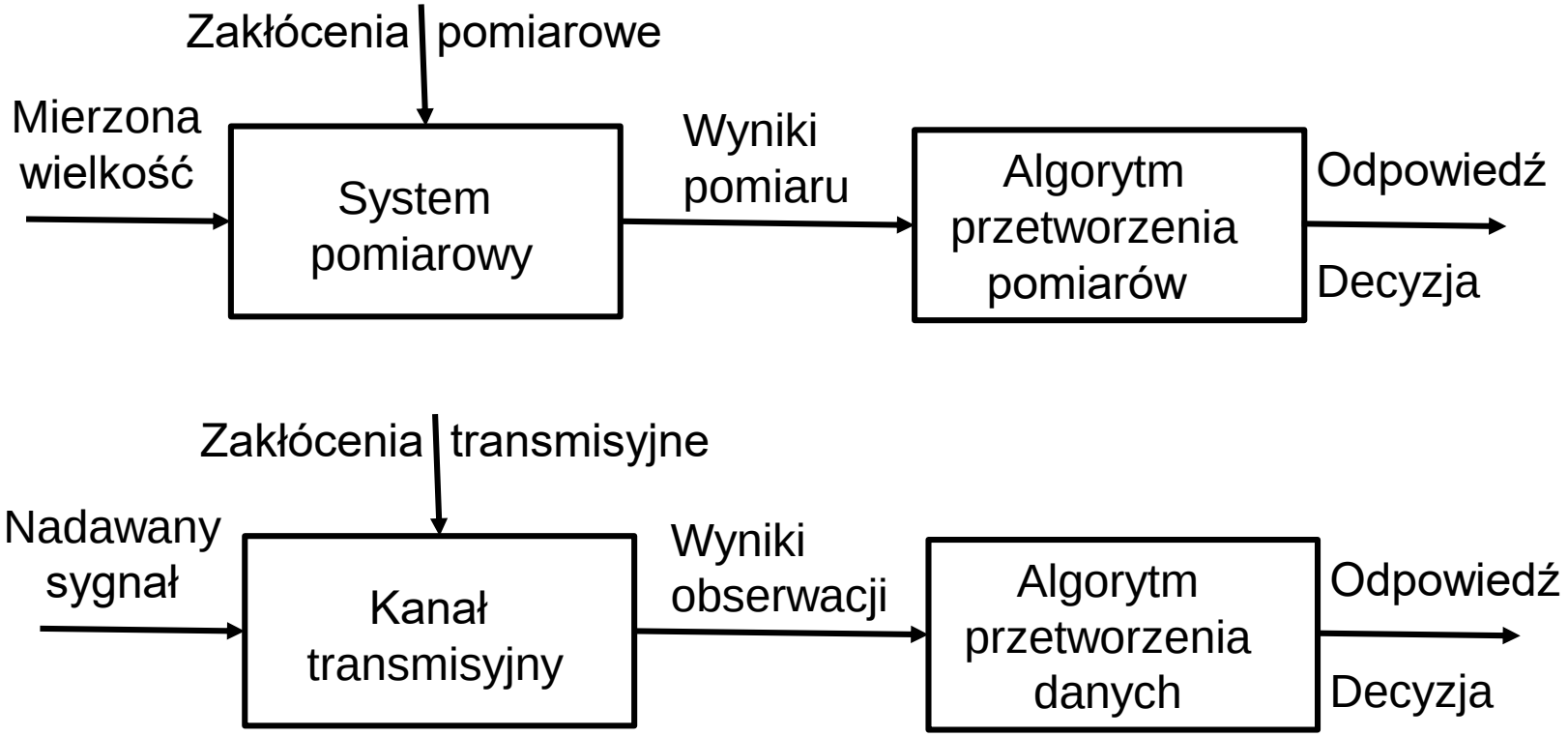
Metody Systemowe i Decyzyjne w Informatyce

Wykład 8a. Optymalny pomiar

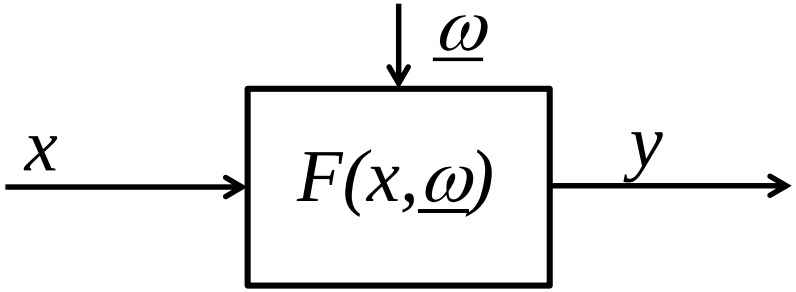
Zakłócony pomiar wielkości fizycznej

- Sformułowanie problemu



Niepewność – zmienna losowa

Podejmowanie decyzji w warunkach niepewności



$$F(x) = \underset{\omega}{E}\left[F(x,\omega)\right]$$

$$\mathscr{D}_x = \underset{\omega}{E}\left[\mathscr{D}_x(\omega)\right] =$$

$$= \left\{ x \in \mathscr{R}^S ;\ \underset{\omega}{E}\left[\varphi_l(x,\omega)\right] = 0,\ l = 1,\ldots,L,\ \underset{\omega}{E}\left[\psi_m(x,\omega)\right] \le 0,\ m = 1,\ldots,M \right\}$$

$$x^* \rightarrow F(x^*) = \min_{x \in D_x} F(x)$$

Podejmowanie decyzji w warunkach niepewności

ω - wartość zmiennej losowej $\underline{\omega}$

$\omega \in \Omega$ - zbiór ciągły, $f_{\omega}(\omega)$ - funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{\omega}$

Wówczas:

$$F(x) = \underset{\underline{\omega}}{E}\left[F(x,\underline{\omega})\right] = \int_{\Omega} F(x,\omega) f_{\omega}(\omega) d\omega$$

$$\mathscr{D}_x = \underset{\underline{\omega}}{E}\left[\mathscr{D}_x(\underline{\omega})\right] =$$

$$= \left\{ x \in \mathscr{R}^S ; \underset{\underline{\omega}}{E}\left[\varphi_l(x,\underline{\omega})\right] = \int_{\Omega} \varphi_l(x,\omega) f_{\omega}(\omega) d\omega = 0, l = 1,\ldots, L, \right.$$

$$\left. \underset{\underline{\omega}}{E}\left[\psi_m(x,\underline{\omega})\right] = \int_{\Omega} \psi_m(x,\omega) f_{\omega}(\omega) d\omega \le 0, m = 1,\ldots, M \right\}$$

Podejmowanie decyzji w warunkach niepewności

ω - wartość zmiennej losowej $\underline{\omega}$

$\omega \in \Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$ - zbiór dyskretny,

$P(\underline{\omega} = \omega_k) = p_k, k = 1, 2, \ldots, K$ - gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{\omega}$

Wówczas:

$$F(x) = \underset{\underline{\omega}}{E}[F(x, \underline{\omega})] = \sum_{k=1}^{K} F(x, \omega_k) p_k$$

$$\mathscr{D}_x = \underset{\underline{\omega}}{E}[\mathscr{D}_x(\underline{\omega})] =$$

$$= \left\{ x \in \mathscr{R}^S ; \underset{\underline{\omega}}{E}[\varphi_l(x, \underline{\omega})] = \sum_{k=1}^{K} \varphi_l(x, \omega_k) p_k = 0, l = 1, \ldots, L, \right.$$

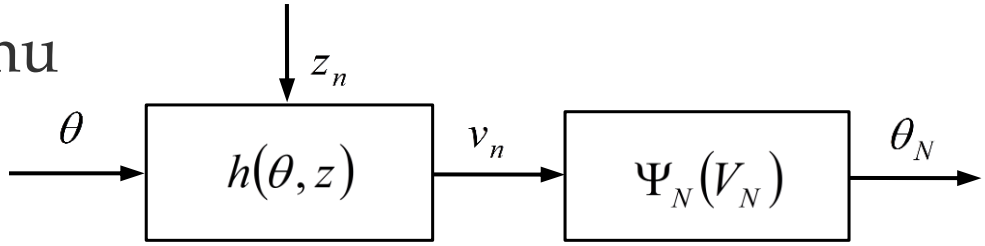
$$\left. \underset{\underline{\omega}}{E}[\psi_m(x, \underline{\omega})] = \sum_{k=1}^{K} \psi_m(x, \omega_k) p_k \leq 0, m = 1, \ldots, M \right\}$$

Zakłócony pomiar wielkości fizycznej

7

Zakłócony pomiar wielkości fizycznej

- Sformułowanie problemu



Opis systemu pomiarowego: $v = h(\theta, z)$

gdzie: $v \in \mathscr{V}$, h – znana funkcja, wzajemnie jednoznaczna względem z

$h : \Theta \times \mathscr{Z} \to \mathscr{V}$, $z = h_z^{-1}(\theta, v)$

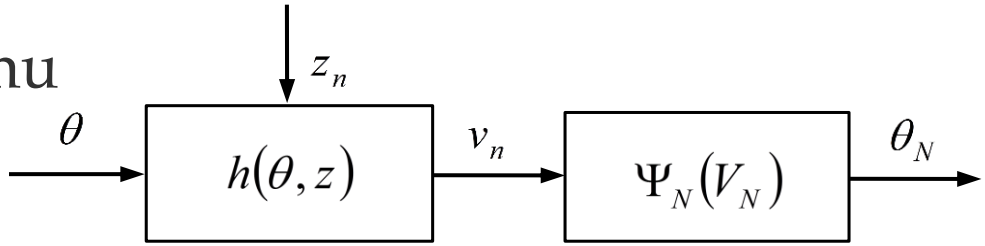
Przykłady funkcji h: $v = h(\theta, z) = \theta + z$

$v = h(\theta, z) = \theta \cdot z$

$\mathscr{V}$ – przestrzeń pomiarów($\dim \theta = \dim z = L$)

Zakłócony pomiar wielkości fizycznej

- Sformułowanie problemu



Zakłócenia pomiarowe:

z_n – wartość zmiennej losowej $\underline{z}$ z przestrzeni $\mathcal{Z}$

$f_z(z)$ – funkcja gęstości rozkładu prawdopodobieństwa zmiennej $\underline{z}$

θ – obserwowany wektor wielkości, wartość zmiennej losowej $\underline{\theta}$, $\theta \in \Theta \subseteq \mathcal{R}^R$

$f_\theta(\theta)$ – funkcja gęstości rozkładu prawdopodobieństwa zmiennej $\underline{\theta}$

Wyniki pomiarów: $V_N = [v_1 \quad v_2 \quad \cdots \quad v_N]$

Zakłócony pomiar wielkości fizycznej

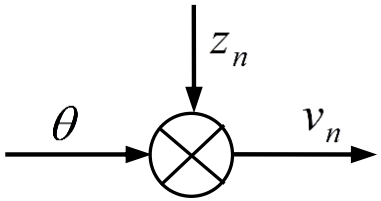
Poszukujemy algorytmu estymacji:

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

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

Metoda najmniejszych kwadratów

Założenia:



$v = h(\theta, z) = \theta + z$ – zakłócenia addytywne

$\underset{z}{E}[z] = 0$ – wartość oczekiwana zakłóceń wynosi zero

$\underset{z}{Var}[z] < \infty$ – skończona wariancja zakłóceń

Empiryczna wariancja zakłóceń:

$$Var_{zN}(V_N, \theta) = \frac{1}{N}\sum_{n=1}^{N}(v_n - \theta)^2$$

Algorytm estymacji otrzymujemy minimalizując empiryczną wariancje zakłóceń:

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

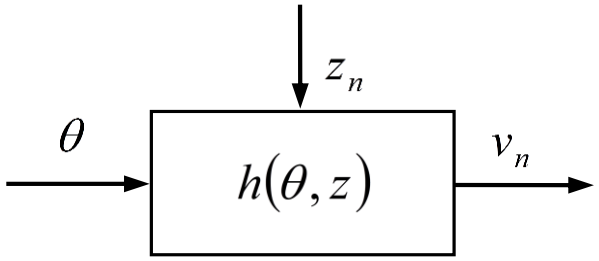
Algorytm estymacji:

$$\theta_N = \frac{1}{N}\sum_{n=1}^{N} v_n$$

Metoda maksymalnej wiarogodności

Założenia:

$\underline{v} = h(\theta, \underline{z})$ – system pomiarowy opisany jest dowolną funkcją *h*, wzajemnie jednoznaczną względem *z*



Postać funkcji gęstości rozkładu prawdopodobieństwa zakłóceń $f_z(z)$ jest znana

Funkcja gęstości rozkładu prawdopodobieństwa obserwowanej zmiennej losowej $\underline{v}$

$$f_v(v,\theta) = f_z\left(h_z^{-1}(\theta, v)\right) \cdot |J_h|, \qquad \text{gdzie: } J_h = \frac{\partial h_z^{-1}(\theta, v)}{\partial v}$$

gdzie: J_h - Jacobian, macierz przekształcenia odwrotnego.

Funkcja wiarogodności ma postać:

$$L_N(V_N, \theta) = \prod_{n=1}^{N} f_v(v_n, \theta) = \prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right) |J_h|,$$

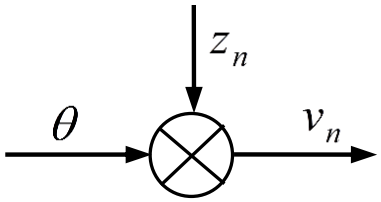
Metoda maksymalnej wiarogodności

Algorytm estymacji otrzymujemy w wyniku maksymalizacji funkcji wiarogodności:

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

Metoda maksymalnej wiarogodności

- Przykład 1



Charakterystyka systemu pomiarowego:

$$v = h(\theta, z) = \theta + z$$

Gęstość rozkładu prawdopodobieństwa zakłóceń:

$$f_z(z) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{(z - m_z)^2}{2\sigma_z^2} \right]$$

Opis systemu pomiarowego:

$$v = h(a, z) = a + z \qquad z = h_z^{-1}(\theta, v) = v - \theta$$

Macierz Jakobiego:

$$J_h = \frac{\partial h_z^{-1}(\theta, v)}{\partial v} = \frac{d}{dv}(v - \theta) = 1$$

Metoda maksymalnej wiarogodności

- Przykład 1

Gęstość rozkładu prawdopodobieństwa :

$$f_v(v,\theta)=\frac{1}{\sigma_z\sqrt{2\pi}}\exp\left[-\frac{(v-\theta-m_z)^2}{2\sigma_z^2}\right]\cdot|1|$$

Funkcja wiarogodności:

$$L_N(V_N,\theta)=\prod_{n=1}^{N}\frac{1}{\sigma_z\sqrt{2\pi}}\exp\left[-\frac{(v_n-\theta-m_z)^2}{2\sigma_z^2}\right]$$

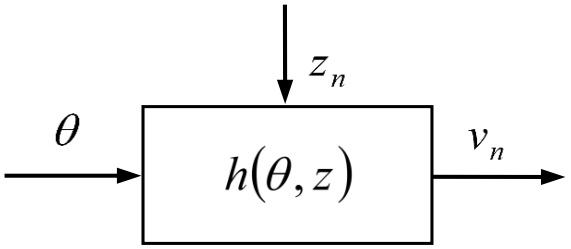
$$L_N(V_N,\theta)=\left(\frac{1}{\sigma_z\sqrt{2\pi}}\right)^N\exp\left[\sum_{n=1}^{N}-\frac{(v_n-\theta-m_z)^2}{2\sigma_z^2}\right]$$

Algorytm estymacji:

$$\theta_N=\Psi_N(V_N)=\frac{1}{N}\sum_{n=1}^{N}(v_n-m_z)$$

Metoda maksymalnej wiarogodności

- Przykład



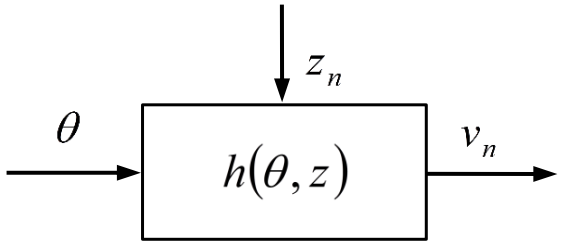
Opis zakłóceń: $$f_z(z) = \begin{cases} 1 & \text{for} \quad z \in [0,1] \\ 0 & \text{for} \quad z \notin [0,1] \end{cases}$$

Opis systemu pomiarowego: $$v = h(\theta, z) = \theta z \quad (\theta > 0)$$

$$z = h_z^{-1}(\theta, v) = \frac{v}{\theta}$$

Metoda maksymalnej wiarogodności

- Przykład



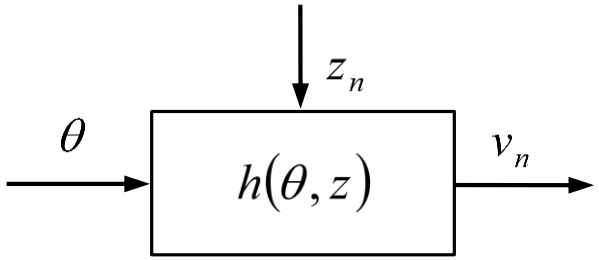
Macierz Jakobiego: $$J_h = \frac{\partial h_z^{-1}(\theta, v)}{\partial v} = \frac{d}{dv}\left(\frac{v}{\theta}\right) = \frac{1}{\theta}$$

Funkcja gęstości rozkładu prawdopodobieństwa obserwowanej zmiennej losowej:

$$f_v(v,\theta) = \begin{cases} \frac{1}{\theta} & \text{for} \quad \frac{v}{\theta} \in [0,1] \\ 0 & \text{for} \quad \frac{v}{\theta} \notin [0,1] \end{cases}$$

Metoda maksymalnej wiarogodności

- Przykład



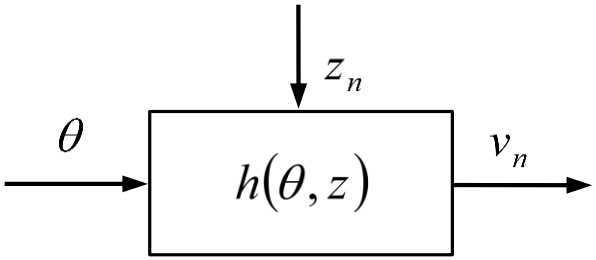
Funkcja wiarogodności:

$$L_N(V_N, \theta) = \begin{cases} \dfrac{1}{\theta^N} & \text{for} \quad \forall n = 1, 2, \ldots, N \quad v_n \in [0, \theta] \\ 0 & \text{for} \quad \exists n = 1, 2, \ldots, N \quad v_n \notin [0, \theta] \end{cases}$$

$$L_N(V_N, \theta) = \begin{cases} \dfrac{1}{\theta^N} & \text{for} \quad \theta \geq \max\limits_{1 \leq n \leq N} \{v_n\} \\ 0 & \text{for} \quad \theta < \max\limits_{1 \leq n \leq N} \{v_n\} \end{cases}$$

Metoda maksymalnej wiarogodności

- Przykład



Algorytm estymacji:

$$\theta_N = \Psi_N(V_N) = \max_{1 \le n \le N} \{v_n\}$$

Interpretacja:

$$\theta_N = \max_{1 \le n \le N} \{v_n\} = \max_{1 \le n \le N} \{\theta\, z_n\} = \theta \max_{1 \le n \le N} \{z_n\}$$

Metody Bayes'a bez obserwacji

- $\underline{\theta}$ - zmienna losowa
- θ – wartość zmiennej losowej
- $\underline{\theta}$ - zmienna losowa ciągła
- $\theta \in \Theta \subseteq \mathscr{R}^R$
- $f_\theta(\theta)$ - gęstość rozkładu prawdopodobieństwa

- $\underline{\theta}$ - zmienna losowa dyskretna
- $\theta \in \Theta \subseteq \{\theta_1, \theta_2, \cdots \theta_m, \cdots \theta_M\}$
- $P(\underline{\theta} = \theta_m) = p_m, m = 1, 2, \cdots, M$

- $\bar{\theta}$ - możliwa decyzja

- $L(\theta, \bar{\theta})$ - funkcja strat – ocena decyzji
- Przykładowe funkcje strat – minimum gdy decyzja jest trafna
- $L(\theta, \bar{\theta}) = (\theta - \bar{\theta})^2$
- $L(\theta, \bar{\theta}) = |\theta - \bar{\theta}|$
- $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$ - delta Dirac'a
- Wylosowana wartość θ nie jest znana

- Ryzyko podjęcia decyzji – wartość oczekiwana funkcji strat
- Dla zmiennych losowych ciągłych
- $R(\bar{\theta}) = \int_{\Theta} L(\theta, \bar{\theta})\, f_{\theta}(\theta) d\theta$
- Dla zmiennych losowych dyskretnych
- $R(\bar{\theta}) = \sum_{m=1}^{M} L(\theta_m, \bar{\theta})\, p_m$

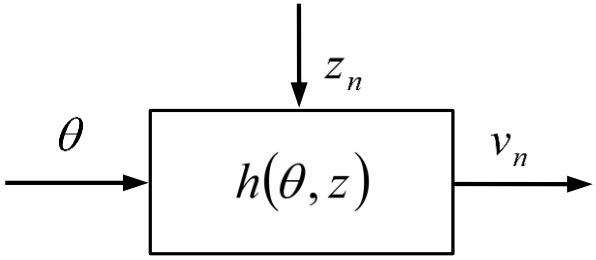
- Decyzja optymalna $-\ \theta^*$
- $R(\theta^*) = \min_{\bar{\theta} \in \Theta} R(\bar{\theta})$

- Dla różnych funkcji strat możemy otrzymać róćne decyzje
- Np.:
- $L(\theta, \bar{\theta}) = (\theta - \bar{\theta})^2$
- $R(\bar{\theta}) = \int_{\Theta} (\theta - \bar{\theta})^2 f_{\theta}(\theta) d\theta$
- $\frac{dR(\bar{\theta})}{d\bar{\theta}} = -2 \int_{\Theta} (\theta - \bar{\theta}) f_{\theta}(\theta) d\theta_{|\theta=\theta^*} = 0$
- $\theta^* \int_{\Theta} f_{\theta}(\theta) d\theta = \int_{\Theta} \theta f_{\theta}(\theta) d\theta$
- $\theta^* = \int_{\Theta} \theta f_{\theta}(\theta) d\theta$ - wartość oczekiwana zmiennej $\underline{\theta}$

- Dla funkcji strat:
- $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$
- Ryzyko
- $R(\bar{\theta}) = -\int_{\Theta} \delta(\theta - \bar{\theta}) f_{\theta}(\theta) d\theta = -f_{\theta}(\bar{\theta})$
- $R(\theta^*) = \min_{\bar{\theta} \in \Theta} R(\bar{\theta}) = \min_{\bar{\theta} \in \Theta} \left(-f_{\theta}(\bar{\theta})\right) = -\max_{\bar{\theta} \in \Theta} \left(f_{\theta}(\bar{\theta})\right)$

- Optymalna decyzja maksymalizuje gęstość rozkładu prawdopodobieństwa
- Uwaga: dla funkcji strat: $L(\theta, \bar{\theta}) = |\theta - \bar{\theta}|$
- Optymalna decyzja - θ^* - mediana

Metody Bayes'a z obserwacją



Założenia:

$\underline{v} = h(\theta, \underline{z})$ – system pomiarowy opisany jest dowolną funkcją, wzajemnie jednoznaczną względem z

θ – obserwowany wektor wielkości, wartość zmiennej losowej $\underline{\theta}$, $\theta \in \Theta \subseteq \mathscr{R}^R$

Postacie funkcji gęstości rozkładów prawdopodobieństwa $f_z(z)$ oraz $f_\theta(\theta)$ są znane .

Dana jest funkcja strat $L(\theta, \bar{\theta})$, gdzie $\bar{\theta}$ jest wartością estymaty obserwowanej wielkości.

Ryzyko: $$R(\overline{\Psi}) \overset{df}{=} \underset{\underline{\theta}, \underline{V_N}}{E}\left[L\left(\underline{\theta}, \underline{\bar{\theta}} = \overline{\Psi}(\underline{V_N})\right)\right] = \int_{\mathscr{V}_N}\int_{\Theta} L\left(\theta, \overline{\Psi}(V_N)\right) f(\theta, V_N)\, d\theta\, dV_N$$

gdzie: $f(\theta, V_N)$ łączny rozkład prawdopodobieństwa

$$f(\theta, V_N) = f'(\theta | V_N) f''(V_N)$$

gdzie f' jest warunkową, a f'' brzegową funkcją gęstości rozkładu prawdopodobieństwa

Metody Bayes'a

Problem: $\Psi_N \rightarrow R(\Psi_N) = \min_{\overline{\Psi}} R(\overline{\Psi})$

$$R(\overline{\Psi}) = \int_{\mathscr{Y}_N} \int_{\Theta} L(\theta, \overline{\Psi}(V_N)) f'(\theta | V_N) d\theta \; f''(V_N) dV_N$$

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

gdzie: r - ryzyko warunkowe

Problem sprowadza się do następującego zadania :

$$\theta_N = \Psi_N(V_N) \rightarrow r(\theta_N, V_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, V_N)$$

Szczególny przypadek

- $L(\theta,\bar{\theta})=(\theta-\bar{\theta})^2$
- $r(\bar{\theta},V_N)=\int_\Omega (\theta-\bar{\theta})^2 f'(\theta/V_N)d\theta$
- $\left.\frac{dr(\bar{\theta},V_N)}{d\bar{\theta}}\right|_{\bar{\theta}=\theta^*} = -2\int_\Omega (\theta-\theta^*) f'(\theta/V_N)d\theta = 0 \ /\text{-}2$
- $\int_\Omega \theta^* f'(\theta/V_N)d\theta = \int_\Omega \theta f'(\theta/V_N)d\theta$
- $\theta^* \int_\Omega f'(\theta/V_N)d\theta = \int_\Omega \theta f'(\theta/V_N)d\theta$
- Stąd
- $\theta^* = \int_\Omega \theta f'(\theta/V_N)d\theta$ – wartość oczekiwana a'posteriori
- Metoda średniej a'posteriori

Szczególny przypadek

- $L(\theta,\bar{\theta})=-\delta(\theta-\bar{\theta})$
- $r(\bar{\theta},V_N)=\int_\Omega -\delta(\theta-\bar{\theta})\,f'(\theta/V_N)d\theta$
- $r(\bar{\theta},V_N)=-f'(\theta/V_N)$
- $\min_{\bar{\theta},} r(\bar{\theta},V_N)=\min_{\bar{\theta},}(-f'(\theta/V_N))=-\max_{\bar{\theta}} f'(\theta/V_N)$
- $\theta^*\rightarrow f'(\theta^*/V_N)=\max_{\bar{\theta}} f'(\theta/V_N)$

- Metoda maksymalnego prawdopodobieństwa a'posteriori

Metody Bayes'a

Gdzie funkcja gęstości rozkładu prawdopodobieństwa a posteriori:

$$f'(\theta|V_N) = \frac{f_\theta(\theta)\, f_{VN}(V_N|\theta)}{\int_\Theta f_\theta(\theta)\, f_{VN}(V_N|\theta)d\theta} = \frac{f_\theta(\theta)\prod_{n=1}^{N} f_v(v_n|\theta)}{\int_\Theta f_\theta(\theta)\prod_{n=1}^{N} f_v(v_n|\theta)d\theta} =$$

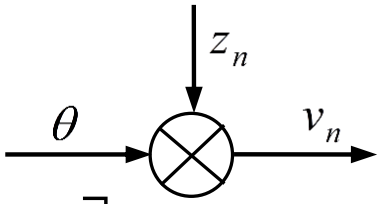
$$= \frac{f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right)|J_h|}{\int_\Theta f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right)|J_h|d\theta} = \frac{f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right)|J_h|}{const}$$

$$\int_\Theta f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right)|J_h|d\theta = const \quad \textit{dla danej serii pomiarowej } V_N$$

$$f'(\theta|V_N) \quad \alpha \quad f_\theta(\theta)\prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right)|J_h|$$

Metody Bayes'a

- Przykład



Opis zakłóceń:

$$f_z(z) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{z^2}{2\sigma_z^2} \right]$$

Rozkład prawdopodobieństwa:

$$f_\theta(\theta) = \frac{1}{\sigma_\theta \sqrt{2\pi}} \exp\left[-\frac{(\theta - m_\theta)^2}{2\sigma_\theta^2} \right]$$

Opis systemu pomiarowego:

$$v = h(\theta, z) = \theta + z, \qquad z = h_z^{-1}(\theta, v) = v - \theta$$

Funkcja strat:

$$L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$$

Metody Bayes'a

- Przykład

Macierz Jakobiego:

$$J_h = \frac{\partial h_z^{-1}(\theta, v)}{\partial v} = \frac{d}{dv}(v - \theta) = 1$$

Funkcja gęstości rozkładu prawdopodobieństwa obserwowanej zmiennej:

$$f_v(v|\theta) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{(v-\theta)^2}{2\sigma_z^2} \right] \cdot |1|$$

Funkcja gęstości rozkładu prawdopodobieństwa a posteriori

$$f'(\theta|V_N) \quad \alpha \quad f_\theta(\bar{\theta}) \prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right) |J_h| =$$

$$= \frac{1}{\sigma_\theta \sqrt{2\pi}} \exp\left[-\frac{(\bar{\theta} - m_\theta)^2}{2\sigma_\theta^2} \right] \prod_{n=1}^{N} \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{(v_n - \bar{\theta})^2}{2\sigma_z^2} \right]$$

Metody Bayes'a

- Przykład

Funkcja gęstości rozkładu prawdopodobieństwa a posteriori po przekształceniach:

$$f'(\theta|V_N) \quad \alpha \quad f_\theta(\bar{\theta}) \prod_{n=1}^{N} f_z\left(h_z^{-1}(\bar{\theta}, v_n)\right)|J_h| =$$

$$= \frac{1}{\sigma_\theta\sqrt{2\pi}}\left(\frac{1}{\sigma_z\sqrt{2\pi}}\right)^N \exp\left[-\frac{(\bar{\theta}-m_\theta)^2}{2\sigma_\theta^2} - \sum_{n=1}^{N}\frac{(v_n-\bar{\theta})^2}{2\sigma_z^2}\right]$$

Dla funkcji strat: $L(\theta,\bar{\theta}) = -\delta(\theta-\bar{\theta})$ ryzyko warunkowe:

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

$$\alpha \quad -f_\theta(\bar{\theta}) \prod_{n=1}^{N} f_z\left(h_z^{-1}(\bar{\theta}, v_n)\right)|J_h| = -\frac{1}{\sigma_\theta\sqrt{2\pi}}\left(\frac{1}{\sigma_z\sqrt{2\pi}}\right)^N \exp\left[-\frac{(\bar{\theta}-m_\theta)^2}{2\sigma_\theta^2} - \sum_{n=1}^{N}\frac{(v_n-\bar{\theta})^2}{2\sigma_z^2}\right]$$

Metody Bayes'a

- Przykład

W wyniku minimalizacji ryzyka warunkowego otrzymujemy algorytm estymacji:

$$\theta_N = \Psi_N(V_N) = \frac{m_\theta + \left(\frac{\sigma_\theta}{\sigma_z}\right)^2 \sum_{n=1}^{N} v_n}{1 + \left(\frac{\sigma_\theta}{\sigma_z}\right)^2 N}$$

Dyskusja:

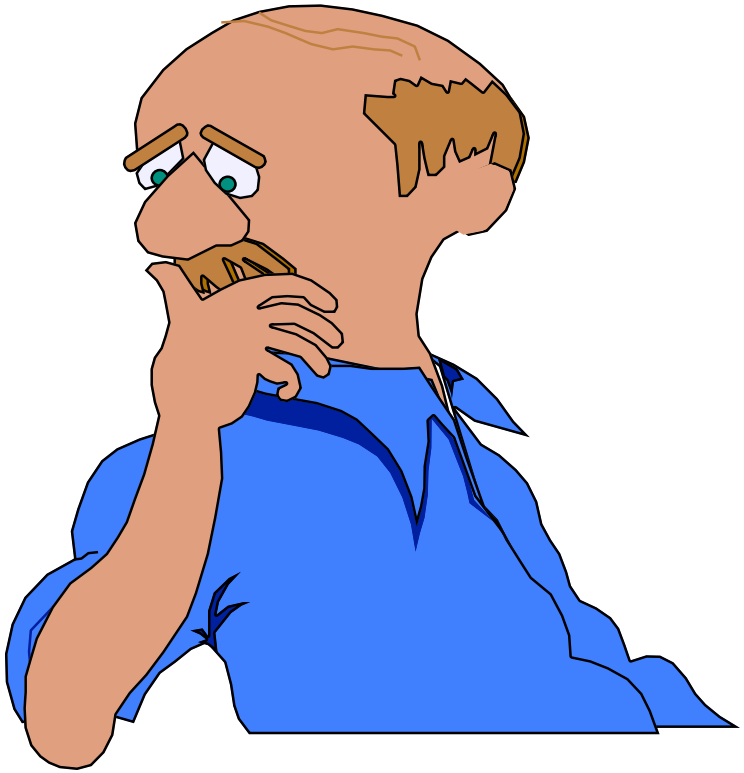
1° N – mała liczba pomiarów

$(\sigma_z >> \sigma_\theta)$ – pomiary bezwartościowe

$$\theta_N \approx m_\theta$$

2° $N \to \infty$

$(\sigma_z << \sigma_\theta)$ - dobre pomiary

$$\theta_N \approx \frac{1}{N} \sum_{n=1}^{N} v_n$$

Dziękuję za uwagę