Worksheet 5

Xuanxi Zhang

1 Condition numbers and pivoted LU

1.1

Solve the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ with

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

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.

1.2

Verify the following LU decomposition of a matrix A without pivoting:

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

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

Now the LU factorization of A with pivoting is

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

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

In real application, we often use Pivot LU in everystep, not only for the case that current diagonal element is 0. We usually swap current row with the biggest one in the column. This make the algorithm more stable and accurate.

2 Projectors

A projector is a square matrix P that satisfies

$$P^2 = P.$$

2.1

Assume P is a projector, show that $I - P$ is also a projector.

2.2

We can show that

$$\text{range}(I - P) = \text{null}(P);$$
$$\text{null}(I - P) = \text{range}(P);$$
$$\text{range}(P) \cap \text{null}(P) = 0.$$

If $\text{range}(P)$ and $\text{null}(P)$ are orthogonal, then we call P an orthogonal projector.

Remark: An orthogonal projector P is not an orthogonal matrix! Why?

2.3

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

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

2.4

A special case of orthogonal projection is the projection onto a vector:

$$P_v = \frac{vv^\top}{v^\top v}.$$

Show that it is indeed an orthogonal projector with range span(v).

2.5

Another orthogonal projection is

$$P_{\perp v} = I - \frac{vv^\top}{v^\top v}.$$

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*).

3.1

Formulate, assuming the polynomial model, the least squares problem for the parameters $x = [a, b, c]^T$ by specifying the matrices A and the vector b:

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

3.2

Same as above, but for the exponential model.

3.3

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 $A^T Ax = A^T b$.

3.4

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:

$$\|Ax - b\|_2^2 \geq \|b_2\|_2^2,$$

where $b_2 = \hat{Q}^\top b$. We also found that this inequality is equality if x solves the least squares problem. Thus, the norm or 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.