Identification of Interconnected Systems by Instrumental Variables Method

Grzegorz Mzyk

Institute of Computer Engineering, Control and Robotics
Wrocław University of Technology
Poland

3th-May-2012

Structure of the presentation

1. Identification of single-element systems
   – *MISO linear static element*
   – *SISO linear dynamic element*
   Least squares (LS) method and **instrumental variables (IV)** method

Structure of the presentation

1. Identification of single-element systems
   – *MISO linear static element*
   – *SISO linear dynamic element*
   Least squares (LS) method and **instrumental variables (IV)** method
2. Interconnected linear static systems
   – LS-based estimate and limit properties
   – IV-based estimate and limit properties
   – generation of instrumental variables

Structure of the presentation

1. Identification of single-element systems
   – *MISO linear static element*
   – *SISO linear dynamic element*
   Least squares (LS) method and **instrumental variables (IV)** method
2. Interconnected linear static systems
   – LS-based estimate and limit properties
   – IV-based estimate and limit properties
   – generation of instrumental variables
3. Nonlinear dynamic block-oriented systems
   – Hammerstein system
   – NARMAX system

MISO linear static block

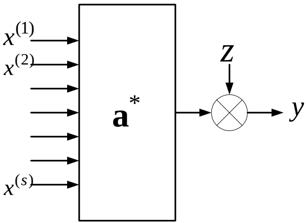

Figure: *MISO linear static block*

$$\mathbf{a}^* = \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_s^* \end{bmatrix}$$

Assumptions:

$\mathbf{E}z = 0$, $\text{var} z < \infty$

$x^{(i)}$, z – **independent !!!**

$$\mathbf{X}_N = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix} = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & .. & x_1^{(s)} \\ x_2^{(1)} & x_2^{(2)} & .. & x_2^{(s)} \\ \vdots & \vdots & \vdots & \vdots \\ x_N^{(1)} & x_N^{(2)} & .. & x_N^{(s)} \end{bmatrix}$$

MISO linear static block (continued)

$$\mathbf{Y}_N = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \mathbf{Z}_N = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}$$

Measurement equation

$$\mathbf{Y}_N = \mathbf{X}_N \mathbf{a}^* + \mathbf{Z}_N$$

Model

$$\overline{\mathbf{Y}}_N(\mathbf{a}) = \mathbf{X}_N \mathbf{a}$$

Least squares criterion

$$\left\| \mathbf{Y}_N - \overline{\mathbf{Y}}_N(\mathbf{a}) \right\|_2^2 \to \min_{\mathbf{a}}$$

Normal equation

$$\mathbf{X}_N^T \mathbf{X}_N \mathbf{a} = \mathbf{X}_N^T \mathbf{Y}_N$$

Uniqueness of the solution

$$\operatorname{rank} \mathbf{X}_N = s$$

LS estimate

$$\widehat{\mathbf{a}}_N = \left(\mathbf{X}_N^T \mathbf{X}_N \right)^{-1} \mathbf{X}_N^T \mathbf{Y}_N = \mathbf{X}_N^{+} \mathbf{Y}_N$$

$$\widehat{\mathbf{a}}_N \xrightarrow{p.1} \mathbf{a}^*, \text{ as } N \to \infty$$

FIR linear dynamics

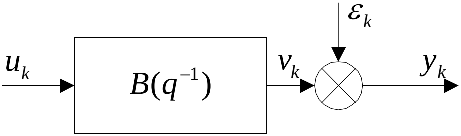

Figure: *Linear dynamic object MA(s)*

$$
\begin{aligned}
v_k &= b_0^* u_k + \ldots + b_s^* u_{k-s} \\
y_k &= v_k + \varepsilon_k \\
y_k &= b_0^* u_k + \ldots + b_s^* u_{k-s} + z_k \\
z_k &= \varepsilon_k
\end{aligned}
$$

FIR linear dynamics (2)

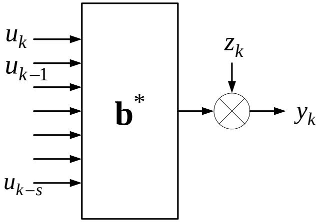

Figure: *MA object*

$$\mathbf{b}^* = \begin{bmatrix} b_0^* \\ b_1^* \\ \vdots \\ b_s^* \end{bmatrix}$$

Assumptions:

$\mathbf{E}z = 0$, var$z < \infty$

$\{u_k\}$, $\{z_k\}$ – **independent !!!**

$$\Phi_N = \begin{bmatrix} \boldsymbol{\phi}_1^T \\ \boldsymbol{\phi}_2^T \\ \vdots \\ \boldsymbol{\phi}_N^T \end{bmatrix} = \begin{bmatrix} u_1 & u_0 & .. & u_{1-s} \\ u_2 & u_1 & .. & u_{2-s} \\ \vdots & \vdots & \vdots & \vdots \\ u_N & u_{N-1} & .. & u_{N-s} \end{bmatrix}$$

$$\mathbf{Y}_N = \Phi_N \mathbf{b}^* + \mathbf{Z}_N \qquad \widehat{\mathbf{b}}_N = \left(\Phi_N^T \Phi_N\right)^{-1} \Phi_N^T \mathbf{Y}_N$$

IIR linear dynamics

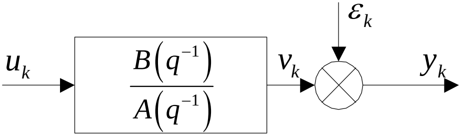

Figure: *Linear dynamic object ARMA(s,p)*

$$v_k = b_0^* u_k + \ldots + b_s^* u_{k-s} + a_1^* v_{k-1} + \ldots. + a_p^* v_{k-p}$$

$$y_k = v_k + \varepsilon_k$$

$$y_k = b_0^* u_k + \ldots + b_s^* u_{k-s} + a_1^* y_{k-1} + \ldots. + a_p^* y_{k-p} + z_k$$

$$z_k = \varepsilon_k - a_1^* \varepsilon_{k-1} - \ldots - a_p^* \varepsilon_{k-p}$$

IIR linear dynamics (2)

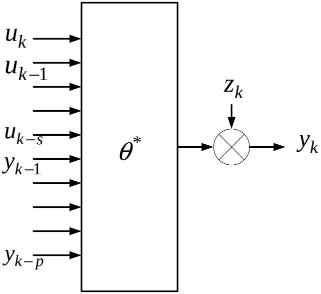

Figure: *ARMA object*

$$\theta^* = \begin{bmatrix} b_0^* \\ b_1^* \\ \vdots \\ b_s^* \\ a_1^* \\ a_2^* \\ \vdots \\ a_p^* \end{bmatrix}$$

$\mathbf{E}z = 0$, $\text{var}z < \infty$

u_{k-i}, z_k – independent

y_{k-i}, z_k – **correlated !!!**

IIR linear dynamics (3)

$$\Phi_N = \begin{bmatrix} \phi_1^T \\ \phi_2^T \\ \vdots \\ \phi_N^T \end{bmatrix} =$$

$$\begin{bmatrix} u_1 & u_0 & \cdots & u_{1-s} & y_0 & y_{-1} & \cdots & y_{1-p} \\ u_2 & u_1 & \cdots & u_{2-s} & y_1 & y_0 & \cdots & y_{2-p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_N & u_{N-1} & \cdots & u_{N-s} & y_{N-1} & y_{N-2} & \cdots & y_{N-p} \end{bmatrix}$$

$$\mathbf{Y}_N = \Phi_N \theta^* + \mathbf{Z}_N \qquad \widehat{\theta}_N = \left(\Phi_N^T \Phi_N\right)^{-1} \Phi_N^T \mathbf{Y}_N$$

Instrumental variables approach

$$\widehat{\theta}_N^{IV} = \left(\Psi_N^T \Phi_N\right)^{-1} \Psi_N^T Y_N$$

- Consistency conditions
  (a) $\dim \Psi_N = \dim \Phi_N$, i.e. $\Psi_N = (\boldsymbol{\psi}_1, ..., \boldsymbol{\psi}_N)^T$,
  $\dim \boldsymbol{\psi}_k = s + p + 1$
  (b) $Plim_{N\to\infty} \left(\frac{1}{N}\Psi_N^T \Phi_N\right)$ exists and is not singular
  (c) $Plim_{N\to\infty} \left(\frac{1}{N}\Psi_N^T \mathbf{Z}_N\right) = 0$

$$\widehat{\theta}_N^{IV} \xrightarrow{p} \theta^*, \text{ as } N \to \infty$$

Instrumental variables approach

$$\widehat{\theta}_N^{IV} = \left(\Psi_N^T \Phi_N\right)^{-1} \Psi_N^T Y_N$$

- Consistency conditions
  (a) $\dim \Psi_N = \dim \Phi_N$, i.e. $\Psi_N = (\boldsymbol{\psi}_1, ..., \boldsymbol{\psi}_N)^T$, $\dim \boldsymbol{\psi}_k = s + p + 1$
  (b) $Plim_{N\to\infty} \left(\frac{1}{N}\Psi_N^T \Phi_N\right)$ exists and is not singular
  (c) $Plim_{N\to\infty} \left(\frac{1}{N}\Psi_N^T \mathbf{Z}_N\right) = 0$

$$\widehat{\theta}_N^{IV} \xrightarrow{p} \theta^*, \text{ as } N \to \infty$$

- Generation of the instruments $\boldsymbol{\psi}_k$
  1) $\boldsymbol{\psi}_k = (u_k, u_{k-1}, ..., u_{k-s}, u_{k-s-1}, ..., u_{k-s-p})^T$
  2) $\boldsymbol{\psi}_k = \left(u_k, u_{k-1}, ..., u_{k-s}, \overline{y}_{k-1}, ..., \overline{y}_{k-p}\right)^T$, where $\overline{y}_{k-i}$ – model output (e.g. L.S.)

Interconnected systems

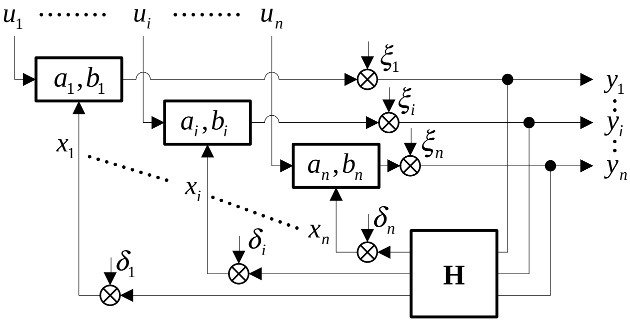

Figure: Interconnected *MIMO* linear static system

Identification of i-th element

$$\begin{aligned} y_i &= a_i x_i + b_i u_i + \xi_i \qquad (i = 1, 2, ..., n) \\ x_i &= H_i y + \delta_i \end{aligned}$$

$$Y_{iN} = (a_i, b_i) W_{iN} + \xi_i$$

$$\begin{aligned} Y_{iN} &= [y_i^{(1)}, y_i^{(2)}, ..., y_i^{(N)}] \\ W_{iN} &= [w_i^{(1)}, w_i^{(2)}, ..., w_i^{(N)}], \quad \text{where } w_i = (x_i, u_i)^T, \end{aligned}$$

Least squares based approach

$$(\widehat{a}_i^{l.s.}, \widehat{b}_i^{l.s.}) = Y_{iN}\widetilde{W}_{iN}^T\left(\widetilde{W}_{iN}\widetilde{W}_{iN}^T\right)^{-1}$$

$$\widetilde{W}_{iN} = [\widetilde{w}_i^{(1)}, \widetilde{w}_i^{(2)}, ..., \widetilde{w}_i^{(N)}]$$

$$\widetilde{w}_i = (\widetilde{x}_i, u_i)^T, \text{ where } \widetilde{x}_i = H_i y = x_i - \delta_i.$$

The estimation error

$$(\widehat{a}_i^{l.s.}, \widehat{b}_i^{l.s.}) - (a_i, b_i) = \Theta_{iN}\widetilde{W}_{iN}\left(\widetilde{W}_{iN}\widetilde{W}_{iN}^T\right)^{-1}$$

does not tend to zero, as $N \longrightarrow \infty$.

Instrumental variables estimate

$$(\widehat{a}_i^{i.v.}, \widehat{b}_i^{i.v.}) = Y_{iN}\Psi_{iN}^T\left(\widetilde{W}_{iN}^T\Psi_{iN}^T\right)^{-1}$$

$$\Psi_{iN} = [\psi_i^{(1)}, \psi_i^{(2)}, ..., \psi_i^{(N)}]$$

$$\psi_i^{(k)} = \left(\psi_{i,1}^{(k)}, \psi_{i,2}^{(k)}\right)^T$$

Theorem

The optimal instruments with respect to the value of

$$Q(\Psi_{iN}) = \|\Delta(\Psi_{iN})\| = \lambda_{\max}\left(\Delta(\Psi_{iN})\Delta^T(\Psi_{iN})\right)$$

has the form

$$\psi_i^* = \overline{w}_i = (\overline{x}_i, u_i)^T, \text{ where } \overline{x}_i = E(x_i|u) = H_iKu.$$

Hammerstein system

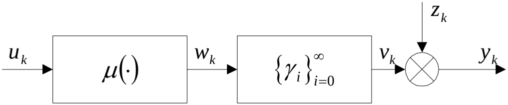

Figure: Hammerstein system

A1: $|u_k| \leqslant u_{\max}$, $\exists$ p.d.f. $\nu(u)$

A2:

$$|\mu(u)| \leqslant w_{\max}$$

A3:

$$\sum_{i=0}^{\infty} |\gamma_i| < \infty$$

A4: $\mu(u_0)$ is known for some u_0 and $\gamma_0 = 1$

A5:

$$z_k = \sum_{i=0}^{\infty} \omega_i \varepsilon_{k-i}$$

$\{\varepsilon_k\}$ – i.i.d. process,
independent of $\{u_k\}$, $E\varepsilon_k = 0$,
$|\varepsilon_k| \leqslant \varepsilon_{\max}$
$\{\omega_i\}_{i=0}^{\infty}$ – unknown,
$\sum_{i=0}^{\infty} |\omega_i| < \infty$

Nonparametric regression

Regression function

$$E\left(y_k | u_k = u\right) = \mu\left(u\right)$$

Kernel estimate

$$\hat{\mu}(u) = \sum_{k=1}^{N} K\left(\frac{u-u_k}{h(N)}\right) \cdot y_k \Bigg/ \sum_{k=1}^{N} K\left(\frac{u-u_k}{h(N)}\right)$$

Orthogonal estimate (wavelet-based)

$$\hat{\mu}\left(u\right) = \sum_{n=0}^{2^M-1} \hat{\alpha}_{Mn}\varphi_{Mn}\left(u\right) + \sum_{m=M}^{K-1} \sum_{n=0}^{2^m-1} \hat{\beta}_{mn}\psi_{mn}\left(u\right)$$

$$\hat{\alpha}_{Mn} = \sum_{l=1}^{k} y_l \int_{u_{l-1}}^{u_l} \varphi_{Mn}\left(u\right) du \qquad \hat{\beta}_{mn} = \sum_{l=1}^{k} y_l \cdot \int_{u_{l-1}}^{u_l} \psi_{mn}\left(u\right) du$$

Parametric knowledge

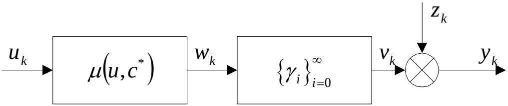

Figure: Hammerstein system (parametric model of the static nonlinearity)

- given $\mu(u,c)$, such that $\mu(u,c^*) = \mu(u)$, where $c^* = (c_1^*, c_2^*, \ldots, c_m^*)^T$ – true parameters
- $\mu(u,c)$ – differentiable with respect to c
- for each $u \in [-u_{\max}, u_{\max}]$ it holds that

$$\|\nabla_c \mu(u,c)\| \leqslant G_{\max} < \infty, \qquad c \in \mathcal{O}(c^*)$$

- c^* is identifiable, i.e. there exits the sequence $\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_{N_0}$ such that

$$\mu(\overline{u}_n, c) = \mu(\overline{u}_n, c^*),\ n = 1, 2, \ldots, N_0 \Rightarrow c = c^*$$

Nonlinearity estimation

$$Q_{N_0}(c) = \sum_{n=1}^{N_0} \left[w_n - \mu(\overline{u}_n, c)\right]^2 \qquad c^* = \arg\min_c Q_{N_0}(c)$$

Stage 1:

$$\widehat{w}_{n,M} = \widehat{R}_M(\overline{u}_n) - \widehat{R}_M(0)$$

Stage 2:

$$\widehat{Q}_{N_0,M}(c) = \sum_{n=1}^{N_0} \left[\widehat{w}_{n,M} - \mu(\overline{u}_n, c)\right]^2 \to \min_c$$

Limit properties

Theorem

If

$$\widehat{R}_M(\overline{u}_n) = R(\overline{u}_n) + O(M^{-\tau}) \text{ in probability, as } M \to \infty$$

then

$$\widehat{c}_{N_0,M} = c^* + O(M^{-\tau}) \text{ in probability, as } M \to \infty.$$

Two-stage identification of linear dynamics

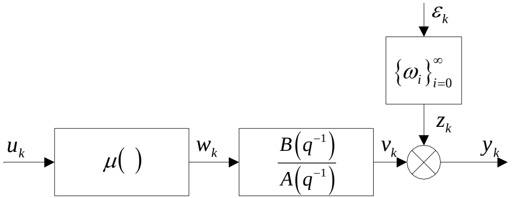

Figure: Hammerstein system (parametric model of the linear component)

$$
\begin{aligned}
v_k &= b_0 w_k + \ldots + b_s w_{k-s} + a_1 v_{k-1} + \ldots. + a_p v_{k-p} \\
\theta &= (b_0, b_1, \ldots, b_s, a_1, a_2, \ldots, a_p)^T \\
\vartheta_k &= (w_k, w_{k-1}, \ldots, w_{k-s}, y_{k-1}, y_{k-2}, \ldots, y_{k-p})^T \\
y_k &= \vartheta_k^T \theta + \overline{z}_k, \qquad \overline{z}_k = z_k - a_1 z_{k-1} - \ldots - a_p z_{k-p} \\
Y_N &= \Theta_N \theta + Z_N, \quad \Theta_N = (\vartheta_1, \ldots, \vartheta_N)^T, \quad Z_N = (\overline{z}_1, \ldots, \overline{z}_N)^T
\end{aligned}
$$

Nonparametric instrumental variables

$$\widehat{\theta}^{(IV)}_{N,M} = (\widehat{\Psi}^T_{N,M}\widehat{\Theta}_{N,M})^{-1}\widehat{\Psi}^T_{N,M}Y_N$$

where

$$\begin{aligned}
\widehat{\Theta}_{N,M} &= (\widehat{\vartheta}_{1,M}, \ldots, \widehat{\vartheta}_{N,M})^T \\
\widehat{\vartheta}_{k,M} &= (\widehat{w}_{k,M}, \ldots, \widehat{w}_{k-s,M}, y_{k-1}, \ldots, y_{k-p})^T \\
\widehat{\Psi}_{N,M} &= (\widehat{\psi}_{1,M}, \ldots, \widehat{\psi}_{N,M})^T \\
\widehat{\psi}_{k,M} &= (\widehat{w}_{k,M}, \ldots, \widehat{w}_{k-s,M}, \widehat{w}_{k-s-1,M}, \ldots, \widehat{w}_{k-s-p,M})^T
\end{aligned}$$

Limit properties

Theorem

It holds that

$$\widehat{\theta}_{N,M}^{(IV)} \longrightarrow \theta \text{ in probability}$$

as $N, M \longrightarrow \infty$, *provided that* $NM^{-\tau} \longrightarrow 0$. *In particular, for* $M \sim N^{(1+\alpha)/\tau}$, $\alpha > 0$, *the asymptotic convergence rate is*

$$\left\| \widehat{\theta}_{N,M}^{(IV)} - \theta \right\| = O(N^{-\min(\frac{1}{2},\alpha)}) \text{ in probability.}$$

Optimal instruments

$$\Delta_N^{(IV)}(\Psi_N) \triangleq \widehat{\theta}_N^{IV} - \theta^*$$

$$Q(\Psi_N) \triangleq \max_{\|Z_N^*\|_2 \leq 1} \left\| \Delta_N^{(IV)}(\Psi_N) \right\|_2^2$$

$$Z_N^* \triangleq \frac{\frac{1}{\sqrt{N}} Z_N}{\overline{z}_{\max}}$$

Theorem

It holds that

$$\lim_{N\to\infty} Q(\Psi_N) \geqslant \lim_{N\to\infty} Q(\Psi_N^*) \text{ with probability } 1$$

where $\Psi_N^* = \left(\psi_1^*, \psi_2^*, \ldots, \psi_N^*\right)^T$, *and*
$\psi_k^* = \left(w_k, \ldots, w_{k-s}, v_{k-1}, \ldots, v_{k-p}\right)^T$.

NARMAX system

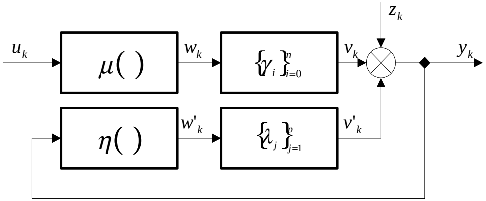

Figure: The NARMAX system

$$
\begin{aligned}
\Lambda &= (\lambda_1, .., \lambda_p)^T \\
\Gamma &= (\gamma_0, ..., \gamma_n)^T \\
c &= (c_1, ..., c_m)^T \\
d &= (d_1, ..., d_q)^T
\end{aligned}
$$

Over-parametrization method

$$\begin{aligned}\theta &= (\gamma_0 c_1, \ldots, \gamma_0 c_m, \ldots, \gamma_n c_1, \ldots, \gamma_n c_m, \\ &\quad \lambda_1 d_1, \ldots, \lambda_1 d_q, \ldots, \lambda_p d_1, \ldots, \lambda_p d_q)^T \\ &= (\theta_1, \ldots, \theta_{(n+1)m}, \theta_{(n+1)m+1}, \ldots, \theta_{(n+1)m+pq})^T\end{aligned}$$

$$\begin{aligned}\phi_k &= (f_1(u_k), \ldots, f_m(u_k), \ldots, f_1(u_{k-n}), \ldots, f_m(u_{k-n}), \\ &\quad g_1(y_{k-1}), \ldots, g_q(y_{k-1}), \ldots, g_1(y_{k-p}), \ldots, g_q(y_{k-p}))^T\end{aligned}$$

$$y_k = \phi_k^T \theta + z_k \qquad Y_N = \Phi_N \theta + Z_N$$

Two-stage estimate

Stage 1. *Instrumental variables*

$$\widehat{\theta}_N^{(\mathrm{IV})} = (\Psi_N^T \Phi_N)^{-1} \Psi_N^T Y_N$$

$\widehat{\Theta}_{\Lambda d}^{(\mathrm{IV})}$, and $\widehat{\Theta}_{\Gamma c}^{(\mathrm{IV})}$ of the matrices $\Theta_{\Lambda d} = \Lambda d^T$ and $\Theta_{\Gamma c} = \Gamma c^T$

Stage 2. *Singular value decomposition (S.V.D.)*

$$\widehat{\Theta}_{\Lambda d}^{(\mathrm{IV})} = \sum_{i=1}^{\min(p,q)} \delta_i \widehat{\xi}_i \widehat{\zeta}_i^T \qquad \widehat{\Theta}_{\Gamma c}^{(\mathrm{IV})} = \sum_{i=1}^{\min(n,m)} \sigma_i \widehat{\mu}_i \widehat{\nu}_i^T$$

$$\begin{aligned}
\widehat{\Lambda}_N^{(\mathrm{IV})} &= sgn(\widehat{\xi}_1[\kappa_{\xi_1}])\widehat{\xi}_1 \\
\widehat{\Gamma}_N^{(\mathrm{IV})} &= sgn(\widehat{\mu}_1[\kappa_{\mu_1}])\widehat{\mu}_1 \\
\widehat{c}_N^{(\mathrm{IV})} &= sgn(\widehat{\mu}_1[\kappa_{\mu_1}])\sigma_1\widehat{\nu}_1 \\
\widehat{d}_N^{(\mathrm{IV})} &= sgn(\widehat{\xi}_1[\kappa_{\xi_1}])\delta_1\widehat{\zeta}_1
\end{aligned}$$

Theorem

If $\det\{\mathbf{E}\psi_k\phi_k^T\} \neq 0$ *and* $\mathbf{E}\psi_k z_k = cov(\psi_k, z_k) = 0$ *then it holds that*

$$\begin{aligned} \widehat{\Lambda}_N^{(\mathrm{IV})} &\to \Lambda \\ \widehat{\Gamma}_N^{(\mathrm{IV})} &\to \Gamma \\ \widehat{c}_N^{(\mathrm{IV})} &\to c \\ \widehat{d}_N^{(\mathrm{IV})} &\to d \end{aligned}$$

with probability 1 *as* $N \to \infty$.

Summary

1. Consistent estimates under correlated excitations and disturbances

Summary

1. Consistent estimates under correlated excitations and disturbances
2. Problem decomposition with the use of nonparametric methods

Summary

1. Consistent estimates under correlated excitations and disturbances
2. Problem decomposition with the use of nonparametric methods
3. Broad class of models (non-linear-in-parameters static blocks + I.I.R. linear filters)