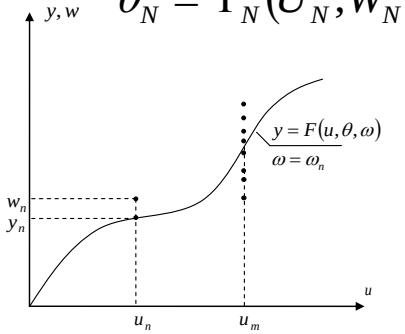
Losowy



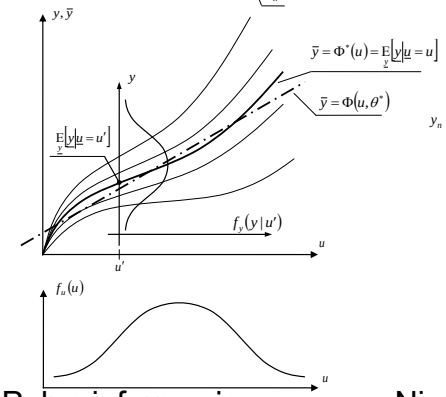
Metody:

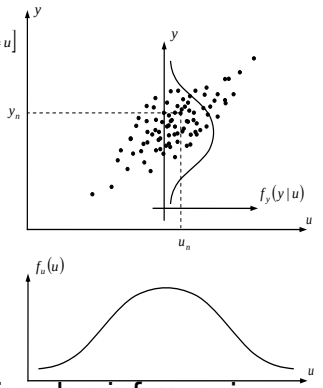
- Najmniejszych kwadratów
- Maksymalnej wiarogodności
- Bayes'a





$$\theta_N = \Psi_N(U_N, W_N)$$





Pełna informacja

- Regresja I rodzaju
- Regresja II rodzaju

Niepełna informacja

- Estymacja wskaźnika jakości
- Estymacja parametrów rozkładu
- Estymacja rozkładu

Dziękuję za uwagę

