

Computer Science

Jerzy Świątek

Systems Modelling and Analysis

Choose yourself and new technologies

L.3b. Estimation of plant parameter with noisy measurements



HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

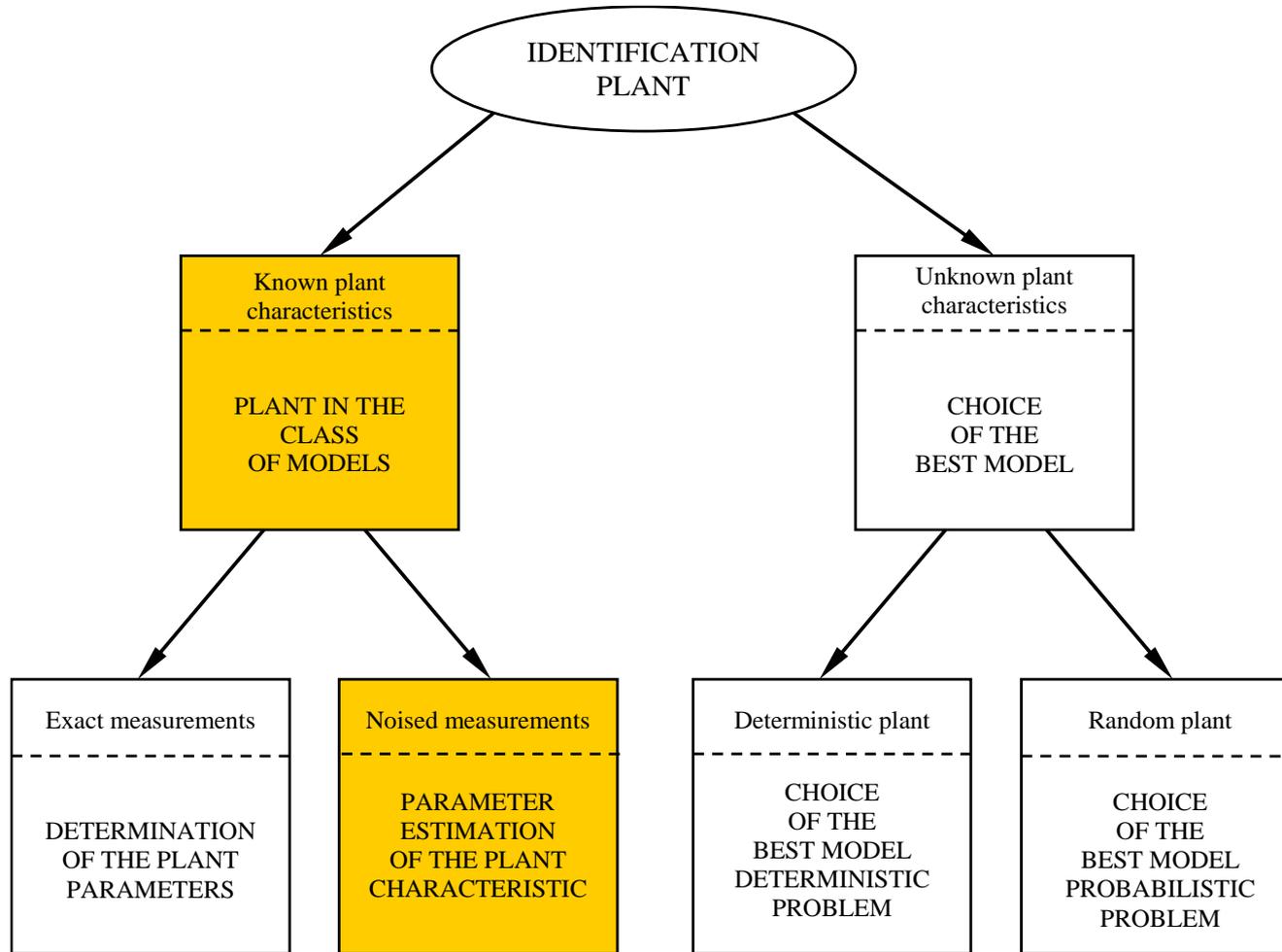


Wrocław University of Technology

EUROPEAN
SOCIAL FUND



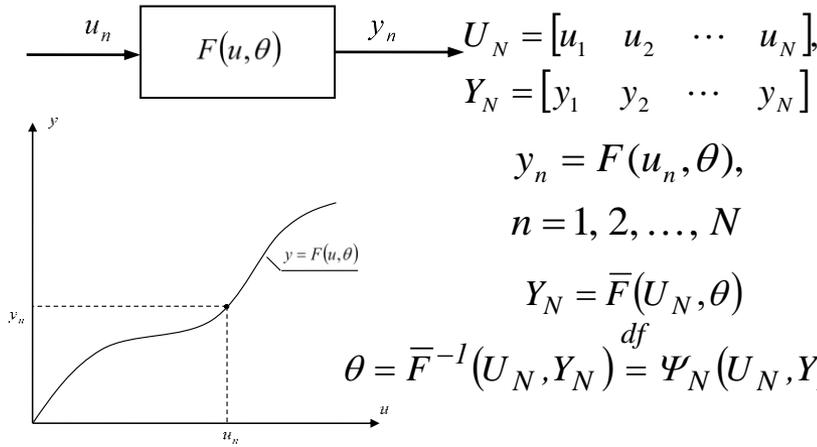
Project co-financed from the EU European Social Fund





Plant in the class of model

Deterministyczny



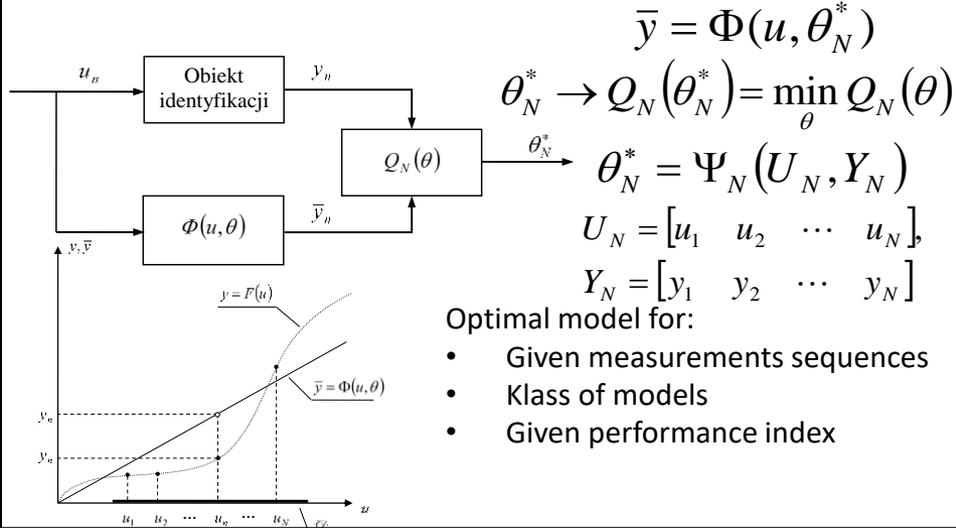
$$y_n = F(u_n, \theta),$$

$$n = 1, 2, \dots, N$$

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

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

Choice of the best model



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

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

$$\theta_N^* = \Psi_N(U_N, Y_N)$$

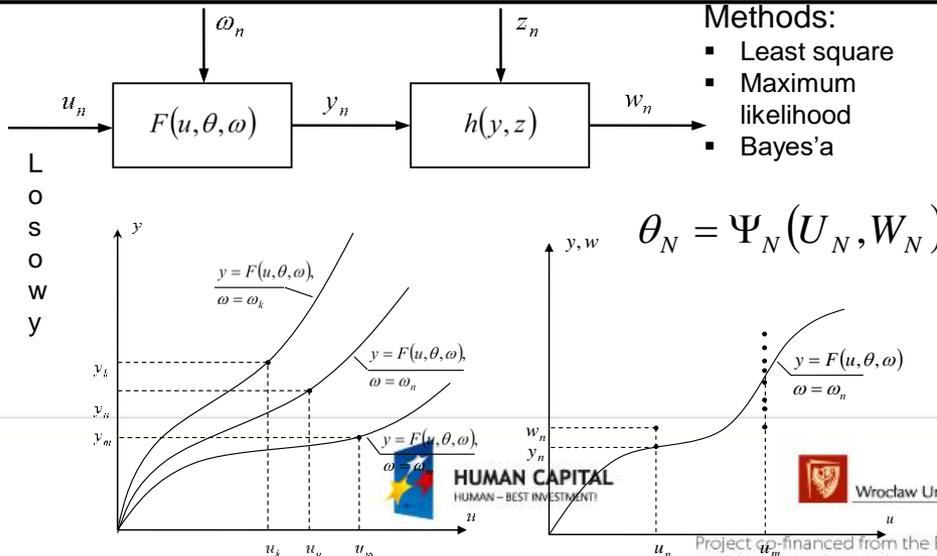
$$U_N = [u_1 \ u_2 \ \dots \ u_N]$$

$$Y_N = [y_1 \ y_2 \ \dots \ y_N]$$

Optimal model for:

- Given measurements sequences
- Class of models
- Given performance index

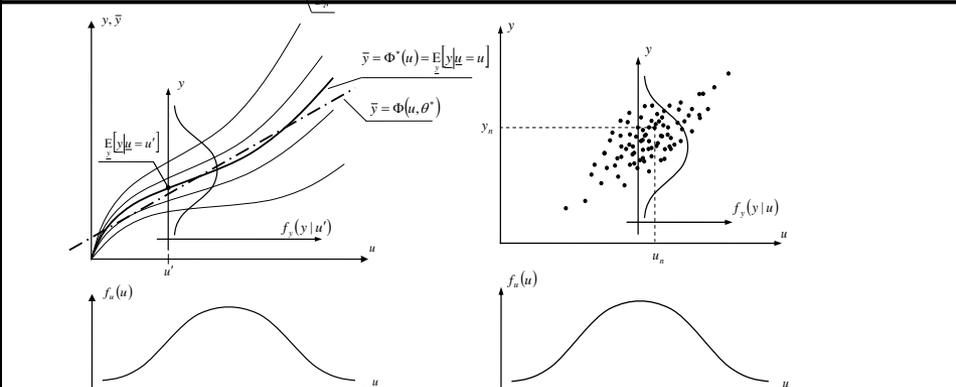
Losowy



Methods:

- Least square
- Maximum likelihood
- Bayes'a

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



Full probabilistic knowledge

- First type regression
- Second type regression

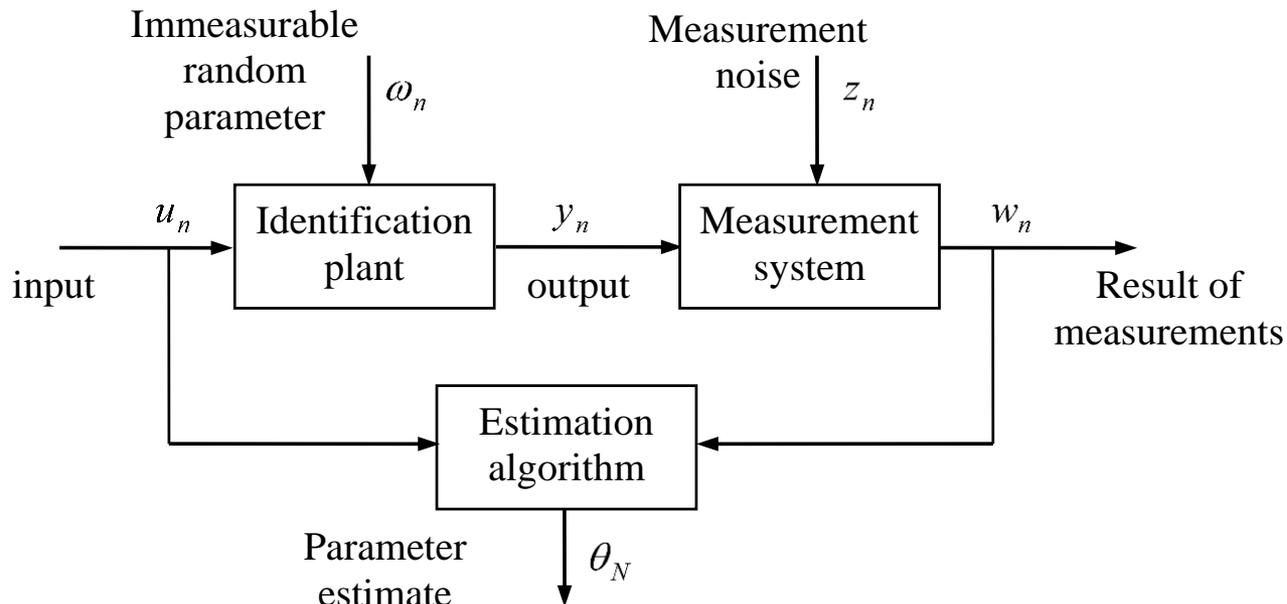
Unknown probabilistic knowledge

- Performance index estimation
- Parameter estimation of the probability distribution
- Probability distribution estimation





Plant parameter estimation problem





Noised measurements of the physical values

- Problem formulation

Measurement noise:

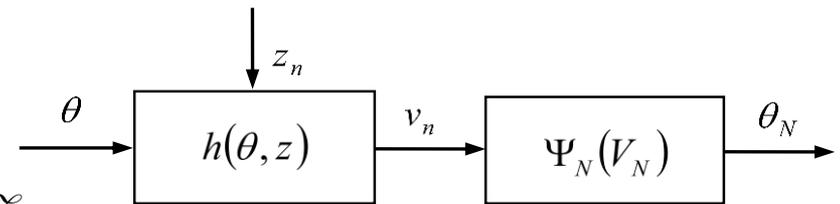
z_n – value of random variable z from the space \mathcal{Z}

$f_z(z)$ – probability density function

θ – observed vector of parameters, value of random variable $\underline{\theta}$, $\theta \in \Theta \subseteq \mathcal{R}^R$

$f_\theta(\theta)$ – probability density function

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





Noised measurements of the physical values

General form of estimation algorithm:

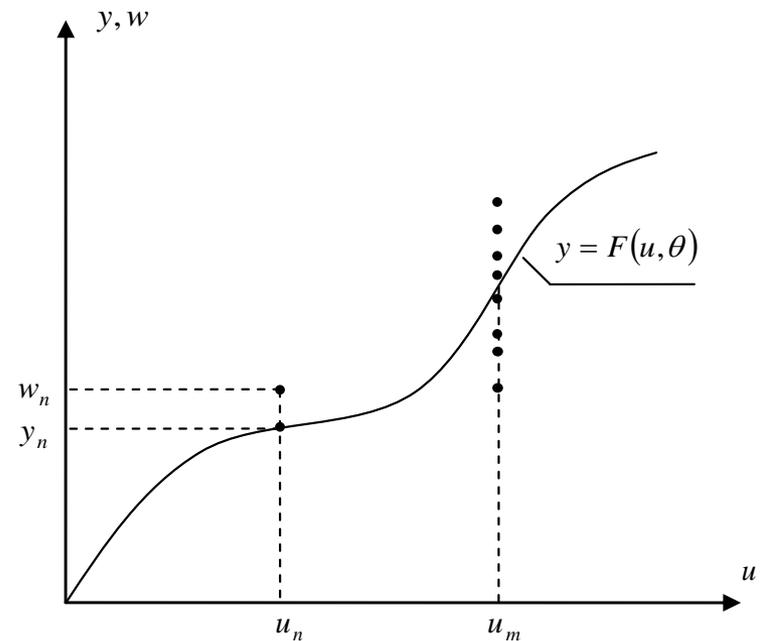
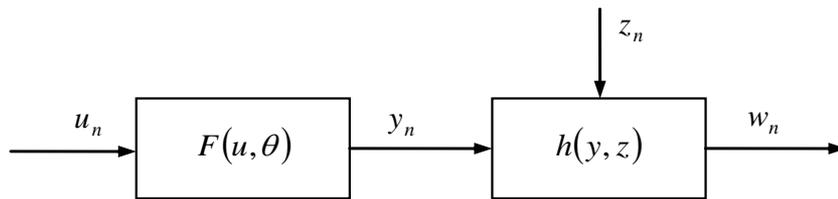
$$\theta_N = \Psi_N(V_N)$$

- Solution:
 - Least square method
 - Maximum likelihood method
 - Bayesian method



Deterministic plant, noised measurements of the plant output

- Noised measurements of the identification plant known static characteristics





Deterministic plant, noised measurements of the plant output

- Problem formulation

Measurement system description: $w = h(y, z)$

where: $w \in \mathcal{W}$, h – known one-to-one function

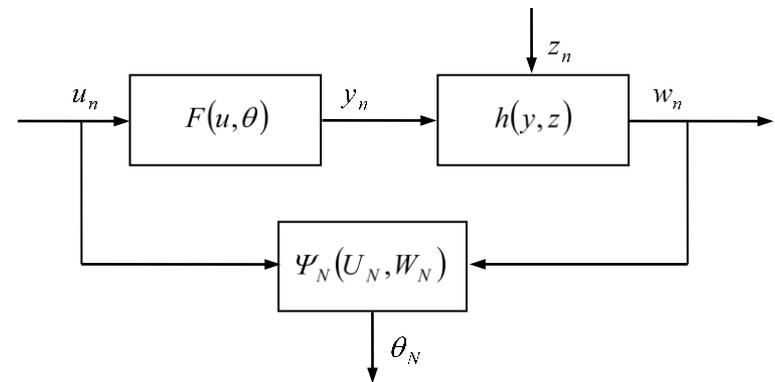
$$h: \mathcal{Y} \times \mathcal{Z} \rightarrow \mathcal{W}, \quad z = h_z^{-1}(y, w)$$

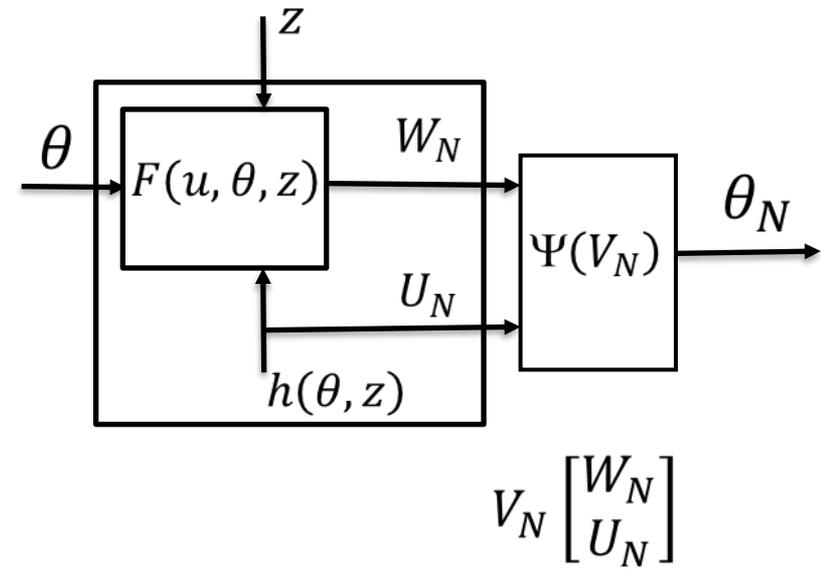
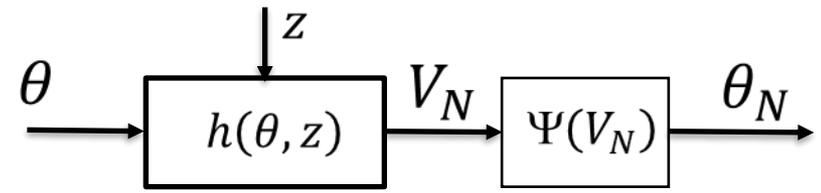
\mathcal{W} – measurements domain ($\dim y = \dim z = L$)

Measurement noise:

z_n – value of random variable \underline{z} from the space \mathcal{Z} $f_z(z)$ – probability density function

Measurements: $U_N = [u_1 \quad u_2 \quad \dots \quad u_N]$, $W_N = [w_1 \quad w_2 \quad \dots \quad w_N]$







Deterministic plant, noised measurements of the plant output

General form of estimation algorithm:

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

- Solution:
 - Least square method
 - Maximum likelihood method
 - Bayes' method



Least square method

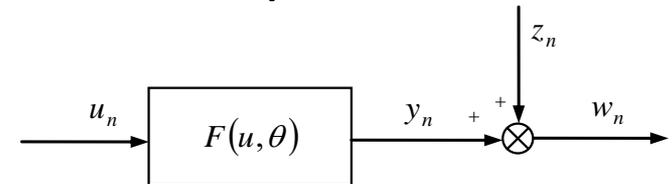
(one dimensional case $L = 1$)

Assumptions:

$$w = h(y, z) = y + z \quad - \text{additive noise}$$

$$E[\underline{z}] = 0 \quad - \text{expected value of the noise signal is zero}$$

$$Var[\underline{z}] < \infty \quad - \text{variance of noise is not infinite}$$



Calculations:

Least square method minimizes empirical variance of noise signal:

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

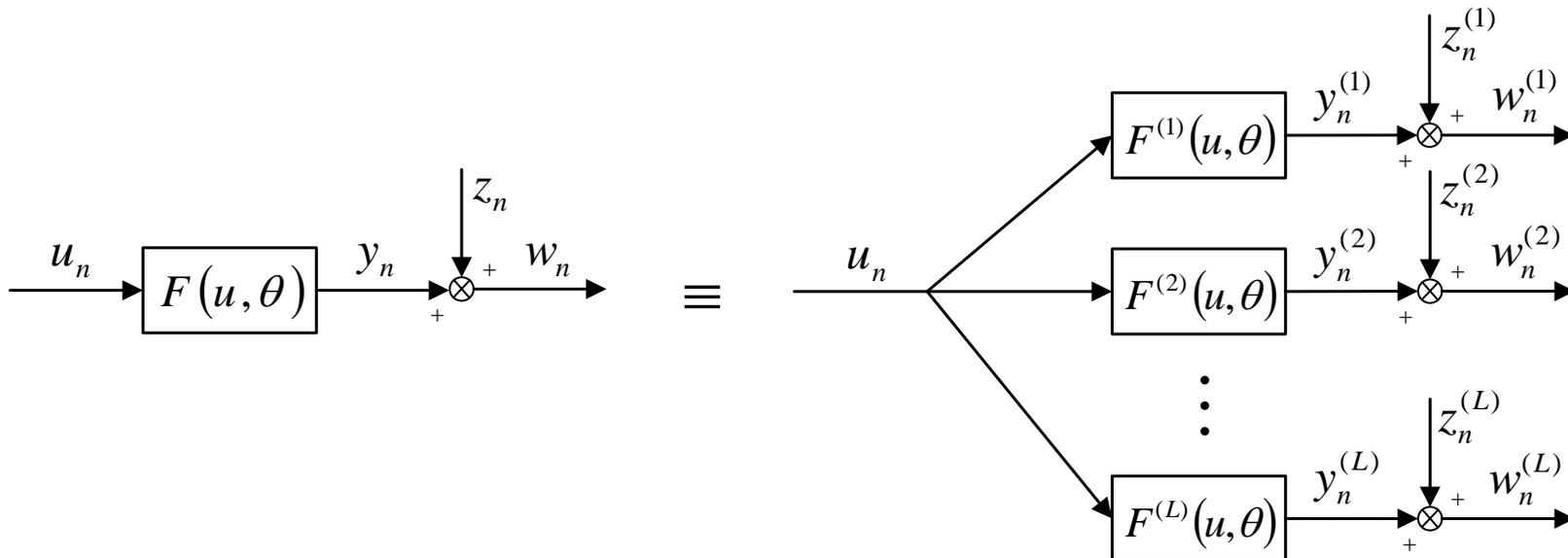
Estimation algorithm has the form:

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



Least square method

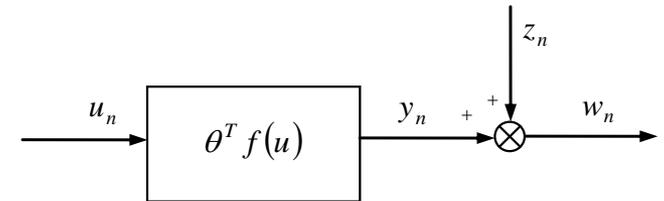
Decomposition of L – dimensional identification plant into L single output plants:





Least square method

- Example



Empirical variance of noise signal:

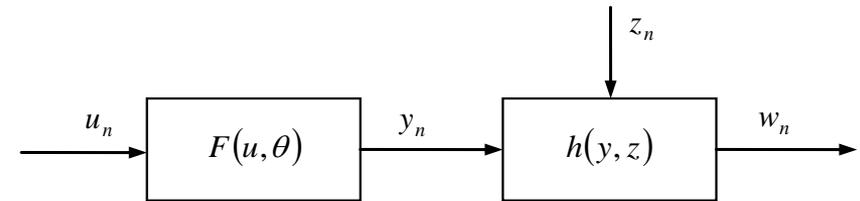
$$\text{Var}_{zN}(U_N, W_N, \theta) = \frac{1}{N} \sum_{n=1}^N (w_n - y_n)^2 = \frac{1}{N} \sum_{n=1}^N (w_n - \theta^T f(u_n))^2$$

Estimation algorithm:

$$\theta_N = \Psi_N(U_N, W_N) = \left[\sum_{n=1}^N f(u_n) f^T(u_n) \right]^{-1} \sum_{n=1}^N w_n f(u_n)$$



Maximum likelihood method



Assumptions:

$\underline{w} = h(y, \underline{z})$ – measurement system is described by any one-to-one invertible function

z_n – independent value of random variable \underline{z} with known probability density function $f_z(\underline{z})$

For a given $u_n, n = 1, 2, \dots, N$ noised output is measured $w_n, n = 1, 2, \dots, N$

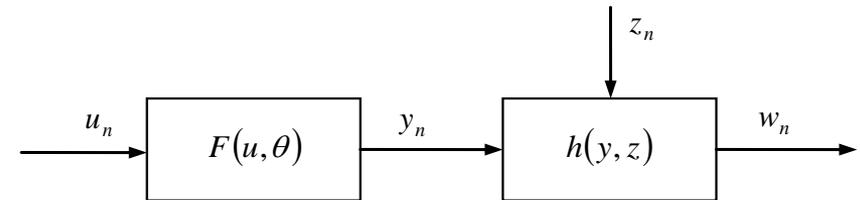
Calculations:

Taking into account plant description $\underline{w} = h(F(u, \theta), \underline{z})$, probability density function of random variable $\underline{w} - f_w(w, \theta; u)$ with parameters θ has the form: $f_w(w, \theta; u) = f_z(h_z^{-1}(F(u, \theta), w)) \cdot |J_h|$

where J_h is Jacobi matrix: $J_h = \frac{\partial h_z^{-1}(y, w)}{\partial w}$



Maximum likelihood method



Likelihood function:

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

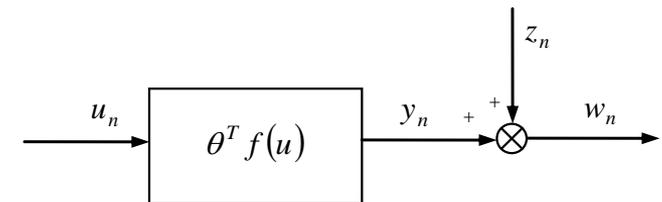
Estimation algorithm:

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



Maximum likelihood method

- Example 1



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

Measurement system description: $w = h(y, z) = y + z$

$$z = h_z^{-1}(F(u, \theta), w) = w - \theta^T f(u)$$

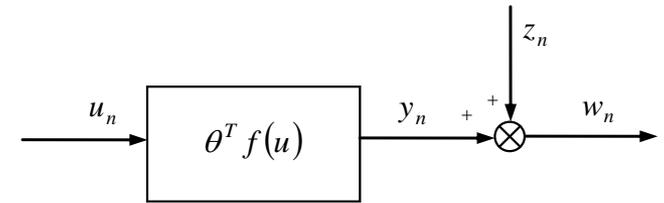
Jacobi matrix of the inverse transform :

$$J_h = \frac{\partial h_z^{-1}(F(u, \theta), w)}{\partial w} = \frac{d}{dw} (w - \theta^T f(u)) = 1$$



Maximum likelihood method

- Example 1



Probability density function : $f_w(w, \theta; u) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(w - \theta^T f(u) - m_z)^2}{2\sigma_z^2} \right] \cdot |1|$

Likelihood function : $L_N(W_N, \theta; U_N) = \prod_{n=1}^N \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(w_n - \theta^T f(u_n) - m_z)^2}{2\sigma_z^2} \right]$

$$L_N(W_N, \theta; U_N) = \left(\frac{1}{\sigma_z \sqrt{2\pi}} \right)^N \exp \left[\sum_{n=1}^N -\frac{(w_n - \theta^T f(u_n) - m_z)^2}{2\sigma_z^2} \right]$$

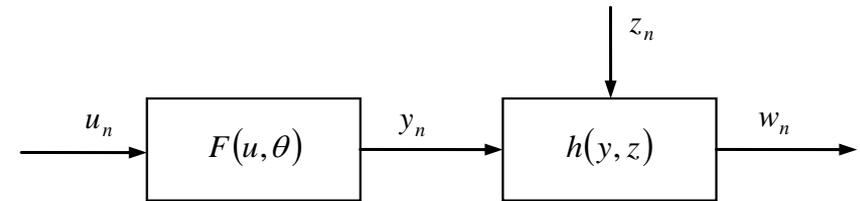
Estimation algorithm:

$$\theta_N = \Psi_N(U_N, W_N) = \left[\sum_{n=1}^N f(u_n) f^T(u_n) \right]^{-1} \sum_{n=1}^N (w_n - m_z) f(u_n)$$



Maximum likelihood method

- Example 2



Plant characteristic: $y = F(u, \theta) = \theta u \quad (\theta > 0)$

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

Additional assumption: $\forall n = 1, 2, \dots, N \quad u_n > 0$

Measurement system description: $w = h(y, z) = h(F(u, \theta), z) = \theta u z$

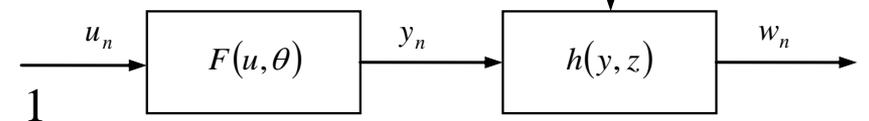
$$z = h_z^{-1}(F(u, \theta), w) = \frac{w}{\theta u}$$



Maximum likelihood method

- Example 2

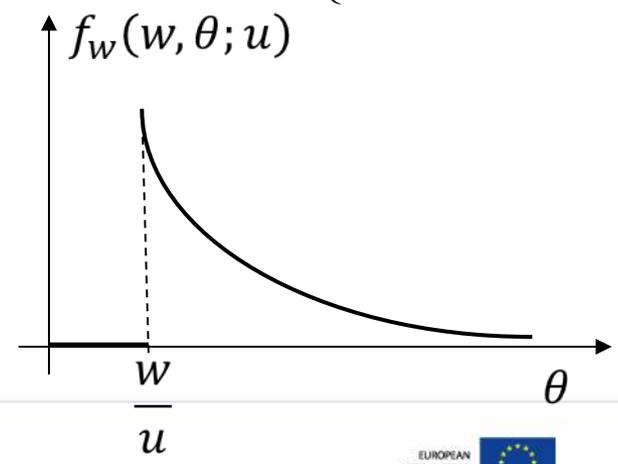
Jacobi matrix: $J_h = \frac{\partial h_z^{-1}(F(u, \theta), w)}{\partial w} = \frac{d}{dw} \left(\frac{w}{\theta u} \right) = \frac{1}{\theta u}$



Probability density function of the observed value:

$$f_w(w, \theta; u) = \begin{cases} \frac{1}{\theta u} & \text{for } \frac{w}{\theta u} \in [0, 1] \\ 0 & \text{for } \frac{w}{\theta u} \notin [0, 1] \end{cases}$$

$$f_w(w, \theta; u) = \begin{cases} \frac{1}{\theta u} & \text{for } \frac{w}{u} \geq \theta \\ 0 & \text{for } \frac{w}{u} < \theta \end{cases}$$



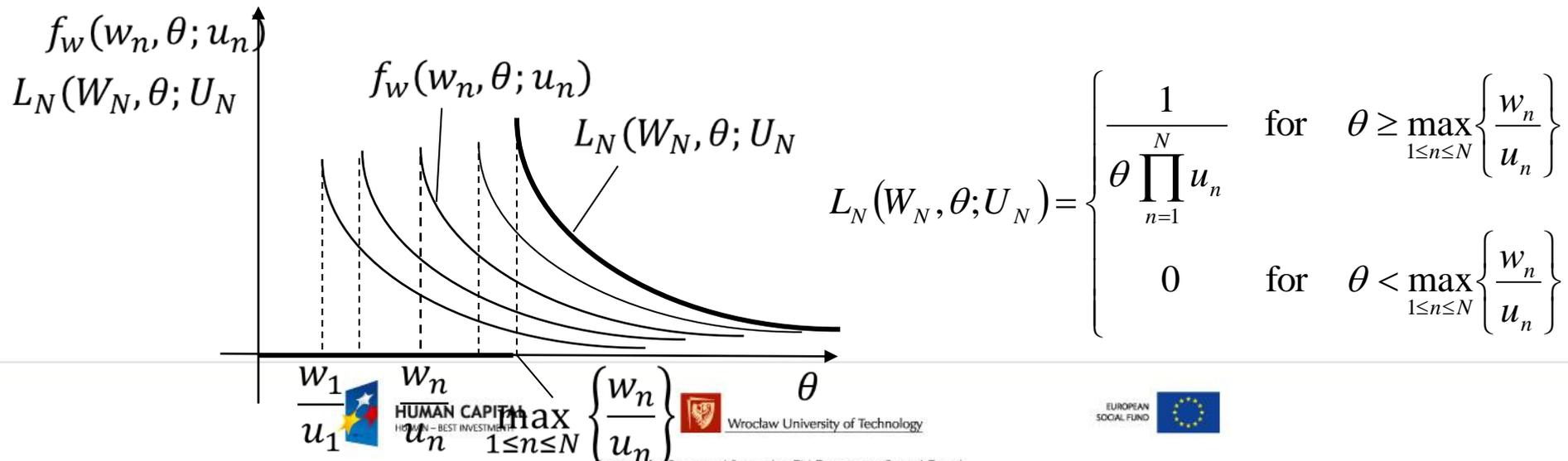


Maximum likelihood method

• Example 2

Likelihood function :

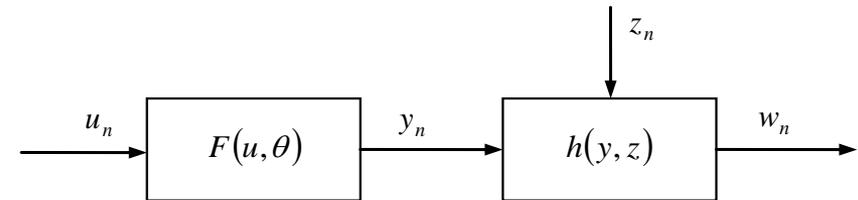
$$L_N(W_N, \theta; U_N) = \begin{cases} \frac{1}{\theta \prod_{n=1}^N u_n} & \text{for } \forall n = 1, 2, \dots, N \quad \frac{w_n}{u_n} \in [0, \theta] \\ 0 & \text{for } \exists n = 1, 2, \dots, N \quad \frac{w_n}{u_n} \notin [0, \theta] \end{cases}$$





Maximum likelihood method

- Example 2



Likelihood function :

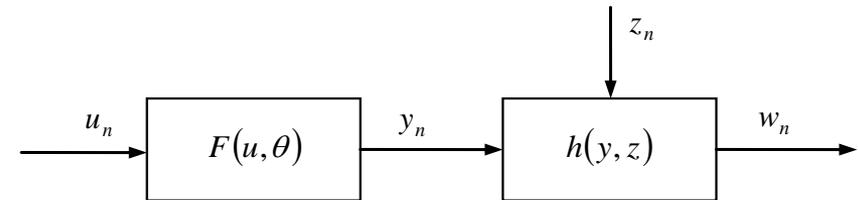
$$L_N(W_N, \theta; U_N) = \begin{cases} \frac{1}{\theta \prod_{n=1}^N u_n} & \text{for } \forall n = 1, 2, \dots, N \quad \frac{w_n}{u_n} \in [0, \theta] \\ 0 & \text{for } \exists n = 1, 2, \dots, N \quad \frac{w_n}{u_n} \notin [0, \theta] \end{cases}$$

$$L_N(W_N, \theta; U_N) = \begin{cases} \frac{1}{\theta \prod_{n=1}^N u_n} & \text{for } \theta \geq \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\} \\ 0 & \text{for } \theta < \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\} \end{cases}$$



Maximum likelihood method

- Example 2



Estimation algorithm:

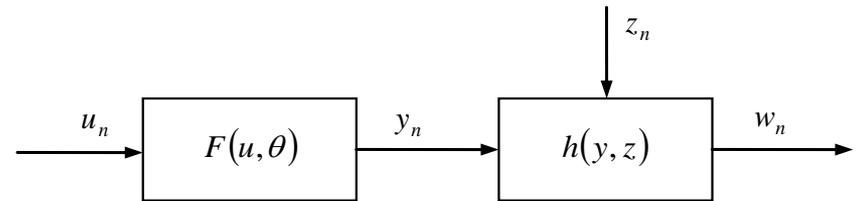
$$\theta_N = \Psi_N(U_N, W_N) = \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\}$$

Interpretation:

$$\theta_N = \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\} = \max_{1 \leq n \leq N} \left\{ \frac{\theta u_n z_n}{u_n} \right\} = \theta \max_{1 \leq n \leq N} \{z_n\}$$



Bayesian method's



Assumptions:

$w = h(y, z)$ – measurement system is described by any one-to-one invertible function

z_n – independent value of random variable z with known probability density function $f_z(z)$

Mathematical formulas describing probability density functions $f_z(z)$ and $f_\theta(\theta)$ are given.

The loss function $L(\theta, \bar{\theta})$ is defined.

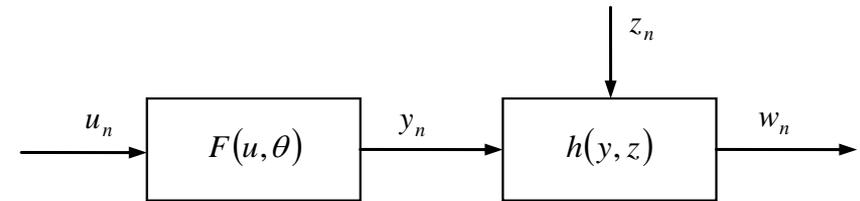
For a given $u_n, n = 1, 2, \dots, N$ noised output is measured $w_n, n = 1, 2, \dots, N$

θ – plant parameters vector, value of random variable $\theta, \theta \in \Theta \subseteq \mathcal{R}^R$



Bayesian method's

Calculations:



Taking into account plant description: $\underline{w} = h(F(u, \theta), \underline{z})$

$$\text{Risk: } R(\bar{\Psi}) \stackrel{\text{df}}{=} E_{\underline{\theta}, \underline{W}_N} [L(\underline{\theta}, \bar{\theta} = \bar{\Psi}(U_N, \underline{W}_N))] = \int_{\mathcal{W}_N} \int_{\Theta} L(\theta, \bar{\Psi}(U_N, W_N)) f(\theta, W_N; U_N) d\theta dW_N$$

where $f(\theta, W_N; U_N)$ is joint probability density function:

$$f(\theta, W_N; U_N) = f'(\theta | W_N; U_N) f''(W_N; U_N)$$

The problem: $\Psi_N \rightarrow R(\Psi_N) = \min_{\bar{\Psi}} R(\bar{\Psi})$



Bayesian method's

Joint probability density function: $f(\theta, W_N; U_N) = f'(\theta|W_N; U_N) f''(W_N; U_N)$

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

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

Conditional probability density function of observed value: $f_w(w|\theta; u) = f_z(h_z^{-1}(F(u, \theta), w)) |J_h|$

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



Bayesian method's

Risk:

$$R(\bar{\Psi}) = \int_{\mathcal{W}_N} \int_{\Theta} L(\theta, \bar{\Psi}(U_N, W_N)) f'(\theta | W_N; U_N) d\theta f''(W_N; U_N) dW_N$$

Conditional risk:

$$r(\bar{\theta}, W_N; U_N) \stackrel{\text{df}}{=} E_{\underline{\theta}} [L(\underline{\theta}, \bar{\theta}) | W_N; U_N] = \int_{\Theta} L(\theta, \bar{\theta} = \bar{\Psi}(U_N, W_N)) f'(\theta | W_N; U_N) d\theta$$

Estimation algorithm:

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



Bayesian method's

- Mean a posteriori value method

Special case of the loss function: $L(\theta, \bar{\theta}) = [\theta - \bar{\theta}]^T [\theta - \bar{\theta}]$

$$\begin{aligned} \text{Conditional risk: } r(\bar{\theta}, W_N; U_N) &\stackrel{df}{=} E_{\underline{\theta}} \left[[\underline{\theta} - \bar{\theta}]^T [\underline{\theta} - \bar{\theta}] | W_N; U_N \right] = \\ &= E_{\underline{\theta}} \left[\underline{\theta}^T \underline{\theta} | W_N; U_N \right] - 2\bar{\theta} E_{\underline{\theta}} \left[\underline{\theta}^T | W_N; U_N \right] + \bar{\theta}^T \bar{\theta} \end{aligned}$$

$$\left. \text{grad}_{\bar{\theta}} r(\bar{\theta}, W_N; U_N) \right|_{\bar{\theta} = \theta_N} = -2 E_{\underline{\theta}} \left[\underline{\theta} | W_N; U_N \right] + 2\theta_N = 0_N$$

Estimation algorithm:

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



Bayesian method's

- Maximum a posteriori probability

Special case of the loss function: $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$

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

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

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

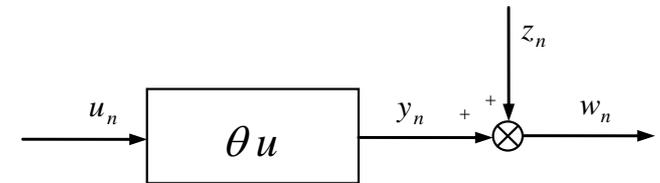
$\theta_N = \Psi_N(U_N, W_N)$ such that:

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



Bayesian method's

- Example



Plant characteristic: $y = F(u, \theta) = \theta u$

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

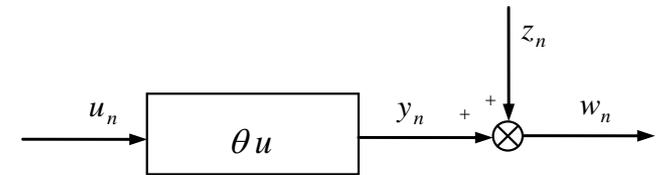
A priori distribution:
$$f_\theta(\theta) = \frac{1}{\sigma_\theta \sqrt{2\pi}} \exp\left[-\frac{(\theta - m_\theta)^2}{2\sigma_\theta^2}\right]$$

Measurement system description: $w = h(y, z) = y + z, \quad z = h_z^{-1}(F(u, \theta), w) = w - \theta u$



Bayesian method's

- Example



Jacobi matrix:
$$J_h = \frac{\partial h_z^{-1}(F(u, \theta), w)}{\partial w} = \frac{d}{dw} (w - \theta u) = 1$$

Probability density function of the observed value:
$$f_w(w|\theta; u) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{(w - \theta u)^2}{2\sigma_z^2}\right] \cdot |1|$$

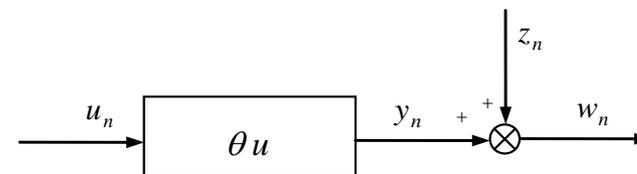
A posteriori probability density function:

$$\begin{aligned} f'(\theta|W_N; U_N) &\propto f_\theta(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \theta), w_n)) |J_h| = \\ &= \frac{1}{\sigma_\theta \sqrt{2\pi}} \exp\left[-\frac{(\bar{\theta} - m_\theta)^2}{2\sigma_\theta^2}\right] \prod_{n=1}^N \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{(w_n - \bar{\theta} u_n)^2}{2\sigma_z^2}\right] \end{aligned}$$



Bayesian method's

- Example



A posteriori probability density function after transformation:

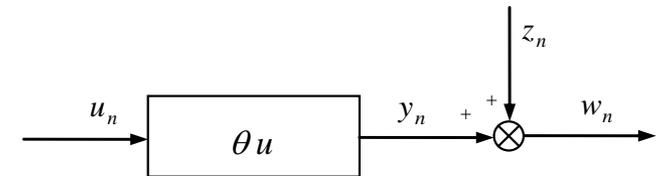
$$f'(\bar{\theta}|V_N) \propto f_{\theta}(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \bar{\theta}), w_n)) |J_h| =$$

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



Bayesian method's

- Example



Estimation algorithm:

$$\theta_N = \Psi_N(U_N, W_N) = \frac{m_\theta + \left(\frac{\sigma_\theta}{\sigma_z}\right)^2 \sum_{n=1}^N w_n u_n}{1 + \left(\frac{\sigma_\theta}{\sigma_z}\right)^2 \sum_{n=1}^N u_n^2}$$

Discussion:

1° N – small number

$(\sigma_z \gg \sigma_\theta)$ – poor measurements

$$\theta_N \approx m_\theta$$

2° $N \rightarrow \infty$

$(\sigma_z \ll \sigma_\theta)$ – good measurements

$$\theta_N \approx \frac{\sum_{n=1}^N w_n u_n}{\sum_{n=1}^N u_n^2}$$



Thank you for attention

