Assignment 2 - Numerical Linear Algebra

Prof.: Bernardo Freitas Paulo da Costa

TA: Beatriz Lúcia Teixeira de Souza

Student: Arthur Rabello Oliveira

26/04/2025

Contents

1. Problem 1: Traditional Least Squares (a - f)

1.1. Linear Regression (a)

We have a set of equally spaced points $S := \{t_i = \frac{i}{m}\}, i = 0, 1, ..., m$, we will find the best *line* $f(t) = \alpha + \beta t$ that approximates the points $(t_i, b_i) \in \mathbb{R}^2$

The system of equations to be solved is to be given as a function of t_i, b_i, m.

Solution:

Approximating 2 points in $\mathbb{R}^2$ by a line is trivial, now approximating more than 2 points is a task that requires linear algebra. To see this, we will analyze the following example to build intuition for the general case:

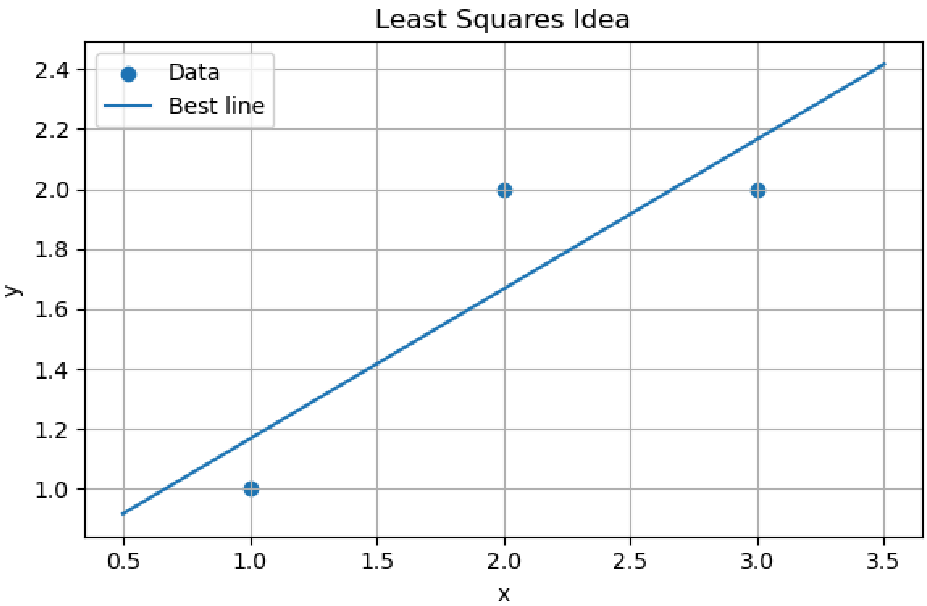

Figure 1: A glimpse into what we want to see

Given the points $(1, 1), (2, 2), (3, 2) \in \mathbb{R}^2$, we would like a *line* $f(t) = y(t) = \alpha + \beta t$ that best approximates these 3 points, in other words, since we know that the line does not pass through all of the 3 points, we would like to find the *closest* line to the line that would pass throught the 3 points, so the system:

$$\begin{aligned} f(1) &= \alpha + \beta = 1 \\ f(2) &= \alpha + 2\beta = 2 \\ f(3) &= \alpha + 3\beta = 2 \end{aligned} \tag{1}$$

Clearly has no solution, (the line does not cross the 3 points), but it has a *closest solution*, which we can find throught **projections**, the system is:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}}_{b} \tag{2}$$

Let $x^* \neq x$ be a solution to the system, we want to **minimize** the error produced by approximating the points through a line, so if the **error** is $e = Ax - b$, we want the smaller error *square* possible (that is why least squares). We square the error to avoid and detect outliers, so:

$$e_1^2 + e_2^2 + e_3^2 \tag{3}$$

Is what we want to minimize, where e_i is the error (distance) from the ith point to the line:

Figure 2: The errors (distances)

In this case, we will project this system into the column space of the matrix A, giving us the closest solution, and the least squares solutions is when $\hat{x}$ minimizes $\|Ax - b\|^2$, this occurs when the residual $e = Ax - b$ is orthogonal to $C(A)$, since $N(A^T) \perp C(A)$ and the dimensions sum up the left dimension of the matrix, so:

$$\begin{gathered} A^T A\hat{x} = A^T b \\ = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix} \end{gathered} \tag{4}$$

So the system to find $\hat{x} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ becomes:

$$\begin{gathered} 3\alpha + 5\beta = 5 \\ 6\alpha + 14\beta = 11 \end{gathered} \tag{5}$$

Notice that with the *errors* e_i^2 as:

$$\begin{gathered} e_1^2 = (\alpha + \beta - 1)^2 \\ e_2^2 = (\alpha + 2\beta - 2)^2 \\ e_3^2 = (\alpha + 3\beta - 2)^2 \end{gathered} \tag{6}$$

The system in eq. (5) is *precisely* what is obtained after using partial derivatives to minimize the erros sum as a function of (α, β):

$$\begin{gathered} f(\alpha,\beta) = (\alpha+\beta-1)^2 + (\alpha+2\beta-2)^2 + (\alpha+3\beta-2)^2 \\ = 3\alpha^2 + 14\beta^2 + 12\alpha\beta - 10\alpha - 22\beta + 9, \\ \frac{\partial f}{\partial \alpha} = \frac{\partial f}{\partial \beta} = 0 \Leftrightarrow 6\alpha + 12d - 10 = 28\beta + 12\alpha - 22 = 0 \Leftrightarrow \begin{cases} 3c + 6d = 11 \\ 6c + 14d = 11 \end{cases} \end{gathered} \tag{7}$$

This new system has a solution in $\hat{\alpha} = \frac{2}{3}, \hat{\beta} = \frac{1}{2}$, so the equation of the optimal line, obtained through *linear regression* (or least squares) is:

$$y(t) = \frac{2}{3} + \frac{1}{2}t. \tag{8}$$

If we have $n > 3$ points to approximate throught a line, the reasoning is analogous:

With $S := \{t_i = \frac{i}{m}\}, i = 0, 1, ..., m$, we will find the best *line* $f(t) = \alpha + \beta t$ that approximates the points $(t_i, b_i) \in \mathbb{R}^2$

The system of equations to be solved be given as a function of t_i, b_i, m.

We want to find the extended system as we did in eq. (7), so our line is:

$$f(t) = \alpha + \beta t \tag{9}$$

That best approximates the points $(0, b_0), (\frac{1}{m}, b_1), ..., (1, b_m)$. The system is:

$$\begin{gathered} f(0) = b_0 = \alpha, \\ f\left(\frac{1}{m}\right) = b_1 = \alpha + \frac{\beta}{m}, \\ f\left(\frac{2}{m}\right) = b_2 = \alpha + \frac{2}{m}\beta \\ \dots \\ f(1) = b_m = \alpha + \beta \end{gathered} \tag{10}$$

And the $Ax = b$ matrices alternative:

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 1 & \frac{1}{m} \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}}_{b} \tag{11}$$

Projecting into $C(A)$, we have:

$$\begin{gathered} A^T A x = A^T b \\ = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & \frac{1}{m} & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & \frac{1}{m} \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} m+1 & \frac{m+1}{2} \\ \frac{m+1}{2} & \frac{(m+1)(2m+2)}{6m} \end{pmatrix} \cdot \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \\ = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & \frac{1}{m} & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} b_1 + b_2 + \dots + b_m \\ \frac{1}{m}[b_2 + 2b_3 + \dots + (m-1)b_m] \end{pmatrix} \end{gathered} \tag{12}$$

So the new system to be solved is:

$$\begin{pmatrix} m+1 & \frac{m+1}{2} \\ \frac{m+1}{2} & \frac{(m+1)(2m+2)}{6m} \end{pmatrix} \cdot \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} b_1 + b_2 + \ldots + b_m \\ \frac{1}{m}[b_2 + 2b_3 + \ldots + (m-1)b_m] \end{pmatrix} \quad (13)$$

Or:

$$\begin{pmatrix} m+1 & \sum_{i=1}^{m} t_i \\ \sum_{i=1}^{m} t_i & \sum_{i=1}^{m} t_i^2 \end{pmatrix} \cdot \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} b_i \\ \sum_{i=1}^{m} \frac{i}{m} \cdot b_i \end{pmatrix} \quad (14)$$

1.2. The Conditioning number of a matrix (b)

1.2.1. Conditioning Number of matrices

1.2.2. Application

1.3. More Regression: A Polynomial Perspective (c)

1.4. Finding the matrix A through Python (d)

1.5. How Perturbations Affect The Conditioning Number (e)

1.6. Another Set of Points (f)

2. Least Squares Algorithms

2.1. The SVD, QR factorizations and the normal approach (a)

2.2. Testing the Algorithms and analyzing plots (b)

2.3. The polynomial approach: An efficiency analysis (c)