Data mining for system identification

— applications to process identification

André Carvalho Bittencourt

Automatic Control, Linköping University, Sweden

February 23, 2015

Outline

1. Problem formulation
2. Theoretical guiding principles
   - Modeling
   - Data
   - Estimation
3. Tests and outline of algorithm
4. Mining data from an entire plant
5. Concluding remarks

Mining data for system identification

Historical database

- continuous vars r, u, y
- discrete mode variable m
- 195 control loops
- 5 types of process variables
- 4 samples per minute
- data pts 3.1K/min, 4.5M/day

Can we extract useful intervals of data for sysid?

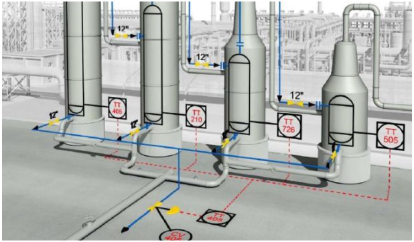
Mining data for system identification

Historical database

- continuous vars r, u, y
- discrete mode variable m
- 195 control loops
- 5 types of process variables
- 4 samples per minute
- data pts 3.1K/min, 4.5M/day

Can we extract useful intervals of data for sysid?

Requisites and assumptions



Requisites

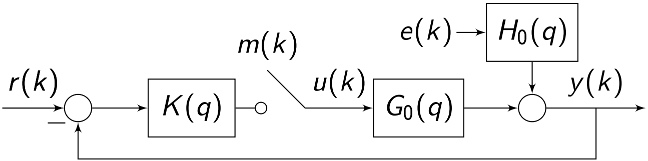
- minimal knowledge
- flexible
- fast
- measure of quality

Assumptions

- SISO loops
- linear dynamics
- real-valued poles*

Approach

- take guidance from the theory
- use flexible models
- and recursive solutions



System $\mathcal{S}$

$$y(k) = G_0(q)u(k) + H_0(q)e(k)$$

Model structure $\mathcal{M}(\theta)$

$$y(k) = G(q,\theta)u(k) + H(q,\theta)e(k)$$

Optimal one-step ahead predictor

$$\hat{y}(k|\theta) = H(q,\theta)^{-1}G(q,\theta)u(k) + \left(1 - H(q,\theta)^{-1}\right)y(k)$$

Model set $\mathcal{M}$

$$\{M(\theta) : \theta \in D_\theta\}$$

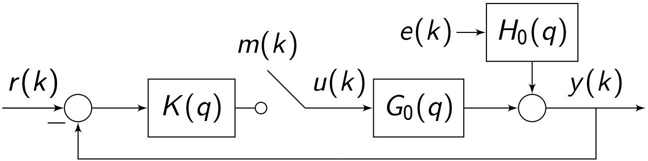
$\mathcal{S} \in \mathcal{M}$

$G_0(z) = G(z,\theta'), H_0(z) = H(z,\theta')$ for $\theta' \in D_\theta$. $\Leftrightarrow \mathcal{S} = \mathcal{M}(\theta')$

"true set" $D_T(\mathcal{S},\mathcal{M})$

$$\{\theta \in D_\theta : \mathcal{S} = \mathcal{M}(\theta)\}$$

Modeling - definitions



System $\mathcal{S}$

$$y(k) = G_0(q)u(k) + H_0(q)e(k)$$

Model set $\mathcal{M}$

$$\{M(\theta) : \theta \in D_\theta\}$$

Model structure $\mathcal{M}(\theta)$

$$y(k) = G(q,\theta)u(k) + H(q,\theta)e(k)$$

$\mathcal{S} \in \mathcal{M}$

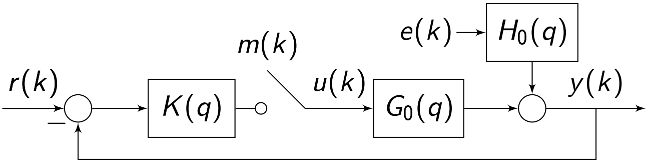
$G_0(z) = G(z,\theta'), H_0(z) = H(z,\theta')$ for $\theta' \in D_\theta$. $\Leftrightarrow \mathcal{S} = \mathcal{M}(\theta')$

Optimal one-step ahead predictor

$$\hat{y}(k|\theta) = H(q,\theta)^{-1}G(q,\theta)u(k) + \left(1 - H(q,\theta)^{-1}\right)y(k)$$

"true set" $D_T(\mathcal{S},\mathcal{M})$

$$\{\theta \in D_\theta : \mathcal{S} = \mathcal{M}(\theta)\}$$



System $\mathcal{S}$

$$y(k) = G_0(q)u(k) + H_0(q)e(k)$$

Model set $\mathcal{M}$

$$\{M(\theta) : \theta \in D_\theta\}$$

Model structure $\mathcal{M}(\theta)$

$$y(k) = G(q,\theta)u(k) + H(q,\theta)e(k)$$

$\mathcal{S} \in \mathcal{M}$

$$G_0(z) = G(z,\theta'),\ H_0(z) = H(z,\theta')$$

for $\theta' \in D_\theta$. $\Leftrightarrow \mathcal{S} = \mathcal{M}(\theta')$

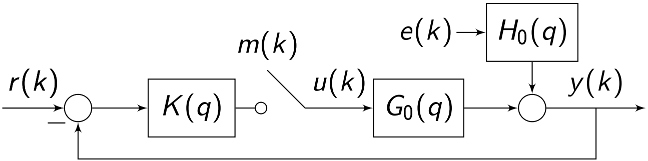
Optimal one-step ahead predictor

$$\hat{y}(k|\theta) = H(q,\theta)^{-1}G(q,\theta)u(k) + \left(1 - H(q,\theta)^{-1}\right)y(k)$$

"true set" $D_T(\mathcal{S},\mathcal{M})$

$$\{\theta \in D_\theta : \mathcal{S} = \mathcal{M}(\theta)\}$$

Modeling - definitions



System $\mathcal{S}$

$$y(k) = G_0(q)u(k) + H_0(q)e(k)$$

Model structure $\mathcal{M}(\theta)$

$$y(k) = G(q,\theta)u(k) + H(q,\theta)e(k)$$

Optimal one-step ahead predictor

$$\hat{y}(k|\theta) = H(q,\theta)^{-1}G(q,\theta)u(k) + \left(1 - H(q,\theta)^{-1}\right)y(k)$$

Model set $\mathcal{M}$

$$\{M(\theta) : \theta \in D_\theta\}$$

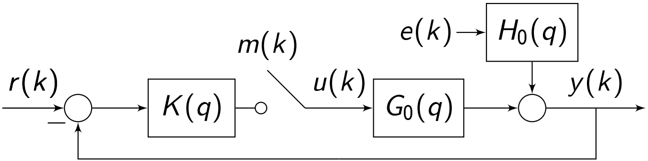
$\mathcal{S} \in \mathcal{M}$

$$G_0(z) = G(z,\theta'), H_0(z) = H(z,\theta')$$
for $\theta' \in D_\theta$. $\Leftrightarrow \mathcal{S} = \mathcal{M}(\theta')$

"true set" $D_T(\mathcal{S},\mathcal{M})$

$$\{\theta \in D_\theta : \mathcal{S} = \mathcal{M}(\theta)\}$$

Modeling - definitions



System $\mathcal{S}$

$$y(k) = G_0(q)u(k) + H_0(q)e(k)$$

Model structure $\mathcal{M}(\theta)$

$$y(k) = G(q,\theta)u(k) + H(q,\theta)e(k)$$

Optimal one-step ahead predictor

$$\hat{y}(k|\theta) = H(q,\theta)^{-1}G(q,\theta)u(k) + \left(1 - H(q,\theta)^{-1}\right)y(k)$$

Model set $\mathcal{M}$

$$\{M(\theta) : \theta \in D_\theta\}$$

$\mathcal{S} \in \mathcal{M}$

$$G_0(z) = G(z,\theta'), H_0(z) = H(z,\theta')$$

for $\theta' \in D_\theta$. $\Leftrightarrow \mathcal{S} = \mathcal{M}(\theta')$

"true set" $D_T(\mathcal{S},\mathcal{M})$

$$\{\theta \in D_\theta : \mathcal{S} = \mathcal{M}(\theta)\}$$

Properties of models

One-to-one relation $T(q,\theta) \leftrightarrow W(q,\theta)$

$$y(k) = [G(q,\theta) \quad H(q,\theta)] \begin{bmatrix} u(k) \\ e(k) \end{bmatrix} = T(q,\theta)x(k)$$

$$\hat{y}(k|\theta) = [H(q,\theta)^{-1}G(q,\theta) \quad (1 - H(q,\theta)^{-1})] \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} = W(q,\theta)z(k)$$

Identifiability at θ'

Whether no other θ gives the same freq resp:
$W(z,\theta) = W(z,\theta'), \forall z \implies \theta = \theta'$

"True parameter" θ_0

If $\mathcal{S} \in \mathcal{M}$ and $\mathcal{M}$ is *globally identifiable*, then $D_T(\mathcal{S},\mathcal{M}) = \theta_0$.

Remarks

- choose $\mathcal{S} \in \mathcal{M}$
- what about the data?

One-to-one relation $T(q,\theta) \leftrightarrow W(q,\theta)$

$$y(k) = \begin{bmatrix} G(q,\theta) & H(q,\theta) \end{bmatrix} \begin{bmatrix} u(k) \\ e(k) \end{bmatrix} = T(q,\theta)x(k)$$

$$\hat{y}(k|\theta) = \begin{bmatrix} H(q,\theta)^{-1}G(q,\theta) & (1 - H(q,\theta)^{-1}) \end{bmatrix} \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} = W(q,\theta)z(k)$$

Identifiability at θ'

Whether no other θ gives the same freq resp:
$W(z,\theta) = W(z,\theta'), \forall z \implies \theta = \theta'$

"True parameter" θ_0

If $\mathcal{S} \in \mathcal{M}$ and $\mathcal{M}$ is *globally identifiable*, then $D_T(\mathcal{S},\mathcal{M}) = \theta_0$.

Remarks

- choose $\mathcal{S} \in \mathcal{M}$
- what about the data?

Properties of models

One-to-one relation $T(q,\theta) \leftrightarrow W(q,\theta)$

$$y(k) = \begin{bmatrix} G(q,\theta) & H(q,\theta) \end{bmatrix} \begin{bmatrix} u(k) \\ e(k) \end{bmatrix} = T(q,\theta)x(k)$$

$$\hat{y}(k|\theta) = \begin{bmatrix} H(q,\theta)^{-1}G(q,\theta) & (1 - H(q,\theta)^{-1}) \end{bmatrix} \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} = W(q,\theta)z(k)$$

Identifiability at θ'

Whether no other θ gives the same freq resp:
$W(z,\theta) = W(z,\theta'), \forall z \implies \theta = \theta'$

"True parameter" θ_0

If $\mathcal{S} \in \mathcal{M}$ and $\mathcal{M}$ is *globally identifiable*,
then $D_T(\mathcal{S},\mathcal{M}) = \theta_0$.

Remarks

- choose $\mathcal{S} \in \mathcal{M}$
- what about the data?

Properties of models

One-to-one relation $T(q,\theta) \leftrightarrow W(q,\theta)$

$$y(k) = \begin{bmatrix} G(q,\theta) & H(q,\theta) \end{bmatrix} \begin{bmatrix} u(k) \\ e(k) \end{bmatrix} = T(q,\theta)x(k)$$

$$\hat{y}(k|\theta) = \begin{bmatrix} H(q,\theta)^{-1}G(q,\theta) & (1 - H(q,\theta)^{-1}) \end{bmatrix} \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} = W(q,\theta)z(k)$$

Identifiability at θ'

Whether no other θ gives the same freq resp:
$W(z,\theta) = W(z,\theta'), \forall z \implies \theta = \theta'$

"True parameter" θ_0

If $\mathcal{S} \in \mathcal{M}$ and $\mathcal{M}$ is *globally identifiable*,
then $D_T(\mathcal{S},\mathcal{M}) = \theta_0$.

Remarks

- choose $\mathcal{S} \in \mathcal{M}$
- what about the data?

Flexible model structures - $\mathcal{S} \in \mathcal{M}$

ARX: delay expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i z^{-i} \triangleq B_n(z)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i z^{-i} \triangleq A_n(z)$$

- exact if $n \to \infty$
- convergence depends on T_s
- unknown delays

L-ARX: Laguerre expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i L_i(z,\alpha) \triangleq \tilde{B}_n(z,\alpha)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i L_i(z,\alpha) \triangleq \tilde{A}_n(z,\alpha)$$

- exact if $n \to \infty$
- convergence depends on α
- more efficient with delays
- real poles

$$L_i(q,\alpha) = \frac{\sqrt{(1-\alpha^2)T_s}}{q-\alpha}\left(\frac{1-\alpha q}{q-\alpha}\right)^{i-1}$$

Remarks

- describe “any” linear dynamics
- globally identifiable
- linear regressions
- Laguerre better for delays
- choice of n and α

- $W_u(z,\boldsymbol{\theta_b}) = \tilde{B}_{n_b}(z,\alpha)$
- $W_y(z,\boldsymbol{\theta_a}) = A_{n_a}(z)$
- integrate $u(k)$ if integrating plant
- guess of largest delay and time cte

Flexible model structures - $\mathcal{S} \in \mathcal{M}$

ARX: delay expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i z^{-i} \triangleq B_n(z)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i z^{-i} \triangleq A_n(z)$$

- exact if $n \to \infty$
- convergence depends on T_s
- unknown delays

L-ARX: Laguerre expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i L_i(z,\alpha) \triangleq \tilde{B}_n(z,\alpha)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i L_i(z,\alpha) \triangleq \tilde{A}_n(z,\alpha)$$

- exact if $n \to \infty$
- convergence depends on α
- more efficient with delays
- real poles

$$L_i(q,\alpha) = \frac{\sqrt{(1-\alpha^2)T_s}}{q-\alpha}\left(\frac{1-\alpha q}{q-\alpha}\right)^{i-1}$$

Remarks

- describe "any" linear dynamics
- globally identifiable
- linear regressions
- Laguerre better for delays
- choice of n and α

- $W_u(z,\boldsymbol{\theta_b}) = \tilde{B}_{n_b}(z,\alpha)$
- $W_y(z,\boldsymbol{\theta_a}) = A_{n_a}(z)$
- integrate $u(k)$ if integrating plant
- guess of largest delay and time cte

ARX: delay expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i z^{-i} \triangleq B_n(z)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i z^{-i} \triangleq A_n(z)$$

- exact if $n \to \infty$
- convergence depends on T_s
- unknown delays

L-ARX: Laguerre expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i L_i(z,\alpha) \triangleq \tilde{B}_n(z,\alpha)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i L_i(z,\alpha) \triangleq \tilde{A}_n(z,\alpha)$$

- exact if $n \to \infty$
- convergence depends on α
- more efficient with delays
- real poles

$$L_i(q,\alpha) = \frac{\sqrt{(1-\alpha^2)T_s}}{q-\alpha}\left(\frac{1-\alpha q}{q-\alpha}\right)^{i-1}$$

Remarks

- describe “any” linear dynamics
- globally identifiable
- linear regressions
- Laguerre better for delays
- choice of n and α

- $W_u(z,\boldsymbol{\theta_b}) = \tilde{B}_{n_b}(z,\alpha)$
- $W_y(z,\boldsymbol{\theta_a}) = A_{n_a}(z)$
- integrate $u(k)$ if integrating plant
- guess of largest delay and time cte

Flexible model structures - $\mathcal{S} \in \mathcal{M}$

ARX: delay expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i z^{-i} \triangleq B_n(z)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i z^{-i} \triangleq A_n(z)$$

- exact if $n \to \infty$
- convergence depends on T_s
- unknown delays

L-ARX: Laguerre expansion

$$W_u(z,\theta) \approx \sum_{i=1}^{n} b_i L_i(z,\alpha) \triangleq \tilde{B}_n(z,\alpha)$$
$$W_y(z,\theta) \approx \sum_{i=1}^{n} a_i L_i(z,\alpha) \triangleq \tilde{A}_n(z,\alpha)$$

- exact if $n \to \infty$
- convergence depends on α
- more efficient with delays
- real poles

$$L_i(q,\alpha) = \frac{\sqrt{(1-\alpha^2)T_s}}{q-\alpha}\left(\frac{1-\alpha q}{q-\alpha}\right)^{i-1}$$

Remarks

- describe “any” linear dynamics
- globally identifiable
- linear regressions
- Laguerre better for delays
- choice of n and α

- $W_u(z,\boldsymbol{\theta_b}) = \tilde{B}_{n_b}(z,\alpha)$
- $W_y(z,\boldsymbol{\theta_a}) = A_{n_a}(z)$
- integrate $u(k)$ if integrating plant
- guess of largest delay and time cte

Properties of the data

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[\left(W(z,\theta_1) - W(z,\theta_2)\right) z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b

closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \not\equiv 0$, servo iff $r(k)$ is SRn_r $n_r \geq \min(n_a - n_x, n_b - n_y)$

Properties of the data

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[(W(z,\theta_1) - W(z,\theta_2))\, z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b

closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection
  iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \not\equiv 0$, servo
  iff $r(k)$ is SRn_r
  $n_r \geq \min(n_a - n_x, n_b - n_y)$

Properties of the data

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[(W(z,\theta_1) - W(z,\theta_2))\, z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b

closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \not\equiv 0$, servo iff $r(k)$ is SRn_r, $n_r \geq \min(n_a - n_x, n_b - n_y)$

Properties of the data

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[\left(W(z,\theta_1) - W(z,\theta_2)\right) z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b

closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection
  iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \not\equiv 0$, servo
  iff $r(k)$ is SRn_r
  $n_r \geq \min(n_a - n_x, n_b - n_y)$

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[\left(W(z,\theta_1) - W(z,\theta_2)\right) z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b

closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection
  iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \not\equiv 0$, servo
  iff $r(k)$ is SRn_r
  $n_r \geq \min(n_a - n_x, n_b - n_y)$

Step signal example

Let $u(k) = \Delta(k)$,
$\phi(k) = [\Delta(k-1), \ldots, \Delta(k-n)]$
$\bar{E}[\phi(k)\phi(k)^T] = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, SR1!

Properties of the data

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[\left(W(z,\theta_1) - W(z,\theta_2)\right) z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b
closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \neq 0$, servo iff $r(k)$ is SRn_r $n_r \geq \min(n_a - n_x, n_b - n_y)$

Remarks

- how to verify? finite sample?
- disturbance rejection?
- look at the estimate!

Informative enough $\{z(k)\}_1^N$

Distinguishes non-equiv. models
$\bar{E}\left[\left(W(z,\theta_1) - W(z,\theta_2)\right) z(k)\right]^2 = 0$
$\Rightarrow W(e^{\imath\omega},\theta_1) \equiv W(e^{\imath\omega},\theta_2)$

Persistent excitation (PE) of $\phi(k)$

Full rank information matrix
$\bar{E}\left[\phi(k)\phi(k)^T\right] > 0$

n-Suff. rich signal $u(k)$ (SRn)

if $\phi(k) = [u(k-1), \cdots, u(k-n)]$ is PE.

ARX informative enough data

open: iff $u(k)$ is SRn_b
closed: let $K(q) = \frac{X(q)}{Y(q)}$

- $r(k) \equiv 0$, dist. rejection
  iff $(n_x - n_a, n_y - n_b) \geq 0$
- $r(k) \not\equiv 0$, servo
  iff $r(k)$ is SRn_r
  $n_r \geq \min(n_a - n_x, n_b - n_y)$

Remarks

- how to verify? finite sample?
- disturbance rejection?
- look at the estimate!

The RLS estimate

$$\hat{\theta}_k = \arg\min_{\theta \in D_\theta} V_k(\theta) = \arg\min_{\theta \in D_\theta} \sum_{i=1}^{k} \lambda^{k-i} \varepsilon^2(k, \theta)$$

Consistent, i.e. $\hat{\theta}_\infty \in D_T(\mathcal{S}, \mathcal{M})(= \theta_0)$, if

- $\mathcal{S} \in \mathcal{M}$, $\mathcal{M}(\theta)$ is identifiable (globally)
- $\lambda \to 1$, Z^N is informative enough

Remarks

- open, excitation in u
- servo, excitation in r through feedback
-

Frequency representation

The RLS estimate

$$\hat{\theta}_k = \arg\min_{\theta \in D_\theta} V_k(\theta) = \arg\min_{\theta \in D_\theta} \sum_{i=1}^{k} \lambda^{k-i} \varepsilon^2(k,\theta)$$

Consistent, i.e. $\hat{\theta}_\infty \in D_T(\mathcal{S},\mathcal{M}) (= \theta_0)$, if

- $\mathcal{S} \in \mathcal{M}$, $\mathcal{M}(\theta)$ is identifiable (globally)
- $\lambda \to 1$, Z^N is informative enough

Remarks

- open, excitation in u
- servo, excitation in r through feedback
-

Frequency representation

For constant $\lambda^{k-i} = \frac{1}{2N}$ and $N \to \infty$. $\bar{V}(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \Phi_\varepsilon(\omega,\theta) d\omega$ and

$$\Phi_\varepsilon = \frac{|G_0 + B_\theta - G_\theta|^2}{|H_\theta|^2} \Phi_u + \frac{|H_0 - H_\theta|^2}{|H_\theta|^2} \left(\gamma_0 - \frac{\Phi_{ue}}{\Phi_u} \right) + \gamma_0$$

$$B_\theta = (H_0 - H_\theta) \frac{\Phi_{ue}}{\Phi_u}$$

The RLS estimate

$$\hat{\theta}_k = \arg\min_{\theta \in D_\theta} V_k(\theta) = \arg\min_{\theta \in D_\theta} \sum_{i=1}^{k} \lambda^{k-i} \varepsilon^2(k, \theta)$$

Consistent, i.e. $\hat{\theta}_\infty \in D_T(\mathcal{S}, \mathcal{M}) (= \theta_0)$, if

- $\mathcal{S} \in \mathcal{M}$, $\mathcal{M}(\theta)$ is identifiable (globally)
- $\lambda \to 1$, Z^N is informative enough

Remarks

- open, excitation in u
- servo, excitation in r through feedback
-

Frequency representation

Open-loop, $\Phi_{ue} \equiv 0$. Unbiased estimate.

$$B_\theta = 0$$

The RLS estimate

$$\hat{\theta}_k = \arg\min_{\theta \in D_\theta} V_k(\theta) = \arg\min_{\theta \in D_\theta} \sum_{i=1}^{k} \lambda^{k-i} \varepsilon^2(k,\theta)$$

Consistent, i.e. $\hat{\theta}_\infty \in D_T(\mathcal{S},\mathcal{M})(=\theta_0)$, if

- $\mathcal{S} \in \mathcal{M}$, $\mathcal{M}(\theta)$ is identifiable (globally)
- $\lambda \to 1$, Z^N is informative enough

Remarks

- open, excitation in u
- servo, excitation in r through feedback

Frequency representation

Servo, $\Phi_u = \Phi_u^r + \Phi_u^e$ and $|\Phi_{ue}|^2 = \Phi_u^e \gamma_0$. Excitation in r through feedback!

$$|B_\theta|^2 = |H_0 - H_\theta|^2 \frac{\gamma_0 \Phi_u^e}{(\Phi_u^r + \Phi_u^e)^2}$$

Recursive least squares

The RLS estimate

$$\hat{\theta}_k = \arg\min_{\theta \in D_\theta} V_k(\theta) = \arg\min_{\theta \in D_\theta} \sum_{i=1}^{k} \lambda^{k-i} \varepsilon^2(k,\theta)$$

Consistent, i.e. $\hat{\theta}_\infty \in D_T(\mathcal{S}, \mathcal{M})(= \theta_0)$, if

- $\mathcal{S} \in \mathcal{M}$, $\mathcal{M}(\theta)$ is identifiable (globally)
- $\lambda \to 1$, Z^N is informative enough

Remarks

- open, excitation in u
- servo, excitation in r through feedback
- dist, excitation in e through feedback

Frequency representation

Dist. rejection, $\Phi_u^r \equiv 0$. **Noise must affect the input!**

$$|B_\theta|^2 = |H_0 - H_\theta|^2 \frac{\gamma_0}{\Phi_u^e}$$

The RLS estimate

$$\hat{\theta}_k = \arg\min_{\theta \in D_\theta} V_k(\theta) = \arg\min_{\theta \in D_\theta} \sum_{i=1}^{k} \lambda^{k-i} \varepsilon^2(k, \theta)$$

Consistent, i.e. $\hat{\theta}_\infty \in D_T(\mathcal{S}, \mathcal{M})(= \theta_0)$, if

- $\mathcal{S} \in \mathcal{M}$, $\mathcal{M}(\theta)$ is identifiable (globally)
- $\lambda \to 1$, Z^N is informative enough

Remarks

- open, excitation in u
- servo, excitation in r through feedback
- ~~dist, excitation in e through feedback~~

Frequency representation

Dist. rejection, $\Phi_u^r \equiv 0$. **Noise must affect the input!**

$$|B_\theta|^2 = |H_0 - H_\theta|^2 \frac{\gamma_0}{\Phi_u^e}$$

RLS for linear regressions

Linear regression

(L-)ARX: $\hat{y}(k|\theta) = \varphi(k)^T\theta$

Recursive solution

$$\hat{\theta}_k - \hat{\theta}_{k-1} = \bar{R}(k)^{-1}\varphi(k)\varepsilon(k, \hat{\theta}_{k-1})$$

$$\bar{R}(k) = \lambda\bar{R}(k-1) + \varphi(k)\varphi(k)^T$$

$$V_k(\hat{\theta}_k) = \lambda V_{k-1}(\hat{\theta}_{k-1}) + \varepsilon(k, \hat{\theta}_{k-1})\varepsilon(k, \hat{\theta}_k)$$

Remarks

- closed-form recursive sol
- invertibility of R
- QR-RLS solution

Asymptotics

Let $\lambda \to 1$ and $k \to \infty$

$$\sqrt{1-\lambda}(\hat{\theta}_k - \theta_0) \in As\mathcal{N}(0, P_\theta),$$

$$P_\theta \triangleq \frac{1}{2}\gamma_0 \bar{E}\left[\varphi(k)\varphi(k)^T\right]^{-1}$$

$$\Sigma_\theta = \frac{(1-\lambda)}{2}\gamma_0 \bar{R}^{-1}$$

Finite sample estimates

$$\hat{\gamma}_k = (1-\lambda)V_k(\hat{\theta}_k)$$

$$\hat{\bar{R}}_k = (1-\lambda)\bar{R}(k)$$

$$\hat{\Sigma}_k = \frac{(1-\lambda)}{2}V_k(\hat{\theta}_k)\bar{R}(k)^{-1}$$

RLS for linear regressions

Linear regression

(L-)ARX: $\hat{y}(k|\theta) = \varphi(k)^T \theta$

Recursive solution

$$\hat{\theta}_k - \hat{\theta}_{k-1} = \bar{R}(k)^{-1}\varphi(k)\varepsilon(k, \hat{\theta}_{k-1})$$
$$\bar{R}(k) = \lambda \bar{R}(k-1) + \varphi(k)\varphi(k)^T$$
$$V_k(\hat{\theta}_k) = \lambda V_{k-1}(\hat{\theta}_{k-1}) + \varepsilon(k, \hat{\theta}_{k-1})\varepsilon(k, \hat{\theta}_k)$$

Asymptotics

Let $\lambda \to 1$ and $k \to \infty$

$$\sqrt{1-\lambda}(\hat{\theta}_k - \theta_0) \in As\mathcal{N}(0, P_\theta),$$
$$P_\theta \triangleq \frac{1}{2}\gamma_0 \bar{E}\left[\varphi(k)\varphi(k)^T\right]^{-1}$$
$$\Sigma_\theta = \frac{(1-\lambda)}{2}\gamma_0 \bar{R}^{-1}$$

Remarks

- closed-form recursive sol
- invertibility of R
- QR-RLS solution

Finite sample estimates

$$\hat{\gamma}_k = (1-\lambda)V_k(\hat{\theta}_k)$$
$$\hat{\bar{R}}_k = (1-\lambda)\bar{R}(k)$$
$$\hat{\Sigma}_k = \frac{(1-\lambda)}{2}V_k(\hat{\theta}_k)\bar{R}(k)^{-1}$$

RLS for linear regressions

Linear regression

(L-)ARX: $\hat{y}(k|\theta) = \varphi(k)^T\theta$

Recursive solution

$$\begin{aligned}\hat{\theta}_k - \hat{\theta}_{k-1} &= \bar{R}(k)^{-1}\varphi(k)\varepsilon(k,\hat{\theta}_{k-1})\\ \bar{R}(k) &= \lambda\bar{R}(k-1) + \varphi(k)\varphi(k)^T\\ V_k(\hat{\theta}_k) &= \lambda V_{k-1}(\hat{\theta}_{k-1})\\ &\quad + \varepsilon(k,\hat{\theta}_{k-1})\varepsilon(k,\hat{\theta}_k)\end{aligned}$$

Remarks

- closed-form recursive sol
- invertibility of R
- QR-RLS solution

Asymptotics

Let $\lambda \to 1$ and $k \to \infty$

$$\sqrt{1-\lambda}(\hat{\theta}_k - \theta_0) \in As\mathcal{N}(0, P_\theta),$$

$$P_\theta \triangleq \frac{1}{2}\gamma_0 \bar{E}\left[\varphi(k)\varphi(k)^T\right]^{-1}$$

$$\Sigma_\theta = \frac{(1-\lambda)}{2}\gamma_0 \bar{R}^{-1}$$

Finite sample estimates

$$\hat{\gamma}_k = (1-\lambda)V_k(\hat{\theta}_k)$$

$$\hat{\bar{R}}_k = (1-\lambda)\bar{R}(k)$$

$$\hat{\Sigma}_k = \frac{(1-\lambda)}{2}V_k(\hat{\theta}_k)\bar{R}(k)^{-1}$$

Linear regression

(L-)ARX: $\hat{y}(k|\theta) = \varphi(k)^T\theta$

Recursive solution

$$\begin{aligned}\hat{\theta}_k - \hat{\theta}_{k-1} &= \bar{R}(k)^{-1}\varphi(k)\varepsilon(k,\hat{\theta}_{k-1})\\ \bar{R}(k) &= \lambda\bar{R}(k-1) + \varphi(k)\varphi(k)^T\\ V_k(\hat{\theta}_k) &= \lambda V_{k-1}(\hat{\theta}_{k-1})\\ &\quad + \varepsilon(k,\hat{\theta}_{k-1})\varepsilon(k,\hat{\theta}_k)\end{aligned}$$

Remarks

- closed-form recursive sol
- invertibility of R
- QR-RLS solution

Asymptotics

Let $\lambda \to 1$ and $k \to \infty$

$$\sqrt{1-\lambda}(\hat{\theta}_k - \theta_0) \in As\mathcal{N}(0, P_\theta),$$

$$P_\theta \triangleq \frac{1}{2}\gamma_0 \bar{E}\left[\varphi(k)\varphi(k)^T\right]^{-1}$$

$$\Sigma_\theta = \frac{(1-\lambda)}{2}\gamma_0 \bar{R}^{-1}$$

Finite sample estimates

$$\hat{\gamma}_k = (1-\lambda)V_k(\hat{\theta}_k)$$

$$\hat{\bar{R}}_k = (1-\lambda)\bar{R}(k)$$

$$\hat{\Sigma}_k = \frac{(1-\lambda)}{2}V_k(\hat{\theta}_k)\bar{R}(k)^{-1}$$

RLS for linear regressions

Linear regression

(L-)ARX: $\hat{y}(k|\theta) = \varphi(k)^T\theta$

Recursive solution

$$\hat{\theta}_k - \hat{\theta}_{k-1} = \bar{R}(k)^{-1}\varphi(k)\varepsilon(k, \hat{\theta}_{k-1})$$
$$\bar{R}(k) = \lambda\bar{R}(k-1) + \varphi(k)\varphi(k)^T$$
$$V_k(\hat{\theta}_k) = \lambda V_{k-1}(\hat{\theta}_{k-1}) + \varepsilon(k, \hat{\theta}_{k-1})\varepsilon(k, \hat{\theta}_k)$$

Remarks

- closed-form recursive sol
- invertibility of R
- QR-RLS solution

Asymptotics

Let $\lambda \to 1$ and $k \to \infty$

$$\sqrt{1-\lambda}(\hat{\theta}_k - \theta_0) \in As\mathcal{N}(0, P_\theta),$$
$$P_\theta \triangleq \frac{1}{2}\gamma_0 \bar{E}\left[\varphi(k)\varphi(k)^T\right]^{-1}$$
$$\Sigma_\theta = \frac{(1-\lambda)}{2}\gamma_0 \bar{R}^{-1}$$

Finite sample estimates

$$\hat{\gamma}_k = (1-\lambda)V_k(\hat{\theta}_k)$$
$$\hat{\bar{R}}_k = (1-\lambda)\bar{R}(k)$$
$$\hat{\Sigma}_k = \frac{(1-\lambda)}{2}V_k(\hat{\theta}_k)\bar{R}(k)^{-1}$$

0: Normalize the data

T_1: step changes in $u(k)$ or $r(k)$

That is how the plant is operated!

$|u(k)| > \eta_1$ or $|r(k)| > \eta_1$

T_2: variability in $y(k)$

$y(k)$ should vary after step

Monitor variance of y $\gamma_y(k) > \eta_2$

T_3: conditioning of info matrix

Check whether $\bar{R}(k)$ is invertible

$\kappa_2^{-1}(\bar{R}(k)) = \frac{\sigma_{\min}(\bar{R}(k))}{\sigma_{\max}(\bar{R}(k))} > \eta_3$

T_4: Granger causality test

Can $u(k)$ help predict $y(k)$?

For $\hat{y}(k|\theta) = \varphi_u(k)^T \theta_k^b + \varphi_y(k)^T \theta_k^a$

under $\mathcal{H}_0 : \theta^b = 0$

$s(k) = (\hat{\theta}_k^b)^T \, (\Sigma_\theta^b)^{-1} \, \hat{\theta}_k^b \in As\,\mathcal{X}_{n_b}$

$s(k)$ is used as a **quality measure**!

5: Logical conditions

- T_{i+1} only computed if T_i passed
- Exit if any test fails
- Accept interval if T_4 passed
- Interval goes from T_1 to exit

Tests

0: Normalize the data

T_1: step changes in $u(k)$ or $r(k)$

That is how the plant is operated!

$|u(k)| > \eta_1$ or $|r(k)| > \eta_1$

T_2: variability in $y(k)$

$y(k)$ should vary after step

Monitor variance of y $\gamma_y(k) > \eta_2$

T_3: conditioning of info matrix

Check whether $\bar{R}(k)$ is invertible

$$\kappa_2^{-1}(\bar{R}(k)) = \frac{\sigma_{\min}(\bar{R}(k))}{\sigma_{\max}(\bar{R}(k))} > \eta_3$$

T_4: Granger causality test

Can $u(k)$ help predict $y(k)$?

For $\hat{y}(k|\theta) = \varphi_u(k)^T \theta_k^b + \varphi_y(k)^T \theta_k^a$

under $\mathcal{H}_0 : \theta^b = 0$

$$s(k) = (\hat{\theta}_k^b)^T (\Sigma_\theta^b)^{-1} \hat{\theta}_k^b \in As\,\mathcal{X}_{n_b}$$

$s(k)$ is used as a **quality measure**!

5: Logical conditions

- T_{i+1} only computed if T_i passed
- Exit if any test fails
- Accept interval if T_4 passed
- Interval goes from T_1 to exit

Tests

0: Normalize the data

T_1: step changes in $u(k)$ or $r(k)$

That is how the plant is operated!

$|u(k)| > \eta_1$ or $|r(k)| > \eta_1$

T_2: variability in $y(k)$

$y(k)$ should vary after step

Monitor variance of y $\gamma_y(k) > \eta_2$

T_3: conditioning of info matrix

Check whether $\bar{R}(k)$ is invertible

$$\kappa_2^{-1}(\bar{R}(k)) = \frac{\sigma_{\min}(\bar{R}(k))}{\sigma_{\max}(\bar{R}(k))} > \eta_3$$

T_4: Granger causality test

Can $u(k)$ help predict $y(k)$?

For $\hat{y}(k|\theta) = \varphi_u(k)^T\theta_k^b + \varphi_y(k)^T\theta_k^a$

under $\mathcal{H}_0 : \theta^b = 0$

$$s(k) = (\hat{\theta}_k^b)^T\,(\Sigma_\theta^b)^{-1}\,\hat{\theta}_k^b \in As\,\mathcal{X}_{n_b}$$

$s(k)$ is used as a **quality measure!**

5: Logical conditions

- T_{i+1} only computed if T_i passed
- Exit if any test fails
- Accept interval if T_4 passed
- Interval goes from T_1 to exit

0: Normalize the data

T_1: step changes in $u(k)$ or $r(k)$

That is how the plant is operated!

$|u(k)| > \eta_1$ or $|r(k)| > \eta_1$

T_2: variability in $y(k)$

$y(k)$ should vary after step

Monitor variance of y $\gamma_y(k) > \eta_2$

T_3: conditioning of info matrix

Check whether $\bar{R}(k)$ is invertible

$$\kappa_2^{-1}(\bar{R}(k)) = \frac{\sigma_{\min}(\bar{R}(k))}{\sigma_{\max}(\bar{R}(k))} > \eta_3$$

T_4: Granger causality test

Can $u(k)$ help predict $y(k)$?

For $\hat{y}(k|\theta) = \varphi_u(k)^T\theta_k^b + \varphi_y(k)^T\theta_k^a$

under $\mathcal{H}_0 : \theta^b = 0$

$$s(k) = (\hat{\theta}_k^b)^T\,(\Sigma_\theta^b)^{-1}\,\hat{\theta}_k^b \in As\,\mathcal{X}_{n_b}$$

$s(k)$ is used as a **quality measure!**

5: Logical conditions

- T_{i+1} only computed if T_i passed
- Exit if any test fails
- Accept interval if T_4 passed
- Interval goes from T_1 to exit

0: Normalize the data

T_1: step changes in $u(k)$ or $r(k)$

That is how the plant is operated!

$|u(k)| > \eta_1$ or $|r(k)| > \eta_1$

T_2: variability in $y(k)$

$y(k)$ should vary after step

Monitor variance of y $\gamma_y(k) > \eta_2$

T_3: conditioning of info matrix

Check whether $\bar{R}(k)$ is invertible

$$\kappa_2^{-1}(\bar{R}(k)) = \frac{\sigma_{\min}(\bar{R}(k))}{\sigma_{\max}(\bar{R}(k))} > \eta_3$$

T_4: Granger causality test

Can $u(k)$ help predict $y(k)$?

For $\hat{y}(k|\theta) = \varphi_u(k)^T \theta_k^b + \varphi_y(k)^T \theta_k^a$

under $\mathcal{H}_0 : \theta^b = 0$

$$s(k) = (\hat{\theta}_k^b)^T \, (\Sigma_\theta^b)^{-1} \, \hat{\theta}_k^b \in As\, \mathcal{X}_{n_b}$$

$s(k)$ is used as a **quality measure!**

5: Logical conditions

- T_{i+1} only computed if T_i passed
- Exit if any test fails
- Accept interval if T_4 passed
- Interval goes from T_1 to exit

FIR Example, Granger causality test

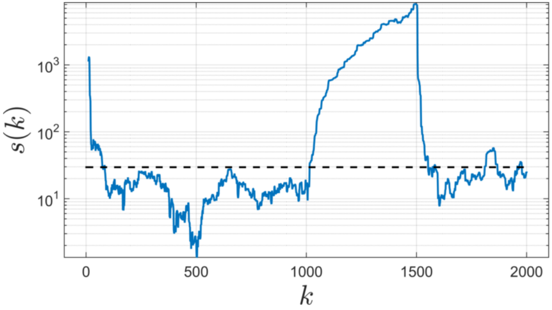
$$y(k) = B_{10}(q)\,(u(k)+d(k)) + e(k)$$

$$\theta \neq 0, k > 1000$$

$$d(k) \neq 0, \begin{cases} 500 < k < 1000 \\ 1500 < k < 2000 \end{cases}$$

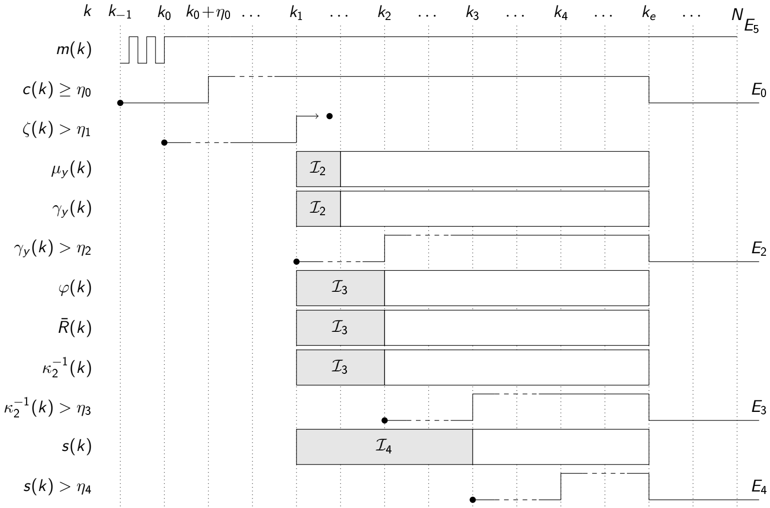
Remarks

- fails as a causality test
- but works well to select data!
- statistical significance of (any) parameter



SNR_u=10, SNR_d=30

Outline of the algorithm



Mining data from an entire plant

Plant

- 195 control loops
- 37 months of data
- 1.15G samples

Evaluation

- 170 minutes to run
- selects 1.46% of all samples
- finds all "bump" tests
- every test is important
- quality measure supports further analysis

	Mode of operation type count (%)		Average length	
Loop type	open	closed	open	closed
Density (i)	14.59	1.20	76	88
Flow	1.37	5.00	199	419
Level (i)	3.51	0.25	72	127
Pressure (i)	5.00	3.00	64	108
Temperature	0.80	0.01	67	76

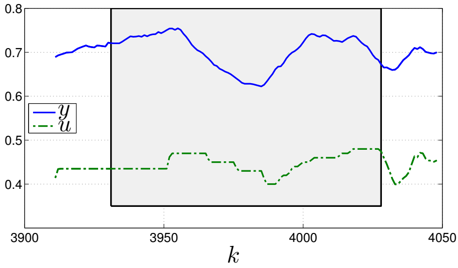
Plant

- 195 control loops
- 37 months of data
- 1.15G samples

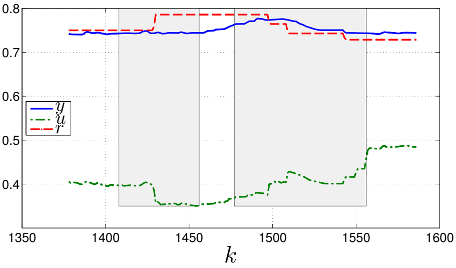
Evaluation

- 170 minutes to run
- selects 1.46% of all samples
- finds all “bump” tests
- every test is important
- quality measure supports further analysis

Level, open



Density, closed



Summary

Requisites

- range of variables (normalization)
- mode operation type (change in u or r)
- integrating plant (finite gain models)
- guess of largest delay (tuning)
- guess of largest time cte (tuning)
- 7 tuning vars, 5 thresholds for entire plant

Extensions

- Kautz polynominals (complex poles)
- finding the topology
- MIMO case

LINKÖPINGS UNIVERSITET

LINKÖPINGS UNIVERSITET