

Problemsolving 2

Anton Lindbro

February 15, 2025

0.1

In order to compute the condition number of the matrix A in the norms asked for we need to multiply the norm of the matrix with the norm of the inverse matrix for the respective norms asked for. So we begin by computing the inverse of the matrix A.

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & -1 & 4 \end{pmatrix}$$
$$A^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 3 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

1-norm

$$\|A\|_1 = \max\left(\sum_{j=1}^n |a_{ij}|\right) = \max(6, 2, 6) = 6$$
$$\|A^{-1}\|_1 = \frac{1}{2} \max\left(\sum_{j=1}^n |a_{ij}|\right) = \frac{1}{2} \max(3, 9, 3) = \frac{9}{2}$$
$$\text{cond}(A) = \|A\|_1 \cdot \|A^{-1}\|_1 = 6 \cdot \frac{9}{2} = 27$$

infinity-norm

For this we need the transpose of the matrix A and the inverse of the transpose of A.

$$A^T = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & -1 \\ 1 & 1 & 4 \end{pmatrix}$$

$$(A^T)^{-1} = (A^{-1})^T = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 \\ 3 & 5 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Now we can compute the condition number for the infinity-norm.

$$\|A\|_\infty = \|A^T\|_1 = \max(4, 2, 8) = 8$$

$$\|A^{-1}\|_\infty = \|(A^{-1})^T\|_1 = \frac{1}{2} \max(5, 7, 3) = \frac{7}{2}$$

$$\text{cond}(A) = \|A\|_\infty \cdot \|A^{-1}\|_\infty = 8 \cdot \frac{7}{2} = 28$$

Frobenius-norm

$$\|A\|_F = \sqrt{\sum (a_{ij})^2} = \sqrt{4 + 1 + 1 + 1 + 0 + 1 + 9 + 1 + 16} = \sqrt{34}$$

$$\|A^{-1}\|_F = \frac{1}{2} \sqrt{\sum (a_{ij})^2} = \frac{1}{2} \sqrt{1 + 9 + 1 + 1 + 25 + 1 + 1 + 1 + 1} = \frac{1}{2} \sqrt{41}$$

$$\text{cond}(A) = \|A\|_F \cdot \|A^{-1}\|_F = \sqrt{34} \cdot \frac{1}{2} \sqrt{41} = \frac{1}{2} \sqrt{1394} = 18,7$$

0.2

1

Now we have a dataset and need to do a least squares fit to the data. We can do this by solving the normal equations. In order to get the normal equations we need to set up the matrix A and the vector a. To do this we first need to make an ansatz. We are doing a linear fit so we assume that the data can be described by a linear function. This means that we can write the data as $y = a_1x + a_2$. We can

then write this as a matrix equation $Aa = y$ where $A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$, $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. To get

a good approximation we need to minimize the error $e = Aa - y$ this we can do by $\min(\|Aa - y\|_2)$ finding this minimum is done by solving the normal equations $A^T Aa = A^T y$. In order to do this for our particular data set we need to set up the matrix A and the vector y.

$$A = \begin{pmatrix} 9 & 1 \\ 10 & 1 \\ 11 & 1 \\ 12 & 1 \\ 13 & 1 \end{pmatrix}, y = \begin{pmatrix} 5.0 \\ 5.0 \\ 4.0 \\ 3.0 \\ 1.0 \end{pmatrix}$$

Now we compute $A^T A$ and $A^T y$.

$$A^T A = \begin{pmatrix} 9 & 10 & 11 & 12 & 13 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 9 & 1 \\ 10 & 1 \\ 11 & 1 \\ 12 & 1 \\ 13 & 1 \end{pmatrix} = \begin{pmatrix} 615 & 55 \\ 55 & 5 \end{pmatrix}$$

$$A^T y = \begin{pmatrix} 9 & 10 & 11 & 12 & 13 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5.0 \\ 5.0 \\ 4.0 \\ 3.0 \\ 1.0 \end{pmatrix} = \begin{pmatrix} 188 \\ 18 \end{pmatrix}$$

Now we can solve the normal equations.

$$A^T A a = A^T y$$

$$\begin{pmatrix} 615 & 55 \\ 55 & 5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 188 \\ 18 \end{pmatrix} \Rightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -1.0 \\ 14.6 \end{pmatrix}$$

Plotting the data and the fit we get the following plot.

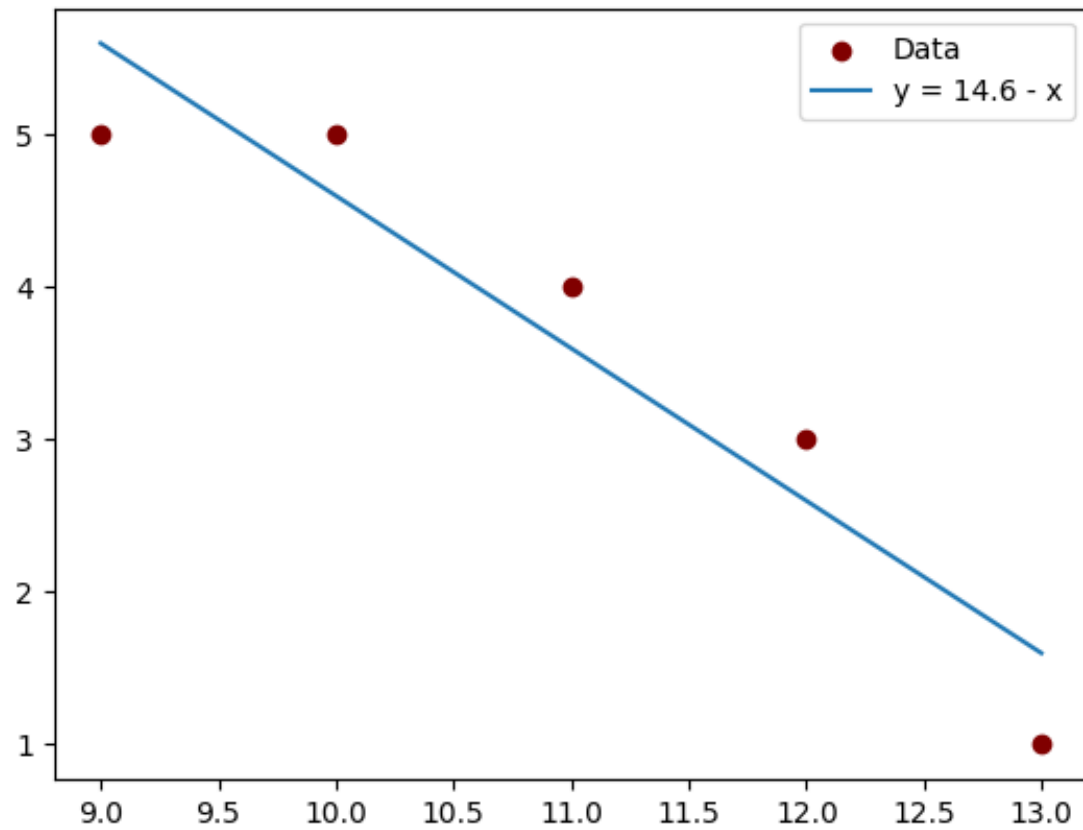


Figure 1: Linear regression of the data.

2

Now we calculate the residual vector

$$e = y - Aa = \begin{pmatrix} 5.0 \\ 5.0 \\ 4.0 \\ 3.0 \\ 1.0 \end{pmatrix} - \begin{pmatrix} 9 & 1 \\ 10 & 1 \\ 11 & 1 \\ 12 & 1 \\ 13 & 1 \end{pmatrix} \begin{pmatrix} -1.0 \\ 14.6 \end{pmatrix} = \begin{pmatrix} -0.6 \\ 0.4 \\ 0.4 \\ 0.4 \\ -0.6 \end{pmatrix}$$

So the components of the residual vector correspond to the distances between the data points and the fit. With this the residual is $\|e\|_2 = 1.1$

3

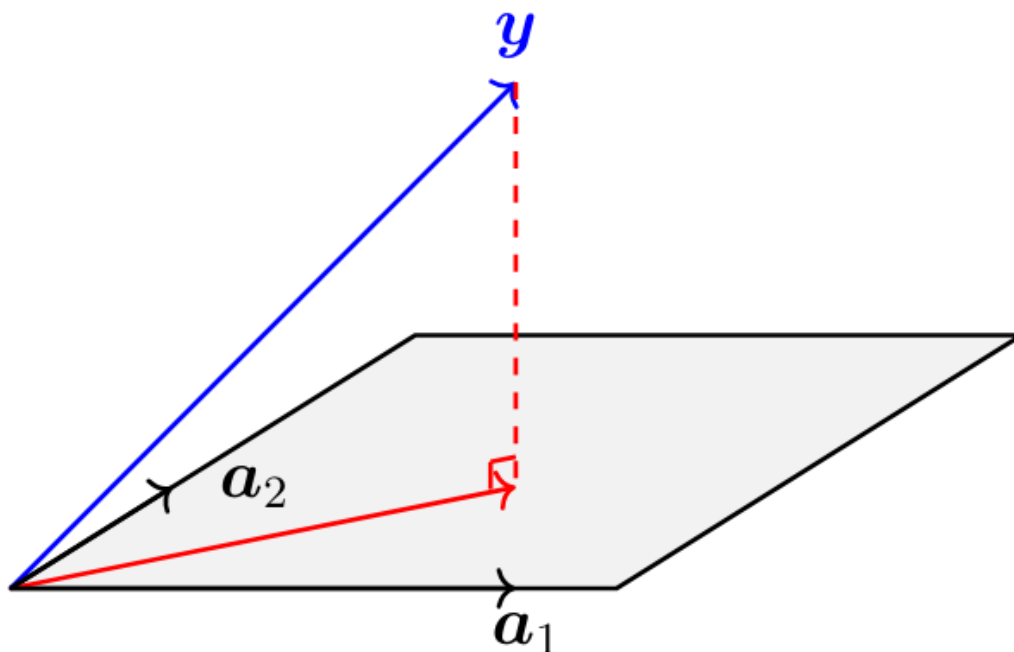


Figure 2: Vector representation of least squares

Relating this image above with the calculations just done we have that the y vector in the picture corresponds to our y vector or our dependent variable, the red vector is our best fit and the dashed red line is our residual vector. The grey plane is the possible fits we can do. The length of the residual vector would be as calculated above 1.1 or $\|blue - red\|_2$.

0.3

1

We are again doing a linear fit using least squares but now we are shifting the data so center it. This gives us a different matrix B .

$$B = \begin{pmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$$

We continue with the same y vector as before. Now we compute $B^T B$ and $B^T y$.

$$B^T B = \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}$$

$$B^T y = \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5.0 \\ 5.0 \\ 4.0 \\ 3.0 \\ 1.0 \end{pmatrix} = \begin{pmatrix} -10 \\ 18 \end{pmatrix}$$

Now we solve the normal equations.

$$B^T B a = B^T y$$

$$\begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -10 \\ 18 \end{pmatrix} \Rightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -1.0 \\ 3.6 \end{pmatrix}$$

Giving the polynomial $y = -1.0x + 3.6$.

2

Now we calculate the 2 norm conditions of the matrices in the normal equations from 2 and 3.

$$\text{cond}(A^T A)_2 = \|A^T A\|_2 \cdot \|(A^T A)^{-1}\|_2 = 7686$$

$$\text{cond}(B^T B)_2 = \|B^T B\|_2 \cdot \|(B^T B)^{-1}\|_2 = 2.0$$

3

Using the given relative error in y we can use the condition number to estimate the relative error in the solution. The relative error in the solution is given by $\frac{\|\delta a\|_2}{\|a\|_2} \leq \text{cond}(A^T A)_2 \cdot \frac{\|\delta y\|_2}{\|y\|_2}$.

$$\frac{\|\delta a\|_2}{\|a\|_2} \leq 7686 \cdot 10^{-3} = 7,686 = 769\%$$

$$\frac{\|\delta a\|_2}{\|a\|_2} \leq 2.0 \cdot 10^{-3} = 0.004 = 0.4\%$$

Top begin for the non centered and bottom being for the centered data.

0.4

When fitting a second order polynomial we make that anzats $p(x) = a_1x^2 + a_2x + a_3$. As before we get a matric equation and for the data given it will look like this

$$\begin{pmatrix} 81 & 9 & 1 \\ 100 & 10 & 1 \\ 121 & 11 & 1 \\ 144 & 12 & 1 \\ 169 & 13 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 5.0 \\ 5.0 \\ 4.0 \\ 3.0 \\ 1.0 \end{pmatrix}$$

Using this we can form the normal equations.

$$\begin{pmatrix} 81 & 100 & 121 & 144 & 169 \\ 9 & 10 & 11 & 12 & 13 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 81 & 9 & 1 \\ 100 & 10 & 1 \\ 121 & 11 & 1 \\ 144 & 12 & 1 \\ 169 & 13 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 81 & 100 & 121 & 144 & 169 \\ 9 & 10 & 11 & 12 & 13 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5.0 \\ 5.0 \\ 4.0 \\ 3.0 \\ 1.0 \end{pmatrix}$$

Computing the matrix multiplications gives the normal equations.

$$\begin{pmatrix} 80499 & 6985 & 615 \\ 6985 & 615 & 55 \\ 615 & 55 & 5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1990 \\ 188 \\ 18 \end{pmatrix}$$

0.5

Now we need to solve the least squares problem using QR factorization. We start by computing the QR factorization of the matrix A using python, using the same A as in 0.2.

$$Q = \begin{pmatrix} -0.36 & -0.68 \\ -0.40 & -0.37 \\ -0.44 & -0.05 \\ -0.48 & 0.26 \\ -0.52 & 0.57 \end{pmatrix}, R = \begin{pmatrix} -24.8 & -2.22 \\ 0 & -0.29 \end{pmatrix}$$

Solving the normal equations using the QR factorization we get the following.

$$Ra = Q^T y \Rightarrow a = \begin{pmatrix} -1.0 \\ 14.6 \end{pmatrix}$$

Since this method gives the same coefficients as the normal equations we can be confident that the QR factorization is correct and it will result in the same figure, the same residual vector and the same residual.

0.6

$$A^{(1)} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 2 & -4 \end{pmatrix}$$

We have $A^{(1)}$ and want to find $A^{(2)}$ using the householder method. We start by $x = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$ giving

$$v = x + \|x\|_2 \hat{\mathbf{1}} = \begin{pmatrix} 6.742 \\ 1 \\ 2 \end{pmatrix}$$

$$H = \begin{pmatrix} -0.802 & -0.267 & -0.535 \\ -0.267 & 0.96 & -0.079 \\ -0.535 & -0.079 & 0.841 \end{pmatrix} \Rightarrow A^{(2)} = \begin{pmatrix} 2 & 1 & 3 \\ 0.0 & -3.742 & 2.673 \\ 0.0 & 0.0 & 1.545 \\ 0.0 & 0.0 & -2.91 \end{pmatrix}$$