Mínimos Quadrados

Sistemas Lineares Sobredeterminados

Nelson Kuhl

IME/USP

3 de setembro de 2020

Exemplo

Considere o seguinte problema. Um corpo foi lançado verticalmente e a sua altura foi medida em determinados instantes, gerando a seguinte tabela [1]:

Tempo (s)	0.0	0.1	0.2	0.3	0.4	0.5
Altura (m)	1.67203	1.79792	2.37791	2.66408	2.11245	2.43969

...	0.6	0.7	0.8	0.9	1.0
...	1.88843	1.59447	1.79634	1.0781	0.21066

A partir destes dados, deseja-se estimar a altura e a velocidade iniciais, bem como a aceleração da gravidade.

[1] Dados retirados de DeVries, *A First Course in Computational Physics*.

Exemplo

Desprezando-se a resistência do ar, sabe-se da cinemática que a altura $y(t)$ como função do tempo t é descrita pela expressão

$$y(t) = a_0 + a_1 t + a_2 t^2 \tag{1}$$

onde a_0 e a_1 são as altura e velovidade iniciais, respectivamente, e

$$a_2 = -\frac{g}{2}$$

onde g é a aceleração da gravidade. Para determinar os coeficientes a partir dos dados, basta então resolver as equações $y(t_i) = h_i$, $1 \leq i \leq 11$, ou

$$a_0 + a_1 t_i + a_2 t_i^2 = h_i, \quad 1 \leq i \leq 11, \tag{2}$$

onde t_i e h_i são os instantes de tempo e as alturas, respectivamente, dados na tabela.

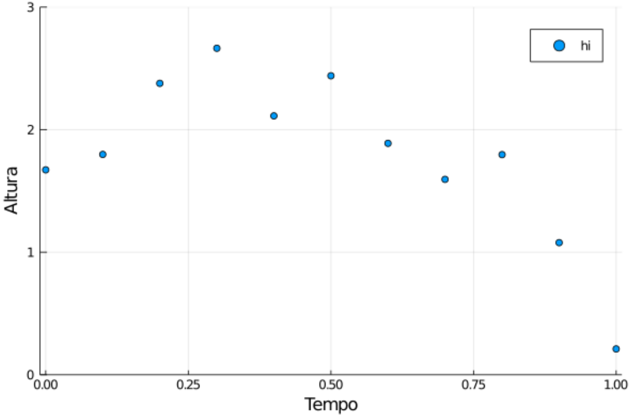
Exemplo

Note que há uma sobredeterminação nas equações (2) pois temos 11 equações para a determinação das 3 incógnitas a_0, a_1 e a_2. É possível resolver o problema? Observe que:

- A equação (1) descreve a altura em função do tempo como uma parábola;
- resolver (2) significa obtermos os coeficientes de uma parábola que passa pelos 11 pontos da tabela.

Exemplo

Vejamos os dados da tabela em um gráfico:



Mesmo sem fazer as contas, é evidente que não existe uma parábola que passa por estes pontos. O que podemos fazer então?

Exemplo

Esta incompatibilidade dos dados com o modelo é comum em problemas de identificação de parâmetros e pode ser devida a erros de medida e à imperfeição do modelo. Entretanto, é possível obter informações relevantes.

Exemplo

Se observarmos atentamente as equações (2), vemos que elas formam um sistema linear sobredeterminado com 11 equações e 3 incógnitas, que pode ser representado matricialmente como

$$\begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{11} & t_{11}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ \vdots \\ h_{11} \end{bmatrix}. \tag{3}$$

Sistemas lineares sobredeterminados em geral não têm solução, mas é possível generalizar a noção de solução.

Sistemas Lineares Sobredeterminados

O exemplo que acabamos de ver é um caso particular de problemas de identificação de parâmetros, os quais em muitas situações são formulados como sistemas lineares sobredeterminados.

Definição

Um sistema linear sobredeterminado com m equações e n incógnitas, onde

$$m > n,$$

é uma equação da forma

$$Ax = b$$

onde são dados a matriz $m \times n$ dos coeficientes A e o vetor $b \in \mathbb{R}^m$, e deseja-se obter $x \in \mathbb{R}^n$ que satisfaz a equação acima.

Sistemas Lineares Sobredeterminados

Sistemas Lineares Sobredeterminados

- A matriz A pode ser associada a uma transformação linear entre os espaços $\mathbb{R}^n$ e $\mathbb{R}^m$;

Sistemas Lineares Sobredeterminados

- A matriz A pode ser associada a uma transformação linear entre os espaços $\mathbb{R}^n$ e $\mathbb{R}^m$;
- logo, a sua imagem tem dimensão no máximo n e é portanto um subespaço vetorial próprio de $\mathbb{R}^m$, pois $m > n$;

Sistemas Lineares Sobredeterminados

- A matriz A pode ser associada a uma transformação linear entre os espaços $\mathbb{R}^n$ e $\mathbb{R}^m$;
- logo, a sua imagem tem dimensão no máximo n e é portanto um subespaço vetorial próprio de $\mathbb{R}^m$, pois $m > n$;
- como $b \in \mathbb{R}^m$, este vetor não estará em geral na imagem de A;

Sistemas Lineares Sobredeterminados

- A matriz A pode ser associada a uma transformação linear entre os espaços $\mathbb{R}^n$ e $\mathbb{R}^m$;
- logo, a sua imagem tem dimensão no máximo n e é portanto um subespaço vetorial próprio de $\mathbb{R}^m$, pois $m > n$;
- como $b \in \mathbb{R}^m$, este vetor não estará em geral na imagem de A;
- logo, em geral, não existirá $x \in \mathbb{R}^n$ tal que $Ax = b$. Esta equação terá solução somente no caso especial em que b pertencer à imagem de A.

Sistemas Lineares Sobredeterminados

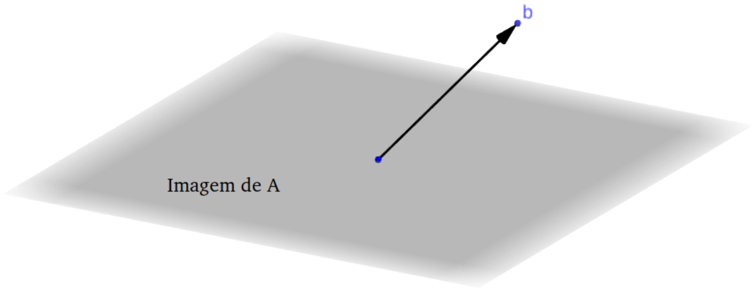


Figura: Caso em que b não pertence à imagem de A.

Na figura acima o plano cinza representa a imagem de A. Para termos solução do sistema sobredeterminado, b deveria de pertencer a este espaço.

Sistemas Lineares Sobredeterminados

Uma maneira de razoável de lidar esta dificuldade é determinar o vetor da imagem de A mais próximo de b.

Nova Formulação

Determinar $\bar{x} \in \mathbb{R}^n$ tal que

$$\|b - A\bar{x}\| \leq \|b - Ax\|, \quad \text{para todo } x \in \mathbb{R}^n,$$

onde

$$\|b - Ax\| = \sqrt{\sum_{i=1}^{m} [b_i - (Ax)_i]^2} \tag{4}$$

é a distância Euclideana entre b e Ax.

Sistemas Lineares Sobredeterminados

Sistemas Lineares Sobredeterminados

- Geometricamente, esta nova formulação significa determinar $\bar{x}$ tal que $A\bar{x}$ **é igual à projeção ortogonal de** b **sobre a imagem de** A;

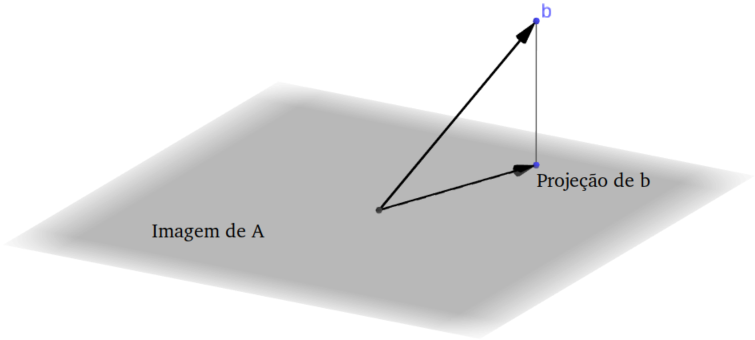
Sistemas Lineares Sobredeterminados

- Geometricamente, esta nova formulação significa determinar $\bar{x}$ tal que $A\bar{x}$ **é igual à projeção ortogonal de** b **sobre a imagem de** A;
- este problema **sempre tem solução** e, se as colunas de A forem vetores linearmente independentes do $\mathbb{R}^m$, então a solução é única;

Sistemas Lineares Sobredeterminados

- Geometricamente, esta nova formulação significa determinar $\bar{x}$ tal que $A\bar{x}$ **é igual à projeção ortogonal de** b **sobre a imagem de** A;
- este problema **sempre tem solução** e, se as colunas de A forem vetores linearmente independentes do $\mathbb{R}^m$, então a solução é única;
- se b pertencer à imagem de A, então a solução desta nova formulação coincide com a solução no sentido usual do sistema linear.

Sistemas Lineares Sobredeterminados



A projeção ortogonal de b na imagem de A é o vetor deste subespaço mais próximo de b.

Sistemas Lineares Sobredeterminados

Esta nova formulação estende o conceito de solução de um sistema linear. Ela é chamada de solução no sentido de mínimos quadrados, pois ela minimiza o **erro quadrático** definido por (4). Resta determinar como calculá-la. Para tanto, note que:

se A é uma matriz $m \times n$ e $x \in \mathbb{R}^n$, então

$$
\begin{aligned}
Ax &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\
&= \sum_{j=1}^{n} x_j a^{(j)},
\end{aligned}
$$

onde $a^{(j)}$ é o j-ésimo vetor coluna de A ($a_i^{(j)} = a_{ij}$, $1 \leq i \leq m$).

Sistemas Lineares Sobredeterminados

Isto nos lembra que os vetores coluna de A geram a sua imagem. Portanto, determinar $\bar{x}$ tal que $A\bar{x}$ é igual à projeção ortogonal de b sobre a imagem de A equivale a determinar $\bar{x}$ tal que $b - A\bar{x}$ **é ortogonal a** $a^{(j)}$, $1 \leq j \leq n$.

Ou seja, $\bar{x}$ deve satisfazer as n equações

$$(a^{(j)})^T(b - Ax) = 0, \quad 1 \leq j \leq n,$$

onde o superescrito T denota o transposto. Desenvolvendo a expressão acima obtemos o sistema linear (exercício)

$$A^TAx = A^Tb \tag{5}$$

conhecido como **Sistema Normal**.

Sistemas Lineares Sobredeterminados

Sistemas Lineares Sobredeterminados

- O sitema normal é um sistema linear de n equações para as n incógnitas x_i, $1 \leq i \leq n$;

Sistemas Lineares Sobredeterminados

- O sitema normal é um sistema linear de n equações para as n incógnitas x_i, $1 \leq i \leq n$;
- o sistema normal *sempre* tem solução. Se os vetores coluna de A forem linearmente independentes, esta solução é única. Consideraremos somente este caso;

Sistemas Lineares Sobredeterminados

- O sitema normal é um sistema linear de n equações para as n incógnitas x_i, $1 \leq i \leq n$;
- o sistema normal *sempre* tem solução. Se os vetores coluna de A forem linearmente independentes, esta solução é única. Consideraremos somente este caso;
- a matriz $A^T A$ do sistema normal é simétrica definida positiva.

Solução do Exemplo

Solução do Exemplo

Voltemos agora ao exemplo. Da expressão (3) obtemos

$$A^T A = \begin{bmatrix} 11 & \sum_{i=1}^{11} t_i & \sum_{i=1}^{11} t_i^2 \\ \sum_{i=1}^{11} t_i & \sum_{i=1}^{11} t_i^2 & \sum_{i=1}^{11} t_i^3 \\ \sum_{i=1}^{11} t_i^2 & \sum_{i=1}^{11} t_i^3 & \sum_{i=1}^{11} t_i^4 \end{bmatrix} \text{ e } A^T b = \begin{bmatrix} \sum_{i=1}^{11} h_i \\ \sum_{i=1}^{11} t_i h_i \\ \sum_{i=1}^{11} t_i^2 h_i \end{bmatrix}.$$

Solução do Exemplo

Voltemos agora ao exemplo. Da expressão (3) obtemos

$$A^T A = \begin{bmatrix} 11 & \sum_{i=1}^{11} t_i & \sum_{i=1}^{11} t_i^2 \\ \sum_{i=1}^{11} t_i & \sum_{i=1}^{11} t_i^2 & \sum_{i=1}^{11} t_i^3 \\ \sum_{i=1}^{11} t_i^2 & \sum_{i=1}^{11} t_i^3 & \sum_{i=1}^{11} t_i^4 \end{bmatrix} \text{ e } A^T b = \begin{bmatrix} \sum_{i=1}^{11} h_i \\ \sum_{i=1}^{11} t_i h_i \\ \sum_{i=1}^{11} t_i^2 h_i \end{bmatrix}.$$

Usando os dados da tabela, concluimos que a_0, a_1 e a_2 satisfazem

$$\begin{bmatrix} 11.0 & 5.5 & 3.85 \\ 5.5 & 3.85 & 3.025 \\ 3.85 & 3.025 & 2.5333 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 19.63208 \\ 8.386632 \\ 4.995481 \end{bmatrix},$$

cuja solução é $a_0 = 1.65415$, $a_1 = 3.90262$ e $a_2 = -5.20209$. A altura e a velocidade iniciais são dadas por a_0 e a_1, respectivamente, e de a_2 obtemos que a aceleração da gravidade é $g = 10.40418\text{m}/\text{s}^2$.

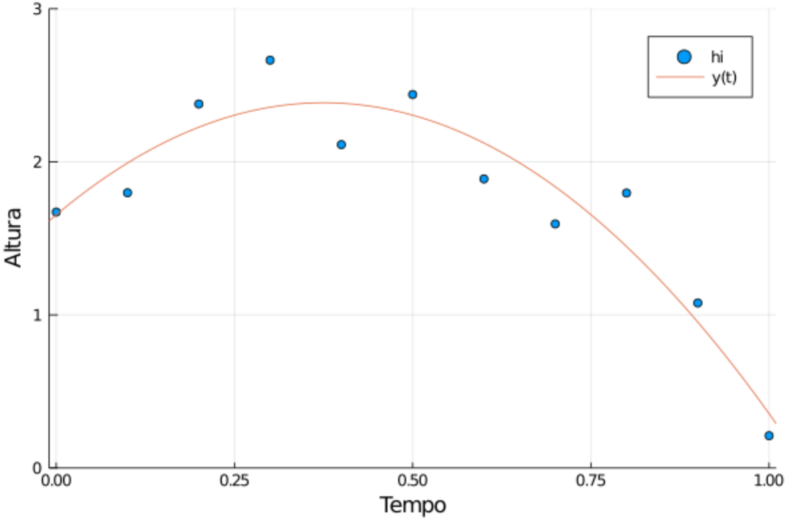


Figura: Pontos e parábola obtida