1 Condition numbers and pivoted LU

(1.a) Solve the matrix equation $\boldsymbol{Ax} = \boldsymbol{b}$ with

$$\boldsymbol{A} := \begin{bmatrix} 1 & 0 \\ 10^4 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

What is $\kappa_\infty(A)$?

Consider a small perturbation $\Delta\boldsymbol{b} = [10^{-3}, 0]^\top$ being added to the right-hand side, and solve again. Repeat with $\Delta\boldsymbol{b} = [0, 10^{-3}]^\top$. You should see that small perturbation can, but does not have to have a large effect even for badly conditioned systems.

(1.b) Verify the following LU decomposition of a matrix A without pivoting:

$$\boldsymbol{A} := \begin{bmatrix} 10^{-4} & 1 \\ 1 & 1 \end{bmatrix} = \boldsymbol{LU} = \begin{bmatrix} 1 & 0 \\ 10^4 & 1 \end{bmatrix} \begin{bmatrix} 10^{-4} & 1 \\ 0 & 1 - 10^4 \end{bmatrix}$$

We have seen in the previous problem that solving a system with the matrix $\boldsymbol{L}$ is sensitive to errors, i.e., it is poorly conditioned. However, the original $\boldsymbol{A}$ matrix is well-conditioned.

Now the LU factorization of $\boldsymbol{A}$ with pivoting is

$$\boldsymbol{PA} = \begin{bmatrix} 1 & 1 \\ 10^{-4} & 1 \end{bmatrix} = \boldsymbol{LU} = \begin{bmatrix} 1 & 0 \\ 10^{-4} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 - 10^{-4} \end{bmatrix}$$

We see that the LU factors with pivoting are better conditioned.

2 Projectors

A projector is a square matrix $\boldsymbol{P}$ that satisfies

$$\boldsymbol{P}^2 = \boldsymbol{P}.$$

(2.a) Assume $\boldsymbol{P}$ is a projector, and show that $\boldsymbol{I} - \boldsymbol{P}$ is also a projector.

(2.b) We can show that

$$\begin{aligned} &\mathrm{range}(\boldsymbol{I} - \boldsymbol{P}) = \mathrm{null}(\boldsymbol{P}); \\ &\mathrm{null}(\boldsymbol{I} - \boldsymbol{P}) = \mathrm{range}(\boldsymbol{P}); \\ &\mathrm{range}(\boldsymbol{P}) \cap \mathrm{null}(\boldsymbol{P}) = 0. \end{aligned}$$

An orthogonal projector is a projector whose has the subspaces range($\boldsymbol{P}$) and null($\boldsymbol{P}$) orthogonal.

n.b.: An orthogonal projector $\boldsymbol{P}$ is not an orthogonal matrix! Why?

(2.c) Show that if $\boldsymbol{P} = \boldsymbol{P}^\top$ symmetric, the projector $\boldsymbol{P}$ is orthogonal (Hint: take one vector in range($\boldsymbol{P}$) and one in null($\boldsymbol{P}$), show that they must be orthogonal to each other).

The reverse direction holds as well. Therefore the two definitions are equivalent.

(2.d) A special case of orthogonal projection is the projection onto a vector:

$$\boldsymbol{P}_v = \frac{\boldsymbol{v}\boldsymbol{v}^\top}{\boldsymbol{v}^\top\boldsymbol{v}}.$$

Show that it is indeed an orthogonal projector with range span($\boldsymbol{v}$).

(2.e) Another orthogonal projection is

$$\boldsymbol{P}_{\perp v} = I - \frac{\boldsymbol{v}\boldsymbol{v}^\top}{\boldsymbol{v}^\top\boldsymbol{v}}.$$

What is its null space? What is its range?

3 Least squares and infections disease

Let us assume an infectious disease with the following reported new infections I_i on each day t_i, for $i = 1, \ldots, 10$. Using least squares fitting, we would like to understand the nature of this growth.

Table 1: Number of new infections I_i on days t_i.

t_i:	1	2	3	4	5	6	7	8	9	10
I_i:	14	20	21	24	15	45	67	150	422	987

We consider two models to describe the connection between time (i.e., days) t and the number of new infections, both with 3 unknown parameters (a, b, c):

$$I(t) = a + bt + ct^2 \qquad \text{(polynomial model)}$$

$$I(t) = a + bt + c\exp(t) \qquad \text{(exponential model)}$$

Our goal is to figure out which model describes the progression of the infections better, and we use least squares fitting to figure that out. Note that if a model would fit the data perfectly, $I(t_i) = I_i$ for all i. In general, you will not be able to find parameters that satisfy this, and thus have to use least squares fitting (sometimes this is also called *regression*).

(3.a) Formulate, assuming the polynomial model, the least squares problem for the parameters $\boldsymbol{x} = [a, b, c]^T$ by specifying the matrices A and the vector $\boldsymbol{b}$:

$$\min_{\boldsymbol{x} \in \mathbb{R}^3} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

(3.b) Same as above, but for the exponential model.

(3.c) Use a QR-factorization in MATLAB or Python to solve these problems and plot the data as points, as well as the model as a line. Repeat using the normal equations $\boldsymbol{A}^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{b}$.

(3.d) To decide which model describes the data better, we need to compute the distance between the model and the data points. Take a look at the proof from class for how the QR factorization can be used to solve least squares problems. In particular, we found that:

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \geq \|\boldsymbol{b}_2\|_2^2,$$

where $\boldsymbol{b}_2 = \hat{\hat{\boldsymbol{Q}}}^\top \boldsymbol{b}$. We also found that this inequality is equality if $\boldsymbol{x}$ solves the least squares problem. Thus, the norm or $\boldsymbol{b}_2$ is a measure of how well the model fits the data. Use this to decide which of the two models above describes the data better.