Modelação Numérica 2017

Aula 22, 16/Maio

- Estimativa de parâmetros e optimização.
- Solução de problemas lineares sub- e sobre-determinados.

http://modnum.ucs.ciencias.ulisboa.pt

Métodos de optimização lineares vs não lineares

- Os métodos de otimização descritos anteriormente (annealing, annealing modificado, genético) permitem resolver qualquer problema de optimização, linear ou não, desde que seja conhecida a função de custo.
- São, no entanto, métodos iterativos que nunca garantem obter a solução correspondente ao mínimo global da função de custo. No caso do annealing modificado (monte carlo downslope in a cooling bath) ele converge para o método de Newton clássico, o que implica que ele converge sempre pelo menos para um mínimo local.
- Quando o modelo a ajustar é linear, é possível obter soluções por um método direto (i.e. não iterativo), o que é em geral mais satisfatório (não é preciso estabelecer critérios de paragem, nem gerar números aleatórios). Existe uma formulação matemática rigorosa e completa por trás das inversões lineares.

Inversões

- Os problemas de inversão podem escrever-se sempre na forma:

$$\vec{o} = M(\vec{p}) + \varepsilon$$

Vector das observações

Modelo (matriz que relaciona as observações com os parâmetros)

Vector dos parâmetros

Erro

Note-se que, de forma geral, o vector das observações e dos parâmetros têm dimensões diferentes.

Inversões lineares

- Os problemas de inversão linear escrevem-se então:

$$\vec{o} = G\vec{p} + \varepsilon$$

Erro

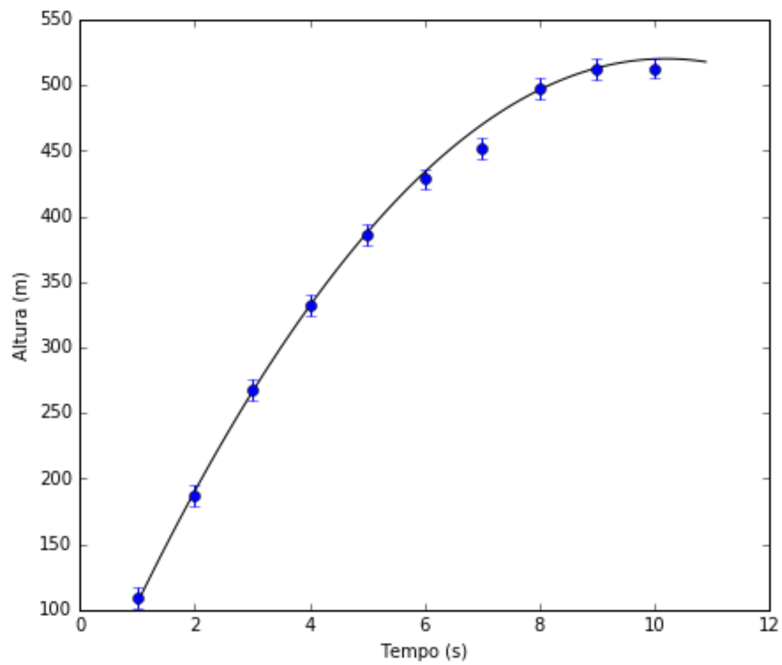
Vector das observações

Modelo (matriz que relaciona as observações com os parâmetros)

Vector dos parâmetros

Exemplo

- Regressão linear. Movimento de um projéctil.



$$\vec{o} = G\vec{p} + \varepsilon$$

Vector das observações

Modelo (matriz que relaciona as observações com os parâmetros)

Vector dos parâmetros

Erro

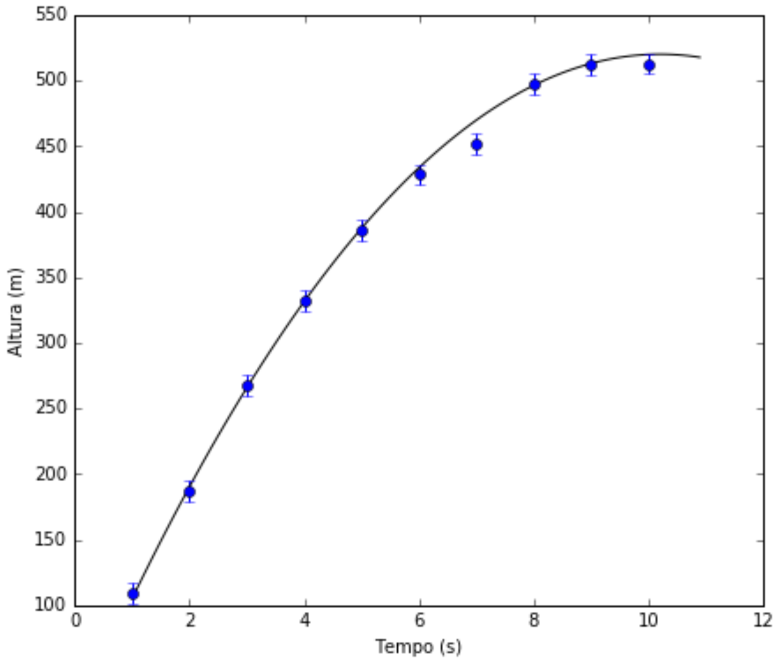
$$y(t) = a + bt - \frac{1}{2}ct^2$$

Vector dos parâmetros

Vector das observações

$$o = p_1 + p_2 t - \frac{1}{2}p_3 t^2$$

Exemplo



Vector das observações

Modelo (matriz que relaciona as observações com os parâmetros)

Vector dos parâmetros

$$\vec{o} = G\vec{p}$$

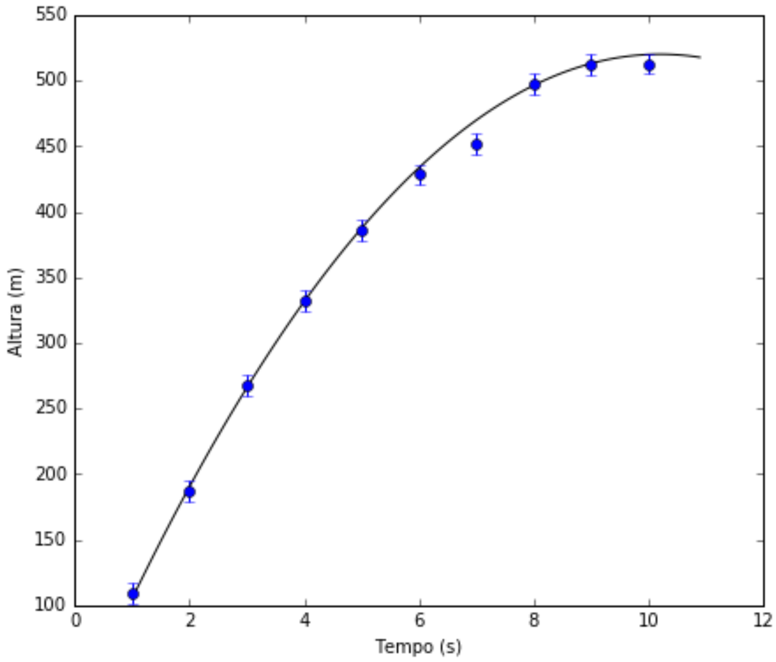
$$y(t) = a + bt - \frac{1}{2}ct^2$$

$$o = p_1 + p_2 t - \frac{1}{2}p_3 t^2$$

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = \quad G \quad \vec{p}$$

Exemplo



n observações
(n=10)

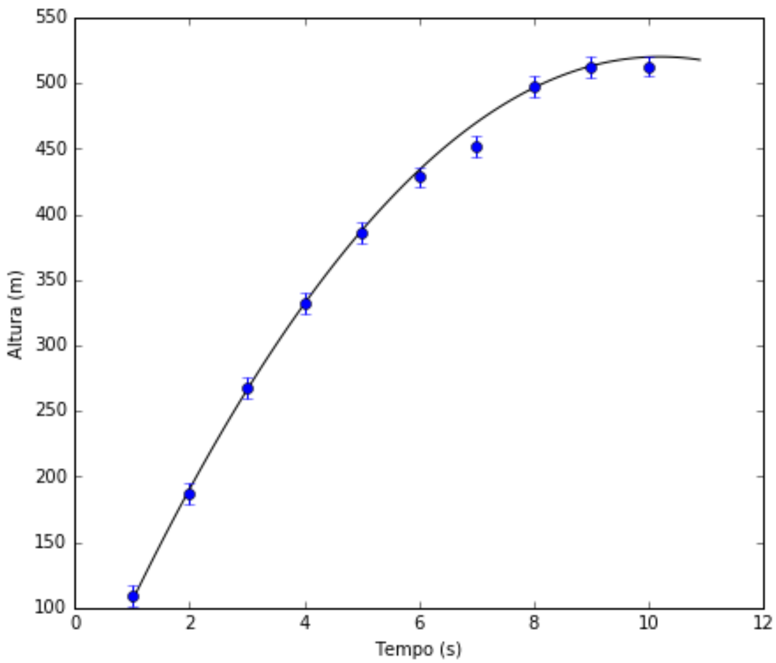
m parâmetros
a estimar
(m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = \quad G \quad \vec{p}$$

Exemplo

- Se m=n, e
- Se a matriz *G* não for singular:
- Então a solução é trivial – consiste no cálculo da matriz inversa: $\vec{p} = G^{-1}\vec{o}$



n observações (n=10)

m parâmetros a estimar (m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = \quad G \quad \vec{p}$$

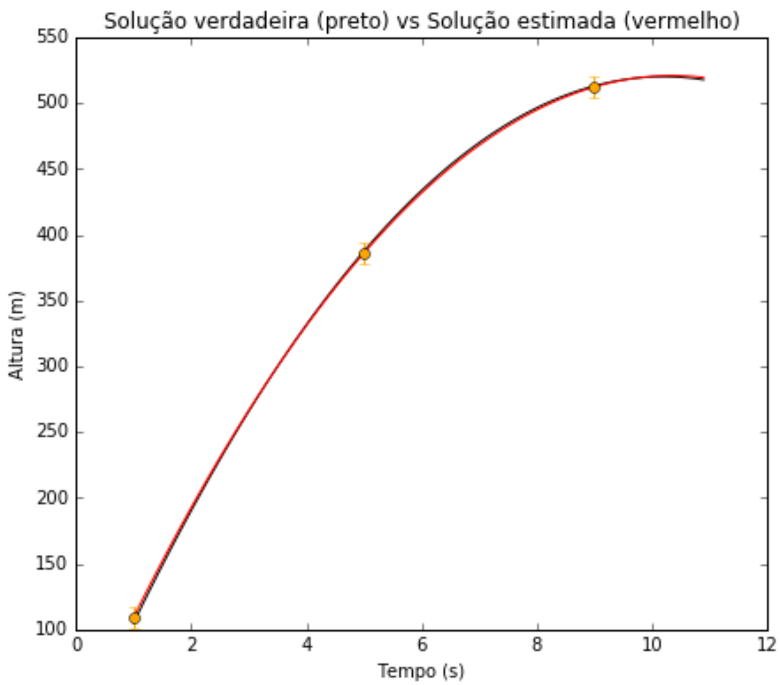
```python
import matplotlib.pyplot as plt
import numpy as np

from numpy.linalg import pinv as pinv

#%%
plt.rcParams['figure.figsize'] = 10, 6

################## Problema directo

ts = np.arange(1., 11.,.1)     # vector de tempo para as observações sintéticas

# parâmetros do problema
p1=10.
p2=100.
p3=9.8

os = p1 + p2*ts - .5*p3*ts**2       # vector de tempo para as observações sintéticas

# plot
plt.close()
plt.rcParams['figure.figsize'] = 7,6

plt.plot(ts,os)
plt.ylabel('Altura (m)')
plt.xlabel('Tempo (s)');
```

```python
#%% ################## Inicialização, caso m=2

t=np.array([  1.,    5.,    9.])      # vector dos tempos em que se fizeram as observações
o=np.array([109.4, 386.1, 512.3])      # vector das observações
erro = 8.*np.ones(3)       # vector dos erros

nobs=len(t)          # nr de observações
npar=3               # nr de parâmetros a estimar

# plot
plt.close()
plt.rcParams['figure.figsize'] = 7,6

plt.plot(ts,os)
plt.errorbar(t,o,yerr=erro, fmt='o')
plt.ylabel('Altura (m)')
plt.xlabel('Tempo (s)');

#%% matriz  G

G=np.zeros([nobs,npar])
G[:,0]=1
G[:,1]=t
G[:,2]=-.5*t**2


#%% solução, caso m=n

Xest = np.dot(np.linalg.inv(G), o)     # solução

pe1=Xest[0]
pe2=Xest[1]
pe3=Xest[2]

Oest = pe1 + pe2*ts - .5*pe3*ts**2
```

```python
# plot
plt.close()
plt.rcParams['figure.figsize'] = 7,6

plt.plot(ts,os, 'k', ts,Oest, 'r')
plt.errorbar(t,o,yerr=erro, fmt='o', color='orange')
plt.ylabel('Altura (m)')
plt.xlabel('Tempo (s)');
plt.title(u'Solução verdadeira (preto) vs Solução estimada (vermelho)');
```

Exemplo

- Se m=n, e
- Se a matriz *G* não for singular:
- Então a solução é trivial – consiste no cálculo da matriz inversa: $\vec{p} = G^{-1}\vec{o}$



n observações (n=3)

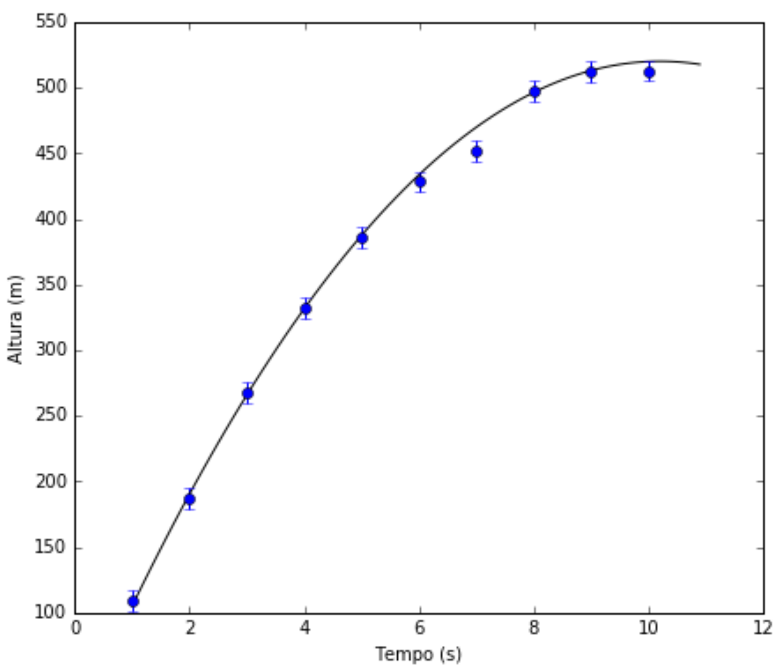
m parâmetros a estimar (m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = \quad G \quad \vec{p}$$

Se m≠n

- Se m>n (mais parâmetros a estimar do que observações):
  - Problema sub-determinado.
  - Existem múltiplas soluções exatas compatíveis com as observações, e com erro nulo.
- Se m>n (menos parâmetros a estimar do que observações):
  - Problema sobre-determinado .
  - É posível estimar o erro da solução.



n observações (n=10)

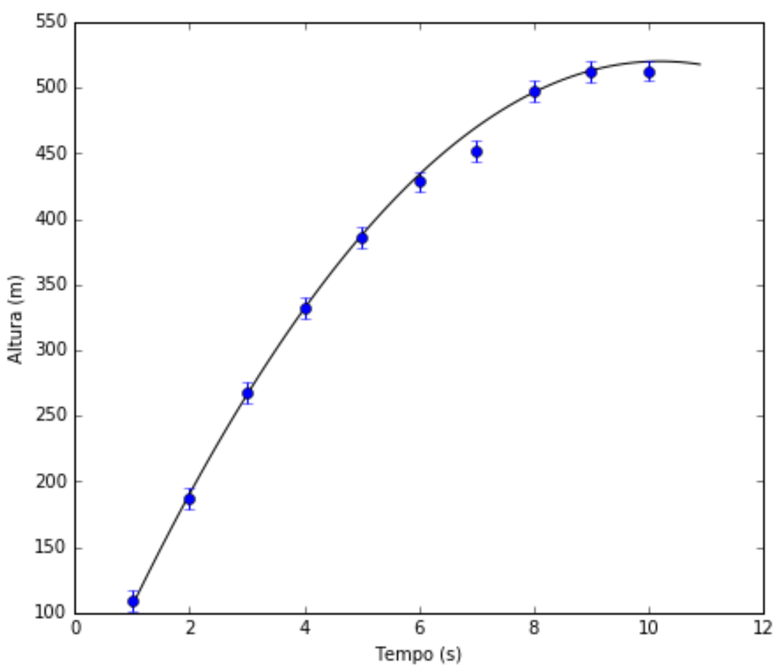
m parâmetros a estimar (m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = \quad G \quad \vec{p}$$

Se m≠n

- Sendo *G(n,m)* é possível determinar univocamente uma matriz *H(m,n)*, tal que: *p = Ho*
- Corresponde à solução com menor norma L2 de entre todas as soluções disponíveis.
- Numericamente, esse cálculo faz-se de forma imediata utilizando a matriz peuso-inverse de Moore-Penrose (np.linalg.pinv).



n observações (n=10)

m parâmetros a estimar (m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

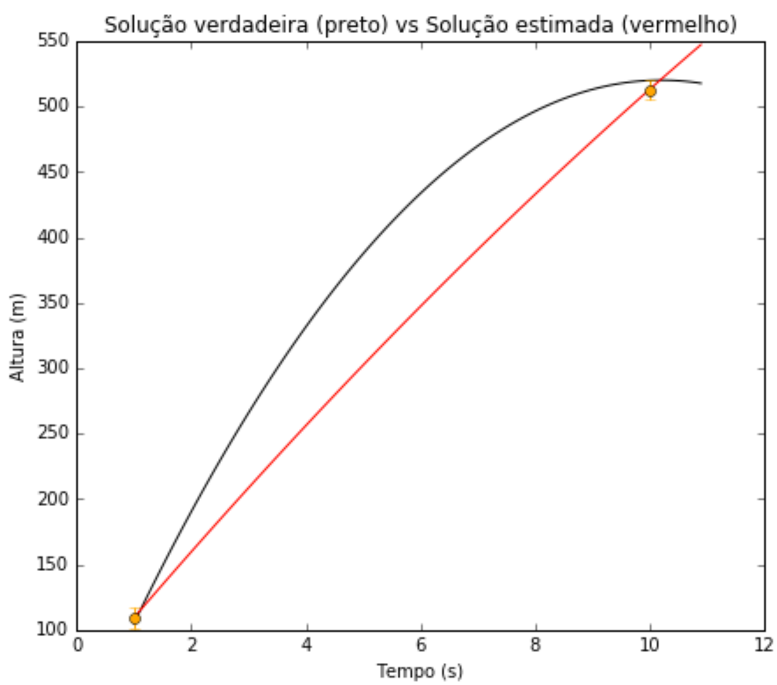
$$\vec{o} = \quad G \quad \vec{p}$$

```python
#%% ################## Inicialização, caso sub ou sobre-determinado

# caso sobre-determinado
t=np.arange(1., 11.)     # vector dos tempos em que se fizeram as observações
o=np.array([109.4, 187.5, 267.5, 331.9, 386.1, 428.4, 452.2, 498.1, 512.3, 513.])     # vector das
erro = 8.*np.ones(10)       # vector dos erros

### caso sub-determinado
#t=np.array([1., 10.])     # vector dos tempos em que se fizeram as observações
#o=np.array([109.4, 513.])     # vector das observações
#erro = 8.*np.ones(2)       # vector dos erros

nobs=len(t)          # nr de observações
npar=3               # nr de parâmetros a estimar

# plot
plt.close()
plt.rcParams['figure.figsize'] = 7,6

plt.plot(ts,os)
plt.errorbar(t,o,yerr=erro, fmt='o')
plt.ylabel('Altura (m)')
plt.xlabel('Tempo (s)');

#%% matriz  G

G=np.zeros([nobs,npar])
G[:,0]=1
G[:,1]=t
G[:,2]=-.5*t**2
```

```python
#%% solução geral

Xest=np.dot(pinv(G),o)     # solução

pe1=Xest[0]
pe2=Xest[1]
pe3=Xest[2]

Oest = pe1 + pe2*ts - .5*pe3*ts**2

# plot
plt.close()
plt.rcParams['figure.figsize'] = 7,6

plt.plot(ts,os, 'k', ts,Oest, 'r')
plt.errorbar(t,o,yerr=erro, fmt='o', color='orange')
plt.ylabel('Altura (m)')
plt.xlabel('Tempo (s)');
plt.title(u'Solução verdadeira (preto) vs Solução estimada (vermelho)');
```

Se m≠n

- Se m>n (mais parâmetros a estimar do que observações):
  - Problema sub-determinado.
  - Existem múltiplas soluções exatas compatíveis com as observações, e com erro nulo.



n observações (n=2)

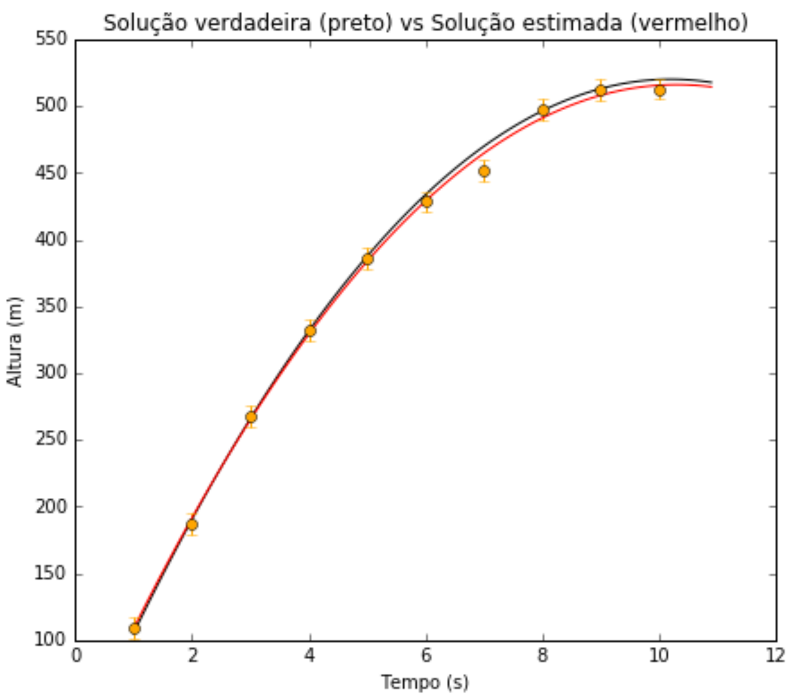
m parâmetros a estimar (m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \dots & \dots & \dots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \dots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = G \vec{p}$$

Se m≠n

- Se m>n (menos parâmetros a estimar do que observações):
  - Problema sobre-determinado .
  - É posível estimar o erro da solução.



n observações (n=10)

m parâmetros a estimar (m=3)

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \ldots \\ o_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & -\frac{1}{2}t_1^2 \\ 1 & t_2 & -\frac{1}{2}t_2^2 \\ 1 & t_3 & -\frac{1}{2}t_3^2 \\ \ldots & \ldots & \ldots \\ 1 & t_n & -\frac{1}{2}t_n^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} p_1 + p_2 t_1 - \frac{1}{2}p_3 t_1^2 \\ p_1 + p_2 t_2 - \frac{1}{2}p_3 t_2^2 \\ p_1 + p_2 t_3 - \frac{1}{2}p_3 t_3^2 \\ \ldots \\ p_1 + p_2 t_n - \frac{1}{2}p_3 t_n^2 \end{bmatrix}$$

$$\vec{o} = G \quad \vec{p}$$