



Identificação de Sistemas Não-Lineares: Algoritmos

Eduardo Mendes

`emmendes@cpdee.ufmg.br`

Departamento de Engenharia Eletrônica
Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, Belo Horizonte, MG, Brasil



Introdução

- ▶ No presente capítulo serão apresentados algoritmos para estimação de parâmetros de representações que são lineares em tais parâmetros. Será suposto que os modelos são não-lineares, o que necessariamente implicará o problema de ter que escolher a sua estrutura dentre uma ampla gama de possibilidades.



Algoritmos MQ Ortogonais

O ponto de partida é a equação matricial

$$\mathbf{y} = \Psi \hat{\boldsymbol{\theta}} + \boldsymbol{\xi},$$

sendo

$$\mathbf{y} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}; \quad \Psi = [\boldsymbol{\psi}_1 \dots \boldsymbol{\psi}_{n_\theta}]; \quad \hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_{n_\theta} \end{bmatrix}; \quad \boldsymbol{\xi} = \begin{bmatrix} \xi(1) \\ \xi(2) \\ \vdots \\ \xi(N) \end{bmatrix}$$

e

$$\boldsymbol{\psi}_i = \begin{bmatrix} \psi_i(1) \\ \psi_i(2) \\ \vdots \\ \psi_i(N) \end{bmatrix}, \quad i = 1, 2, \dots, n_\theta.$$



Matriz de regressores de um modelo NARX

- ▶ Considere o seguinte vetor de regressores:

$$\boldsymbol{\psi}(k-1) = \begin{bmatrix} y(k-1) & y(k-2) & u(k-1) & y(k-1)^2 & y(k-1)y(k-2) \\ y(k-2)^2 & u(k-1)^2 & y(k-1)u(k-1) & y(k-2)u(k-1) \end{bmatrix}^T .$$



Matriz de regressores de um modelo NARX

- ▶ Considere o seguinte vetor de regressores:

$$\psi(k-1) = \begin{bmatrix} y(k-1) & y(k-2) & u(k-1) & y(k-1)^2 & y(k-1)y(k-2) \\ y(k-2)^2 & u(k-1)^2 & y(k-1)u(k-1) & y(k-2)u(k-1) \end{bmatrix}^T .$$

- ▶ As primeiras 4 linhas da matriz de regressores são

$$\Psi = \begin{bmatrix} y(0) & y(-1) & u(0) & \dots & y(0)u(0) & y(-1)u(0) \\ y(1) & y(0) & u(1) & \dots & y(1)u(1) & y(0)u(1) \\ y(2) & y(1) & u(2) & \dots & y(2)u(2) & y(1)u(2) \\ y(3) & y(2) & u(3) & \dots & y(3)u(3) & y(2)u(3) \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \end{bmatrix} .$$



▶ Além disso, tem-se

$$\psi_1 = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix}; \quad \psi_2 = \begin{bmatrix} y(-1) \\ y(0) \\ \vdots \\ y(N-2) \end{bmatrix}; \quad \psi_9 = \begin{bmatrix} y(-1)u(0) \\ y(0)u(1) \\ \vdots \\ y(N-2)u(N-1) \end{bmatrix}$$

e assim sucessivamente.



O método clássico de Gram-Schmidt (CGS)

- ▶ O objetivo neste método é fatorar a matriz de regressores Ψ , que é de posto pleno, da seguinte forma $\Psi = QA$, sendo que

$$A = \begin{bmatrix} 1 & \alpha_{1,2} & \alpha_{1,3} & \dots & \alpha_{1,n_\theta} \\ 0 & 1 & \alpha_{2,3} & \dots & \alpha_{2,n_\theta} \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & 0 & 0 & 1 & \alpha_{n_\theta-1,n_\theta} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

é uma matriz triangular superior de dimensão $n_\theta \times n_\theta$, e Q é uma matriz de dimensão $N \times n_\theta$ com colunas ortogonais tal que $Q^T Q = D$, sendo D uma matriz diagonal, definida positiva.



▶ Portanto, a partir da equação normal, tem-se

$$\Psi^T \Psi \hat{\theta} = \Psi^T \mathbf{y}$$

$$A^T Q^T Q A \hat{\theta} = A^T Q^T \mathbf{y}$$

$$A^T D A \hat{\theta} = A^T Q^T \mathbf{y}$$

$$D A \hat{\theta} = Q^T \mathbf{y}$$

$$A \hat{\theta} = D^{-1} Q^T \mathbf{y},$$

sendo que a última equação fornece os parâmetros $\hat{\mathbf{g}} = A \hat{\theta}$ do modelo no espaço ortogonal.

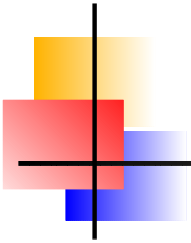
- 
- ▶ Portanto, a partir da equação normal, tem-se

$$\begin{aligned}\Psi^T \Psi \hat{\boldsymbol{\theta}} &= \Psi^T \mathbf{y} \\ A^T Q^T Q A \hat{\boldsymbol{\theta}} &= A^T Q^T \mathbf{y} \\ A^T D A \hat{\boldsymbol{\theta}} &= A^T Q^T \mathbf{y} \\ D A \hat{\boldsymbol{\theta}} &= Q^T \mathbf{y} \\ A \hat{\boldsymbol{\theta}} &= D^{-1} Q^T \mathbf{y},\end{aligned}$$

sendo que a última equação fornece os parâmetros $\hat{\mathbf{g}} = A \hat{\boldsymbol{\theta}}$ do modelo no espaço ortogonal.

- ▶ Portanto, tais parâmetros podem ser determinados usando-se a última equação ou

$$\hat{g}_i = \frac{\langle \mathbf{q}_i, \mathbf{y} \rangle}{\langle \mathbf{q}_i, \mathbf{q}_i \rangle}, \quad i = 1, 2, \dots, n_\theta.$$

- 
-
- ▶ Finalmente, os parâmetros do modelo original (representado no espaço de regressores não ortogonais) podem ser determinados como $\hat{\theta} = A^{-1}\hat{g}$.



O método modificado de Gram-Schmidt (MGS)

- ▶ O método modificado de Gram-Schmidt objetiva resolver problemas de sensibilidade a mal-condicionamento numérico observados no método clássico. A fim de atingir esse alvo, o método MGS determina a matriz A uma linha por vez e ortogonaliza Ψ de tal forma que na i -ésima iteração as colunas $i + 1, \dots, n_\theta$ são feitas ortogonais à i -ésima coluna.



O método de Golub-Householder (GH)

- ▶ Pré-multiplicando o modelo por uma matriz $Q \in \mathbb{R}^{n_\theta \times n_\theta}$, tem-se

$$Qy(k) = Q\psi^T \hat{\theta} + Q\xi(k)$$

$$y^*(k) = \psi^{*T} \hat{\theta} + \xi^*(k).$$



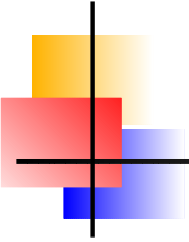
O método de Golub-Householder (GH)

- ▶ Pré-multiplicando o modelo por uma matriz $Q \in \mathbb{R}^{n_\theta \times n_\theta}$, tem-se

$$\begin{aligned}Qy(k) &= Q\psi^\top \hat{\boldsymbol{\theta}} + Q\xi(k) \\y^*(k) &= \psi^{*\top} \hat{\boldsymbol{\theta}} + \xi^*(k).\end{aligned}$$

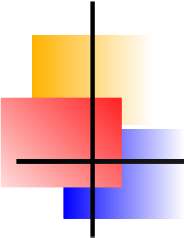
- ▶ A função custo de mínimos quadrados neste caso torna-se

$$\begin{aligned}J_{\text{MQ}}^* &= \sum_{i=1}^N \xi(i)^{*2} = \boldsymbol{\xi}^{*\top} \boldsymbol{\xi}^* \\&= (\mathbf{y}^* - \Psi^* \hat{\boldsymbol{\theta}})^\top (\mathbf{y}^* - \Psi^* \hat{\boldsymbol{\theta}}) \\&= \mathbf{y}^{*\top} \mathbf{y}^* - \mathbf{y}^{*\top} \Psi^* \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^\top \Psi^{*\top} \mathbf{y}^* + \hat{\boldsymbol{\theta}}^\top \Psi^{*\top} \Psi^* \hat{\boldsymbol{\theta}} \\&= \mathbf{y}^\top Q^\top Q \mathbf{y} - \mathbf{y}^\top Q^\top Q \Psi \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^\top \Psi^\top Q^\top Q \mathbf{y} + \hat{\boldsymbol{\theta}}^\top \Psi^\top Q^\top Q \Psi \hat{\boldsymbol{\theta}} \\&= (\mathbf{y} - \Psi \hat{\boldsymbol{\theta}})^\top Q^\top Q (\mathbf{y} - \Psi \hat{\boldsymbol{\theta}}).\end{aligned}$$

- 
-
- ▶ Para que $J_{\text{MQ}}^* = J_{\text{MQ}} = (\mathbf{y} - \Psi\hat{\boldsymbol{\theta}})^\text{T}(\mathbf{y} - \Psi\hat{\boldsymbol{\theta}})$, necessariamente $Q^\text{T}Q = I$, que é a condição para que a matriz Q seja ortonormal. Além disso, impõe-se que

$$\Psi^* = Q\Psi = \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix},$$

sendo que V é triangular superior e $V \in \mathbb{R}^{n_\theta \times n_\theta}$.

- 
- ▶ Para que $J_{\text{MQ}}^* = J_{\text{MQ}} = (\mathbf{y} - \Psi\hat{\boldsymbol{\theta}})^T(\mathbf{y} - \Psi\hat{\boldsymbol{\theta}})$, necessariamente $Q^T Q = I$, que é a condição para que a matriz Q seja ortonormal. Além disso, impõe-se que

$$\Psi^* = Q\Psi = \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix},$$

sendo que V é triangular superior e $V \in \mathbb{R}^{n_\theta \times n_\theta}$.

- ▶ Então,

$$\begin{aligned} \Psi^{*\text{T}} \Psi^* &= [V^T \quad \mathbf{0}^T] \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \\ &= V^T V \\ &= \Psi^T Q^T Q \Psi \\ &= \Psi^T \Psi. \end{aligned}$$



► Reescrevendo J_{MQ}^* , tem-se

$$\begin{aligned} J_{\text{MQ}}^* &= \left(\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \right)^T \left(\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \right) \\ &= \left(\begin{bmatrix} \mathbf{y}_1^{*\text{T}} & \mathbf{y}_2^{*\text{T}} \end{bmatrix} - \hat{\boldsymbol{\theta}}^T \begin{bmatrix} V^T & \mathbf{0} \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \right) \\ &= \mathbf{y}_1^{*\text{T}} \mathbf{y}_1^* + \mathbf{y}_2^{*\text{T}} \mathbf{y}_2^* - \hat{\boldsymbol{\theta}}^T V^T \mathbf{y}_1^* - \mathbf{y}_1^{*\text{T}} V \hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\theta}}^T V^T V \hat{\boldsymbol{\theta}} \\ &= (\mathbf{y}_1^* - V \hat{\boldsymbol{\theta}})^T (\mathbf{y}_1^* - V \hat{\boldsymbol{\theta}}) + \mathbf{y}_2^{*\text{T}} \mathbf{y}_2^*. \end{aligned}$$

- 
- Reescrevendo J_{MQ}^* , tem-se

$$\begin{aligned} J_{\text{MQ}}^* &= \left(\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \right)^T \left(\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \right) \\ &= \left(\begin{bmatrix} \mathbf{y}_1^{*\text{T}} & \mathbf{y}_2^{*\text{T}} \end{bmatrix} - \hat{\boldsymbol{\theta}}^T \begin{bmatrix} V^T & \mathbf{0} \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\theta}} \right) \\ &= \mathbf{y}_1^{*\text{T}} \mathbf{y}_1^* + \mathbf{y}_2^{*\text{T}} \mathbf{y}_2^* - \hat{\boldsymbol{\theta}}^T V^T \mathbf{y}_1^* - \mathbf{y}_1^{*\text{T}} V \hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\theta}}^T V^T V \hat{\boldsymbol{\theta}} \\ &= (\mathbf{y}_1^* - V \hat{\boldsymbol{\theta}})^T (\mathbf{y}_1^* - V \hat{\boldsymbol{\theta}}) + \mathbf{y}_2^{*\text{T}} \mathbf{y}_2^*. \end{aligned}$$

- Claramente, J_{MQ}^* é minimizado quando $\mathbf{y}_1^* = V \hat{\boldsymbol{\theta}}$. Nesse caso, tem-se que $\hat{\boldsymbol{\theta}}_{\text{MQ}} = V^{-1} \mathbf{y}_1^*$ e conseqüentemente $J_{\text{MQ}}^* = \mathbf{y}_2^{*\text{T}} \mathbf{y}_2^*$. Como V é triangular superior, a estimativa

$$\hat{\boldsymbol{\theta}}_{\text{MQ}} = V^{-1} \mathbf{y}_1^*$$

pode ser calculada com relativa facilidade.



A Taxa de Redução de Erro

- ▶ A fim de definir a taxa de redução de erro, será considerado o seguinte modelo NARMAX geral:

$$\begin{aligned}y(k) &= \boldsymbol{\psi}^T(k-1)\hat{\boldsymbol{\theta}} + \xi(k) \\ &= \sum_{i=1}^{n_\theta} \hat{\theta}_i \psi_i(k-1) + \xi(k)\end{aligned}$$



A Taxa de Redução de Erro

- ▶ A fim de definir a taxa de redução de erro, será considerado o seguinte modelo NARMAX geral:

$$\begin{aligned}y(k) &= \boldsymbol{\psi}^T(k-1)\hat{\boldsymbol{\theta}} + \xi(k) \\ &= \sum_{i=1}^{n_\theta} \hat{\theta}_i \psi_i(k-1) + \xi(k)\end{aligned}$$

- ▶ e o modelo auxiliar $y(k) = \sum_{i=1}^{n_\theta} \hat{g}_i w_i(k-1) + \xi(k)$, sendo que os regressores w_i são ortogonais sobre os dados, ou seja,

$$\begin{aligned}\frac{1}{N} \sum_{k=1}^N w_i(k)w_{j+1}(k) &= 0, \\ \overline{w_i(k)w_{j+1}(k)} &= 0, \quad \forall i \neq j,\end{aligned}$$

ou simplesmente $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0$, sendo que $\mathbf{w}_i = [w_i(1) \dots w_i(N)]^T$.

- 
-
- ▶ A soma dos valores quadráticos de $y(k)$ é $\langle \mathbf{y}, \mathbf{y} \rangle$ ou $\mathbf{y}^T \mathbf{y}$.
Portanto,

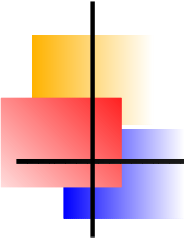
$$y(k)^2 = \left(\sum_{i=1}^{n_\theta} \hat{g}_i w_i(k-1) + \xi(k) \right) \times \left(\sum_{i=1}^{n_\theta} \hat{g}_j w_j(k-1) + \xi(k) \right).$$

- 
- ▶ A soma dos valores quadráticos de $y(k)$ é $\langle \mathbf{y}, \mathbf{y} \rangle$ ou $\mathbf{y}^T \mathbf{y}$. Portanto,

$$y(k)^2 = \left(\sum_{i=1}^{n_\theta} \hat{g}_i w_i(k-1) + \xi(k) \right) \times \left(\sum_{i=1}^{n_\theta} \hat{g}_j w_j(k-1) + \xi(k) \right).$$

- ▶ Tomando-se o valor médio de sobre os dados, tem-se

$$\begin{aligned} \langle \mathbf{y}, \mathbf{y} \rangle &= \sum_{i=1}^{n_\theta} \hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle + \sum_{i=1, j=1, \forall i \neq j}^{n_\theta} 2\hat{g}_i \hat{g}_j \langle \mathbf{w}_i, \mathbf{w}_j \rangle + \\ &\quad + 2 \sum_{i=1}^{n_\theta} \hat{g}_i \langle \mathbf{w}_i, \boldsymbol{\xi} \rangle + \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle \\ &= \sum_{i=1}^{n_\theta} \hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle + \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle. \end{aligned}$$

- 
- ▶ Se o i -ésimo regressor for acrescentado ao modelo, a parcela de $\langle \mathbf{y}, \mathbf{y} \rangle$ que passa a explicar é $\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle$. Portanto, a *taxa de redução de erro* devido à inclusão do i -ésimo regressor expressa como uma fração da soma dos valores quadráticos dos dados é

$$[ERR]_i = \frac{\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle},$$

sendo que a sigla ERR vem do inglês *error reduction ratio*.



Identificação da equação logística

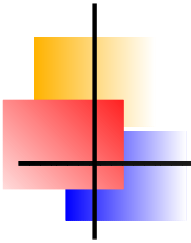

- ▶ O presente exemplo considerará poucos (oito) valores gerados pela equação logística

$y(k) = 3,9 y(k-1) - 3,9 y(k-1)^2$. Inicializando-se tal equação com valores entre 0 e 1, obtém-se uma seqüência de números que é caótica.

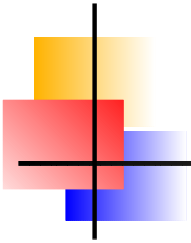



Identificação da equação logística

- ▶ O presente exemplo considerará poucos (oito) valores gerados pela equação logística $y(k) = 3,9 y(k-1) - 3,9 y(k-1)^2$. Inicializando-se tal equação com valores entre 0 e 1, obtém-se uma seqüência de números que é caótica.
- ▶ A fim de obter um modelo a partir dos dados, geram-se todos os possíveis regressores candidatos até segunda ordem com grau de não-linearidade até dois, ou seja, para $n_y = 2$ e $\ell = 2$.

- 
- ▶ Usando-se o algoritmo GH (exlog ) , chega-se à seguinte matriz

$$V = \begin{bmatrix} -2,8284 & -1,2701 & -1,7023 & -1,2874 & -1,7113 & -0,9264 \\ & 0,9446 & 0,8170 & -0,4957 & -0,2807 & 0,4001 \\ & & -0,1641 & 0,4364 & 0,3867 & -0,0263 \\ & & & -0,7097 & -0,6961 & -0,3622 \\ & & & & & -0,0383 \\ & & & & & -0,1232 \end{bmatrix}$$


- 
- ▶ Usando-se o algoritmo GH (exlog ) , chega-se à seguinte matriz

$$V = \begin{bmatrix} -2,8284 & -1,2701 & -1,7023 & -1,2874 & -1,7113 & -0,9264 \\ & 0,9446 & 0,8170 & -0,4957 & -0,2807 & 0,4001 \\ & & -0,1641 & 0,4364 & 0,3867 & -0,0263 \\ & & & -0,7097 & -0,6961 & -0,3622 \\ & & & & & -0,0383 \\ & & & & & -0,1232 \end{bmatrix}$$

- ▶ Pivotando-se a matriz de regressores originais chega-se aos termos: constante, $y(k-1)^2$, $y(k-1)$ e $y(k-2)^2$, com as respectivas taxas de redução de erro: 0,8302; 0,0524; 0,1175 e $6,74 \times 10^{-32} \approx 0$. Estimando-se os parâmetros dos três primeiros regressores obtém-se, respectivamente: $-1,15 \times 10^{-14}$, -3,9 e 3,9. O regressor constante é espúrio e pode ser omitido do modelo.



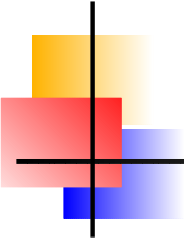
Escolha de agrupamento de termos para oscilador eletrônico

- ▶ Os dados utilizados neste exemplo foram coletados de um oscilador eletrônico não-linear (dsvc1 ). Foram tomados o termo constante e as combinações até terceiro grau dos monômios $y(k - 1)$ e $y(k - 2)$, ou seja, foi considerado $n_y = 2$ e $\ell = 3$, ao todo são dez regressores.



▶ O resultado foi o seguinte:

Regressor	$\hat{\theta}_i$	ERR
$y(k-1)$	2,0318	$9,8444 \times 10^{-1}$
$y(k-2)$	$-9,6099 \times 10^{-1}$	$1,3762 \times 10^{-2}$
$y(k-1)^2 y(k-2)$	6,5580	$9,6521 \times 10^{-4}$
$y(k-1)^2$	$2,5508 \times 10^{-2}$	$1,3457 \times 10^{-5}$
$y(k-1)^3$	-2,5916	$1,8259 \times 10^{-6}$
$y(k-1)y(k-2)^2$	-5,6382	$4,8803 \times 10^{-5}$
$y(k-2)^3$	1,5341	$2,3595 \times 10^{-6}$
$y(k-2)^2$	$3,3658 \times 10^{-2}$	$2,9550 \times 10^{-8}$
$y(k-1)y(k-2)$	$-6,4813 \times 10^{-2}$	$1,1252 \times 10^{-7}$
constante	$-1,6458 \times 10^{-4}$	$1,4065 \times 10^{-8}$

- 
-
- ▶ Os termos quadráticos e o constante são os últimos na lista. O termo $y(k - 1)^2$ que aparece em quarto lugar na lista de prioridades do algoritmo. O coeficiente do agrupamento Ω_{y^2} que é $\Sigma_{y^2} = -5,6471 \times 10^{-3}$. Isso sugere um possível cancelamento, o que, normalmente, pode ser entendido como evidência de que Ω_{y^2} é um agrupamento espúrio.



Identificação de modelo com termo exponencial

- ▶ Foram gerados dados usando-se o seguinte modelo com termo exponencial:

$$y(k) = 0,4 y(k-1) + 0,5 y(k-2) + u(k-1) - 0,3 y(k-1) \exp[-u(k-2)^2] + e(k),$$

sendo que $u(k)$ e $e(k)$ foram escolhidos como variáveis aleatórias independentes, com distribuição gaussiana, média zero e variâncias $\sigma_u^2 = 0,9$ e $\sigma_e^2 = 0,0025$, respectivamente.




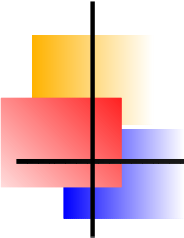
Identificação de modelo com termo exponencial

- ▶ Foram gerados dados usando-se o seguinte modelo com termo exponencial:

$$y(k) = 0,4 y(k-1) + 0,5 y(k-2) + u(k-1) - 0,3 y(k-1) \exp[-u(k-2)^2] + e(k),$$

sendo que $u(k)$ e $e(k)$ foram escolhidos como variáveis aleatórias independentes, com distribuição gaussiana, média zero e variâncias $\sigma_u^2 = 0,9$ e $\sigma_e^2 = 0,0025$, respectivamente.

- ▶ Foram tomados 490 valores, gerados pelo modelo, e um conjunto de 21 regressores candidatos (extended ). Além de todos os multinômios de grau $\ell = 2$ e ordem dois $n_y = n_u = 2$, foram gerados regressores exponenciais da forma $y(k-j) \exp[-x(k-i)^2]$, com $i, j = 1, 2$ e $x(k) = u(k), y(k)$.

- 
-
- ▶ O uso do algoritmo descrito resultou na escolha dos regressores $u(k-1)$, $y(k-2)$, $y(k-1)$ e $y(k-1)\exp[-u(k-2)^2]$ com os seguintes parâmetros, respectivamente: 1,0022; 0,5016; 0,3978 e -0,2928 com taxa de redução de erro de 0,54778; 0,37039; $7,1152 \times 10^{-2}$ e $9,1295 \times 10^{-3}$.



Algoritmos para Modelos Racionais

- ▶ Assume-se que um modelo racional do tipo

$$y(k) = \frac{a(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u))}{b(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u))} + c(e(k-1), \dots, e(k-n_e)) + e(k),$$

sendo que o ruído é modelado como sendo um polinômio, podendo ser linear ou não.



Algoritmos para Modelos Racionais

- ▶ Assume-se que um modelo racional do tipo

$$y(k) = \frac{a(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u))}{b(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u))} + c(e(k-1), \dots, e(k-n_e)) + e(k),$$

sendo que o ruído é modelado como sendo um polinômio, podendo ser linear ou não.

- ▶ Assim sendo, sugere-se o seguinte procedimento:

- 
1. Faça $i = 0$. Monte a matriz de regressores e estime os coeficientes usando MQ

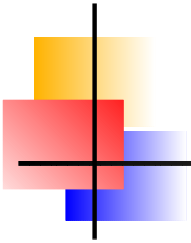
$$\begin{bmatrix} \hat{\theta}_n^i \\ \hat{\theta}_{d1}^i \end{bmatrix} = [\Psi^T \Psi]^{-1} \Psi^T \mathbf{y}^*,$$

sendo que o índice i indica a interação. Além disso,

$$\Psi = \begin{bmatrix} \boldsymbol{\psi}_n^T(k-1) & \boldsymbol{\psi}_{d1}^T(k-1) \\ \vdots & \vdots \\ \boldsymbol{\psi}_n^T(k+N-2) & \boldsymbol{\psi}_{d1}^T(k+N-2) \end{bmatrix}.$$

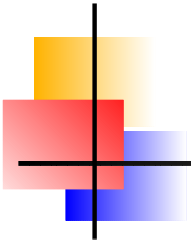
Analogamente, o vetor $\mathbf{y}^* \in \mathbb{R}^{N \times 1}$ é formado tomando $y^*(k)$ sobre os dados, ou seja,

$$\mathbf{y}^{*T} = [y^*(k) \ y^*(k+1) \ \dots \ y^*(k+N-1)].$$

- 
2. Faça $i = i + 1$. Determine os resíduos e a sua variância, respectivamente, como:

$$\xi^i(k) = y(k) - \frac{\psi_n^T(k-1)\hat{\theta}_n^i}{\psi_d^T(k-1) \begin{bmatrix} 1 \\ \hat{\theta}_d^i \end{bmatrix}}$$
$$(\sigma_\xi^2)^i = \frac{1}{N - m_d} \sum_{i=m_d+1}^N (\xi^i(k))^2,$$

sendo que N é o tamanho dos dados e $m_d = \max(n_y, n_u, n_e)$.

- 
3. Usando-se os resíduos determinados no passo 2, atualize $\Psi^T \Psi$ e $\Psi^T \mathbf{y}^*$ usando (-15), a seguir. Além disso, atualize (no caso de $i = 1$, essa matriz deverá ser formada) a seguinte matriz:

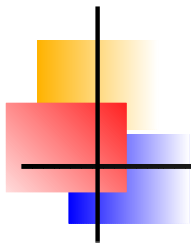
$$\Psi = \begin{bmatrix} \boldsymbol{\psi}_n^T(k-1) & y(k)\boldsymbol{\psi}_{d1}^T(k-1) & \boldsymbol{\psi}_\xi^T(k-1) \\ \vdots & \vdots & \vdots \\ \boldsymbol{\psi}_n^T(k+N-2) & y(k)\boldsymbol{\psi}_{d1}^T(k+N-2) & \boldsymbol{\psi}_\xi^T(k+N-2) \end{bmatrix},$$

sendo que $\boldsymbol{\psi}_\xi$ é o vetor de regressores do modelo de ruído.



4. Determine:

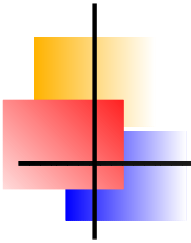
$$\Phi = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots \\ \vdots & & \vdots & \vdots & \\ 0 & \dots & 0 & y(k) \sum_{k=1}^N p_{d2}^2 & \dots \\ \vdots & & \vdots & \vdots & \\ 0 & \dots & 0 & y(k) \sum_{k=1}^N p_{dN_d} p_{d2} & \dots \\ \dots & & 0 & 0 & \dots & 0 \\ & & \vdots & \vdots & \vdots & \\ \dots & y(k) \sum_{k=1}^N p_{d2} p_{dN_d} & 0 & \dots & 0 \\ & & \vdots & \vdots & \vdots & \\ \dots & y(k) \sum_{k=1}^N p_{dN_d}^2 & 0 & \dots & 0 \end{bmatrix}$$



$$\phi = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -y(k) \sum_{k=1}^N p_{d2} p_{d1} \\ -y(k) \sum_{k=1}^N p_{dN_d} p_{d1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

e estime novamente os parâmetros usando

$$\begin{bmatrix} \hat{\theta}_n^i \\ \hat{\theta}_{d1}^i \end{bmatrix} = [\Psi^T \Psi - (\xi_\xi^2)^i \Phi]^{-1} [\Psi^T \mathbf{y}^* - (\xi_\xi^2)^i \phi].$$

- 
-
5. Volte ao passo 2 até atingir convergência (de parâmetros ou de variância de resíduos).



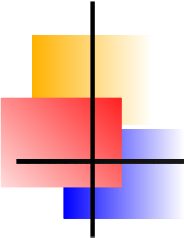
Mapa de conversor CC-CC buck

- ▶ Este exemplo considera um modelo racional estimado a partir de dados gerados por um modelo de um conversor CC-CC do tipo buck. O mapa obtido a partir das equações do circuito tem a seguinte forma:

$$y(k) = \alpha y(k-1) + \frac{h(d_n)^2 \beta E [E - y(k-1)]}{y(k-1)},$$

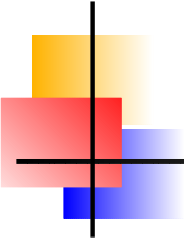
sendo $\alpha = 0,8872$, $\beta = 1,2$ e $E = 33$. d_n é um sinal de tensão que implementa a ação de controle e a saturação $h(d_n)$ é dada por

$$h(d_n) = \begin{cases} 0 & \text{se } d_n < 0, \\ 1 & \text{se } d_n > 1, \\ d_n & \text{caso contrário.} \end{cases}$$

- 
-
- ▶ Modelos polinomiais não resultam em bons modelos para a dinâmica do mapa. Um modelo com estrutura ad hoc é:

$$y(k) = 46,429 \exp[22 - y(k-1)] + \frac{a y(k-1)^2 - b y(k-1) + c}{y(k-1)},$$

sendo $a = 2,6204$, $b = 99,875$ e $c = 1,4171 \times 10^3$.

- 
- ▶ Modelos polinomiais não resultam em bons modelos para a dinâmica do mapa. Um modelo com estrutura ad hoc é:

$$y(k) = 46,429 \exp[22 - y(k-1)] + \frac{a y(k-1)^2 - b y(k-1) + c}{y(k-1)},$$

sendo $a = 2,6204$, $b = 99,875$ e $c = 1,4171 \times 10^3$.

- ▶ Por outro lado, o seguinte modelo racional

$$y(k) = \frac{1}{D} \times \{8,658 + 0,1223 \times 10^{-2} y(k-1)^3 - 0,441 \times 10^{-1} y(k-1)^2\},$$

com

$$D = 1 - 0,8381 \times 10^{-1} y(k-1) + 0,1766 \times 10^{-2} y(k-1)^2,$$

aproxima bem a dinâmica em questão.

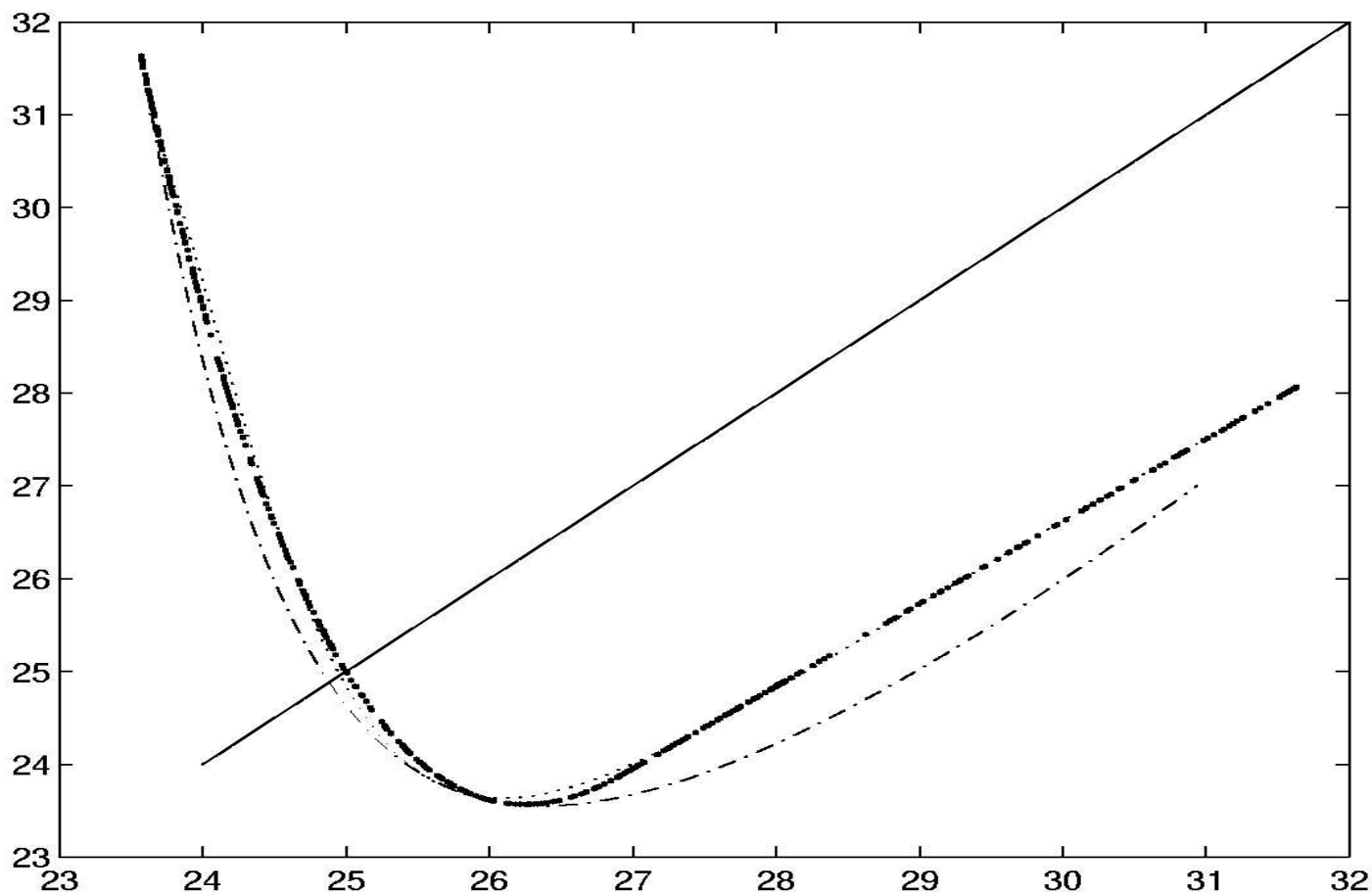
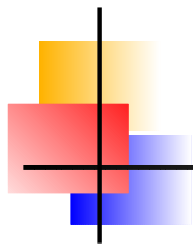


Figura 1: Eixo x é $y(k)$ e eixo y é $y(k + 1)$, ambos em volts. Mapas de primeiro retorno do mapa original: pontilhado denso; do modelo racional: pontilhado leve; do modelo ad-hoc: traço-ponto.