Chap 2 : Linear regreession with least square estimate

Prof M.Tognolini HEIG-VD, Feb. 2022

Exercice 2.1 Elipses

Description

We collect a dataset of measurements (for exemple positions) and they are plotted in a 2 axis graph shown bellow :

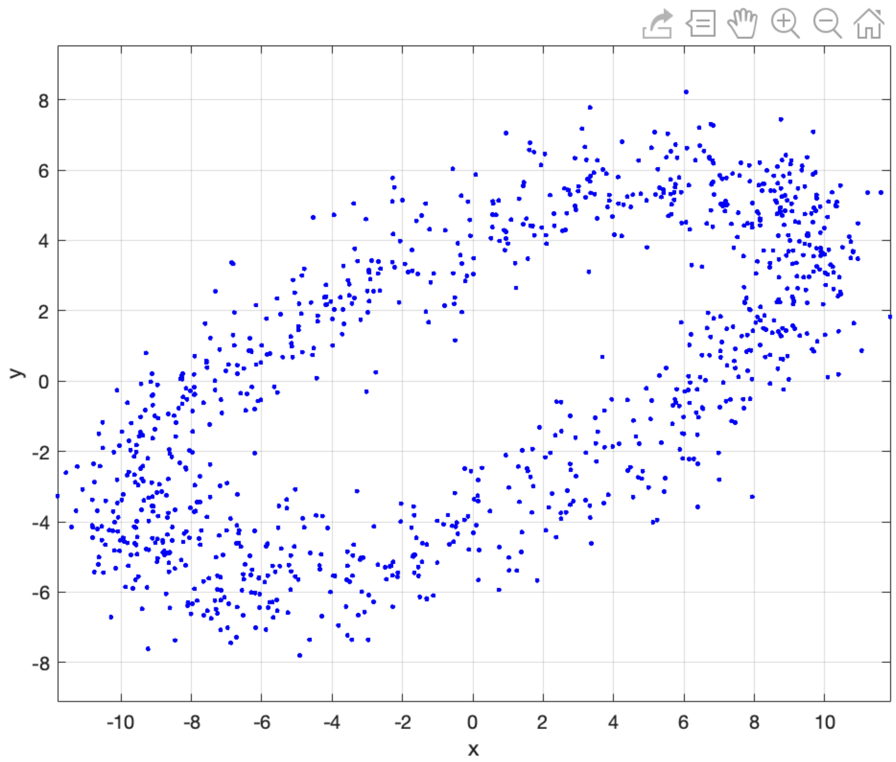


Figure 1

The goal is to find the best Elipse aproximation with the minimum of the L2 norm. and then plot it like in the figure 2:

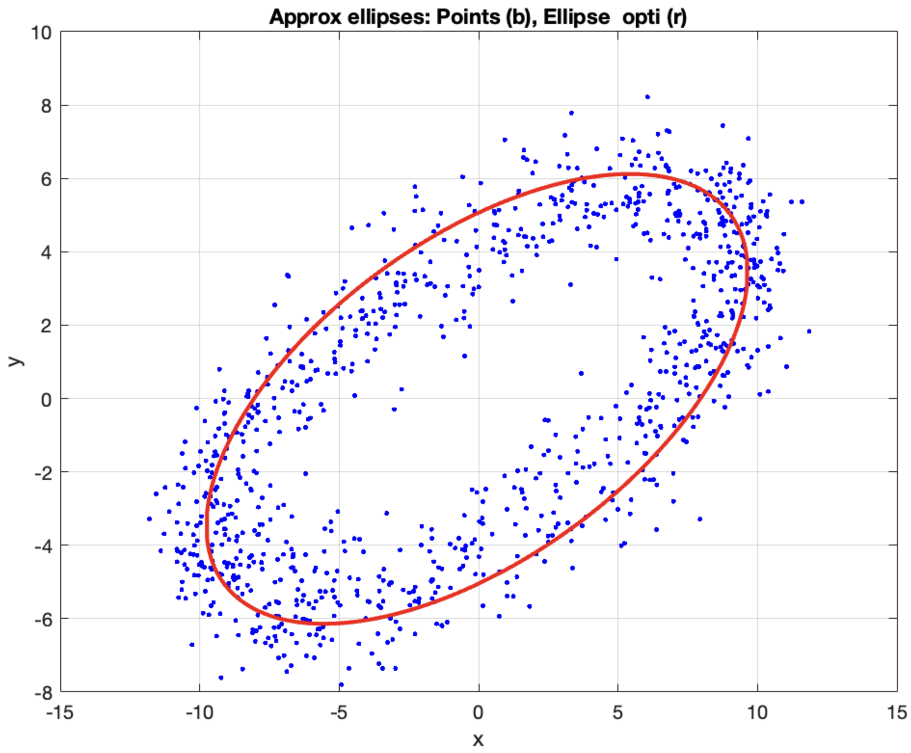


Figure 2

Theory

The eelipses equations are give by :

$$ax^2 + by^2 + cx + dy + exy = 1 \qquad (1)$$

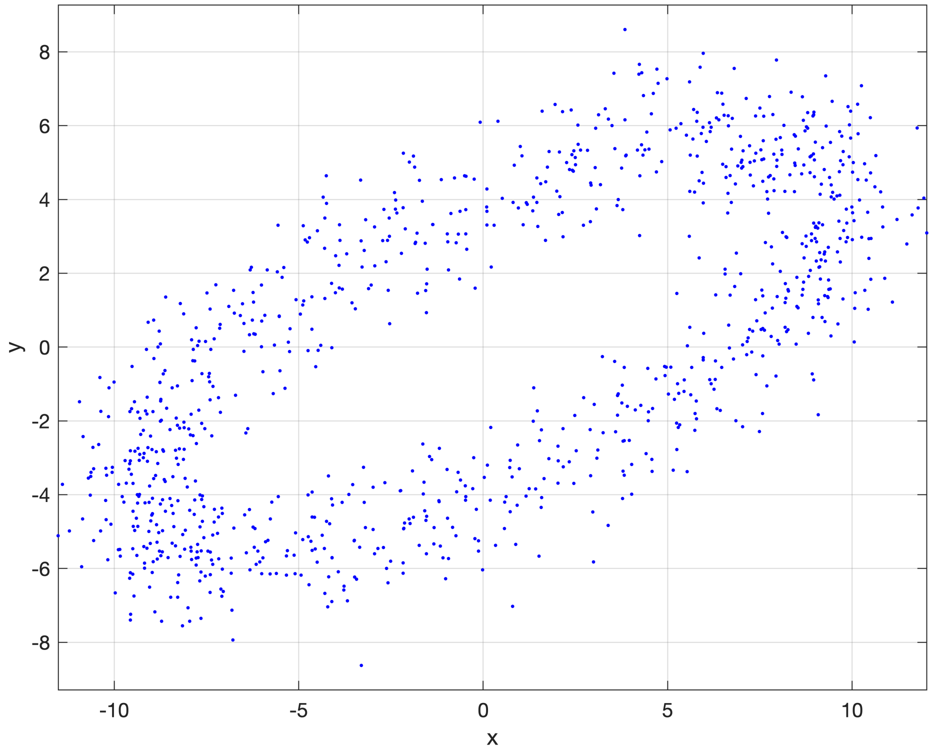
the constants values a,b,c,d,e are the parameters to be found, in this case they are the unknown of the problem to solve. we need only 5 points P(xi,yi) to determine the constants a,b,c,d,e , but here we want to find the best approximation that minimize the L2 norm of the error between the points P(xi,yi) and the best Elipse approximation.

For each point Pi = P(xi,yi) we could write the equation (1), for N points we have the overdetermined system of equations with N equations and 5 unknowns.

Data import and plot

First import the data from the file EllipseData.mat, the points are in the matrix Pi , the column number is the point number i the first row contain the xi values and the second row contans the yi values.

Next plot the points in a graph liike this shown in figure 1



Compute the parameters for the best Elipse

Build the matrix A and the vector b representing the system of N equations like the equation (1):

to find the best parameters Θ we need to solve the over-determined system :

$A\Theta = b$ (2)

It is no solution Θ that satisfies all the conditions (all points) se we could only find an approximation $\hat{\Theta}$, solving eq (2) by pseudo inverse will give $\hat{\Theta}$:

$\hat{\Theta} = (A^T A)^{-1} A^T b$ (3)

Compute the approximation parameters :

```
Thetahat = 5×1
    0.0162
    0.0395
   -0.0019
    0.0003
   -0.0296
```

Compute the best Elipse points

With those parameters it is possible to plot the approximation of the Elipse that fit the measurement points Pi but it is not trivial because y coordinates depends of x coordinates as shown by the Elipse equation (1) so we need compute for each possible x (in the same range than in the dataset) the corresponding y on the best Elipse given by the approximation parameters. From the equation (1) we need to solve $\hat{y} = f(x, \hat{\Theta})$ and this is a second order equation with 2 or 1 or 0 real possible values of y for each x.

Let's re-write the eq (1) like this:

$by^2 + (d + ex)y + (ax^2 + cx - 1) = 0$ we need to exprime y and to simplify the notation we call : $V = d + ex$ and $W = (ax^2 + cx - 1)$ and $Z = b$

the equation would be : $Zy^2 + Vy + W = 0$ so y is find by solving this second order equation:

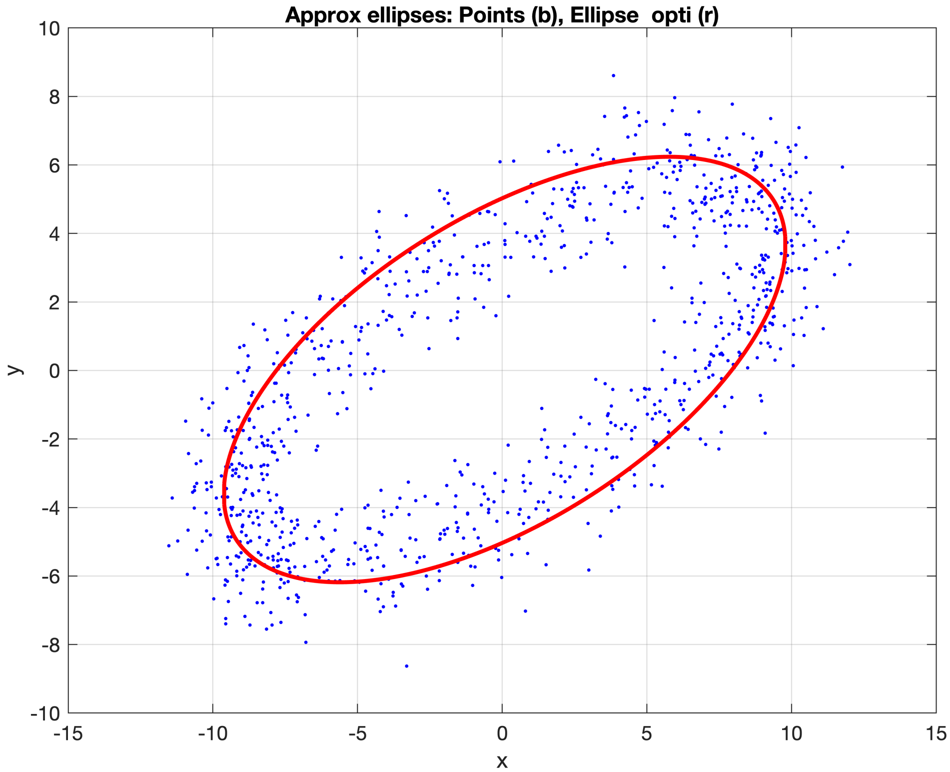
$$y = -V \pm \frac{\sqrt{V^2 - 4ZW}}{2Z} \quad (4)$$

First create a x vectos with the enough points to plot an Elipse trajectory and with the same span than the xi values in the data set:

Create the variables Z, V and W with the data xi and the parameters in the vector $\hat{\Theta}$:

Next compute the $\hat{yi}$ values by solving (4) for positive and negative case (be careful to not take the non-real values):

Plot the best Elipse



Eigen Vectors and eigenvalues application

To find the principal axis of the Elipse we could compute the covariance matrix of the coordinate points :

The Covariance matrix Cp is give by , the size is 2 x 2

$Cp = \frac{1}{N}(Pi^T - m_p)^T(Pi^T - m_p)$ where m_p is the vector of mean valua of the columns of Pi

We obtain the same result with cov() function

```
mp = 1×2
   -0.0037   -0.0265

Cp = 2×2
   46.8203   18.5040
   18.5040   17.1023
```

The covariance matrix id Hermitian (the complex ranspose is equal to himself) that give some particularity , the eigen values and the eigen vectors are real.

Now we compute the eigen vectors and the eigen values :

```
V = 2×2
    0.9017   -0.4324
    0.4324    0.9017
D = 2×2
   55.6929         0
         0    8.2297
```

the eigenvector corresponding to the largest eigenvalue is the bigger axis of the Elipse.

You could notice the two eigenvectors are perpendicular.

The Eigenvector that corresponds to the larges value of the eigenvalue corresponds with the largest axis of the elipse.