Practice 7
Activities sheet

Activity 1. Compute a "least square solution" of $\mathsf{A}\vec{x} = \vec{b}$ from the normal equations and compute the residual error of the approximation.

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix}.$$

The associated system of normal equations is $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$. We can find a solution of this system using, for instance, the \ command:

```
 -->A=[1 1 0; 1 1 0; 1 0 1; 1 0 1];

-->b=[1; 3; 8; 2];

-->x=(A'*A)\(A'*b)
Advertencia :
la matriz esta cerca de la singularidad o mal escalada. rcond =    0.0000D+00
calculando la solución de mínimos cuadrados. (vea Isq).

 x  =

    5.
  - 3.
    0.
```

Now we compute the error of the approximation:

```
-->norm(A*x-b)
 ans  =

    4.472136
```

Observation: In this case the least square solution is not unique because the matrix A^tA has not maximum rank and, then, the system of normal equations has infinitely many solutions. All these solutions are "least square solutions". Let us check it:

```
-->rank(A'*A)
 ans  =

    2.
```

Activity 2. Compute the equation $y = \beta_0 + \beta_1 x$ of the regression line fitting the points $(2, 3)$, $(3, 2)$, $(5, 1)$ and $(6, 0)$, and compute the norm of the residual vector.

Since, in an "ideal" situation, the given points "should belong" to the regression line, we must consider the system of equations (with unknowns β_0 and β_1) obtained by replacing, in $y = \beta_0 + \beta_1 x$, the variable x by the x-coordinate of each point, and the variable y by the y-coordinate:

$$\beta_0 + 2\beta_1 = 3$$
$$\beta_0 + 3\beta_1 = 2$$
$$\beta_0 + 5\beta_1 = 1$$
$$\beta_0 + 6\beta_1 = 0$$

Its coefficient matrix and its vector of independent terms are (using Scilab):

```
-->A=[1 2; 1 3; 1 5; 1 6];

-->b=[3; 2; 1; 0];
```

Now we "solve" the system $\mathsf{A}\vec{x} = \vec{b}$ using the least square method (because it is inconsistent). To do it, we can solve (as in the previous activity) the system of normal equations or we can apply directly the \ command to the system $\mathsf{A}\vec{x} = \vec{b}$. We use the last method because it is easier:

```
-->x=A\b
 x  =

    4.3
  - 0.7
```

Then, the regression line is the one with equation $y = 4.3 - 0.7x$. The error is:

```
-->norm(A*x-b)
 ans  =

    0.3162278
```

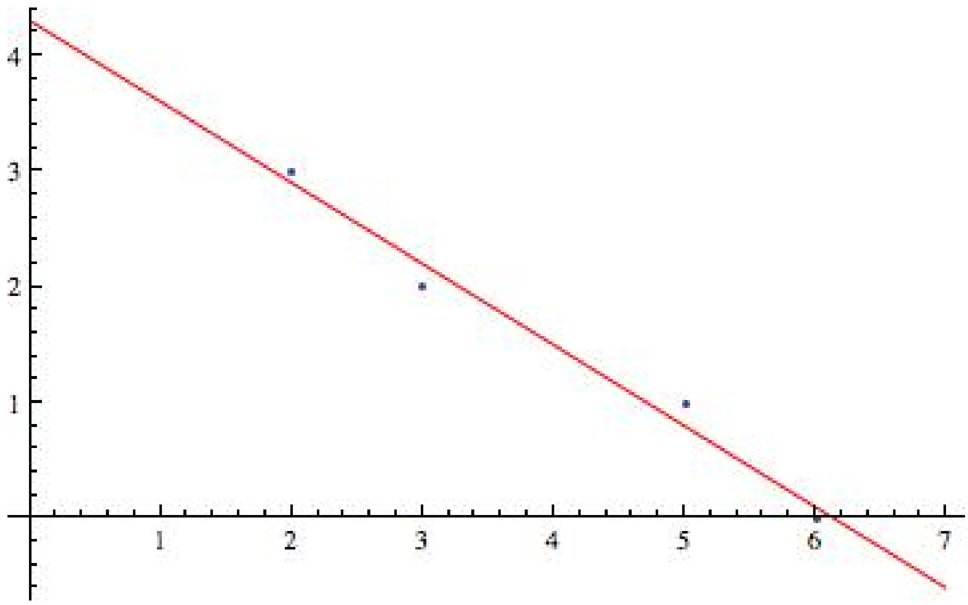
Observation 1: If you prefer it, instead of the \ command you can use also the **lsq** command.

Observation 2: In this case, there "least squares solution" is unique because the system of normal equations $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$ has a unique solution (the matrix $\mathsf{A}^t\mathsf{A}$ has maximum rank):

```
-->A'*A
 ans  =

    4.     16.
```

```
    16.     74.

-->rank(ans)
 ans  =

    2.
```

Notice that this happens because the x-coordinates of all the given points are different.



Activity 3. To measure the performance of the engine of an aircraft during flight, his horizontal position was measured every second from $t = 0$ to $t = 12$. The obtained positions were 0; $8{,}8$; $29{,}9$; $62{,}0$; $104{,}7$; $159{,}1$; $222{,}0$; $294{,}5$; $380{,}4$; $471{,}1$; $571{,}7$; $686{,}8$ i $809{,}2$. Determine the cubic curve $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ fitting the points, obtained by the least squares method. Apply the result to estimate the **speed** of the aircraft when $t = 4.5$ seconds.

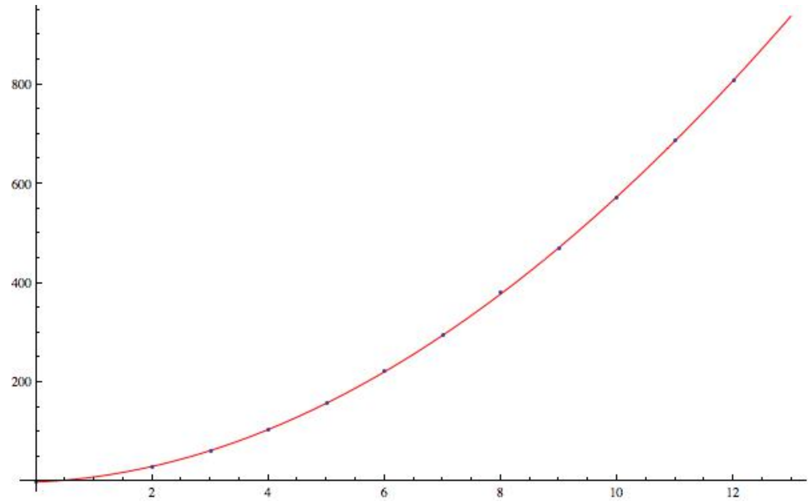
We proceed as in the previous exercise, but now the fitting curve is a cubic curve instead of a line. That is, we must compute the "most approximated" cubic curve $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ to the points $(0, 0)$, (1,8.8)$, (2, 29.9), (3, 62)$ and so on. We obtain, in this way, the equations $0 = \beta_0$, $\beta_0 + \beta_1 + \beta_2 + \beta_3 = 8.8$, $\beta_0 + 2 * \beta_1 + 2^2 \beta_2 + 2^3 \beta_3 = 29.2$, $\beta_0 + 3 * \beta_1 + 3^2 \beta_2 + 3^3 \beta_3 = 62$, and so on .... It is easy to obtain quickly the coefficient matrix and the vector of independent terms of this system using Scilab:

```
-->t=[0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12];

-->A=[ones(13,1) t t^2 t^3]
 A  =

    1.    0.     0.      0.
    1.    1.     1.      1.
    1.    2.     4.      8.
    1.    3.     9.      27.
    1.    4.     16.     64.
    1.    5.     25.     125.
    1.    6.     36.     216.
    1.    7.     49.     343.
    1.    8.     64.     512.
    1.    9.     81.     729.
    1.    10.    100.    1000.
    1.    11.    121.    1331.
    1.    12.    144.    1728.

-->b=[0;8.8;29.9;62;104.7;159.1;222;294.5;380.4;471.1;571.7;686.8;809.2];
```

We have to "solve" by least squares the system $A\vec{x} = \vec{b}$:

```
-->x=A\b
 x  =

  - 0.8557692
    4.702485
    5.5553696
  - 0.0273601
```

Therefore, the fitting cubic curve is $y = -0.8557692 + 4.702485t + 5.5553696t^2 + -0.0273601t^3$.



Since the speed is the derivative of the space with respect to the time, we can estimate the speed when $t = 4.5$ by computing the derivative function of y and evaluating it in $t = 4.5$. This derivative function is

$$y' = 4.702485 + 2 \cdot 5.5553696t - 3 \cdot 0.0273601t^2.$$

Therefore: $y'(4.5) = 53.0386 \;\; m/s$

Activity 4. When the monthly sales of certain product are subject to fluctuations throughout the season, a curve that approximates the sales data could take the form $y = \beta_0 + \beta_1 x + \beta_2 \sin(\pi x/6)$, where x is the time (in months). Determine the least squares curve along 6 months knowing that the fluctuations are: 0.80; 0.66; 0.64; 0.73; 0.78 and 0.67. Compute the norm of the residual vector.

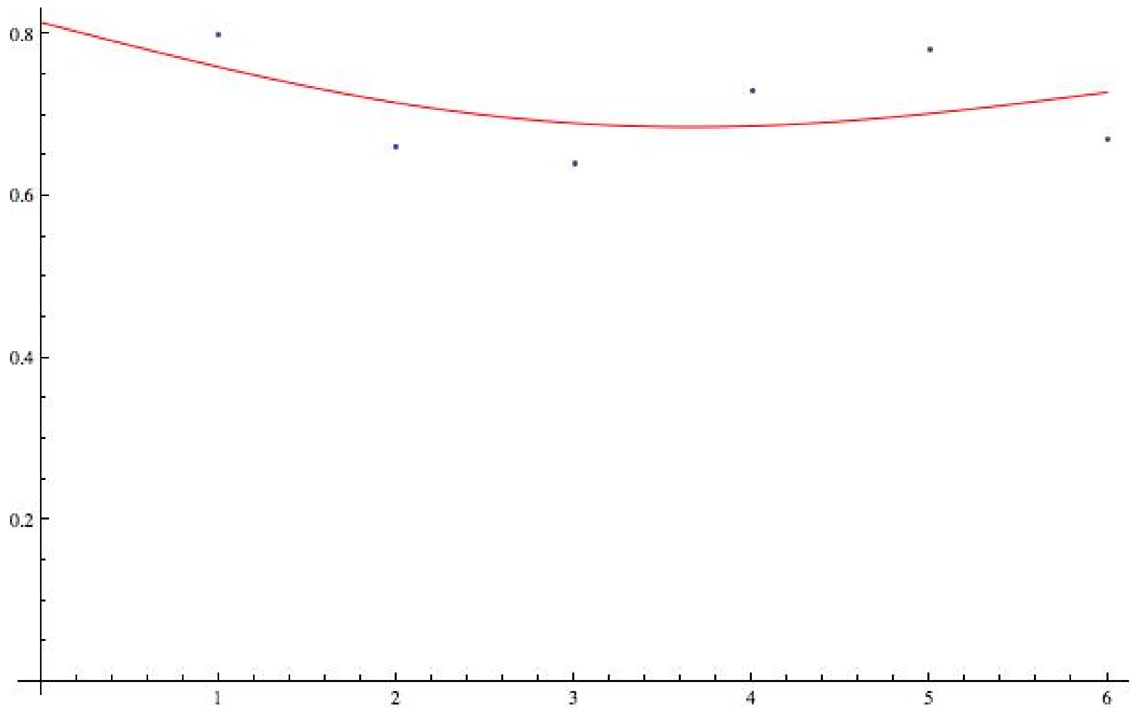
Since the resolution of this activity is totally similar to the previous ones, I think that you need not many explanations here.

```
-->t=[1; 2; 3; 4; 5; 6];

-->A=[ones(6,1) t sin(%pi*t/6)]
 A  =

    1.    1.    0.5
    1.    2.    0.8660254
    1.    3.    1.
    1.    4.    0.8660254
    1.    5.    0.5
    1.    6.    1.225D-16

-->b=[0.8; 0.66; 0.64; 0.73; 0.78; 0.67];

-->x=A\b
 x  =

    0.8144262
  - 0.0144028
  - 0.0814828
```

Therefore, the required curve is $y = 0.8144262 - 0.0144028x - 0.0814828\ \sin(\pi x/6)$.

Finally, we compute the residual error:

```
-->norm(A*x-b)
 ans  =

    0.1362983
```


0.8
0.6
0.4
0.2
1
2
3
4
5
6