

Cálculo Numérico

Notas de aula - Ajuste pelo Método dos Mínimos Quadrados

Prof. Yuri Dumaresq Sobral

Departamento de Matemática
Universidade de Brasília

2025

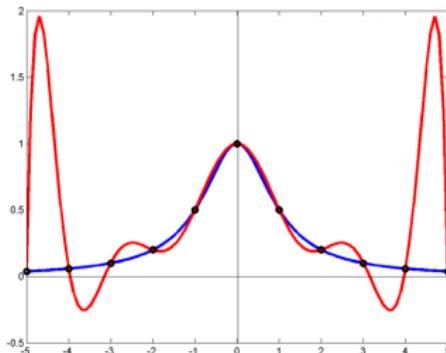
Polinômio interpolador

- Já estudamos o **polinômio interpolador** de um conjunto de pontos. Sua propriedade principal é **passar por todos os pontos** do conjunto!
- Vimos algumas maneiras de **construí-los**. Podemos:
 - Resolver um **sistema linear** com uma matriz de Vandermonde;
 - Podemos calcular **polinômios de Lagrange**;
 - Podemos usar uma **família de polinômios triangulares**.
- **Atenção!** O **polinômio interpolador** é **único**!
- Mas nem tudo são maravilhas... O **erro** cometido por uma aproximação por **polinômio interpolador** $p(t)$ é dado por

$$|f(t) - p(t)| \leq \frac{f^{(n+1)}(\xi(t))}{(n+1)!} \prod_{i=1}^{n+1} (t - t_i), \quad \xi(t) \in (a, b).$$

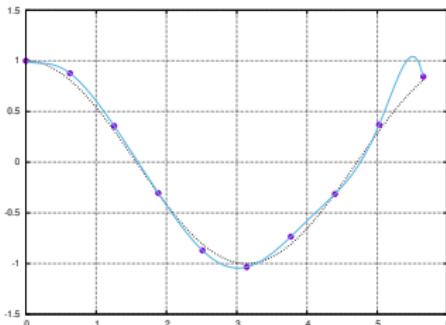
Os termos em **vermelho** podem ser perigosos! Podem **crescer** com $n \rightarrow \infty$. Tanto na **interpolação** quanto na **extrapolação**!

- Um exemplo clássico deste tipo de problema é o Fenômeno de Runge:



- Neste caso, quantos **mais pontos** utilizarmos, **pior** será a aproximação obtida! Este é um problema causado por usarmos **pontos equidistantes** e pelas **derivadas** da função ficarem **ilimitadas**...
- Há uma **cura**: melhorar o **espacamento** entre os pontos! **Nós de Chebychev**. (Para um curso de Cálculo Numérico 2?)
- Além disto... Note que temos um **problema** intrínseco ao **polinômio interpolador**: **NÃO PODEMOS ESCOLHER SEU GRAU!**

- O grau do polinômio interpolador é dado pelo tamanho do conjunto de dados! Isto é grave! Nunca ouvimos falar de uma teoria dada por um polinômio de grau 29, mas podemos facilmente fazer 30 medições em um experimento!
- Vimos no Fenômeno de Runge, por exemplo, que aumentar o número de pontos, isto é, aumentar o grau do polinômio interpolador, trás muitas oscilações! Aproximações ruins!
- O mesmo acontece com ruídos (perturbações) dos dados! Exemplo: uma perturbação de 9% nos dados de $f(x) = \cos(x)$ causa o seguinte efeito no polinômio interpolador:

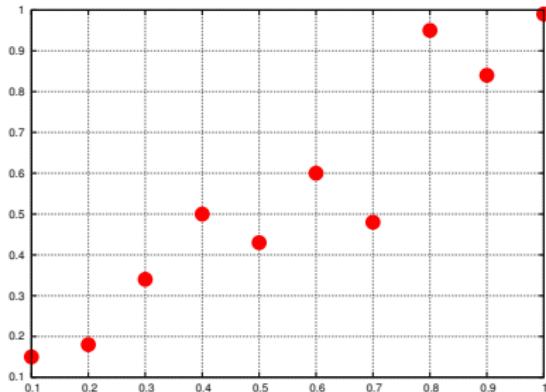


- CONCLUSÃO: Precisamos encontrar OUTRA maneira de encontrar aproximações para (grandes) conjuntos de dados!

Aproximação de funções

Suponha que tenhamos um conjunto de dados (**pares ordenados**) obtidos de um determinado experimento.

t_i	y_i
t_1	y_1
t_2	y_2
t_3	y_3
\vdots	\vdots
t_{n+1}	y_{n+1}



Será que podemos encontrar uma função que represente bem (agora de uma outra maneira) este conjunto de $n + 1$ pontos?

Vamos supor que saibamos, *a priori*, por exemplo, por meio de uma teoria, que y deve variar linearmente com t , isto é:

$$y(t) = \alpha + \beta t$$

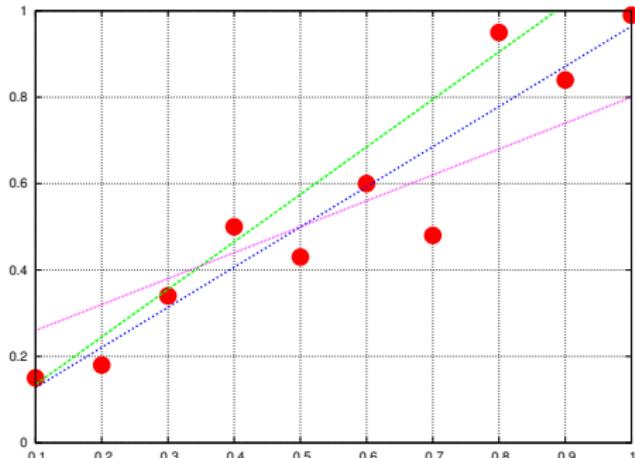
Porém, vemos claramente que os pontos não representam uma reta! Há flutuações nos valores observados!

Devemos pensar em duas possibilidades (**nada a ver com Cálculo Numérico!**):

- ① Problemas no experimento: leitura dos dados imprecisa, impossibilidade de isolar o fenômeno desejado, etc.
- ② A teoria que prevê o comportamento linear de y com t é uma aproximação.

Ajuste de curvas

A pergunta que nos faremos aqui neste curso é: será que conseguimos encontrar os α e β que melhor representam o conjunto de dados disponíveis? Veja que esta é uma pergunta relevante!



A olho nu, podemos pensar em várias possibilidades de retas que se ajustam bem aos pontos do gráfico...

Ajuste de curvas

Se conhecêssemos os valores de α e β , poderíamos calcular, para cada valor de t_i , quanto deveria ser $y_i^{\text{teórico}} = \alpha + \beta t_i$ e, assim, poderíamos determinar o **erro** em cada uma das medições:

$$e_i = y_i - y_i^{\text{teórico}} = y_i - (\alpha + \beta t_i).$$

Se fizéssemos isto para todos os n pontos disponíveis na tabela, obteríamos:

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} y_1 - (\alpha + \beta t_1) \\ y_2 - (\alpha + \beta t_2) \\ y_3 - (\alpha + \beta t_3) \\ \vdots \\ y_n - (\alpha + \beta t_n) \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{b}} - \underbrace{\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{\mathbf{x}}.$$

Portanto,

$$\mathbf{e} = \mathbf{b} - A\mathbf{x}.$$

Idealmente, gostaríamos de encontrar α e β tal que **cada uma** das n medidas tivesse o **menor** erro possível. Porém, isto claramente não é possível, pois só temos 2 variáveis contra n medições...

O que podemos fazer, porém, é minimizar uma quantidade **global**: A norma do vetor \mathbf{e} !

$$\mathcal{E} = |\mathbf{e}| = \sqrt{\mathbf{e}^T \mathbf{e}} = \sqrt{e_1^2 + e_2^2 + e_3^2 + \cdots + e_n^2} = |\mathbf{b} - A\mathbf{x}|$$

Portanto, queremos encontrar o valor de \mathbf{x} tal que o erro global \mathcal{E} seja **mínimo**.

Ajuste de curvas

Note que, se conseguirmos encontrar \mathbf{x} tal que

$$A\mathbf{x} = \mathbf{b},$$

resolvemos o problema exatamente, com erro global $\mathcal{E} = 0$!

Porém... temos um sistema linear de n equações e 2 incógnitas!

$$A_{n \times 2} \mathbf{x}_{2 \times 1} = \mathbf{b}_{n \times 1}.$$

Sistemas deste tipo nem sempre tem soluções: são chamados de sistemas **sobredeterminados**. Podemos resolvê-los exatamente apenas em alguns casos especiais...

Se não pudermos resolvê-lo **exatamente**, será que podemos resolvê-lo **da melhor maneira possível**?

Ajuste de curvas

Vamos voltar ao cálculo do erro global:

$$\mathcal{E}^2 = |\mathbf{b} - A\mathbf{x}|^2 = (\mathbf{b} - A\mathbf{x})^T(\mathbf{b} - A\mathbf{x}) = (\mathbf{b}^T - \mathbf{x}^T A^T)(\mathbf{b} - A\mathbf{x}) = \dots$$

$$\dots = \mathbf{b}^T \mathbf{b} - \mathbf{b}^T A \mathbf{x} - \mathbf{x}^T A^T \mathbf{b} + \mathbf{x}^T A^T A \mathbf{x} = \mathbf{b}^T \mathbf{b} - 2\mathbf{x}^T A^T \mathbf{b} + \mathbf{x}^T A^T A \mathbf{x}$$

pois $\mathbf{x}^T A^T \mathbf{b} = (\mathbf{b}^T A \mathbf{x})^T = \mathbf{b}^T A \mathbf{x}$, já que o resultado é um escalar, uma matriz 1×1 .

$$\mathcal{E}^2 = \underbrace{\mathbf{x}^T A^T A \mathbf{x} - 2\mathbf{x}^T A^T \mathbf{b} + \mathbf{b}^T \mathbf{b}}_{\text{forma quadrática}}.$$

Uma **forma quadrática** é uma generalização da função escalar quadrática:

Ajuste de curvas

$$f(x) = ax^2 + bx + c, \quad x \in I\!\!R$$

que descreve uma parábola! Sabemos que se $a > 0$, então esta parábola terá um mínimo.

Se generalizarmos esta função para várias variáveis, temos:

$$f(x_1, x_2, x_3, \dots, x_n) = \sum_{i,j} \kappa_{ij} x_i x_j - 2 \sum_i \ell_i x_i + c,$$

com κ_{ij} , ℓ_i e c definindo os coeficientes dos termos quadráticos, lineares e constante, respectivamente. Por construção, podemos tomar como simétricos, isto é, $\kappa_{ij} = \kappa_{ji}$.

Em notação matricial:

$$f = \mathbf{x}^T K \mathbf{x} - 2 \mathbf{x}^T \boldsymbol{\ell} + c.$$

Para sabermos se esta função admite um mínimo (global), precisamos garantir que

$$\mathbf{x}^T K \mathbf{x} > 0, \quad \forall \mathbf{x} \in I\!\!R^n.$$

Definição: Dizemos que uma matriz K é **positiva definida** se ela for **simétrica** e se $\mathbf{x}^T K \mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0} \in I\!\!R^n$.

Se tivermos que $\mathbf{x}^T K \mathbf{x} \geq 0, \forall \mathbf{x} \in I\!\!R^n$, dizemos que K é **semi-positiva definida**.

Se K for positiva definida, então K é inversível.

ATENÇÃO: K ser positiva definida **não** quer dizer que K tenha todos seus elementos $\kappa_{ij} > 0$.

Ajuste de curvas

Mas o que isto tem a ver com nosso problema?

Queremos minimizar a forma quadrática de \mathcal{E} em termos de α e β .
O seguinte resultado vai nos dizer como!

Teorema: Seja uma forma quadrática $f = \mathbf{x}^T K \mathbf{x} - 2\mathbf{x}^T \ell + c$.

Então, se K for simétrica positiva definida, f tem **um único**
mínimo dado por $\mathbf{x}^* = K^{-1} \ell$.

De fato, seja $K \mathbf{x}^* = \ell$. Então:

$$f = \mathbf{x}^T K \mathbf{x} - 2\mathbf{x}^T \ell + c = \mathbf{x}^T K \mathbf{x} - 2\mathbf{x}^T K \mathbf{x}^* + c = \dots$$

$$\dots = \mathbf{x}^T K \mathbf{x} - \mathbf{x}^T K \mathbf{x}^* - (\mathbf{x}^T K \mathbf{x}^*)^T + c = \mathbf{x}^T K (\mathbf{x} - \mathbf{x}^*) - \mathbf{x}^{*T} K^T \mathbf{x} + c = \dots$$

$$\dots = \mathbf{x}^T K (\mathbf{x} - \mathbf{x}^*) - \mathbf{x}^{*T} K^T (\mathbf{x} - \mathbf{x}^*) - \mathbf{x}^{*T} K \mathbf{x}^* + c = \dots$$

$$\dots = \underbrace{(\mathbf{x} - \mathbf{x}^*)^T K (\mathbf{x} - \mathbf{x}^*)}_{\text{termo 1}} - \underbrace{\mathbf{x}^{*T} K \mathbf{x}^*}_{\text{termo 2}} + c$$

Ajuste de curvas

Como K é simétrica positiva definida, então o termo 1 é sempre estritamente positivo e só será zero, isto é, **mínimo**, quando $\mathbf{x} - \mathbf{x}^* = \mathbf{0}$, isto é, quando $\mathbf{x} = \mathbf{x}^*$.

O termo 2 não depende de \mathbf{x} e, portanto, não afeta o mínimo de f .

Detalhes: Se K for só semi-positiva definida, então o mínimo não será único (poderemos somar qualquer vetor no kernel de K a \mathbf{x}^*). Para todos os outros casos de K , \mathbf{f} não adimitirá um mínimo global.

Então, voltando ao nosso problema, para minimizarmos \mathcal{E} ,

$$\mathcal{E}^2 = \mathbf{x}^T \textcolor{green}{A^T A} \mathbf{x} - 2\mathbf{x}^T \textcolor{orange}{A^T b} + \mathbf{b}^T \mathbf{b},$$

precisamos saber se a matriz $\textcolor{red}{A^T A}$ se é positiva definida.

Ajuste de curvas

Note que:

$$\mathbf{x}^T A^T A \mathbf{x} = (\mathbf{Ax})^T (\mathbf{Ax}) = |\mathbf{Ax}|^2 \geq 0 \quad \forall \mathbf{x} \in I\!\!R$$

Além disto, note que, por construção, a matriz $K = A^T A$ é sempre simétrica.

Portanto, $K = A^T A$ sempre será pelo menos **semi positiva definida**. Se o $\ker(A) = \{\mathbf{0}\}$, então $K = A^T A$ é positiva definida.

A matriz $K = A^T A$ é chamada de **MATRIZ DE GRAM**.

Então, a solução que minimiza o erro global \mathcal{E} é

$$\mathbf{x}^* = (A^T A)^{-1} A^T \mathbf{b}.$$

Conclusão: Para encontrarmos o α e β que melhor se ajustam ao conjunto de pontos, precisamos resolver o sistema $Ax = b$ da melhor forma, isto é, precisamos encontrar a solução de

$$A^T A x = A^T b \quad \Leftrightarrow \quad x^* = (A^T A)^{-1} A^T b.$$

A solução deste sistema (regularizado) é tal que o erro quadrático global é **mínimo**.

Este é o **Método dos Mínimos Quadrados!**

Vamos ver o que acontece no caso específico que apresentamos no início desta aula:

Mínimos Quadrados

$$A\mathbf{x} = \mathbf{b} \Leftrightarrow \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ t_1 & t_2 & t_3 & \cdots & t_n \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ t_1 & t_2 & t_3 & \cdots & t_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} 1 + 1 + 1 + \cdots + 1 & t_1 + t_2 + t_3 + \cdots + t_n \\ t_1 + t_2 + t_3 + \cdots + t_n & t_1^2 + t_2^2 + t_3^2 + \cdots + t_n^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} =$$

Mínimos Quadrados

$$= \begin{bmatrix} y_1 + y_2 + y_3 + \cdots + y_n \\ t_1 y_1 + t_2 y_2 + t_3 y_3 + \cdots + t_n y_n \end{bmatrix} \Leftrightarrow \begin{bmatrix} n & n\bar{t} \\ n\bar{t} & n\bar{t}^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} n\bar{y} \\ n\bar{t}\bar{y} \end{bmatrix}$$

onde:

$$\bar{t} = \frac{\sum_{i=1}^n t_i}{n} \quad \bar{t}^2 = \frac{\sum_{i=1}^n t_i^2}{n}$$
$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad \bar{t}\bar{y} = \frac{\sum_{i=1}^n t_i y_i}{n}$$

são os valores médios das quantidades envolvidas. Resolvendo o sistema explicitamente, encontramos:

$$\alpha = \bar{y} - \bar{t}\beta \quad \text{e} \quad \beta = \frac{\bar{t}\bar{y} - \bar{y}\bar{t}}{\bar{t}^2 - (\bar{t})^2}$$

Mínimos Quadrados

Exemplo: Ajuste por uma reta $y = \alpha + \beta t$ os dados abaixo:

t	0	1	3	6
y	2	3	7	12

Queremos resolver da melhor maneira possível o sistema:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 7 \\ 12 \end{bmatrix}$$

Então, devemos regularizar o sistema multiplicando-o pela transposta da matriz A :

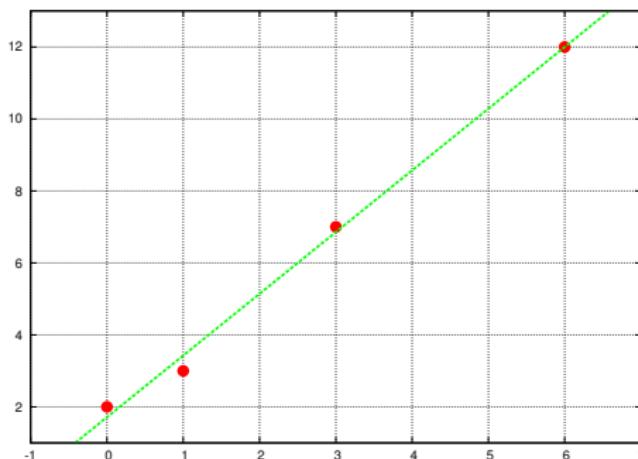
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 7 \\ 12 \end{bmatrix}$$

Mínimos Quadrados

Então, obtemos um sistema regularizado:

$$\begin{bmatrix} 4 & 10 \\ 10 & 46 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 24 \\ 96 \end{bmatrix} \Leftrightarrow \alpha = \beta = \frac{12}{7}$$

Assim, temos que $y = \frac{12}{7}t + \frac{12}{7}$.



Atenção: A matriz de Gram $K = A^T A$ **será inversível** desde que as linhas da matriz A sejam **linearmente independentes**, ou seja, se as **medidas/pontos** forem determinados em valores de t distintos (o que é bastante natural).

Generalização: Claramente podemos generalizar o **MÉTODO DOS MÍNIMOS QUADRADOS** para **polinômios de qualquer grau** ou mesmo para qualquer conjunto de funções $\{h_1(t), h_2(t), h_3(t), \dots, h_m(t)\}$.

O procedimento será sempre o mesmo: para encontrarmos o melhor **melhor** ajuste, para cada um dos n pontos (dados) disponíveis, (t_i, y_i) , faremos:

Mínimos Quadrados

$$y_1 = \alpha_1 h_1(t_1) + \alpha_2 h_2(t_1) + \alpha_3 h_3(t_1) + \cdots + \alpha_m h_m(t_1)$$

$$y_2 = \alpha_1 h_1(t_2) + \alpha_2 h_2(t_2) + \alpha_3 h_3(t_2) + \cdots + \alpha_m h_m(t_2)$$

$$y_3 = \alpha_1 h_1(t_3) + \alpha_2 h_2(t_3) + \alpha_3 h_3(t_3) + \cdots + \alpha_m h_m(t_3)$$

 \vdots \vdots \vdots \vdots \vdots

$$y_n = \alpha_1 h_1(t_n) + \alpha_2 h_2(t_n) + \alpha_3 h_3(t_n) + \cdots + \alpha_m h_m(t_n)$$

Com estas equações, construímos o sistema $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} h_1(t_1) & h_2(t_1) & \cdots & h_m(t_1) \\ h_1(t_2) & h_2(t_2) & \cdots & h_m(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ h_1(t_n) & h_2(t_n) & \cdots & h_m(t_n) \end{bmatrix}_{n \times m} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}_{m \times 1}$$

Mínimos Quadrados

Pelo MÉTODO DOS MÍNIMOS QUADRADOS, encontraremos os melhores valores de $\alpha_1, \dots, \alpha_m$ resolvendo o sistema regularizado:

$$A^T A \mathbf{x} = A^T \mathbf{b} \Leftrightarrow K \mathbf{x} = A^T \mathbf{b} \Leftrightarrow$$

$$\left[\begin{array}{cccc} \sum_{k=1}^n h_1(t_k)h_1(t_k) & \sum_{k=1}^n h_1(t_k)h_2(t_k) & \cdots & \sum_{k=1}^n h_1(t_k)h_m(t_k) \\ \sum_{k=1}^n h_2(t_k)h_1(t_k) & \sum_{k=1}^n h_2(t_k)h_2(t_k) & \cdots & \sum_{k=1}^n h_2(t_k)h_m(t_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^n h_m(t_k)h_1(t_k) & \sum_{k=1}^n h_m(t_k)h_2(t_k) & \cdots & \sum_{k=1}^n h_m(t_k)h_m(t_k) \end{array} \right] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = \left[\begin{array}{cccc} \sum_{k=1}^n h_1(t_k)y_k & \sum_{k=1}^n h_2(t_k)y_k & \cdots & \sum_{k=1}^n h_m(t_k)y_k \end{array} \right]^T$$

PROBLEMA! É possível que, dependendo das funções h_i , as linhas da matriz A deixem de ser **linearmente independentes**, mesmo os dados tendo sido obtidos em diferentes t_i .

Exemplo: $y(t) = \alpha \sin(t) + \beta \cos(t)$

$$A = \begin{bmatrix} \sin(t_1) & \cos(t_1) \\ \sin(t_2) & \cos(t_2) \\ \vdots & \vdots \\ \sin(t_n) & \cos(t_n) \end{bmatrix}$$

É possível que para alguns t_i , t_j , $\sin(t_i) = \sin(t_j)$ e $\cos(t_i) = \cos(t_j)$!

Este é um problema complicado e deve ser analisado caso a caso com cuidado.

Na grande maioria dos casos, combinações **lineares** simples com algumas poucas funções são suficientes para excelentes resultados. Porém, é possível **generalizar** o método que construímos para combinações **não-lineares** de funções!

Generalização: os coeficientes a serem determinados pelo **MÉTODO DOS MÍNIMOS QUADRADOS** podem aparecer, a priori, em diferentes locais na combinação linear que descreve as observações:

$$y_i = \alpha_1 h_1(\beta_1 t_i) + \alpha_2 h_2(\beta_2 t_i) + \alpha_3 h_3(\beta_3 t_i) + \cdots + \alpha_m h_m(\beta_m t_i)$$

Neste caso, porém, o sistema a ser resolvido para determinar os α_i e os β_i é **não-linear** e **sobredeterminado**: factível, mas **muito complicado**! (Também para um curso de Cálculo Numérico 2?)