Universidade Federal do Rio Grande do Sul
Escola de Engenharia
Departamento de Engenharia Elétrica
ENG04037 Sistemas de Controle Digitais

Identificação via Mínimos Quadrados

Prof. Walter Fetter Lages

22 de novembro de 2011

1 Introdução

Para projetar um sistema de controle é usualmente necessário dispor de um modelo da planta. Em alguns casos, estes modelos podem ser obtidos através de leis físicas (enfoque da caixa-branca). Em outros casos, pode não ser possível obter o modelo desta maneira. Nestes casos, talvez seja possível obter-se um modelo através da observação da resposta do sistema à entradas apropriadas (enfoque da caixa-preta), como mostrado na figura 1. Este procedimento é denominado *identificação de sistemas*.

Os principais elementos do problema de estimação paramétrica são:

1. Classe de modelo
   (a) complexidade
   (b) adequação
2. Critério de desempenho
   (a) escolha da estrutura
   (b) escolha da ordem
   (c) estimação dos parâmetros
3. Condições experimentais
   (a) condições de operação
   (b) validade do modelo

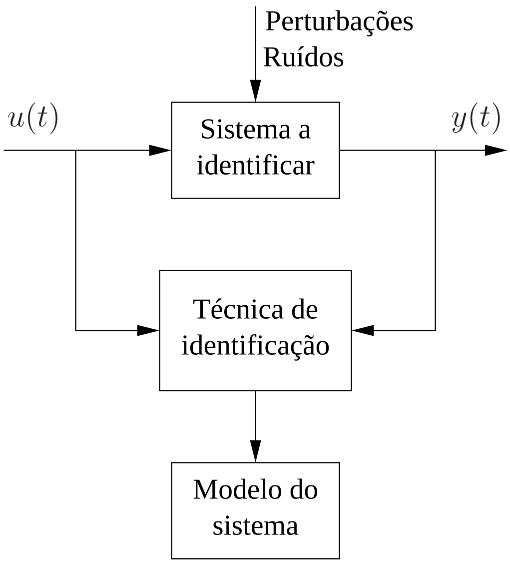

Figura 1: Diagrama de blocos básico para identificação de sistemas.

   (c) qualidade do modelo

4. Algoritmo de Estimação
   (a) *on-line*
   (b) *off-line*
   (c) garantias
      i. convergência da saída
      ii. convergência dos parâmetros
      iii. taxa de convergência
      iv. robustez
         A. à ruído
         B. à dinâmica não modelada
         C. à erros numéricos

5. Uso de conhecimento *a priori*
   (a) restrições estruturais
   (b) valores de parâmetros
   (c) faixa de valores de parâmetros

Neste curso o interesse é basicamente em algoritmos de estimação *on-line* devido a sua maior aplicabilidade para controle adaptativo.

Quando a estimação é realizada *on-line*, torna-se necessário obter uma estimativa atualizada dentro do tempo entre duas amostras sucessivas. Portanto, é altamente desejável que o algoritmo seja simples e facilmente implementável.

Uma classe particularmente interessante de algoritmos *on-line* é quando a estimativa atual $\theta(t)$ é computada em função das estimativas anteriores, seja quando se tem uma estimativa que pode ser computada recursivamente.

2 Método dos Mínimos Quadrados

Seja um sistema representado por um modelo ARX[1]:

$$y(t+1) = a_1 y(t) + \cdots + a_p y(t-p+1) + b_1 u(t) + \cdots + b_q u(t-q+1) + \omega(t+1) \quad (1)$$

onde $\omega(t+1)$ é um ruído gaussiano.

O modelo (1) pode ser reescrito na forma:

$$y(t+1) = \phi^T(t)\theta + \omega(t+1) \quad (2)$$

com

$$\theta = \begin{bmatrix} a_1 \\ \vdots \\ a_p \\ b_1 \\ \vdots \\ b_1 \end{bmatrix} \text{, denominado vetor de parâmetros}$$

e

$$\phi(t) = \begin{bmatrix} y(t) \\ \vdots \\ y(t-p+1) \\ u(t) \\ \vdots \\ u(t-q+1) \end{bmatrix} \text{, denominado vetor de regressão.}$$

O problema de identificação consiste em determinar θ com base nas informações (medidas) de $y(t+1)$ e $\phi(t)$ para $t = 0, 1, \ldots, n$.

[1] AutoRegressive with eXogenous inputs.

Para solucionar este problema pode-se formula-lo com um problema de otimização com custo a minimizar:

$$J(n,\theta) = \frac{1}{n}\sum_{t=0}^{n-1}\left(y(t+1) - \phi^T(t)\theta\right)^2$$

onde $y(t+1) - \phi^T(t)\theta$ é o erro de predição.

Mais formalmente tem-se

$$\hat{\theta}(n) = \underset{\theta}{\text{Arg}} \quad \min J(n,\theta) \tag{3}$$

que pode ser escrita na forma

$$\hat{\theta}(n) = \underset{\theta}{\text{Arg}} \quad \min \ (Y(n) - \Phi(n)\theta)^T (Y(n) - \Phi(n)\theta)$$

com

$$Y(n) = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix}$$

e

$$\Phi(n) = \begin{bmatrix} \phi^T(0) \\ \phi^T(1) \\ \vdots \\ \phi^T(n-1) \end{bmatrix}$$

Assim, (3) pode ser resolvida fazendo-se

$$\left.\frac{\partial J(n,\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}(n)} = 0 = -2\Phi^T(n)Y(n) + 2\Phi^T(n)\Phi(n)\hat{\theta}(n)$$

Logo

$$\hat{\theta}(n) = \left(\Phi^T(n)\Phi(n)\right)^{-1}\Phi^T(n)Y(n) = \left(\sum_{t=0}^{n-1}\phi(t)\phi^T(t)\right)^{-1}\sum_{t=0}^{n-1}\phi(t)y(t+1)$$

Definindo

$$P(n) = \left(\sum_{t=0}^{n} \phi(t)\phi^T(t)\right)^{-1} = \left(\Phi^T(n+1)\Phi(n+1)\right)^{-1}$$

tem-se

$$\hat{\theta}(n+1) = P(n)\sum_{t=0}^{n} \phi(t)y(t+1) \tag{4}$$

e

$$\begin{aligned} P^{-1}(n) &= \sum_{t=0}^{n} \phi(t)\phi^T(t) \\ &= \sum_{t=0}^{n-1} \phi(t)\phi^T(t) + \phi(n)\phi^T(n) \\ &= P^{-1}(n-1) + \phi(n)\phi^T(n) \end{aligned}$$

ou

$$\begin{aligned} P(n) &= \left(P^{-1}(n-1) + \phi(n)\phi^T(n)\right)^{-1} \\ &= P(n-1) - P(n-1)\phi(n)\left(\phi^T(n)P(n-1)\phi(n) + 1\right)^{-1}\phi^T(n)P(n-1) \end{aligned}$$

Obs 1 : *Lema de inversão de matrizes:*

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}$$

Utilizando-se o Lema de Inversão de Matrizes com $A = P^{-1}(n-1)$, $B = \phi(n)$, $C = 1$ e $D = \phi^T(n)$, pode-se calcular

$$P(n) = P(n-1) - \frac{P(n-1)\phi(n)\phi^T(n)P(n-1)}{1 + \phi^T(n)P(n-1)\phi(n)}$$

De (4) tem-se

$$
\begin{aligned}
\hat{\theta}(n+1) &= P(n)\sum_{t=0}^{n}\phi(t)y(t+1) \\
&= P(n)\left(\sum_{t=0}^{n-1}\phi(t)y(t+1)+\phi(n)y(n+1)\right) \\
&= \left(P(n-1)-\frac{P(n-1)\phi(n)\phi^T(n)P(n-1)}{1+\phi^T(n)P(n-1)\phi(n)}\right)\left(\sum_{t=0}^{n-1}\phi(t)y(t+1)+\phi(n)y(n+1)\right)
\end{aligned}
$$

e expandindo-se o produto resulta

$$
\begin{aligned}
\hat{\theta}(n+1) &= P(n-1)\sum_{t=0}^{n-1}\phi(t)y(t+1)+P(n-1)\phi(n)y(n+1) \\
&- \frac{P(n-1)\phi(n)\phi^T(n)P(n-1)}{1+\phi^T(n)P(n-1)\phi(n)}\sum_{t=0}^{n-1}\phi(t)y(t+1) \\
&- \frac{P(n-1)\phi(n)\phi^T(n)P(n-1)}{1+\phi^T(n)P(n-1)\phi(n)}\phi(n)y(n+1)
\end{aligned}
$$

Substituindo-se (4), atrasada de um período de amostragem, chega-se a:

$$
\begin{aligned}
\hat{\theta}(n+1) &= \hat{\theta}(n)+P(n-1)\phi(n)y(n+1) \\
&- \frac{P(n-1)\phi(n)\phi^T(n)}{1+\phi^T(n)P(n-1)\phi(n)}\hat{\theta}(n) \\
&- \frac{P(n-1)\phi(n)\phi^T(n)P(n-1)}{1+\phi^T(n)P(n-1)\phi(n)}\phi(n)y(n+1)
\end{aligned}
$$

e agrupando-se os temos em $\phi(n)y(n+1)$ tem-se

$$
\begin{aligned}
&\hat{\theta}(n+1)=\hat{\theta}(n)+ \\
&+\frac{P(n-1)+P(n-1)\phi^T(n)P(n-1)\phi(n)-P(n-1)\phi(n)\phi^T(n)P(n-1)}{1+\phi^T(n)P(n-1)\phi(n)}\phi(n)y(n+1) \\
-\ &\frac{P(n-1)\phi(n)\phi^T(n)}{1+\phi^T(n)P(n-1)\phi(n)}\hat{\theta}(n)
\end{aligned}
$$

ou

$$
\begin{aligned}
\hat{\theta}(n+1) &= \hat{\theta}(n) + \frac{P(n-1)}{1+\phi^T(n)P(n-1)\phi(n)}\phi(n)y(n+1) \\
&- \frac{P(n-1)\phi(n)\phi^T(n)}{1+\phi^T(n)P(n-1)\phi(n)}\hat{\theta}(n)
\end{aligned}
$$

que pode ser reescrita como

$$
\hat{\theta}(n+1) = \hat{\theta}(n) + \frac{P(n-1)\phi(n)}{1+\phi^T(n)P(n-1)\phi(n)}\left(y(n+1) - \phi^T(n)\hat{\theta}(n)\right)
$$

Portanto, a solução para o problema (3) pode ser obtida de forma recursiva através de

$$
\hat{\theta}(n+1) = \hat{\theta}(n) + K(n)\left(y(n+1) - \phi^T(n)\hat{\theta}(n)\right)
$$

$$
K(n) = \frac{P(n-1)\phi(n)}{1+\phi^T(n)P(n-1)\phi(n)}
$$

$$
P(n) = \left(I - K(n)\phi^T(n)\right)P(n-1)
$$

Algoritmo:
Inicialize $\phi(0)$, $\hat{\theta}(0)$, e $P(-1) = cI$. No instante $n+1$:

1. leia $y(n+1)$
2. Calcule o valor predito da saída $\hat{y}(n+1)$

$$
\hat{y}(n+1) = \phi^T(n)\hat{\theta}(n)
$$

3. Calcule o ganho $K(n)$

$$
K(n) = \frac{P(n-1)\phi(n)}{1+\phi^T(n)P(n-1)\phi(n)}
$$

4. Atualize a estimativa

$$
\hat{\theta}(n+1) = \hat{\theta}(n) + K(n)\left(y(n+1) - \hat{y}(n+1)\right)
$$

5. Armazene $\hat{\theta}(n+1)$ para resumo gráfico

6. Atualize a matriz de covariância

$$P(n) = \left(I - K(n)\phi^T(n)\right) P(n-1)$$

7. Incremente n e retorne ao passo 1

3 Intervalos de Confiança

Suponha que

$$\sqrt{n}\left(\hat{\theta}(n) - \theta\right) \sim \text{ As } N(0, \bar{P})$$

onde $\bar{P}$ é a matriz de covariância do erro de estimação e As denota assintoticamente.

Então o i-ésimo componente de $\hat{\theta}_i(n)$ satisfaz

$$\sqrt{n}\left(\hat{\theta}_i(n) - \theta_i\right) \sim \text{ As } N(0, \bar{P}_{ii})$$

sendo $\bar{P}_{ii}$ o i-ésimo elemento da diagonal da matriz de covariância do erro de estimação, ou seja

$$\underbrace{\sqrt{\frac{n}{\bar{P}_{ii}}}\left(\hat{\theta}_i(n) - \theta_i\right)}_{=Z} \sim \text{ As } N(0, 1)$$

Logo, definindo α como mostrado na figura 2
tem-se

$$P\left(\{\omega : -z\alpha/2 \leq Z \leq z\alpha/2\}\right) \triangleq \int_{-z\alpha/2}^{z\alpha/2} f(z)dz$$
$$= \int_{-\infty}^{\infty} f(z)dz - \int_{-\infty}^{z\alpha/2} f(z)dz - \int_{z\alpha/2}^{\infty} f(z)dz = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

que implica

$$P\left(\{\omega : -z\alpha/2 \leq \sqrt{\frac{n}{\bar{P}_{ii}}}\left(\hat{\theta}_i(n) - \theta_i\right) \leq z\alpha/2\}\right) = 1 - \alpha$$

De onde pode-se escrever

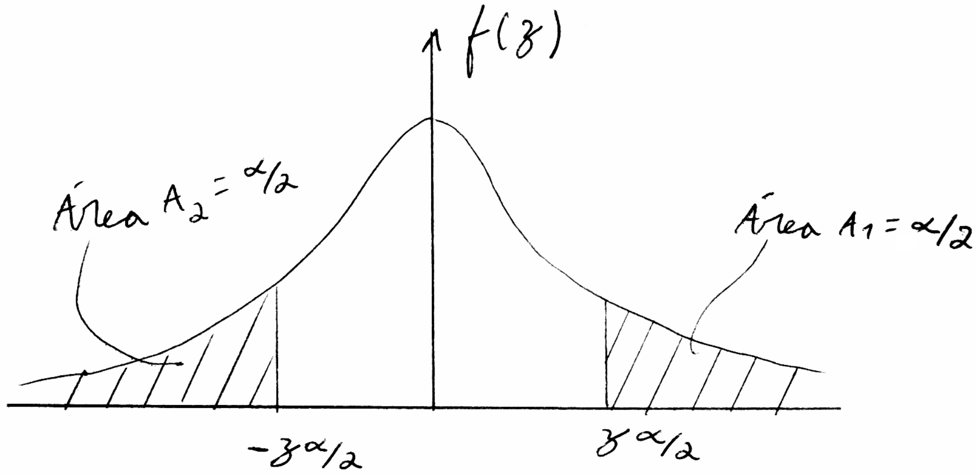

Figura 2: Definição de α.

$$P\left(\left\{\omega : \left(\hat{\theta}_i(n) - \sqrt{\frac{\bar{P}_{ii}}{n}}z\alpha/2\right) \leq \theta_i \leq \left(\hat{\theta}_i(n) + \sqrt{\frac{\bar{P}_{ii}}{n}}z\alpha/2\right)\right\}\right) \approx 1 - \alpha$$

sendo a aproximação tanto melhor quanto maior n.

Logo, para uma dada estimativa $\hat{\theta}_i(n)$, o intervalo de $(1-\alpha)\times 100\%$ de confiança para o parâmetro θ_i é dado por

$$\hat{\theta}_i(n) - \sqrt{\frac{\bar{P}_{ii}}{n}}z\alpha/2 \leq \theta_i \leq \hat{\theta}_i(n) + \sqrt{\frac{\bar{P}_{ii}}{n}}z\alpha/2 \tag{5}$$

Exemplo 1 *Para $\alpha = 0.05$ tem-se que um intervalo de $(1-\alpha)\times 100\% = 95\%$ de confiança corresponde a $z\alpha/2 = z0.025 = 1.96$ (da tabela de densidade gaussiana).*

$\ulcorner$

Obs 2 : *Para o caso ARX, pode ser mostrado que*

$$\bar{P} = \sigma^2 E\left[\phi(t)\phi^T(t)\right]^{-1} \text{ onde } \sigma^2 = E\left[\omega^2(t+1)\right]$$

Em geral, $\bar{P}$ não pode ser calculado a priori, pois depende da variância do ruído e do vetor de parâmetros θ.

Contudo, $\bar{P}$ *pode ser estimado:*

$$\hat{\bar{P}}(n) = \hat{\sigma}^2(n)\left(\frac{1}{n}\sum_{t=0}^{n-1}\phi(t)\phi^T(t)\right)^{-1}$$

com

$$\hat{\sigma}^2(n) = \frac{1}{n}\sum_{t=0}^{n-1} e^2(t+1,\theta)$$

sendo $e(t+1,\theta)$ *o erro de predição.*

Neste caso, deve-se formalmente substituir $z\alpha/2$ *em (5) por* $t\alpha/2$ *e consultar a tabela da densidade* t *de Student. Contudo, para* n *grande (tipicamente* > 100*) pode-se utilizar com erro desprezível, a densidade gaussiana.*

∟

Teorema 1 ([2]) *Se o sistema identificado for exponencialmente estável ("o passado é esquecido exponencialmente"),*

$$\underbrace{\frac{1}{n}\sum_{t=0}^{n-1} l((e(t+1,\theta))}_{V(n,\theta} \rightarrow \underbrace{\frac{1}{n}\sum_{t=0}^{n-1} E\left[l\left(e(t+1,\theta)\right)\right]}_{\omega(n,\theta)}$$

$l\left(e(t+1,\theta)\right)$ *for três vezes diferenciável em relação a* θ *e* e *e estas derivadas satisfizerem algumas condições técnicas (vide [2] para detalhes), então* $\sqrt{n}\left(\hat{\theta}(n)-\theta\right)$ *é assintoticamente normal com média zero e matriz de covariância*

$$\bar{P} = \left(\bar{\omega}''(\theta)\right)^{-1} U \left(\bar{\omega}''(\theta)\right)^{-1}$$

onde

$$\bar{\omega}(\theta) = \lim_{n\to\infty}\omega(n,\theta) \rightarrow U = \lim_{n\to\infty} E\left[nV'(\theta)\left(V'(\theta)\right)^T\right]$$

4 Exercícios

1. Seja um sistema eletromecânico como da figura 3 com os parâmetros:

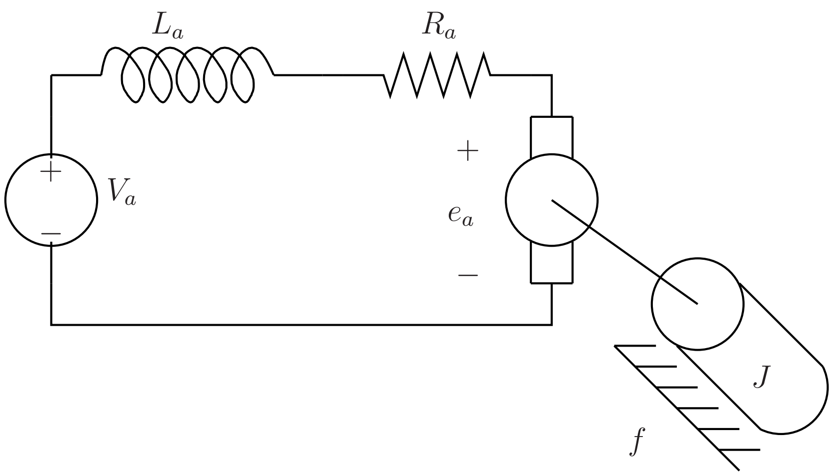

Figura 3: Sistema eletromecânico.

$$
\begin{aligned}
L_a &= 0 \\
R_a &= 0.2\Omega \\
K_a &= 0.06V/rad/s \\
K_T &= 300 \times 10^{-5} kgm/A \\
J &= 7.33 \times 10^{-5} kgm/rad/s^2 \\
f &= 6.67 \times 10^{-4} kgm/rad/s \\
T &= 10ms
\end{aligned}
$$

Simule a saída da planta para uma entrada degrau unitário e supondo que a planta está sujeita a um ruído gaussiano com média zero e variância:

(a) $\sigma^2 = 0.01$

(b) $\sigma^2 = 0.1$

Apresentar os gráficos de $y(t)$.

2. Para a mesma planta do item 1 simule o algoritmo de identificação RLS para uma entrada degrau unitário e supondo que a planta está sujeita a um ruído gaussiano com média zero e variância

   (a) $\sigma^2 = 0.01$

   (b) $\sigma^2 = 0.1$

Apresentar os gráficos das estimativas $\hat{a}_1(n+1)$ e $\hat{b}_1(n+1)$. $y(t)$, $\hat{y}(t)$ e elementos da diagonal de $P(t)$.

Referências

[1] L. A. Aguirre. *Introdução à Identificação de Sistemas: Técnicas Lineares e Não Lineares Aplicadas a Sistemas Reais*. Editora UFMG, Belo Horizonte, MG, 2000.

[2] L. Ljung and P. E. Caines. Asymptotic normality of prediction error estimators for approximate system models. *Automatica*, 3:29–46, 1979.

A Lema de Inversão de Matrizes

Lema 1 (Lema de Inversão de Matrizes[1]) *Sejam A, B, C e D matrizes tais que A, C e $(A+BCD)$ sejam inversíveis, então*

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1} \tag{6}$$

Prova 1 *Pré-multiplicando-se ambos os lados de (6) por $(A+BCD)$ tem-se:*

$$(A+BCD)(A+BCD)^{-1} = (A+BCD)A^{-1} - (A+BCD)A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}$$

$$I = I + BCDA^{-1} - \left(I + BCDA^{-1}\right)B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}$$

$$I = I + BCDA^{-1} - \left(B + BCDA^{-1}B\right)\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}$$

Colocando-se BC em evidência no terceiro termo, tem-se

$$I = I + BCDA^{-1} - BC\left(C^{-1} + DA^{-1}B\right)\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}$$

$$I = I + BCDA^{-1} - BCDA^{-1}$$

e portanto

$$I = I$$