# 5. Modelos baseados em dados

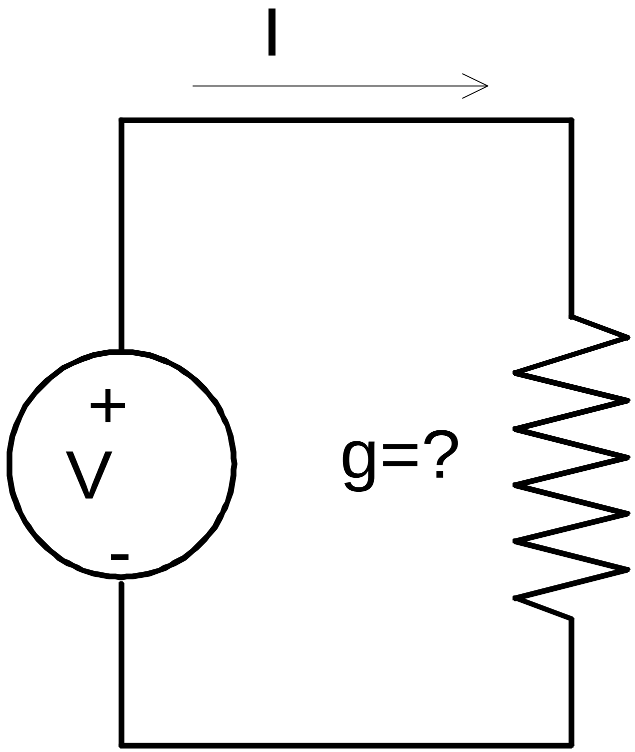
*Objectivo*: Após completar este módulo, o aluno deverá ser capaz de formular e resolver problemas simples de estimação de parâmetros com base no método dos mínimos quadrados e de identificação de sistemas usando a *Control Systems Toolbox* do MATLAB.

# Ajuste de uma reta a dados experimentais

*Uma situação experimental:*

Pretende-se relacionar a corrente *I* com a tensão *V* no circuito da figura.

Para tal são aplicados diversos valores de tensão à resistência e registados os dados




| Tensão [volt] | Corrente [mA] |
|---|---|
| $V_1=1$ | $I_1=2.1$ |
| $V_2=2$ | $I_2=3.9$ |
| $V_3=3$ | $I_3=6.2$ |
| $V_4=4$ | $I_4=7.9$ |

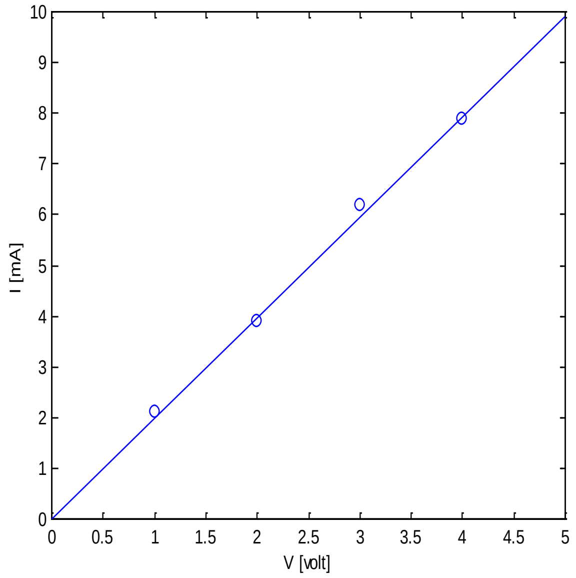
Sob certas condições, a relação teórica (o modelo) existente entre a tensão *V* aplicada à resistência e a corrente *I* é:

$$I = gV$$

em que $g$ é um parâmetro que se pretende estimar a partir dos dados.

Devido aos erros experimentais, os pontos experimentais não se encontram exactamente sobre a reta $I = gV$ mas têm desvios.

*Como decidir qual a recta melhor ajustada?*



De acordo com o **Princípio dos Mínimos Quadrados** é escolhida a recta que minimiza a soma dos quadrados dos desvios.

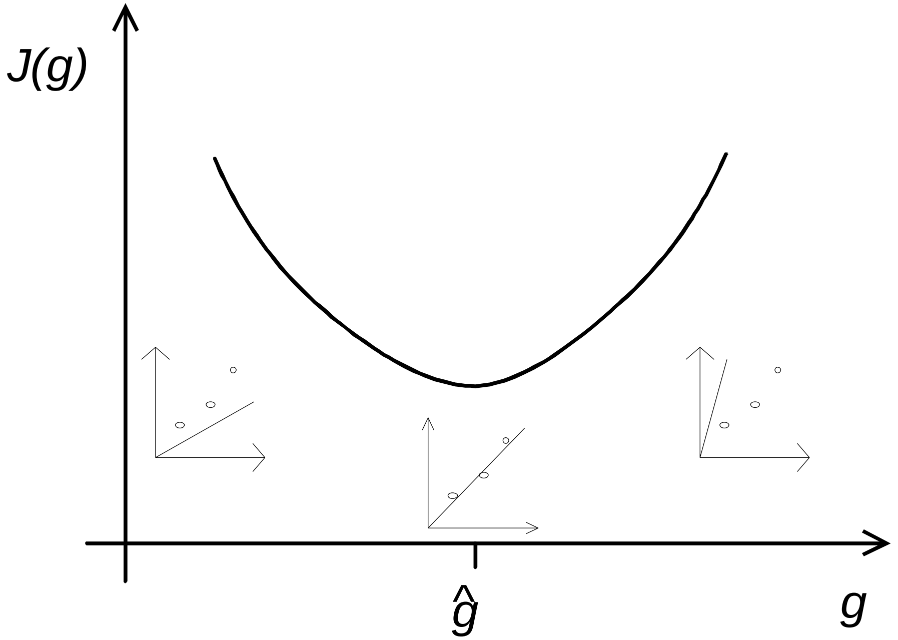
De acordo com este princípio, a estimativa de $g$ é tal que minimiza

$$J(g) = (2.1 - g \times 1)^2 + (3.9 - g \times 2)^2 + (6.2 - g \times 3)^2 + (7.9 - g \times 4)^2$$

O que efetivamente observamos

O que esperamos que seja a corrente quando a tensão é *2* (Depende da estimativa de *g*)

## "custo" $J(g)$ associado ao critério de mínimos quadrados



Como $J(g)$ é uma função quadrática de $g$ , a estimativa de mínimos quadrados verifica a equação

$$\frac{1}{2}\frac{\partial J}{\partial g}|_{g=\hat{g}} = 0$$

ou seja

$$-1\times(2.1-\hat{g}\times 1)-2\times(3.9-\hat{g}\times 2)-3\times(6.2-\hat{g}\times 3)-4\times(7.9-\hat{g}\times 4)=0$$

Esta equação simplifica-se para $60\hat{g}-120.1=0$ sendo a estimativa de mínimos quadrados dada por

$$\hat{g} = 2.00$$

## Ajuste de uma reta a dados experimentais (Caso geral)

Suponhamos que a relação teórica entre duas grandezas $X$ e $Y$ é do tipo

$$Y = \alpha X$$

em que $\alpha$ é um parâmetro desconhecido que se pretende estimar.

Repare-se que, conhecendo uma estimativa de $\alpha$ podemos responder a perguntas do tipo "*Se* $X$ *valer … quanto se espera que valha* $Y$ *?* "

Suponhamos que são observados $n$ pares $(X_i, Y_i), i = 1, \ldots, n$ correspondentes a outros tantos ensaios experimentais. Constrói-se a tabela

| | |
|---|---|
| $X_1$ | $Y_1$ |
| $X_2$ | $Y_2$ |
| $X_3$ | $Y_3$ |
| $X_4$ | $Y_4$ |
| $X_5$ | $Y_5$ |

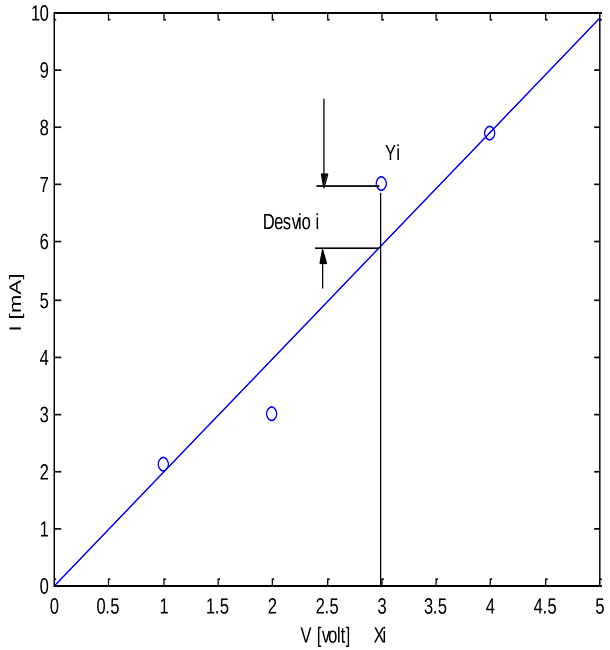
Pretende-se estimar a reta melhor ajustada aos dados experimentais, de acordo com o critério de mínimos quadrados.

De acordo com esta critério, a estimativa é tal que minimiza a soma dos quadrados dos desvios:

$$J(\alpha) = \sum_{i=1}^{n} \left(Y_i - \alpha X_i\right)^2$$

O que efetivamente observamos

O que estamos à espera que seja $Y_i$



A estimativa de mínimos quadrados verifica a equação (eq. "normal"):

$$\frac{1}{2}\frac{\partial J}{\partial \alpha}|_{\alpha=\hat{\alpha}}$$

ou seja

$$\sum_{i=1}^{n} X_i \left(Y_i - \hat{\alpha} X_i\right) = 0$$

esta equação tem por solução

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$$

# Mínimo ou não?

A condição $\left.\frac{1}{2}\frac{dJ}{d\alpha}\right|_{\alpha=\hat{\alpha}}=0$ não garante necessariamente que $J(\alpha)$ seja mínimo para $\alpha=\hat{\alpha}$. É necessário impor uma condição na segunda derivada:

$$\frac{d^2J}{d\alpha^2}=\sum_{i=1}^{n}X_i^2>0$$

Neste exemplo, esta condição é verificada se for feita pelo menos uma medida com $X \neq 0$ (o que tem uma interpretação geométrica imediata).

Veremos a seguir que se estimarmos mais do que um parâmetro a segunda derivada deixa de ser um escalar. A condição de mínimo é então a de que os dados sejam tais que a matriz de segundas derivadas seja definida positiva.

## Bom, ou apenas ótimo?

A estimativa de mínimos quadrados é "óptima" no sentido em que minimiza um funcional de custo. No entanto, o funcional de custo pode não ser o mais adequado.

Como caricatura, pode dizer-se que os bons relógios são os que estão parados pois dão horas absolutamente certas duas vezes por dia.

Um outro exemplo é o de um caçador que vê dois pombos. Se disparar para o ponto que minimiza a distância média quadrática aos *dois* pombos…

Isto sugere que por vezes são necessários outros critérios.

## Outros critérios de Estimação

Os exemplos anterios sugerem a utilidade de utilizar critérios que ultrapassem as limitações dos Mínimos Quadrados. Um dos mais utilizados em Estimação é o critério de **Máxima Verosimilhança**.

Em presença de ruído colorido os mínimos quadrados dão uma estimativa polarizada.

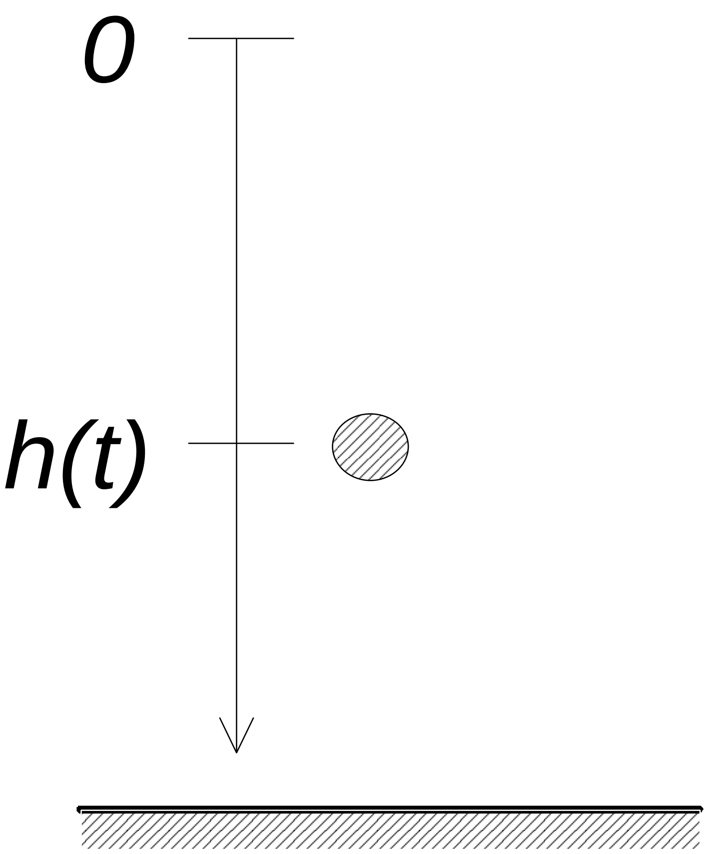
**Carl Frederich Gauss (1777-1855)** utilizou pela primeira vez o critério de mínimos quadrados para a estimação de parâmetros em equações.

Em 1801, o astrónomo italiano Piazzi observou pela primeira vez um pequeno planeta denominado Ceres. Infelizmente, a duração das observações era muito curta devido a Ceres se esconder atrás do Sol, pelo que estas eram insuficientes para estimar os parâmetros da sua órbita pelos métodos tradicionais. Recorrendo ao critério dos mínimos quadrados, Gauss efectuou uma estimativa (bastante diferente das obtidas pelos métodos clássicos) que foi brilhantemente confirmada pelas observações experimentais.

Qual a estimativa de mínimos quadrados da aceleração da gravidade $g$?

Modelo: $h = g\dfrac{t^2}{2} + e$

| $t$ [$s$] | $h$ [$m$] |
|---|---|
| 1 | 8.49 |
| 2 | 20.05 |
| 3 | 50.65 |
| 4 | 72.19 |
| 5 | 129.85 |
| 6 | 171.56 |

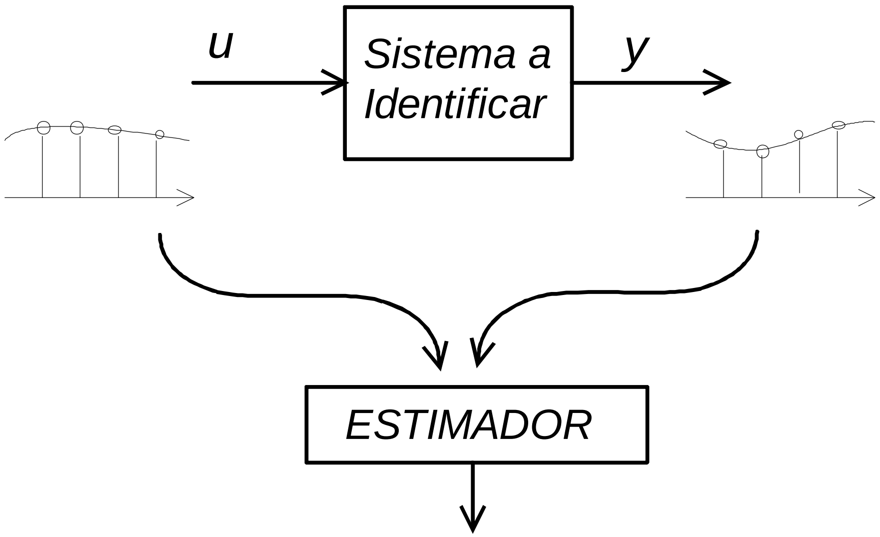


# Estimação de parâmetros em equações de diferenças

## Modelo:

$$y(t)+a_1y(t-1)+\ldots+a_ny(t-n)=b_0u(t-1)+\ldots+b_mu(t-m-1)$$

**Problema:** A partir das amostras de $u$ e $y$, estimar os parâmetros $a_i\,,b_j$



## Exemplo

Considere-se o sistema

$$y(k+1) = ay(k) + bu(k) + e(k+1)$$

Dados recolhidos

$$\sum_{k=1}^{1000} y^2(k) = 30 \qquad \sum_{k=1}^{1000} u^2(k) = 50 \qquad \sum_{k=1}^{1000} y(k+1)y(k) = 1$$

$$\sum_{k=1}^{1000} y(k+1)u(k) = 36 \qquad \sum_{k=1}^{1000} y(k)u(k) = 20$$

Determinar as estimativas de mínimos quadrados dos parâmetros $a$ e $b$

- Escrever a funcional de mínimos quadrados
- Calcular as derivadas parciais em ordem aos parâmetros e igualar a zero

$$J = \frac{1}{2}\sum_{k=1}^{N}\left[y(k+1) - ay(k) - bu(k)\right]^2$$

$$\frac{\partial J}{\partial a} = -\sum_{k=1}^{N} y(k)\left[y(k+1) - ay(k) - bu(k)\right] = 0$$

$$\frac{\partial J}{\partial b} = -\sum_{k=1}^{N} u(k)\left[y(k+1) - ay(k) - bu(k)\right] = 0$$

$$\hat{a}\sum_{k=1}^{N} y^2(k) + \hat{b}\sum_{k=1}^{N} y(k)u(k) = \sum_{k=1}^{N} y(k)y(k+1) \qquad 30\hat{a} + 20\hat{b} = 1$$

$$\hat{a}\sum_{k=1}^{N} u(k)y(k) + \hat{b}\sum_{k=1}^{N} u^2(k) = \sum_{k=1}^{N} u(k)y(k+1) \qquad 20\hat{a} + 50\hat{b} = 36$$

$$\hat{a} = -0.61 \qquad \hat{b} = 0.964$$

## Notação matricial

Se quisermos resolver o problema de estimação para um número arbitrário de parâmetros temos de usar a notação matricial.

$$y(t)+a_1 y(t-1)+\ldots+a_n y(t-n)=b_0 u(t-1)+\ldots+b_m u(t-1-m)+e(t)$$

Define-se o **regressor** , $\varphi$ ,como

$$\varphi'(t-1)=\left[-y(t-1)\ \ldots\ -y(t-n)\ u(t-1)\ u(t-1-m)\right]$$

e o vector de parâmetros a estimar, $\theta$ ,como

$$\theta'=\left[a_1\ \ldots\ a_n\ b_1\ \ldots\ b_m\right]$$

O modelo de escreve-se na forma do modelo de regressão:

$$y(t)=\varphi'(t-1)\theta+e(t)$$

# Critério de Mínimos Quadrados

Dadas $N$ observações, estimar o vetor de parâmetros $\theta_o$ por um vector $\hat{\theta}$ por forma a que o funcional seguinte seja mínimo:

$$J(\theta) = \frac{1}{2N} \sum_{t=1}^{N} \left[ y(t) - \theta' \varphi(t-1) \right]^2$$

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} \varphi'(0)\theta \\ \varphi'(1)\theta \\ \vdots \\ \varphi'(N-1)\theta \end{bmatrix} + \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix}$$

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} \varphi'(0) \\ \varphi'(1) \\ \vdots \\ \varphi'(N-1) \end{bmatrix}\theta + \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix}$$

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} \varphi'(0) \\ \varphi'(1) \\ \vdots \\ \varphi'(N-1) \end{bmatrix} \theta + \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix}$$

$\bar{y}[N \times 1]$ $\Phi[N \times n_p]$ $[n_p \times 1]$ $\bar{\varepsilon}[N \times 1]$

O conjunto das $N$ observações satisfaz:

$$\bar{y} = \Phi\theta + \bar{\varepsilon}$$

Funcional de mínimos quadrados escrito matricialmente:

$$J(\theta) = \frac{1}{2N}\left\|\bar{\varepsilon}\right\|^2 = \frac{1}{2N}\bar{\varepsilon}'\bar{\varepsilon}$$

Como

$$\bar{\varepsilon} = \bar{y} - \Phi\hat{\theta}$$

Vem

$$J(\theta) = \frac{1}{2N}\left(\bar{y}' - \theta'\Phi'\right)\left(\bar{y} - \Phi\theta\right)$$

$$J(\theta) = \frac{1}{2N}\left(\bar{y}'\bar{y} - 2\bar{y}'\Phi\theta + \theta'\Phi'\Phi\theta\right)$$

O funcional de mínimos quadrados é uma **forma quadrática** em $\theta$

A estimativa $\hat{\theta}$ de mínimos quadrados satisfaz

$$\nabla_\theta J(\theta)\Big|_{\theta=\hat{\theta}} = 0$$

Gradiente da forma quadrática

$$\nabla_x \left(x'Ax\right) = 2x'A$$

Recordando

$$J(\theta) = \frac{1}{2N}\left(\bar{y}'\bar{y} - 2\bar{y}'\Phi\theta + \theta'\Phi'\Phi\theta\right)$$

vem

$$\nabla_\theta J = \frac{1}{2N}\left(-2\bar{y}'\Phi + 2\theta'\Phi'\Phi\right)$$

A estimativa de mínimos quadrados satisfaz pois a equação

$$\nabla_{\theta} J = \frac{1}{2N}\left(-2\bar{y}'\Phi + 2\theta'\Phi'\Phi\right)$$

ou seja

$$\hat{\theta}'\Phi'\Phi = \bar{y}'\Phi$$

ou, transpondo (a transposta do produto é o produto das transpostas por ordem inversa: $(MN)^T = N^T M^T$):

$$\Phi'\Phi\hat{\theta} = \Phi'\bar{y}$$

## Equação Normal

Em conclusão, a estimativa de mínimos quadrados do vetor de parâmetros $\theta$ do modelo de regressão linear

$$y(t) = \varphi'(t-1)\theta + e(t)$$

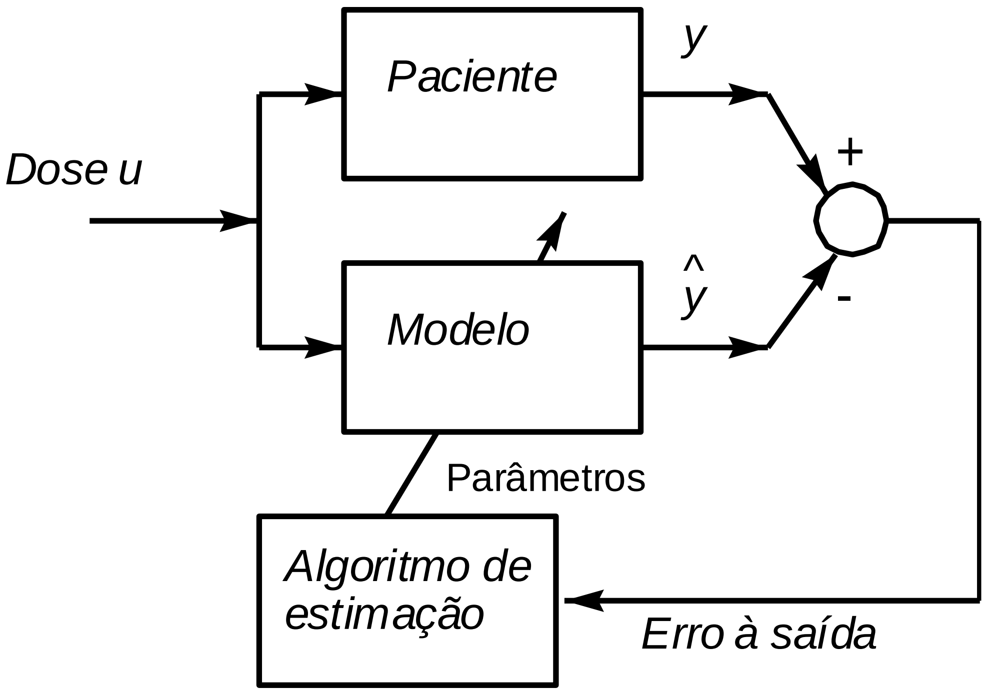
satisfaz a equação matricial (dita *equação normal*)

$$\Phi'\Phi\hat{\theta} = \Phi'\bar{y}$$

Se existir a inversa de $\Phi'\Phi$ a estimativa de mínimos quadrados existe e é única, sendo dada por

$$\hat{\theta} = \left(\Phi'\Phi\right)^{-1}\Phi'\bar{y}$$

# Dependência não linear dos parâmetros




Modelo não linear

dependente dos parâmetros

$$\dot{x} = f(x, u, \theta)$$

$$y = h(x)$$

$$J(\theta) = \sum_{k=1}^{N} \frac{1}{\sigma_k^2} (y_k - \hat{y}_k(\theta))^2$$

## Exemplo: Modelo de infecção por HIV-1

$$\begin{cases} \dfrac{dT}{dt} = s - dT - \beta T v \\ \dfrac{dT^*}{dt} = \beta T v - \mu_2 T^* \\ \dfrac{dv}{dt} = kT^* - \mu_1 v \end{cases}$$

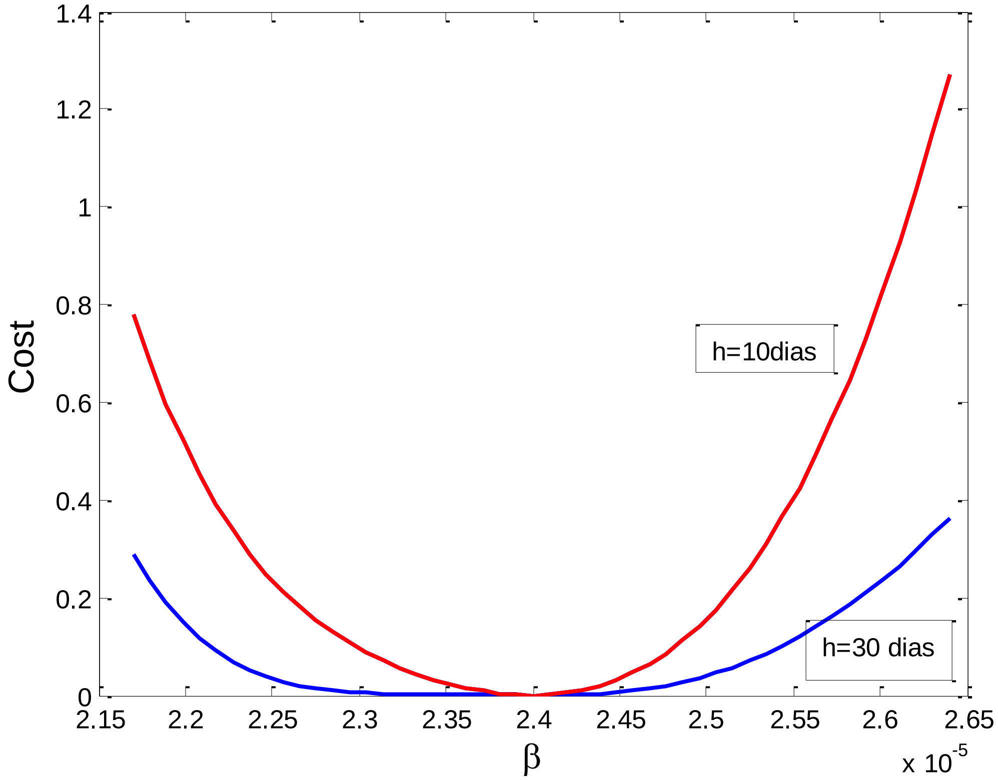
Observam-se $v$ e $T + T^*$. A partir destas observações pretende-se estimar os parâmetros $s, d, \beta, k, \mu_1, \mu_2$.

Calcular o funcional de mínimos quadrados para um dado valor dos parâmetros implica resolver as equações diferenciais do modelo.

Não há solução em forma fechada.

É necessário usar métodos numéricos de otimização.

# Efeito dos dados no funcional de custo




A sensibilidade da função de custo depende das condições experimentais de recolha de dados. Ex.: No modelo do HIV, espaçando mais os dados a sensibilidade da função em ordem ao parâmetro a estimar diminui.

## Inclusão de informação *a priori* sobre os parâmetros

$$J(\theta) = \sum_{i=1}^{n_p} \frac{1}{\bar{\sigma}_i^2} (\theta_i - \bar{\theta}_i)^2 + \sum_{k=1}^{N} \frac{1}{\sigma_k^2} (y_k - \hat{y}_k(\theta))^2$$

$\bar{\theta}_i, i = 1, \dots, n_p$ são as estimativas *a priori* (antes de usar os dados)

$\bar{\sigma}_i^2$ são as variâncias dos erros das estimativas *a priori* (medem o grau de confiança nas estimativas *a priori*).

Esta expressão pressupõe uma distribuição Gaussiana dos erros de estimação.

*Exercício*: Deduza a expressão das estimativas para o caso em que há informação *a priori* sobre o valor dos parâmetros e o modelo é linear ($y_k = \varphi_k^T \theta + e_k$).

$$y_k = \varphi_k^T \theta + e_k$$

$$\begin{bmatrix} y_1 \\ \cdots \\ y_N \end{bmatrix} = \begin{bmatrix} \varphi_0^T \\ \cdots \\ \varphi_{N-1}^T \end{bmatrix} \theta + \begin{bmatrix} e_1 \\ \cdots \\ e_N \end{bmatrix} \quad \text{ou} \quad Y = \Phi\theta + \varepsilon$$

$$J(\theta) = \alpha \parallel \theta - \bar{\theta} \parallel^2 + \parallel Y - \Phi\theta \parallel^2$$

Cálculo da norma euclidiana

$$\parallel x \parallel^2 = x^T x = \sum_{i=1}^{n} x_i^2$$

Regras para cálculo do gradiente

$$\frac{\partial}{\partial x} b^T x = b^T \qquad \frac{\partial}{\partial x} x^T A x = 2x^T A$$

$$J(\theta) = \alpha \parallel \theta - \bar{\theta} \parallel^2 + \parallel Y - \Phi\theta \parallel^2$$

$$J(\theta) = \alpha(\theta - \bar{\theta})^T(\theta - \bar{\theta}) + (Y^T - \theta^T\Phi^T)(Y - \Phi\theta)$$

$$J(\theta) = \alpha\theta^T\theta - 2\alpha\bar{\theta}^T\theta + \alpha\bar{\theta}^T\bar{\theta} + Y^TY - 2Y^T\Phi\theta + \theta^T\Phi^T\Phi\theta$$

$$\frac{\partial J}{\partial \theta} = 2\alpha\theta^T - 2\alpha\bar{\theta}^T - 2Y^T\Phi + 2\theta^T\Phi^T\Phi = 0$$

Equação normal regularizada

$$(\Phi^T\Phi + \alpha I)\theta = \Phi^TY + \alpha\bar{\theta}$$

Estimador com prior (regularizado)

$$\theta = (\Phi^T\Phi + \alpha I)^{-1}(\Phi^TY + \alpha\overline{\theta)}$$

## LASSO

Podemos considerar outros tipos de regularização.

Uma possibilidade é penalizar o número de elementos não nulos do vetor de parâmetros, (norma $l_0$, de facto não é uma norma) por forma a “explicar” os dados com um número mais reduzido de parâmetros.

A minimização deste funcional leva a uma carga computacional incomportável pelo que se usa a aproximação do termo de penalização pela norma $l_1$ do vetor de parâmetros (soma dos módulos dos componentes). O funcional a minimizar passa a ser

$$J(\theta) = \| Y - \Phi\theta \|^2 + \lambda \| \theta \|_1$$

Que corresponde ao algoritmo LASSO (*least absolute shrinkage and selection operator*).

O LASSO dá origem a estimativas esparsas, pois tende a reduzir o número de parâmetros.

Este tipo de abordagem tem grande importância em Processamento de Sinais e Controlo. Em particular em imagem permite obter resultados espetaculares.

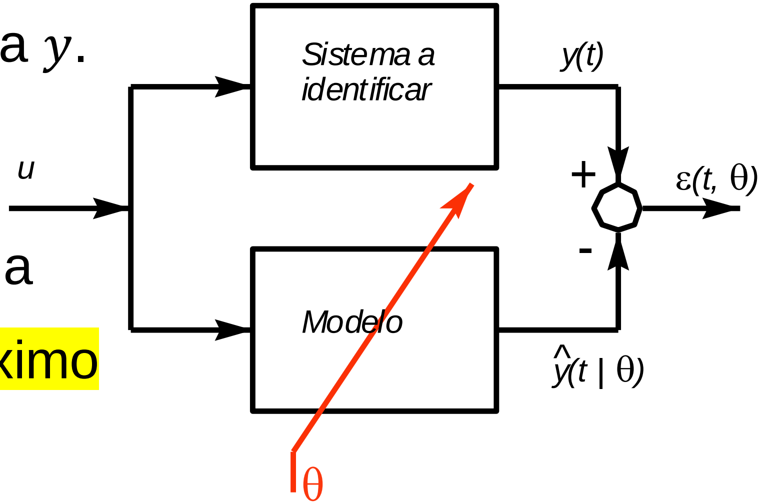
Há o problema de o termo associado à norma $l_1$ não ser diferenciável, o que requer conceitos da otimização numérica de funções não diferenciáveis como sub-gradiente e *proximal operators*.

Pelo menos para problemas simples, a otimização numérica associada ao LASSO pode ser feita com base no ambiente cvx, que corre no MATLAB e é de utilização livre. Ver http://cvxr.com/cvx/

# Métodos de Erro de Predição

Ref.: L. Ljung. *System Identification – Theory for the user*. Prentice-Hall, 1987. Cap. 7.

Considere-se um sistema com entrada $u$ e saída $y$.




O modelo depende de um vetor de parâmetros $\theta$ que pretendemos ajustar por forma a que a saída do modelo $\hat{y}(t,\theta)$ esteja o mais próximo possível da saída do sistema a identificar.

A melhor estimativa de $\theta$ é a que faz com que $\hat{y}(t,\theta)$ preveja melhor o valor da saída $y(t)$ do sistema (critério), pelo que os métodos de estimação de $\theta$ baseados no erro de predição se dizem Métodos de Erro de Predição (em Inglês: *Prediction Error Methods – PEM*).

Erro de predição $\varepsilon(t,\theta) = y(t) - \hat{y}(t,\theta)$

Erro de predição filtrado $\varepsilon_F(t,\theta) = L(q)\varepsilon(t,\theta)$

Dados entrada/saída $Z^N = [\, y(1), u(1), y(2), u(2), \dots, y(N), u(N)]$

Critério de ajuste $V(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} l(\varepsilon_F(t,\theta))$

Estimativa $\hat{\theta} = \hat{\theta}(Z^N) = \arg min_\theta \; V(\theta, Z^N)$

Podemos considerar vários

- Tipos de modelos (equação de diferenças, estado, …)
- Critérios de ajuste

Consoante as escolhas feitas, assim o algoritmo resultante para a estimativa de $\theta$ e as suas propriedades.

**Exemplo – Mínimos quadrados**:

$$y(t)+a_1y(t-1)+\ldots+a_ny(t-n)=b_0u(t-1)+\ldots+b_mu(t-1-m)+e(t)$$

$$\varphi'(t-1):=[-y(t-1)\quad -y(t-2)\ \ldots\ \ u(t-1)\ \ldots u(t-m)]$$

$$\theta_o{}'=[a_1\quad a_2\ \ \ldots\ a_n\ b_0\ \ldots\ b_m]$$

Preditor: $\hat{y}(t)=\varphi'(t-1)\theta$

Erro de predição: $\varepsilon(t,\theta)=y(t)-\hat{y}(t,\theta)$

Funcional de custo: $V(\theta,Z^N)=\frac{1}{N}\sum_{t=1}^{N}\varepsilon^2(t,\theta)$

Uma outra possibilidade é usar modelos de estado, em tempo contínuo ou discreto. *System Identification Toolbox* (MATLAB), função *ssest*

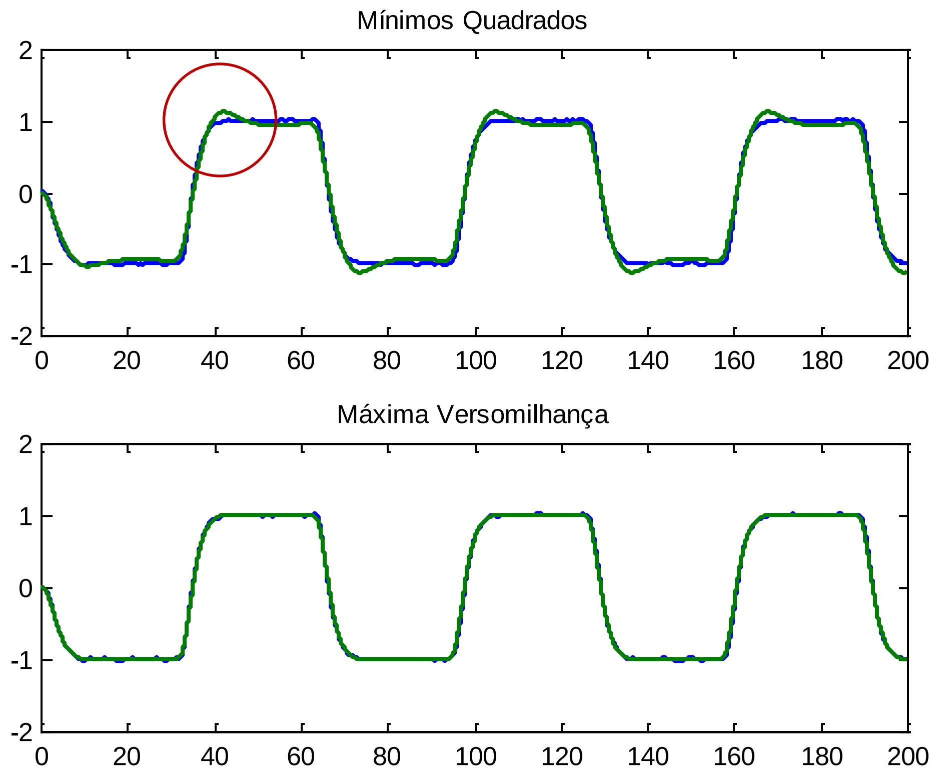
# *Estimação com o MATLAB*

```
z=[y u];
na=4;
nb=1;
nc=4;
nk=1;
nn=[na nb nc nk];
th=armax(z,nn);
yh=idsim(u,th);
plot(t,[y yh]);
[Phi,Gamma,C,D,K,X0]=th2ss(th);
T=0.05;
[A,B]=d2c(Phi,Gamma,T);
[num,den]=ss2tf(A,B,C,D);
```

A estimação de Mínimos Quadrados faz-se com a função *arx*

A estimação em presença de ruído colorido faz-se com a função *armax*

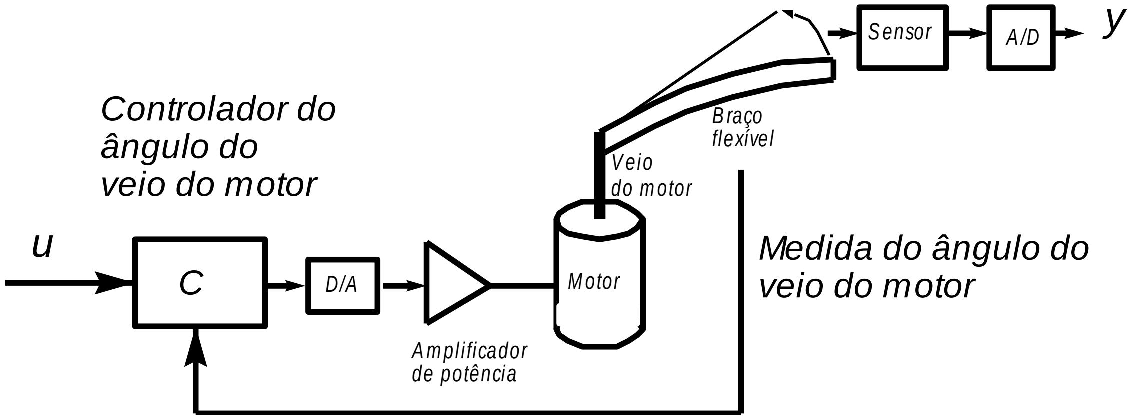
# Um exemplo simples (dados simulados)




Dado que o ruído é colorido, a estimativa de mínimos quadrados é polarizada, o que leva a um modelo que tem uma resposta diferente do "sistema real".

A utilização de um método capaz de lidar com o ruído colorido evita este problema.

# Exemplo: Identificação da dinâmica de uma barra flexível



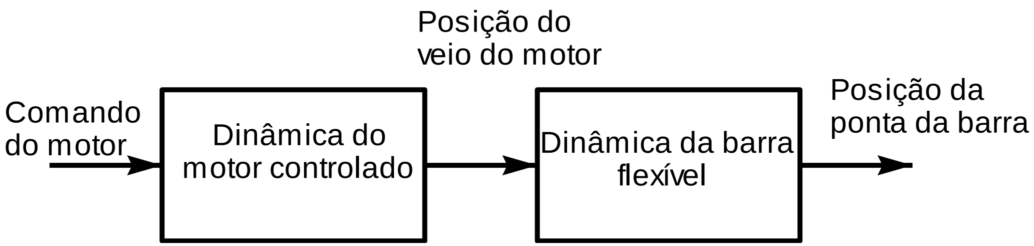




Diagrama de blocos simplificado: os movimentos flexíveis da barra afectam o movimento do motor.

Pretende-se identificar um modelo a partir de dados experimentais.

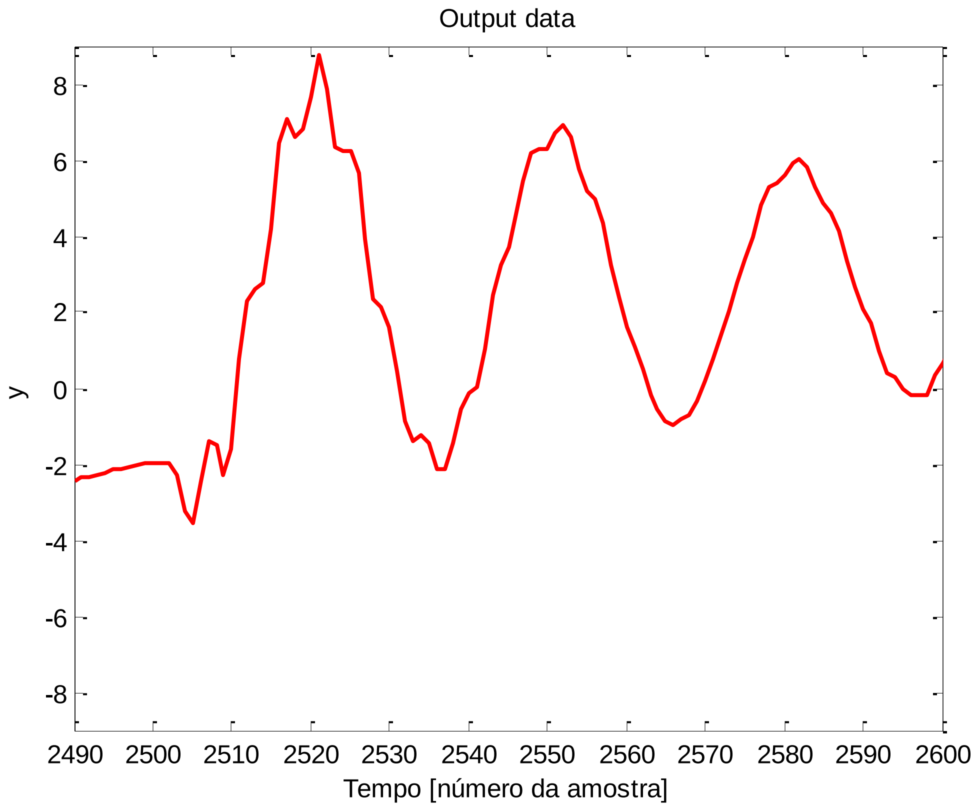
Dama com Arminho
(Retrato de Cecilia Gallerani)
1495-1490

Olhar e ver: o que podemos ver neste quadro?

Tal como olhamos para um quadro e tentamos perceber a mensagem que ele nos transmite, também devemos olhar para os dados e perceber quais as suas caraterísticas:

- Os dados fazem sentido?
- Há saturações que roubem informação?
- A média é nula?
- Qual a “escala de tempo” dos dados?
- Há comportamentos oscilatórios? Em que bandas de frequência?
- Há comportamentos rápidos e lentos?
- Há comportamentos não lineares?
- ……………………..

# Dinâmica da cadeia aberta (resposta ao escalão)



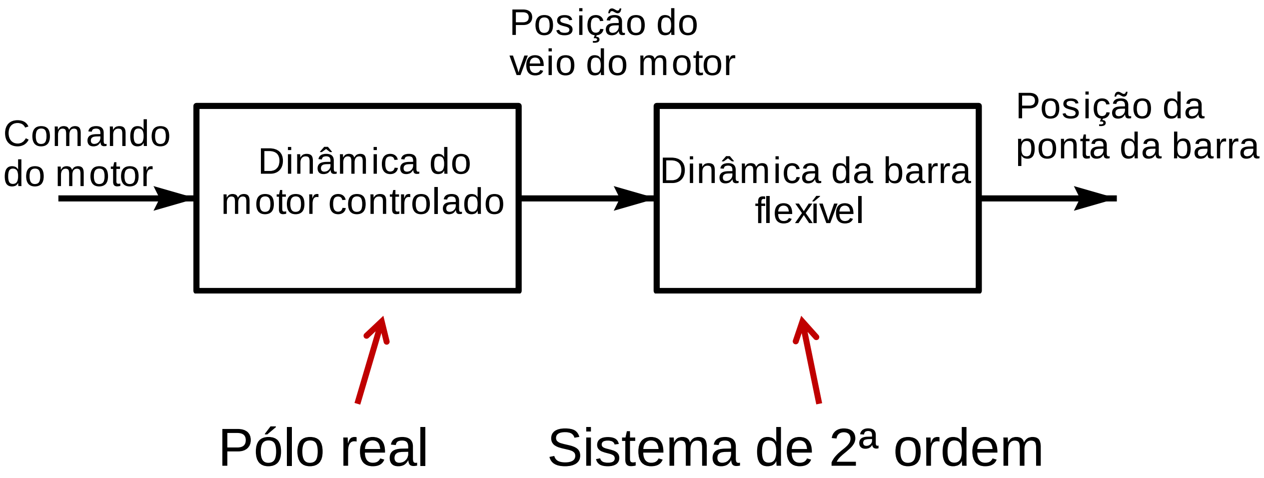

O que podemos concluir da observação desta resposta, que seja útil para construir um modelo?

## Características da resposta

- Reposta oscilatória com pouco amortecimento. (pólos complexos conjugados)
- Observa-se um modo oscilatório numa frequência dominante, mas também um modo oscilatório de mais alta frequência, com menor amplitude. (adicionar pólos complexos conjugados na alta frequência)
- Há um efeito de resposta inversa, que vai originar um zero fora do círculo unitário (zero de fase não mínima). Este efeito é devido à reacção da barra sobre o veio do motor (o modelo em cascata **não** prevê este efeito!). (zeros)

# Qual a ordem a escolher para o modelo?

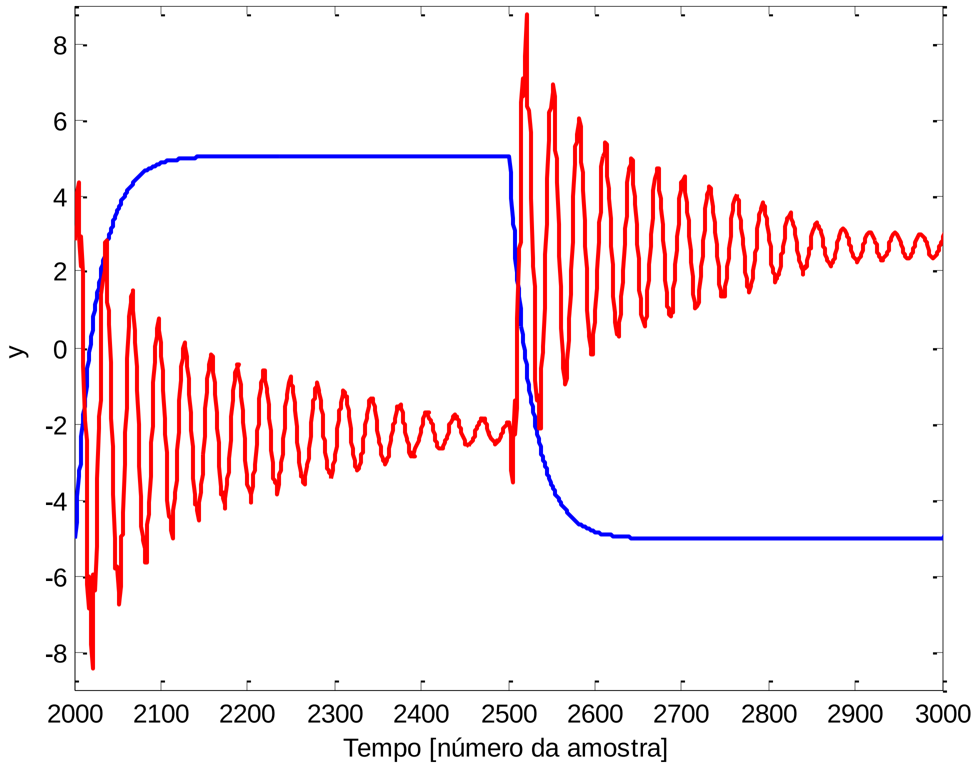
- Começar com uma ordem baixa que capte a dinâmica do modelo.
- Ir aumento progressivamente a ordem até atingir bons resultados.




Vamos começar com n=1+2=3 e m=2 (3 pólos e 2 zeros).

Admitimos um atraso puro de 1.

# Mínimos quadrados, arx(3,2)

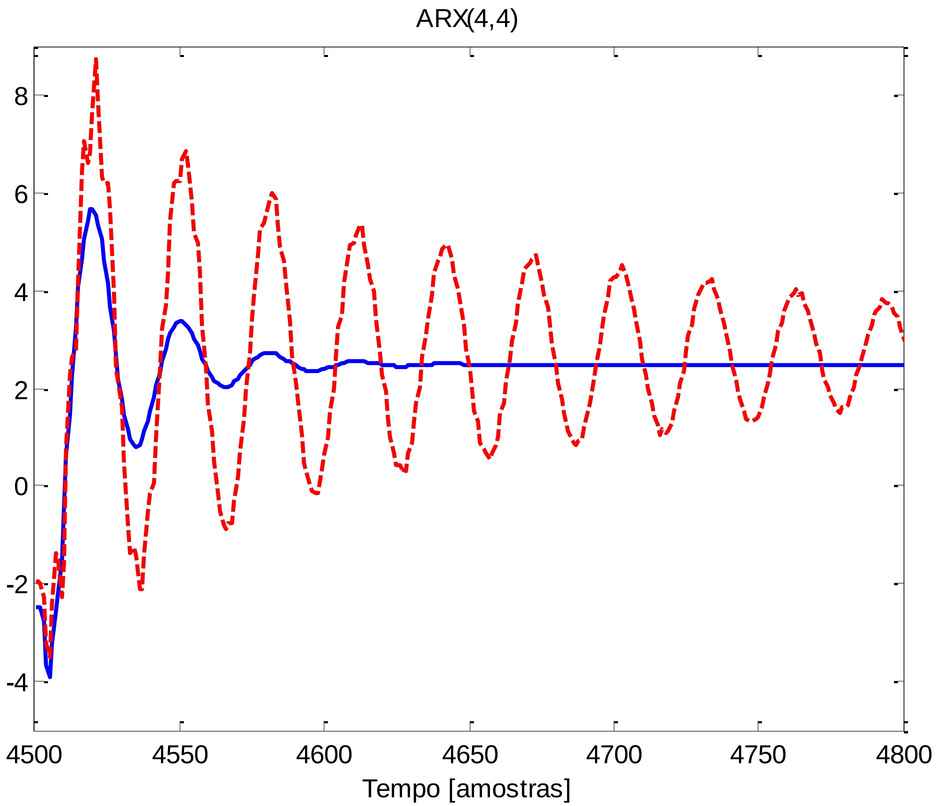



O resultado é um desastre!

A resposta do modelo estimado não é oscilatória.

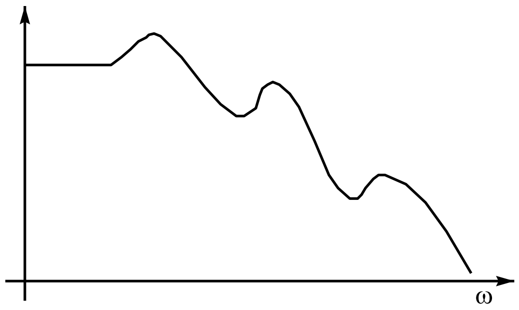
Nem sequer o ganho estático tem o sinal certo (isto é possivelmente devido ao efeito de resposta inversa).

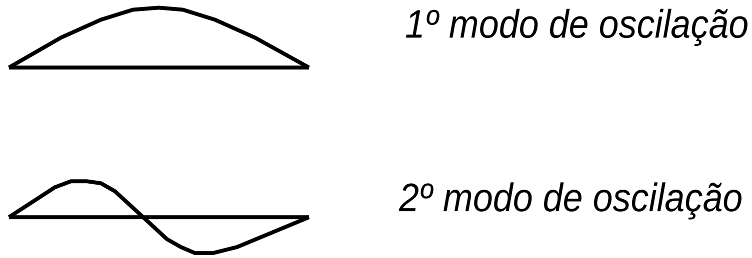
# Mínimos quadrados, arx(4,4)




Aumentando a ordem a resposta do modelo melhora (é replicado o efeito de fase não mínima, o modelo tem uma resposta oscilatória), mas o amortecimento é ainda muito grande.

# Interpretação em termos da dinâmica não modelada




A resposta em frequência inclui vários modos de oscilação a frequências sucessivamente mais altas.

Estes modos correspondem a picos na resposta em frequência.

Com os instrumentos de que dispomos, apenas podemos ver os 2 primeiros modos de oscilação, mas o modelo matemático da barra prevê mesmo a existência de um número infinito de modos, com amplitudes cada vez mais baixas.

Na ausência de ruído, o movimento da ponta da barra satisfaz o modelo

$$A(q)y(t) = B(q)u(t)$$

Dado que impomos uma ordem aos polinómios $A$ e $B$, o algoritmo de identificação estima um polinómio correspondente a essa ordem ($A_n$, $B_n$), que difere do polinómio verdadeiro por um termo de erro ($\Delta A$, $\Delta B$):

$$(A_n + \Delta A)y = (B_n + \Delta)u$$

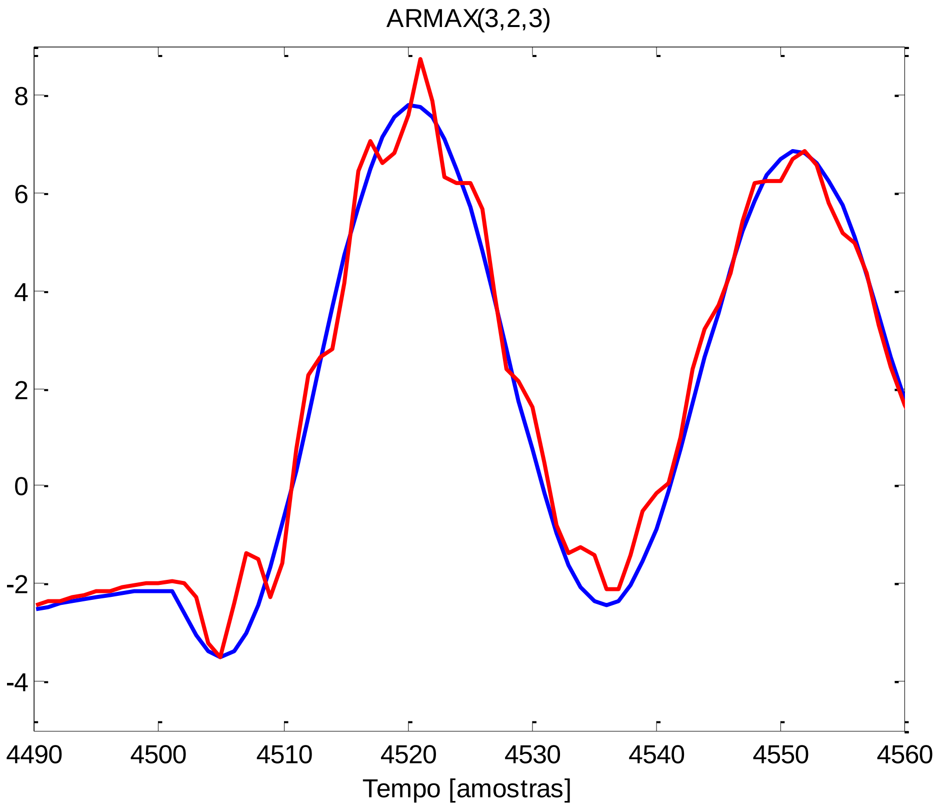
ou seja

$$A_n y = B_n u + v \qquad \text{em que} \qquad v = -A_n y + B_n u$$

Devido aos modos de ressonância de alta frequência, o termo $v$ é visto pelo algoritmo como um "ruído" fortemente correlacionado (ruído colorido).

Temos de usar um método que resista ao ruído colorido!

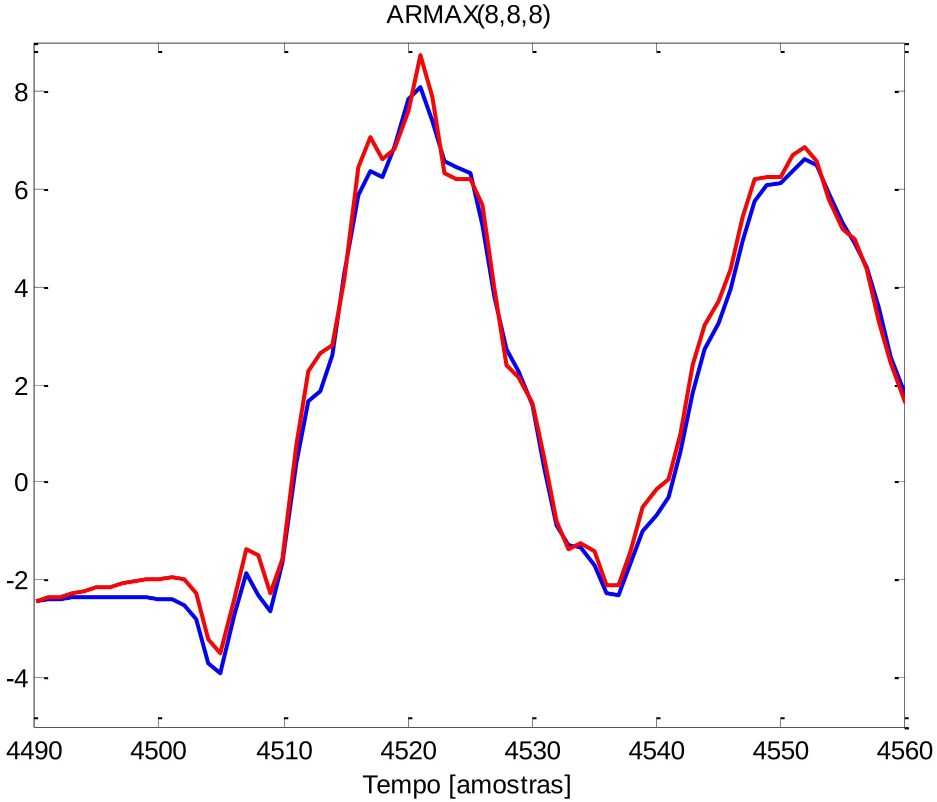
# Estimação do modelo ARMAX (3,2,3)




Com a estimação de máxima verosimilhança, o modelo estimado consegue já seguir razoavelmente o modo de oscilação fundamental.

Dependendo do desempenho requerido, este modelo pode já ser suficiente para o projeto do controlador.

# Estimação do modelo ARMAX(8,8,8)



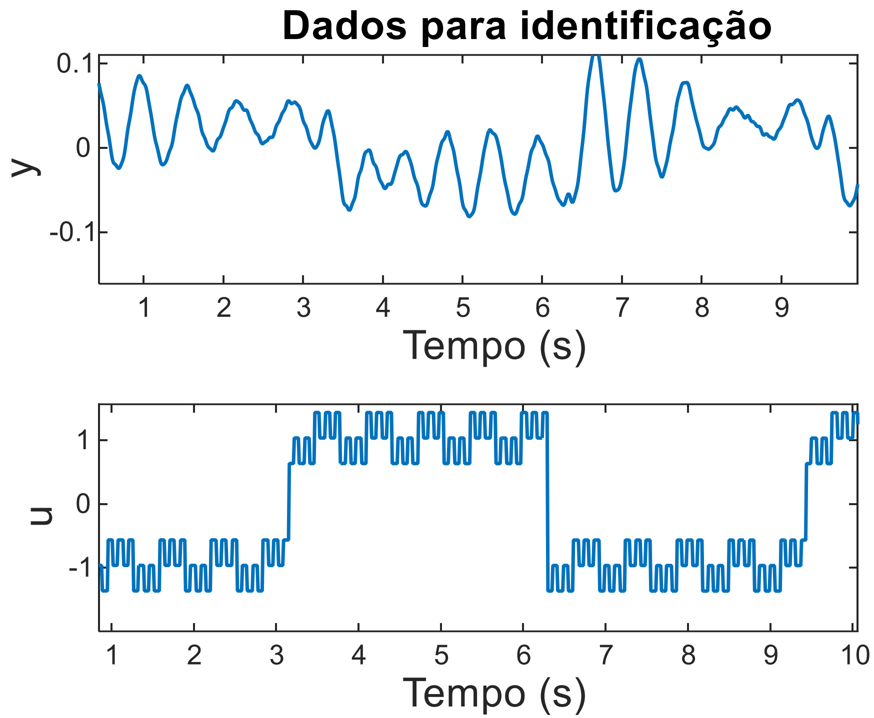

Aumentando (bastante!) a ordem do modelo consegue capturar-se também o 2º modo.

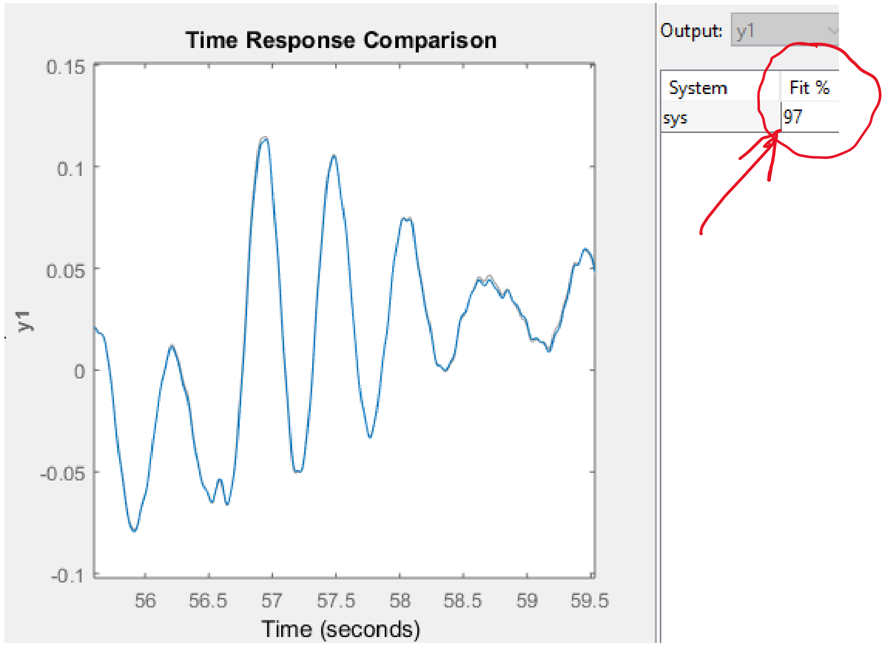
## Estimação de parâmetros do modelo de estado

Uma outra possibilidade é usar modelos de estado, em tempo contínuo ou discreto. *System Identification Toolbox* (MATLAB), função *ssest* (algoritmo PEM com modelo de estado).

A função *ssest* permite identificar uma grande variedade de modelos.

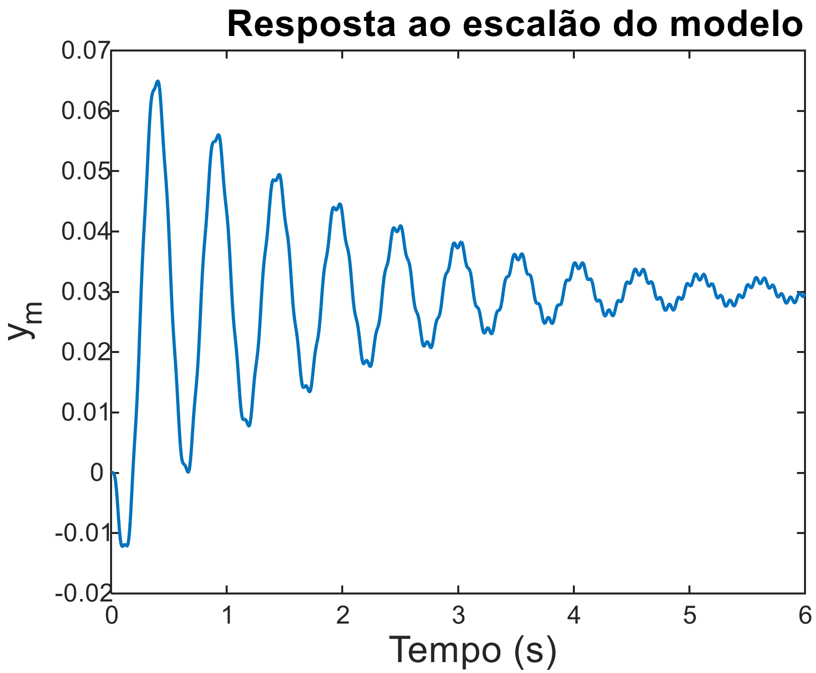
Exemplo com dados simulados, semelhante à barra flexível.

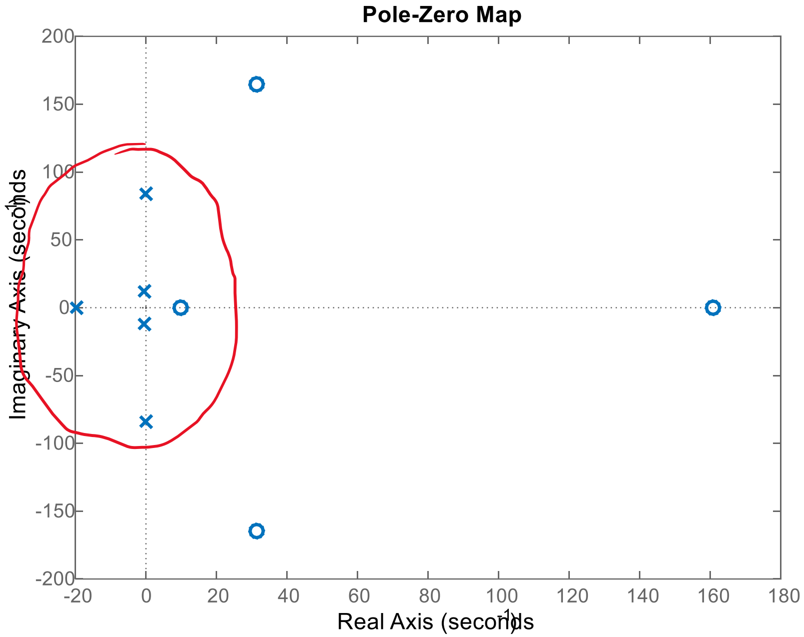



Ordem assumida para o modelo de estado n=5

# Comprtamento do modelo



# Escolha da ordem

Os valores singulares de Hankel traduzem a energia dos modos associados aos diferentes estados. Considerar ordens mais elevadas aumenta a complexidade do modelo mas não aumenta significativamente a sua capacidade de replicar a saída do processo.

## Aspetos práticos importantes da identificação

- É essencial ter bons dados. Atenção às condições experimentais:
  - Sinal de entrada (não muito rápido e não muito lento, na banda em que o sistema responde
  - não muito fraco para não ser "limpo" pelas bandas mortas, e não muito forte para não excitar não linearidades)
- Filtrar o ruído de alta frequência;
- Trabalhar com sinais de média nula. Remover as tendências com a função *detrend*