Capítulo 8 - Estimadores Recursivos

Eduardo Mendes

`emmendes@cpdee.ufmg.br`

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

Introdução

Será que é possível utilizar os dados seqüencialmente para atualizar o vetor de parâmetros de um determinado modelo? A resposta é afirmativa e o procedimento é conhecido como estimação recursiva.

No presente capítulo serão estudados alguns algoritmos recursivos. Algumas pequenas modificações serão feitas na nomenclatura. Neste capítulo, o subíndice k indicará a iteração do respectivo algoritmo recursivo.

Atualização Recursiva

- Seja o sistema $y(k) = \boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\theta} + e(k)$, sendo E$[e(k)]$=0 e cov$[e(k)] = R$. Um modelo para o sistema acima pode ser escrito da seguinte maneira:

  $$y(k) = \boldsymbol{\psi}_k^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}}_k + \xi(k),$$

  sendo que o vetor de regressores $\boldsymbol{\psi}_k(k-1)$ é formado na iteração k com informação disponível até a iteração $k-1$ e $\hat{\boldsymbol{\theta}}_k$ indica o vetor de parâmetros estimado na iteração k.

Atualização Recursiva

- Seja o sistema $y(k) = \boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\theta} + e(k)$, sendo E$[e(k)]$=0 e cov$[e(k)] = R$. Um modelo para o sistema acima pode ser escrito da seguinte maneira:

$$y(k) = \boldsymbol{\psi}_k^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}}_k + \xi(k),$$

  sendo que o vetor de regressores $\boldsymbol{\psi}_k(k-1)$ é formado na iteração k com informação disponível até a iteração $k-1$ e $\hat{\boldsymbol{\theta}}_k$ indica o vetor de parâmetros estimado na iteração k.

- Propõe-se, então, escrever

$$\hat{\boldsymbol{\theta}}_k = J_k\hat{\boldsymbol{\theta}}_{k-1} + K_k y(k),$$

  e J_k e K_k são matrizes que deverão ser determinadas de tal maneira a garantir que $\hat{\boldsymbol{\theta}}_k$ seja uma boa estimativa.

Serão usadas as seguintes restrições:

i) $\hat{\boldsymbol{\theta}}_k$ deve ser não polarizado, ou seja, $\mathrm{E}[\hat{\boldsymbol{\theta}}_k] - \boldsymbol{\theta} = 0$;

Serão usadas as seguintes restrições:

i) $\hat{\boldsymbol{\theta}}_k$ deve ser não polarizado, ou seja, $\mathrm{E}[\hat{\boldsymbol{\theta}}_k] - \boldsymbol{\theta} = 0$;

ii) $\mathrm{cov}(\hat{\boldsymbol{\theta}}_k)$ deve ser tão pequena quanto possível.

Atualização recursiva não polarizada

- Impondo-se não-polarização, tem-se

$$
\begin{aligned}
\boldsymbol{\theta} &= \mathrm{E}[\hat{\boldsymbol{\theta}}_k] \\
&= J_k \mathrm{E}[\hat{\boldsymbol{\theta}}_{k-1}] + K_k \mathrm{E}[y(k)] \\
&= J_k \boldsymbol{\theta} + K_k \mathrm{E}[\boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_k] + K_k \mathrm{E}[\xi(k)].
\end{aligned}
$$

Atualização recursiva não polarizada

- Impondo-se não-polarização, tem-se

$$\begin{aligned}\boldsymbol{\theta} &= \mathrm{E}[\hat{\boldsymbol{\theta}}_k] \\ &= J_k \mathrm{E}[\hat{\boldsymbol{\theta}}_{k-1}] + K_k \mathrm{E}[y(k)] \\ &= J_k \boldsymbol{\theta} + K_k \mathrm{E}[\boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_k] + K_k \mathrm{E}[\xi(k)].\end{aligned}$$

- Considerando a média dos resíduos nula e os regressores determinísticos, tem-se

$$\begin{aligned}\boldsymbol{\theta} &= J_k \boldsymbol{\theta} + K_k \boldsymbol{\psi}_k^{\mathrm{T}} \boldsymbol{\theta} \\ I &= J_k + K_k \boldsymbol{\psi}_k^{\mathrm{T}} \\ J_k &= I - K_k \boldsymbol{\psi}_k^{\mathrm{T}}.\end{aligned}$$

- Portanto,

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= (I - K_k \boldsymbol{\psi}_k^{\mathrm{T}})\hat{\boldsymbol{\theta}}_{k-1} + K_k y(k) \\
\hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k(y(k) - \boldsymbol{\psi}_k^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{k-1}).
\end{aligned}$$

- Portanto,

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= (I - K_k \boldsymbol{\psi}_k^{\mathrm{T}})\hat{\boldsymbol{\theta}}_{k-1} + K_k y(k) \\
\hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k (y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_{k-1}).
\end{aligned}
$$

- A *inovação* é

$$
\eta(k) = y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_{k-1}.
$$

Atualização recursiva não polarizada de mínima covariância

- A seguir, $P_k = \mathrm{cov}[\hat{\boldsymbol{\theta}}_k]$ indicará a matriz de covariância calculada na iteração k.

$$\begin{aligned} P_k &= \mathrm{E}\left[(\hat{\boldsymbol{\theta}}_k - \mathrm{E}[\hat{\boldsymbol{\theta}}_k])(\hat{\boldsymbol{\theta}}_k - \mathrm{E}[\hat{\boldsymbol{\theta}}_k])^{\mathrm{T}}\right] \\ &= \mathrm{E}\left[\left((I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})\hat{\boldsymbol{\theta}}_{k-1} + K_k y(k) - \boldsymbol{\theta}_k\right)(\bullet)^{\mathrm{T}}\right], \end{aligned}$$

Atualização recursiva não polarizada de mínima covariância

- A seguir, $P_k = \mathrm{cov}[\hat{\boldsymbol{\theta}}_k]$ indicará a matriz de covariância calculada na iteração k.

$$\begin{aligned} P_k &= \mathrm{E}\left[(\hat{\boldsymbol{\theta}}_k - \mathrm{E}[\hat{\boldsymbol{\theta}}_k])(\hat{\boldsymbol{\theta}}_k - \mathrm{E}[\hat{\boldsymbol{\theta}}_k])^{\mathrm{T}}\right] \\ &= \mathrm{E}\left[\left((I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})\hat{\boldsymbol{\theta}}_{k-1} + K_k y(k) - \boldsymbol{\theta}_k\right)(\bullet)^{\mathrm{T}}\right], \end{aligned}$$

- Somando-se e subtraindo-se a matriz $K_k\boldsymbol{\psi}_k^{\mathrm{T}}\boldsymbol{\theta}$ nos parênteses da equação (-4), chega-se a

$$\begin{aligned} P_k &= \mathrm{E}\left[\left((I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta}) + K_k(y(k) - \boldsymbol{\psi}_k^{\mathrm{T}}\boldsymbol{\theta}_k)\right)(\bullet)^{\mathrm{T}}\right] \\ &= \mathrm{E}\left[\left((I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta}) + K_k e(k)\right)(\bullet)^{\mathrm{T}}\right]. \end{aligned}$$

- Considerando-se que o ruído $e(k)$ não está correlacionado com o vetor de regressores $\boldsymbol{\psi}_k$ nem com $(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})$,pode-se escrever

$$\begin{aligned} P_k &= \mathrm{E}\left[(I - K_k \boldsymbol{\psi}_k^{\mathrm{T}})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})^{\mathrm{T}}(I - K_k \boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}}\right] \\ &\quad + \mathrm{E}\left[K_k e(k) e(k)^{\mathrm{T}} K_k^{\mathrm{T}}\right]. \end{aligned}$$

- Considerando-se que o ruído $e(k)$ não está correlacionado com o vetor de regressores ψ_k nem com $(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})$,pode-se escrever

$$\begin{aligned} P_k &= \mathrm{E}\left[(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})^{\mathrm{T}}(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}}\right] \\ &\quad + \mathrm{E}\left[K_k e(k) e(k)^{\mathrm{T}} K_k^{\mathrm{T}}\right]. \end{aligned}$$

- Considerando-se os regressores determinísticos, tem-se

$$\begin{aligned} P_k &= (I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})\mathrm{E}\left[(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})^{\mathrm{T}}\right](I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} \\ &\quad + K_k\mathrm{E}\left[e(k)e(k)^{\mathrm{T}}\right]K_k^{\mathrm{T}}. \end{aligned}$$

- Considerando-se que o ruído $e(k)$ não está correlacionado com o vetor de regressores $\boldsymbol{\psi}_k$ nem com $(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})$,pode-se escrever

$$\begin{aligned} P_k &= \mathrm{E}\left[(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})^{\mathrm{T}}(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}}\right] \\ &\quad + \mathrm{E}\left[K_k e(k) e(k)^{\mathrm{T}} K_k^{\mathrm{T}}\right]. \end{aligned}$$

- Considerando-se os regressores determinísticos, tem-se

$$\begin{aligned} P_k &= (I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})\mathrm{E}\left[(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{k-1} - \boldsymbol{\theta})^{\mathrm{T}}\right](I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} \\ &\quad + K_k\mathrm{E}\left[e(k) e(k)^{\mathrm{T}}\right] K_k^{\mathrm{T}}. \end{aligned}$$

- Finalmente, supondo que o ruído tem média nula, tem-se

$$P_k = (I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})P_{k-1}(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} + K_k R K_k^{\mathrm{T}},$$

sendo $\mathrm{cov}[e(k)] = R = \mathrm{E}\left[e(k)e(k)^{\mathrm{T}}\right]$.

- No caso monovariável, $e(k)$ é um escalar e, portanto, $R = \sigma_e^2$ é simplesmente a variância do ruído.

- No caso monovariável, $e(k)$ é um escalar e, portanto, $R = \sigma_e^2$ é simplesmente a variância do ruído.
- P_k é uma função discreta e nesse caso deseja-se minimizar a variação ΔP_k provocada por uma variação ΔK_k. Assim, chamando $P_k^* = P_k + \Delta P_k$ o novo valor de P_k devido a uma variação em K_k, tem-se

$$
\begin{aligned}
\Delta P_k \quad = \quad & (I - (K_k + \Delta K_k)\boldsymbol{\psi}_k^{\mathrm{T}})\, P_{k-1}\, (I - (K_k + \Delta K_k)\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} + \\
& (K_k + \Delta K_k)R(K_k + \Delta K_k)^{\mathrm{T}} - \\
& (I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})P_{k-1}(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} - K_k R K_k^{\mathrm{T}}.
\end{aligned}
$$

- No caso monovariável, $e(k)$ é um escalar e, portanto, $R = \sigma_e^2$ é simplesmente a variância do ruído.
- P_k é uma função discreta e nesse caso deseja-se minimizar a variação ΔP_k provocada por uma variação ΔK_k. Assim, chamando $P_k^* = P_k + \Delta P_k$ o novo valor de P_k devido a uma variação em K_k, tem-se

$$\begin{aligned}
\Delta P_k \quad = \quad & (I - (K_k + \Delta K_k)\boldsymbol{\psi}_k^{\mathrm{T}})\, P_{k-1}\, (I - (K_k + \Delta K_k)\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} + \\
& (K_k + \Delta K_k)R(K_k + \Delta K_k)^{\mathrm{T}} - \\
& (I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})P_{k-1}(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} - K_k R K_k^{\mathrm{T}}.
\end{aligned}$$

- Desprezando-se os termos quadráticos em ΔK_k e igualando-se a expressão resultante a zero, obtém-se

$$\begin{aligned}
\Delta P_k \quad \approx \quad & \Delta K_k \left[-\boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1}(I - \boldsymbol{\psi}_k K_k^{\mathrm{T}}) + R K_k^{\mathrm{T}}\right] \\
& - \left[(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})P_{k-1}\boldsymbol{\psi}_k - K_k R\right] \Delta K_k^{\mathrm{T}}.
\end{aligned}$$

- Para que $\Delta P_k = 0$, ambos os termos entre colchetes devem ser nulos. Essas condições geram as seguintes equações:

$$\begin{aligned} 0 &= -\psi_k^{\mathrm{T}} P_{k-1}(I - \psi_k K_k^{\mathrm{T}}) + R K_k^{\mathrm{T}} \\ 0 &= (I - K_k \psi_k^{\mathrm{T}}) P_{k-1} \psi_k - K_k R. \end{aligned}$$

- Para que $\Delta P_k = 0$, ambos os termos entre colchetes devem ser nulos. Essas condições geram as seguintes equações:

$$\begin{aligned} 0 &= -\psi_k^{\mathrm{T}} P_{k-1}(I - \psi_k K_k^{\mathrm{T}}) + R K_k^{\mathrm{T}} \\ 0 &= (I - K_k \psi_k^{\mathrm{T}}) P_{k-1} \psi_k - K_k R. \end{aligned}$$

- Finalmente, da última equação, tem-se

$$\begin{aligned} K_k(\psi_k^{\mathrm{T}} P_{k-1} \psi_k + R) &= P_{k-1} \psi_k \\ K_k &= P_{k-1} \psi_k (\psi_k^{\mathrm{T}} P_{k-1} \psi_k + R)^{-1}. \end{aligned}$$

- Portanto, um algoritmo recursivo para a estimação do vetor $\boldsymbol{\theta}$ é

$$\left.\begin{aligned}
K_k &= P_{k-1}\boldsymbol{\psi}_k(\boldsymbol{\psi}_k^{\mathrm{T}}P_{k-1}\boldsymbol{\psi}_k + R)^{-1};\\
\hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k(y(k) - \boldsymbol{\psi}_k^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{k-1});\\
P_k &= (I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})P_{k-1}(I - K_k\boldsymbol{\psi}_k^{\mathrm{T}})^{\mathrm{T}} + K_kRK_k^{\mathrm{T}}.
\end{aligned}\right\}$$

- A dimensão da matriz $[\psi_k^{\mathrm{T}} P_{k-1} \psi_k + R]$ é igual ao número de saídas medidas.
  Na primeira iteração tem-se

$$
\begin{aligned}
y(1) &= \boldsymbol{\psi}_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_1 \\
&= \boldsymbol{\psi}_1^{\mathrm{T}} \left[\hat{\boldsymbol{\theta}}_0 + K_1 \left(y(1) - \boldsymbol{\psi}_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 \right) \right] \\
&= \boldsymbol{\psi}_1^{\mathrm{T}} \left[\hat{\boldsymbol{\theta}}_0 + P_0 \boldsymbol{\psi}_1 (\boldsymbol{\psi}_1^{\mathrm{T}} P_0 \boldsymbol{\psi}_1 + R)^{-1} \left(y(1) - \boldsymbol{\psi}_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 \right) \right].
\end{aligned}
$$

- A dimensão da matriz $[\psi_k^{\mathrm{T}} P_{k-1} \psi_k + R]$ é igual ao número de saídas medidas.
  Na primeira iteração tem-se

$$
\begin{aligned}
y(1) &= \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_1 \\
&= \psi_1^{\mathrm{T}} \left[\hat{\boldsymbol{\theta}}_0 + K_1 \left(y(1) - \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 \right) \right] \\
&= \psi_1^{\mathrm{T}} \left[\hat{\boldsymbol{\theta}}_0 + P_0 \psi_1 (\psi_1^{\mathrm{T}} P_0 \psi_1 + R)^{-1} \left(y(1) - \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 \right) \right].
\end{aligned}
$$

- Se $\psi_1^{\mathrm{T}} P_0 \psi_1 \gg R$ (isso pode ser conseguido para valores elevados de P_0) a última equação pode ser aproximada por:

$$
\begin{aligned}
\hat{y}(1) &\approx \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 + \psi_1^{\mathrm{T}} P_0 \psi_1 \left(\psi_1^{\mathrm{T}} P_0 \psi_1 \right)^{-1} y(1) \\
&\quad - \psi_1^{\mathrm{T}} P_0 \psi_1 \left(\psi_1^{\mathrm{T}} P_0 \psi_1 \right)^{-1} \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 \\
&\approx \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 + y(1) - \psi_1^{\mathrm{T}} \hat{\boldsymbol{\theta}}_0 = y(1).
\end{aligned}
$$

Estimador Recursivo MQ

O ponto de partida é o modelo $y(k) = \boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}} + \xi(k)$ e uma seqüência de dados, representados em $y(k)$ e $\boldsymbol{\psi}(k-1)$, $k = 1, \ldots$

- O objetivo nesta seção é derivar um algoritmo recursivo baseado no estimador MQ que pode ser escrito da seguinte forma

$$\hat{\boldsymbol{\theta}}_{\mathrm{MQ}\,k} = \left[\sum_{i=1}^{k} \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right]^{-1} \left[\sum_{i=1}^{k} \boldsymbol{\psi}(i-1)y(i)\right].$$

- A seguir será usada a seguinte notação:

$$
\begin{aligned}
P_k &= \left[\sum_{i=1}^{k} \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right]^{-1}, \\
P_k^{-1} &= \left[\sum_{i=1}^{k-1} \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right] + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1) \\
&= P_{k-1}^{-1} + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1).
\end{aligned}
$$

- A seguir será usada a seguinte notação:

$$
\begin{aligned}
P_k &= \left[\sum_{i=1}^{k} \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right]^{-1}, \\
P_k^{-1} &= \left[\sum_{i=1}^{k-1} \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right] + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1) \\
&= P_{k-1}^{-1} + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1).
\end{aligned}
$$

- O estimador pode ser rescrito

$$
\hat{\boldsymbol{\theta}}_k = P_k \left[\sum_{i=1}^{k-1} \boldsymbol{\psi}(i-1)y(i) + \boldsymbol{\psi}(k-1)y(k)\right].
$$

- Escrevendo-se o estimador para o instante $k-1$, obtém-se

$$\left[\sum_{i=1}^{k-1}\boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right]\hat{\boldsymbol{\theta}}_{k-1}=\left[\sum_{i=1}^{k-1}\boldsymbol{\psi}(i-1)y(i)\right],$$

sendo que o l.e. da eq. pode ser representado em forma compacta como sendo $P_{k-1}^{-1}\hat{\boldsymbol{\theta}}_{k-1}$, resultando em

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= P_k\left[P_{k-1}^{-1}\hat{\boldsymbol{\theta}}_{k-1}+\boldsymbol{\psi}(k-1)y(k)\right]\\
&= P_k\left[\left(P_k^{-1}-\boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1)\right)\hat{\boldsymbol{\theta}}_{k-1}+\boldsymbol{\psi}(k-1)y(k)\right]\\
&= \hat{\boldsymbol{\theta}}_{k-1}-P_k\boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}}_{k-1}+P_k\boldsymbol{\psi}(k-1)y(k)\\
&= \hat{\boldsymbol{\theta}}_{k-1}+P_k\boldsymbol{\psi}(k-1)\left[y(k)-\boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}}_{k-1}\right]\\
&= \hat{\boldsymbol{\theta}}_{k-1}+K_k\eta(k),
\end{aligned}$$

sendo $K_k=P_k\boldsymbol{\psi}(k-1)$ uma matriz de ganho e $\eta(k)=y(k)-\boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}}_{k-1}$ a inovação no instante k.

- Aplicando-se o lema da inversão tem-se

$$\begin{aligned} P_k &= P_{k-1} - \\ &-P_{k-1}\boldsymbol{\psi}(k-1)\left(\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)+1\right)^{-1}\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}, \end{aligned}$$

e, como antes, o termo a ser invertido será um escalar para modelos com apenas uma saída.

- Aplicando-se o lema da inversão tem-se

$$\begin{aligned} P_k &= P_{k-1} - \\ &-P_{k-1}\boldsymbol{\psi}(k-1)\left(\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)+1\right)^{-1}\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}, \end{aligned}$$

e, como antes, o termo a ser invertido será um escalar para modelos com apenas uma saída.

- Finalmente, chega-se a

$$\begin{aligned} K_k &= P_{k-1}\boldsymbol{\psi}(k-1) - \frac{P_{k-1}\boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)}{\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)+1} \\ &= \frac{P_{k-1}\boldsymbol{\psi}(k-1)}{\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)+1}. \end{aligned}$$

- Portanto, o algoritmo recursivo de MQ (RMQ) é

$$\left.\begin{array}{l}
K_k = \dfrac{P_{k-1}\boldsymbol{\psi}_k}{\boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1}\boldsymbol{\psi}_k + 1}; \\[2em]
\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + K_k \left[y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_{k-1} \right]; \\[2em]
P_k = P_{k-1} - K_k \psi_k^{\mathrm{T}} P_{k-1},
\end{array}\right\} .$$

Outros Estimadores Recursivos

Nesta seção serão mencionados os estimadores recursivos EMQ e VI, além do estimador para aproximação estocástica.

- Seja o modelo ARMAX

$$
\begin{aligned}
y(k) &= -\hat{a}_1 y(k-1) \ldots - \hat{a}_{n_y} y(k-n_y) + \hat{b}_1 u(k-\tau_{\mathrm{d}}-1) \ldots \\
&\quad \ldots + \hat{b}_{n_u} u(k-\tau_{\mathrm{d}}-n_u) + \hat{c}_1 \xi(k-1) \ldots + \hat{c}_{n_e} \xi(k-n_e) + \xi(k) \\
&= [y(k-1) \ldots y(k-n_y)\ u(k-\tau_{\mathrm{d}}-1) \ldots u(k-\tau_{\mathrm{d}}-n_u) \\
&\qquad \xi(k-1) \ldots \xi(k-n_e)]\, \hat{\boldsymbol{\theta}} + \xi(k) \\
&= \boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}} + \xi(k).
\end{aligned}
$$

Estimador recursivo estendido de mínimos quadrados

- O *estimador recursivo estendido de mínimos quadrados* é a seguinte:

$$
\left.
\begin{aligned}
K_k &= P_{k-1}\boldsymbol{\psi}_k[\boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1}\boldsymbol{\psi}_k + 1]^{-1};\\
\hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k\left[y(k) - \boldsymbol{\psi}_k^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{k-1}\right];\\
P_k &= P_{k-1} - K_k\boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1};\\
\xi(k) &= y(k) - \boldsymbol{\psi}_k^{\mathrm{T}}\hat{\boldsymbol{\theta}}_k,
\end{aligned}
\right\}
$$

sendo que, na primeira iteração, ψ_k^{T} não contém resíduos $\xi(k)$.

Estimador recursivo de variáveis instrumentais

- Chamando de $\mathbf{z}_k$ o vetor de variáveis instrumentais na k-ésima iteração, o *estimador recursivo de variáveis instrumentais* pode ser implementado como se segue:

$$
\left.
\begin{aligned}
M_k &= M_{k-1} - M_{k-1}\mathbf{z}_k \left[1 + \boldsymbol{\psi}_k^{\mathrm{T}} M_{k-1}\mathbf{z}_k\right]^{-1} \boldsymbol{\psi}_k^{\mathrm{T}} M_{k-1}; \\
K_k &= M_k \mathbf{z}_k; \\
\hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k \left[y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_{k-1}\right]; \\
\xi(k) &= y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_k.
\end{aligned}
\right\}.
$$

Deve estar claro que no estimador VI a matriz M não é simétrica e não mais corresponde à matriz de covariância cov($\hat{\boldsymbol{\theta}}$).

Deve estar claro que no estimador VI a matriz M não é simétrica e não mais corresponde à matriz de covariância $\text{cov}(\hat{\boldsymbol{\theta}})$.

Se a matriz de ganho for $K_k \neq P_k \boldsymbol{\psi}_k$, o estimador continuará sendo não polarizado, porém não será de variância mínima.

Estimador recursivo de aproximação estocástica

Em casos nos quais as limitações de tempo de computação são críticas pode-se usar uma matriz de ganho do tipo $K_k = \gamma_k \boldsymbol{\psi}_k$. O resultado é o *estimador recursivo de aproximação estocástica*

$$\left.\begin{array}{l} K_k = \gamma_k \boldsymbol{\psi}_k; \\ \\ \hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + K_k \left[y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_{k-1} \right]; \\ \\ \xi(k) = y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_k, \end{array}\right\}$$

para uma dada seqüência γ_k. Uma forma de escolher γ_k é $\gamma_k = C\, k^{-\alpha}$; $0,5 \leq \alpha \leq 1$ e C uma constante positiva. A escolha de γ_k deverá ser um compromisso entre tempo de convergência e covariância do vetor estimado.

Estimação de Parâmetros Variantes no Tempo

- Começa-se pelo estimador MQP

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= \left[\sum_{i=1}^{k} w_i(k)\boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1)\right]^{-1}\sum_{i=1}^{k} w_i(k)\boldsymbol{\psi}(i-1)y(i) \\
&= P_k F_k.
\end{aligned}
$$

sendo que $w_i(k)$ é o valor do i-ésimo peso na k-ésima iteração ($k \geq i$).

- A seqüência de pesos deverá satisfazer as seguintes restrições:

$$\left.\begin{aligned} w_k(k) &= 1 \\ w_i(k) &= \lambda w_i(k-1), \end{aligned}\right\}$$

ou seja, o maior peso sempre corresponde ao último valor recebido e é igual a um. Quando um novo dado é recebido, todos os pesos são multiplicados por um fator λ, que na prática recebe valores na faixa $0,95 \leq \lambda \leq 0,99$.

- P_k^{-1} é

$$
\begin{aligned}
P_k^{-1} &= \sum_{i=1}^{k-1} w_i(k)\boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1) + w_k(k)\boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1) \\
&= \sum_{i=1}^{k-1} \lambda w_i(k-1)\boldsymbol{\psi}(i-1)\boldsymbol{\psi}^{\mathrm{T}}(i-1) + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1) \\
&= \lambda P_{k-1}^{-1} + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1),
\end{aligned}
$$

e

$$
\begin{aligned}
F_k &= \sum_{i=1}^{k-1} w_i(k)\boldsymbol{\psi}(i-1)y(i) + w_k(k)\boldsymbol{\psi}(k-1)y(k) \\
&= \lambda F_{k-1} + \boldsymbol{\psi}(k-1)y(k).
\end{aligned}
$$

- Após algumas substituições, chega-se a

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= P_k F_k \\
&= P_k \left[\lambda F_{k-1} + \boldsymbol{\psi}(k-1) y(k)\right] \\
&= P_k \left[\lambda P_{k-1}^{-1} \hat{\boldsymbol{\theta}}_{k-1} + \boldsymbol{\psi}(k-1) y(k)\right] \\
&= P_k \left\{\lambda \left[\frac{P_k^{-1} - \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1)}{\lambda}\right] \hat{\boldsymbol{\theta}}_{k-1} + \boldsymbol{\psi}(k-1) y(k)\right\} \\
&= \hat{\boldsymbol{\theta}}_{k-1} + P_k \boldsymbol{\psi}(k-1) \left[y(k) - \boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}}_{k-1}\right].
\end{aligned}
$$

- Após algumas substituições, chega-se a

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= P_k F_k \\
&= P_k \left[\lambda F_{k-1} + \boldsymbol{\psi}(k-1) y(k)\right] \\
&= P_k \left[\lambda P_{k-1}^{-1} \hat{\boldsymbol{\theta}}_{k-1} + \boldsymbol{\psi}(k-1) y(k)\right] \\
&= P_k \left\{\lambda \left[\frac{P_k^{-1} - \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1)}{\lambda}\right] \hat{\boldsymbol{\theta}}_{k-1} + \boldsymbol{\psi}(k-1) y(k)\right\} \\
&= \hat{\boldsymbol{\theta}}_{k-1} + P_k \boldsymbol{\psi}(k-1) \left[y(k) - \boldsymbol{\psi}^{\mathrm{T}}(k-1) \hat{\boldsymbol{\theta}}_{k-1}\right].
\end{aligned}$$

- Aplicando-se o lema da inversão, fazendo-se $A = \lambda P_{k-1}^{-1}$, $B = D^{\mathrm{T}} = \boldsymbol{\psi}(k-1)$ e $C = 1$, tem-se

$$\begin{aligned}
P_k &= \frac{P_{k-1}}{\lambda} - \frac{P_{k-1}}{\lambda} \boldsymbol{\psi}(k-1) \left[\boldsymbol{\psi}^{\mathrm{T}}(k-1) P_{k-1} \boldsymbol{\psi}(k-1)\right]^{-1} \frac{\lambda \boldsymbol{\psi}^{\mathrm{T}}(k-1) P_{k-1}}{\lambda} \\
&= \frac{1}{\lambda} \left(P_{k-1} - \frac{P_{k-1} \boldsymbol{\psi}(k-1) \boldsymbol{\psi}^{\mathrm{T}}(k-1) P_{k-1}}{\boldsymbol{\psi}^{\mathrm{T}}(k-1) P_{k-1} \boldsymbol{\psi}(k-1) + \lambda}\right).
\end{aligned}$$

- Finalmente,

$$
\begin{aligned}
K_k &= P_k \boldsymbol{\psi}(k-1) \\
&= \frac{1}{\lambda}\left(P_{k-1}\boldsymbol{\psi}(k-1) - \frac{P_{k-1}\boldsymbol{\psi}(k-1)\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)}{\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)+\lambda}\right) \\
&= \frac{P_{k-1}\boldsymbol{\psi}(k-1)}{\boldsymbol{\psi}^{\mathrm{T}}(k-1)P_{k-1}\boldsymbol{\psi}(k-1)+\lambda}.
\end{aligned}
$$

- Portanto, o estimador recursivo de mínimos quadrados com fator de esquecimento λ é

$$
\left.
\begin{array}{l}
K_k = \frac{P_{k-1}\boldsymbol{\psi}_k}{\boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1}\boldsymbol{\psi}_k + \lambda}; \\[2em]
\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + P_k \psi_k \left[y(k) - \boldsymbol{\psi}_k^{\mathrm{T}} \hat{\boldsymbol{\theta}}_{k-1} \right]; \\[2em]
P_k = \frac{1}{\lambda} \left(P_{k-1} - \frac{P_{k-1}\boldsymbol{\psi}_k \boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1}}{\boldsymbol{\psi}_k^{\mathrm{T}} P_{k-1}\boldsymbol{\psi}_k + \lambda} \right),
\end{array}
\right\} .
$$

Estimação recursiva de ganho e constante de tempo

O processo consiste basicamente de uma resistência elétrica dentro de uma caixa metálica sendo que a temperatura externa da caixa é medida e armazenada.

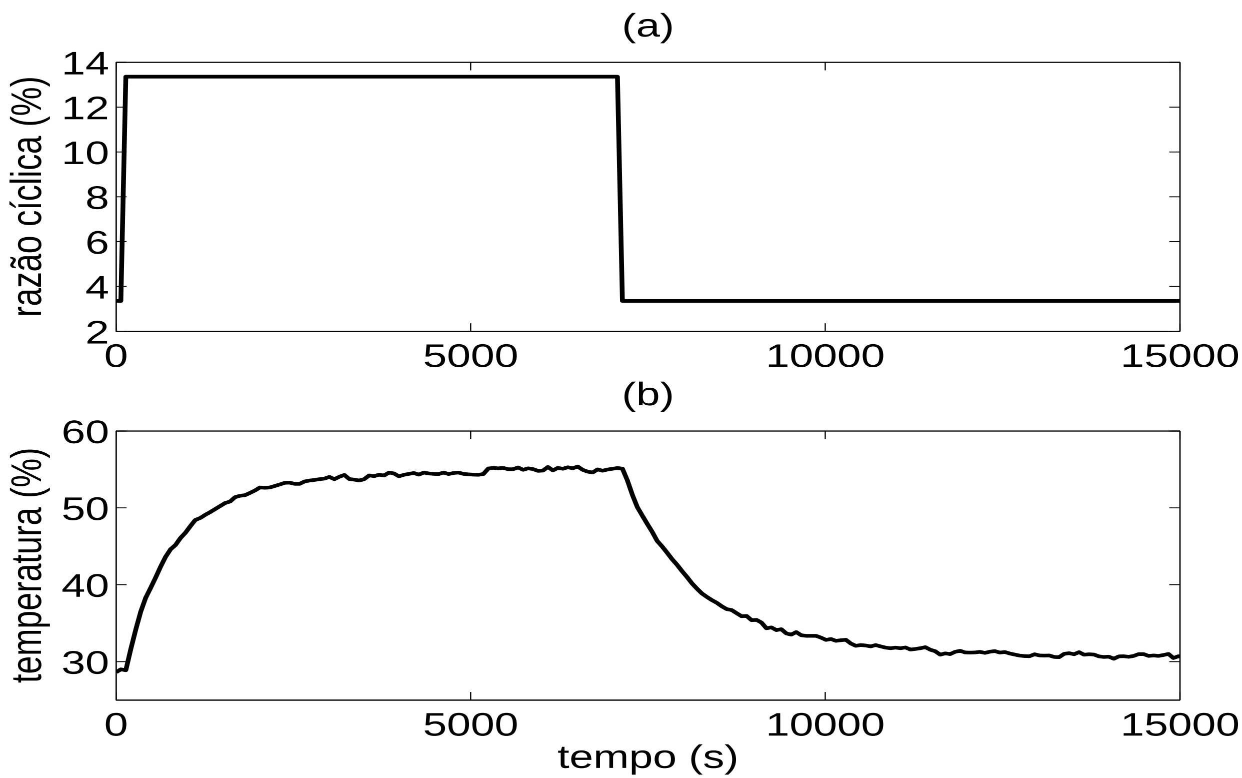

Figura 1: Dados de teste de um processo térmico Dados de teste realizado em malha aberta de um processo térmico. (a) entrada $u(k)$, razão cíclica, e (b) saída $y(k)$, temperatura.

- Uma vez que se deseja estimar ganho e constante de tempo, utilizou-se um modelo de primeira ordem do tipo

$$\frac{Y(s)}{U(s)} = \frac{K}{\tau\, s + 1}.$$

- Uma vez que se deseja estimar ganho e constante de tempo, utilizou-se um modelo de primeira ordem do tipo

$$\frac{Y(s)}{U(s)} = \frac{K}{\tau\, s + 1}.$$

- Mas como o modelo a ser estimado é discreto, utilizou-se a aproximação $\dot{y} = [y(k+1) - y(k)]/T_{\mathrm{d}}$, com $T_{\mathrm{d}} = T_{\mathrm{s}}$, resultando em $y(k) = a_1 y(k-1) + b_1 u(k-1)$, sendo que

$$a_1 = 1 - \frac{T_{\mathrm{s}}}{\tau}; \qquad b_1 = \frac{T_{\mathrm{s}} K}{\tau}.$$

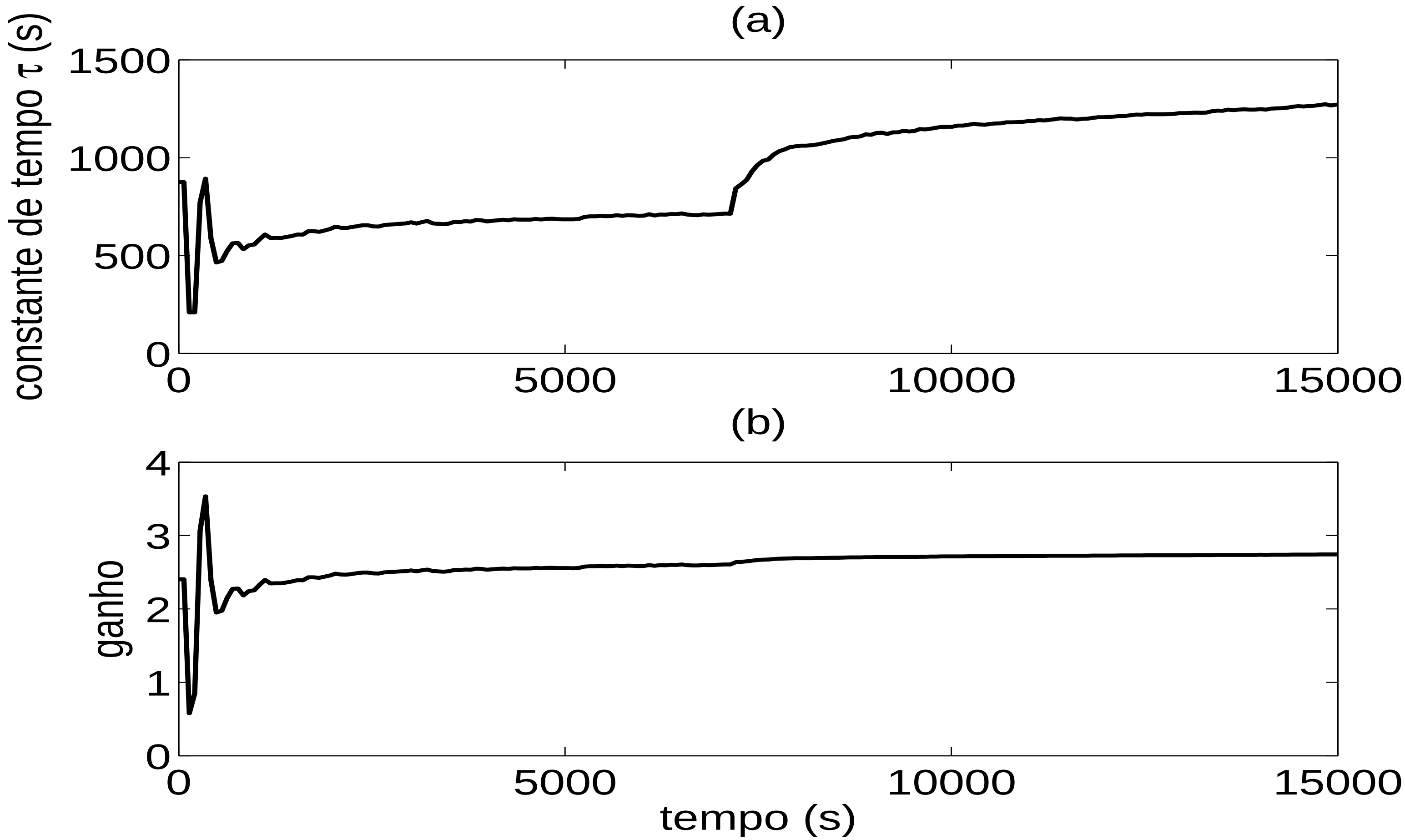

Figura 2: Constante de tempo e ganho recursivamente estimados Constante de tempo e ganho recursivamente estimados a partir dos dados da Figura 1. (a) constante de tempo e (b) ganho.

- A constante de tempo e o ganho estimados na k-ésima iteração são, respectivamente,

$$\hat{\tau}_k = \frac{-T_\mathrm{s}}{\hat{a}_{1k} - 1}; \qquad \hat{K}_k = \frac{\hat{b}_{1k}\tau_k}{T_\mathrm{s}}.$$

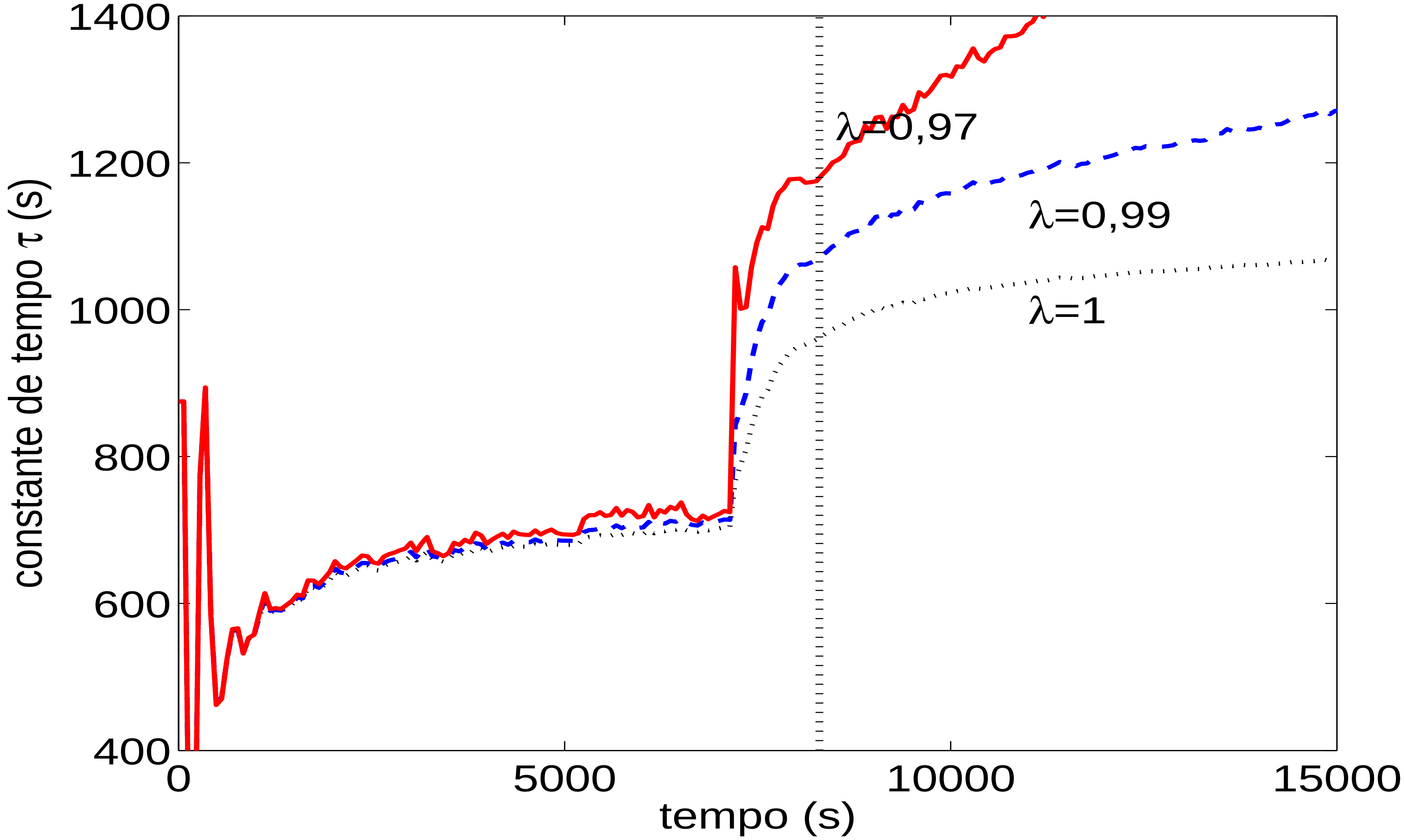

Figura 3: Efeito do fator de esquecimento Efeito do fator de esquecimento na estimação recursiva da constante de tempo a partir dos dados da Figura 1.

Complementos - Estimação de matrizes de estado

Neste exemplo os dados serão gerados pelo seguinte modelo em espaço de estados:

$$\mathbf{x}(k) = \begin{bmatrix} 0,8 & 0,1 & 0,005 \\ 0,0 & 0,7 & 0,1 \\ 0,0 & 0,0 & 0,9 \end{bmatrix} \mathbf{x}(k-1) + $$

$$+ \begin{bmatrix} 1,0 & 0,05 & 0,0025 \\ 0,0 & 1,0 & 0,05 \\ 0,0 & 0,0 & 1,0 \end{bmatrix} \mathbf{u}(k-1) + \mathbf{e}(k),$$

sendo que o vetor de entradas $\mathbf{u}(k)$ é composto por três variáveis aleatórias de média zero, distribuição gaussiana e variância unitária. Semelhantemente, $\mathbf{e}(k)$ é também composto de três variáveis com distribuição gaussiana e variância tal que resulta numa relação sinal ruído igual a 50 dB.

Foi considerada a matriz de observação $C_{\mathrm{d}} = [1\ 1\ 1]$, portanto

$$\begin{aligned} y(k) &= C_{\mathrm{d}}\mathbf{x}(k) \\ \hat{y}(k) &= C_{\mathrm{d}}\hat{\mathbf{x}}(k), \end{aligned}$$

sendo $\hat{\mathbf{x}}(k) = \hat{\Phi}\mathbf{x}(k-1) + \hat{\Gamma}\mathbf{u}(k-1)$.
Os autovalores do modelo original são 0,7; 0,8 e 0,9. Os autovalores da matriz de estado estimada são eig($\hat{\Phi}$)=(0,6953; 0,8143; 0,8923).

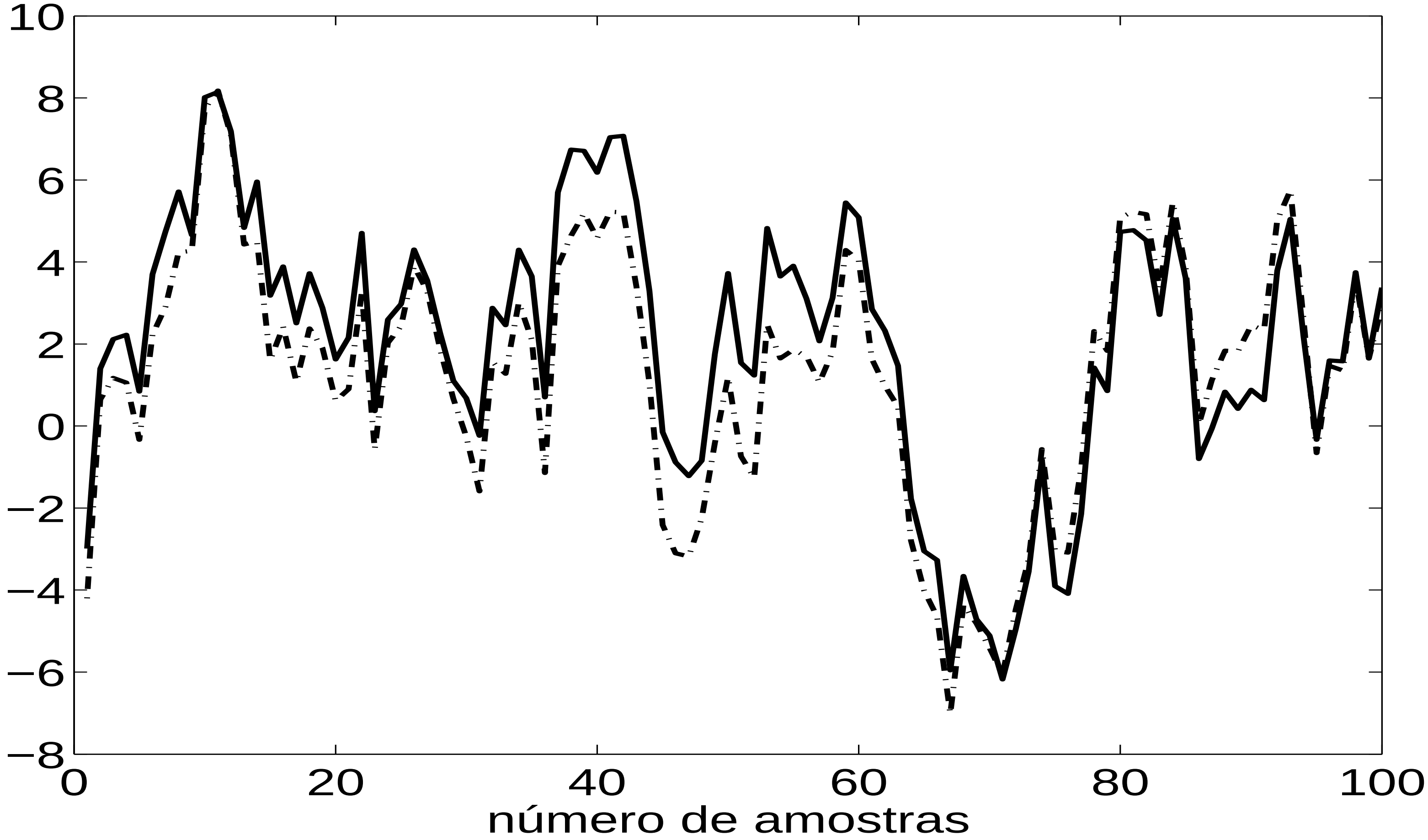

Figura 4: Desempenho de modelo em espaço de estados Saída (—) $y(k)$ e saída estimada $(- \cdot -)$ $\hat{y}(k)$ utilizando o modelo em espaço de estados.