THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics

Mathematical Modelling Project Team
mathmodel@math.cuhk.edu.hk

Exercise (Advanced Linear Algebra)

Last updated: 11/5/2026

Part A: Key Concepts

1 Euclidean norm

For a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)^\top$, its **Euclidean norm** is defined as

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Example

If $\mathbf{x} = (1, 3, 1)$, we have $\|\mathbf{x}\| = \sqrt{1^2 + 3^2 + 1^2} = \sqrt{11}$.

Remark

You can check that Euclidean norm satisfies all the following requirements:

1. $\|\mathbf{x}\| \geq 0$, and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the zero vector.
2. $\|\alpha\mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$, for all $\alpha \in \mathbb{R}$.
3. (Triangle Inequality): $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

With these conditions, we can extend the idea of Euclidean norm and define more general p-norms.

2 vector p-norm

For a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)^\top$, its **$p$-norm** is defined as

$$\|\mathbf{x}\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}, \quad 1 \leq p \leq \infty.$$

Example

For the vector $\mathbf{x} = (2, -3)^\top$, $\|\mathbf{x}\|_3 = \sqrt[3]{|2|^3 + |-3|^3} = \sqrt[3]{35}$.

Remark

$\|\mathbf{x}\|_2$ is nothing but the Euclidean norm. Also, we can see that

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

and

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

You may take the latter one for granted, which can also be derived from definition, but it may need some technique. One can verify that all p-norms satisfy the three requirements of a norm listed above. (Proving the triangle inequality may also require some techniques.)

3 Frobenius norm

If we consider the matrix as a sequence of numbers, the Euclidean norm of this sequence is called the **Frobenius norm**. More precisely, if A is an $m \times n$ matrix and the (i,j)-th entry of A is a_{ij}, then

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}.$$

Example

If $A = \begin{pmatrix} 1 & -3 \\ -2 & 4 \end{pmatrix}$, then $\|A\|_F = \sqrt{1+4+9+16} = \sqrt{30}$.

Remark

But more frequently, we prefer matrix norms that are induced from a vector p-norm.

4 Matrix p-norm

If A is an $m \times n$ matrix, then the **p-norm** of A is defined as:

$$\|A\|_p = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|_p .$$

This definition is hard to read. Let us provide some further explanation. The p-norm of a matrix A is a measure of the maximum extent to which A can amplify the p-norm of any vector. Specifically, it is defined as the maximum value of where $\mathbf{x}$ is a vector with unit p norm. This definition captures the idea of the maximum "stretching" effect the matrix can have on a unit vector measured in the p-norm.

In practice, there are some much simpler formulas that we can use for calculating the matrix p-norms for some specific p.

In particular, it can be proved that for $p = 1$, we have

$$\|A\|_1 = \text{maximum absolute column sum of } A = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

For $p = 2$, we have

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)},$$

where $\lambda_{\max}(A^\top A)$ is the maximum eigenvalue of $A^\top A$.

For $p = \infty$, we have

$$\|A\|_\infty = \text{maximum absolute row sum of } A = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|.$$

5 Linear system

A **linear system** is usually in the following form

$$\begin{matrix} a_{11}x_1+ & a_{12}x_2+ & \cdots+ & a_{1n}x_n & =b_1 \\ a_{21}x_1+ & a_{22}x_2+ & \cdots+ & a_{2n}x_n & =b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1}x_1+ & a_{m2}x_2+ & \cdots+ & a_{mn}x_n & =b_m \end{matrix}$$

This may occur in many applications (often with very large m and n). It is convenient to express the above in the matrix form

$$A\mathbf{x} = \mathbf{b},$$

where A is an $m \times n$ matrix with elements a_{ij}, and $\mathbf{x}$ is a $n \times 1$ vector, $\mathbf{b}$ is a $m \times 1$ vector.
When $m > n$, it is called a **overdetermined** system since the conditions are more than unknowns. When $m < n$, it is called a **underdetermined** system since the conditions are less than unknowns.
For now, we are interested in the case when $m = n$. It follows from the definition of a non-singular matrix that the above linear system has a unique solution if and only if A is non-singular, given by $\mathbf{x} = A^{-1}\mathbf{b}$. Although this seems like a conceptually easy problem, it is actually a hard one when n gets large. Nowadays, linear systems with $n = 1,000,000$ arise routinely in computational problems. And even for small n there are some potential pitfalls. We will see how to solve $A\mathbf{x} = \mathbf{b}$ both efficiently and accurately.

6 Triangular matrix

A matrix L is called **lower triangular** if all entries above the diagonal are zero:

$$L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ l_{n1} & \cdots & \cdots & l_{nn} \end{pmatrix}.$$

Note that the determinant is just

$$\det(L) = l_{11}l_{22}\cdots l_{nn}$$

So the matrix will be non-singular if and only if all of the diagonal elements are non-zero.

Example

We will see that it is easy to solve $L\mathbf{x} = \mathbf{b}$ for $n = 4$.
The system is

$$\begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

which is equivalent to

$$\begin{cases} l_{11}x_1 = b_1, \\ l_{21}x_1 + l_{22}x_2 = b_2, \\ l_{31}x_1 + l_{32}x_2 + l_{33}x_3 = b_3, \\ l_{41}x_1 + l_{42}x_2 + l_{43}x_3 + l_{44}x_4 = b_4. \end{cases}$$

Example

We can solve step-by-step:

$$x_1 = \frac{b_1}{l_{11}}, \quad x_2 = \frac{b_2 - l_{21}x_1}{l_{22}}, \quad x_3 = \frac{b_3 - l_{31}x_1 - l_{32}x_2}{l_{33}}, \quad x_4 = \frac{b_4 - l_{41}x_1 - l_{42}x_2 - l_{43}x_3}{l_{44}}.$$

This works since we know that $l_{11}, l_{22}, l_{33}, l_{44}$ are all non-zero when a solution exists.
In general, any lower triangular system $L\mathbf{x} = \mathbf{b}$ can be solved by:

$$x_j = \frac{b_j - \sum_{k=1}^{j-1} l_{jk}x_k}{l_{jj}}, \quad j = 1, \cdots, n.$$

Similarly, an **upper triangular** matrix U has the form

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix}$$

and an upper-triangular system $U\mathbf{x} = \mathbf{b}$ may be solved by:

$$x_j = \frac{b_j - \sum_{k=j+1}^{n} u_{jk}x_k}{u_{jj}}, \quad j = n, \cdots, 1.$$

7 Gaussian elimination

If our matrix A is not triangular, we can try to transform it to triangular form. **Gaussian elimination** uses elementary row operations to transform the system to upper triangular form $U\mathbf{x} = \mathbf{y}$.
Elementary row operations include swapping rows and adding multiples of one row to another. They will not change the solution x, but will change the matrix A and the right-hand side b.

Let $A^{(1)} = A$ and $b^{(1)} = b$. Then for each k from 1 to $n-1$, compute a new matrix $A^{(k+1)}$ and right-hand side $b^{(k+1)}$ by the following procedure:

1. $m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k+1, \ldots, n.$
2. Use these to remove the unknown x_k from equations $k+1$ to n, leaving:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)}, \quad b_i^{(k+1)} = b_i^{(k)} - m_{ik}b_k^{(k)}, \quad i, j = k+1, \ldots, n.$$

The final matrix $A^{(n)} = U$ will then be upper triangular. The notation may seem messy, so let's see an easy example.

Example

Transform the following system to upper triangular form:

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ x_1 - 2x_2 + 2x_3 = 4, \\ 2x_1 + 12x_2 - 2x_3 = 4. \end{cases}$$

Matrix representation:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & 12 & -2 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix}.$$

Subtract $1 \times$ equation 1 from equation 2, and $2 \times$ equation 1 from equation 3:

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ -4x_2 + x_3 = 4, \\ 8x_2 - 4x_3 = 4. \end{cases}$$

Updated matrices:

$$A^{(2)} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 8 & -4 \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix}, \quad m_{21} = 1, \quad m_{31} = 2.$$

Subtract $-2 \times$ equation 2 from equation 3:

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ -4x_2 + x_3 = 4, \\ -2x_3 = 12. \end{cases}$$

Final upper triangular form:

$$A^{(3)} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 0 & -2 \end{pmatrix}, \quad b^{(3)} = \begin{pmatrix} 0 \\ 4 \\ 12 \end{pmatrix}, \quad m_{32} = -2.$$

The solution is:

$$x_1 = 11, \quad x_2 = -\frac{5}{2}, \quad x_3 = -6.$$

8 LU decomposition

If a $n \times n$ matrix A can be factorized into $A = LU$, where L is a lower triangular matrix with 1 as its diagonal entries, and U is a upper triangular matrix, such decomposition is called **LU decomposition**.

We will see that the sequence of row operations that transforms A to U is equivalent to left-multiplying by a matrix F, so that

$$FA = U, \quad Ux = Fb.$$

To see this, note that step k of Gaussian elimination can be written in the following form

$$A^{(k+1)} = F^{(k)}A^{(k)}, \quad b^{(k+1)} = F^{(k)}b^{(k)}, \quad \text{where} \quad F^{(k)} := \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & \ddots & \ddots & \vdots \\ \vdots & -m_{k+1,k} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & -m_{n,k} & \cdots & 0 & 1 \end{pmatrix}.$$

Multiplying by $F^{(k)}$ has the effect of subtracting m_{ik} times row k from row i, for $i = k+1, \ldots, n$.

Example

You can check in the earlier example that

$$F^{(1)}A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & 12 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 1-1(1) & -2-1(2) & 2-1(1) \\ 2-2(1) & 12-2(2) & -2-2(1) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 8 & -4 \end{pmatrix} = A^{(2)},$$

Example

$$F^{(2)}A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 8 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 8+2(-4) & -4+2(1) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 0 & -2 \end{pmatrix} = A^{(3)} = U.$$

It follows that $U = A^{(n)} = F^{(n-1)}F^{(n-2)} \cdots F^{(1)}A$.
Now the $F^{(k)}$ are invertible, and the inverse is just given by adding rows instead of subtracting:

$$(F^{(k)})^{-1} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & \ddots & \ddots & \vdots \\ \vdots & & m_{k+1,k} & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & m_{n,k} & \cdots & 0 & 1 \end{pmatrix}.$$

So we could write $A = (F^{(1)})^{-1}(F^{(2)})^{-1} \cdots (F^{(n-1)})^{-1}U$.
Since the successive operations don't "interfere" with each other, or by direct computation, we can write

$$(F^{(1)})^{-1}(F^{(2)})^{-1} \cdots (F^{(n-1)})^{-1} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ m_{2,1} & 1 & \ddots & \ddots & \ddots & \vdots \\ m_{3,1} & m_{3,2} & 1 & \ddots & \ddots & \vdots \\ m_{4,1} & m_{4,2} & m_{4,3} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ m_{n,1} & m_{n,2} & m_{n,3} & \cdots & m_{n,n-1} & 1 \end{pmatrix} := L.$$

Thus we have established the LU decomposition.

The system $Ax = b$ becomes $LUx = b$, which we can readily solve by setting $Ux = y$. We first solve $Ly = b$ for y, then $Ux = y$ for x. Both are triangular systems.

Example

Solve the same question as before by LU decomposition.

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & 12 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix}.$$

We apply Gaussian elimination as before, but ignore b (for now), leading to

$$U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 0 & -2 \end{pmatrix}.$$

Example

As we apply the elimination, we record the multipliers so as to construct the matrix

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

Thus we have the decomposition

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & 12 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 0 & -2 \end{pmatrix}.$$

With the matrices L and U, we can readily solve for any right-hand side b. We illustrate for our particular b. Firstly, solve $Ly = b$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix} \implies y_1 = 0, \quad y_2 = 4 - y_1 = 4, \quad y_3 = 4 - 2y_1 + 2y_2 = 12.$$

Notice that y is the right-hand side $b^{(3)}$ constructed earlier. Then, solve $Ux = y$:

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 1 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 12 \end{pmatrix} \implies x_3 = -6, \quad x_2 = -\frac{1}{4}(4 - x_3) = -\frac{5}{2}, \quad x_1 = -2x_2 - x_3 = 11.$$

9 Pivoting

Gaussian elimination and LU decomposition will both fail if we hit a zero on the diagonal. But this does not mean that the matrix A is singular.

Example

The system

$$\begin{pmatrix} 0 & 3 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

obviously has solution $x_1 = x_2 = x_3 = 1$. But Gaussian elimination will fail because $a_{11}^{(1)} = 0$, so we cannot calculate m_{21} and m_{31}.

However, we could avoid the problem by changing the order of the equations to get the equivalent system

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

Now there is no problem with Gaussian elimination.

Swapping rows or columns is called **pivoting**. It is needed if the pivot element is zero, as in the above example. But it is also used to reduce rounding error.

Example

Consider the system

$$\begin{pmatrix} 10^{-4} & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

1. Using Gaussian elimination with exact arithmetic gives

$$m_{21} = -10^4, \quad a_{22}^{(2)} = 2 + 10^4, \quad b_2^{(2)} = 1 + 10^4.$$

   So backward substitution gives the solution

$$x_2 = \frac{1 + 10^4}{2 + 10^4} = 0.9999, \quad x_1 = \frac{1 - x_2}{a_{11}} = 10^4 \left(1 - \frac{1 + 10^4}{2 + 10^4}\right) = \frac{10^4}{2 + 10^4} = 0.9998.$$

2. Now do the calculation in 3-digit arithmetic. We have

$$m_{21} = \mathrm{fl}(-10^4) = -10^4, \quad a_{22}^{(2)} = \mathrm{fl}(2 + 10^4) = 10^4, \quad b_2^{(2)} = \mathrm{fl}(1 + 10^4) = 10^4.$$

   Now backward substitution gives

$$x_2 = \mathrm{fl}\left(\frac{10^4}{10^4}\right) = 1, \quad x_1 = \mathrm{fl}\left(10^4(1 - 1)\right) = 0.$$

   The large value of m_{21} has caused a rounding error which has later led to a loss of significance during the evaluation of x_1.

3. We do the calculation correctly in 3-digit arithmetic if we first swap the equations,

$$\begin{pmatrix} -1 & 2 \\ 10^{-4} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

   Now, $m_{21} = \mathrm{fl}(-10^{-4}) = -10^{-4}$, $a_{22}^{(2)} = \mathrm{fl}(1 + 10^{-4}) = 1$, $b_2^{(2)} = \mathrm{fl}(1 + 10^{-4}) = 1$, and

$$x_2 = \mathrm{fl}\left(\frac{1}{1}\right) = 1, \quad x_1 = \mathrm{fl}\left(-[1 - 2(1)]\right) = 1.$$

   Now both x_1 and x_2 are correct to 3 significant figures.

10 Orthogonality

Recall that the inner product between two column vectors $x, y \in \mathbb{R}^n$ is defined as

$$x \cdot y = x^\top y = \sum_{k=1}^{n} x_k y_k.$$

This is related to the 2-norm since $\|x\|_2 = \sqrt{x^\top x}$. The angle θ between x and y is given by

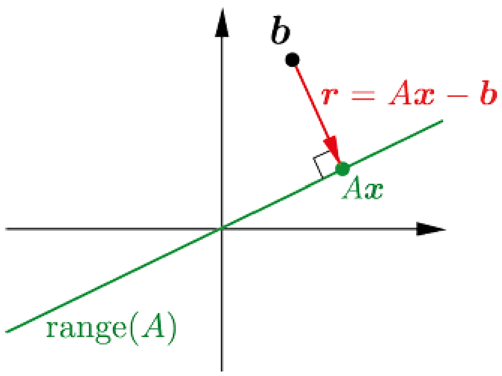
$$x^\top y = \|x\|_2 \|y\|_2 \cos\theta.$$

Two vectors x and y are **orthogonal** if $x^\top y = 0$.
Let $S = \{x_1, x_2, \dots, x_n\}$ be a set of n vectors. Then S is called **orthogonal** if $x_i^\top x_j = 0$ for all $i, j \in \{1, 2, \dots, n\}$ with $i \neq j$.

11 Discrete least-squares problem

In this section, we are interested in cases where A is an $m \times n$ rectangular matrix with $m > n$, then the linear system $Ax = b$ is overdetermined and will usually have no solution. But we can still look for an approximate solution. The **discrete least-squares problem** is to find $\mathbf{x}$ that minimizes the 2-norm $\|A\mathbf{x} - \mathbf{b}\|_2$.

To solve the problem, it is useful to think geometrically. The range of A, written range(A), is the set of all possible vectors $A\mathbf{x} \in \mathbb{R}^m$, where $\mathbf{x} \in \mathbb{R}^n$. This will only be a subspace of $\mathbb{R}^m$, and in particular it will not, in general, contain $\mathbf{b}$. We are therefore looking for $x \in \mathbb{R}^n$ such that $A\mathbf{x}$ is as close as possible to $\mathbf{b}$ in $\mathbb{R}^m$ (as measured by the 2-norm distance).



The distance from $A\mathbf{x}$ to $\mathbf{b}$ is given by $\|r\|_2 = \|A\mathbf{x} - \mathbf{b}\|_2$. Geometrically, we see that $\|r\|_2$ will be minimized by choosing r orthogonal to $A\mathbf{x}$, i.e.,

$$(A\mathbf{x})^\top (A\mathbf{x} - \mathbf{b}) = 0 \iff \mathbf{x}^\top (A^\top A\mathbf{x} - A^\top \mathbf{b}) = 0.$$

This will be satisfied if $\mathbf{x}$ satisfies the $n \times n$ linear system

$$A^\top A x = A^\top \mathbf{b}.$$

Note that $A^\top A$ is a $n \times n$ matrix, so we can solve the system with methods introduced before.

Example

Fit a least-squares straight line to the data $f(-3) = f(0) = 0$, $f(6) = 2$.
Here $n = 1$ (fitting a straight line) and $m = 2$ (3 data points), so $x_0 = -3$, $x_1 = 0$ and $x_2 = 6$.
The overdetermined system is

$$\begin{pmatrix} 1 & -3 \\ 1 & 0 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix},$$

and the normal equations have the form

$$\begin{pmatrix} 3 & x_0 + x_1 + x_2 \\ x_0 + x_1 + x_2 & x_0^2 + x_1^2 + x_2^2 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} f(x_0) + f(x_1) + f(x_2) \\ x_0 f(x_0) + x_1 f(x_1) + x_2 f(x_2) \end{pmatrix}$$

$$\iff \begin{pmatrix} 3 & 3 \\ 3 & 45 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 12 \end{pmatrix} \implies \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} \frac{3}{7} \\ \frac{5}{21} \end{pmatrix}.$$

So the least-squares approximation by straight line is $p_1(x) = \dfrac{3}{7} + \dfrac{5}{21}x$.

12 Eigenvalues and eigenvectors

Let A be a $n \times n$ matrix over $\mathbb{R}$. A nonzero vector $v \in \mathbb{R}^n$ is called an **eigenvector** of A if $Av = \lambda v$ for some scalar λ. The scalar λ is called the **eigenvalue** of A that corresponds to the eigenvector v.

Example

Let $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$, $\quad v_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, $\quad$ and $\quad v_2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.
Since $Av_1 = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = -2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -2v_1$, v_1 is an eigenvector of A. Here $\lambda_1 = -2$ is the eigenvalue corresponding to v_1. Furthermore,

$$Av_2 = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 15 \\ 20 \end{pmatrix} = 5 \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 5v_2,$$

and so v_2 is an eigenvector of A with the corresponding eigenvalue $\lambda_2 = 5$.

Theorem and Proof

Let A be an $n \times n$ matrix over $\mathbb{R}$. Then a scalar λ is an eigenvalue of A if and only if $\det(A - \lambda I_n) = 0$.
Proof. A scalar λ is an eigenvalue of A if and only if there exists a nonzero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$, that is, $(A - \lambda I_n)(v) = 0$. This is true if and only if $A - \lambda I_n$ is not invertible. However, this result is equivalent to the statement that $\det(A - \lambda I_n) = 0$. The polynomial $f(t) = det(A - tI_n)$ is called the **characteristic polynomial** of A.

Example

To find the eigenvalues of A, we compute its characteristic polynomial:

$$\det(A - tI_2) = \det \begin{pmatrix} 1 - t & 1 \\ 4 & 1 - t \end{pmatrix} = t^2 - 2t - 3 = (t - 3)(t + 1).$$

It follows from theorem that the only eigenvalues of A are 3 and -1.

Part B: Basic Questions

1. Consider the vectors $a = (1, -2, 3)^\top$, $b = (2, 0, -1)^\top$, and $c = (0, 1, 4)^\top$. Find the 1-norm, 2-norm and infinity-norm of a, b and c. Verify that for a single vector x, the norms satisfy the ordering $\|x\|_1 \geq \|x\|_2 \geq \|x\|_\infty$.

2. Let $A = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}$. Measured by the 2-norm, a unit vector in $\mathbb{R}^2$ has the form $x = (\cos\theta, \sin\theta)^\top$. Then $Ax = \begin{pmatrix} \sin\theta \\ 3\cos\theta \end{pmatrix}$. Use this fact to find $\|A\|_2$.

3. For the matrix $A = \begin{pmatrix} -7 & 3 & -1 \\ 2 & 4 & 5 \\ -4 & 6 & 0 \end{pmatrix}$, find the value of $\|A\|_1$ and $\|A\|_\infty$.

4. Use Gaussian elimination to solve the following systems of linear equations.

   (a) $\begin{cases} x_1 + 2x_2 - x_3 = -1 \\ 2x_1 + 2x_2 + x_3 = 1 \\ 3x_1 + 5x_2 - 2x_3 = -1 \end{cases}$

   (b) $\begin{cases} x_1 - 2x_2 - x_3 = 1 \\ 2x_1 - 3x_2 + x_3 = 6 \\ 3x_1 - 5x_2 = 7 \\ x_1 + 5x_3 = 9 \end{cases}$

   (c) $\begin{cases} x_1 + 2x_2 + 2x_4 = 6 \\ 3x_1 + 5x_2 - x_3 + 6x_4 = 17 \\ 2x_1 + 4x_2 + x_3 + 2x_4 = 12 \\ 2x_1 - 7x_3 + 11x_4 = 7 \end{cases}$

   (d) $\begin{cases} x_1 - x_2 - 2x_3 + 3x_4 = -63 \\ 2x_1 - x_2 + 6x_3 + 6x_4 = -2 \\ -2x_1 + x_2 - 4x_3 - 3x_4 = 0 \\ 3x_1 - 2x_2 + 9x_3 + 10x_4 = -5 \end{cases}$

5. Let $A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 8 & 14 \\ 2 & 6 & 13 \end{pmatrix} = LU$ where $L = \begin{pmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{pmatrix}$ and $U = \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix}$.

   (a) Find the LU decomposition by direct matrix multiplication.

   (b) Find the LU decomposition by Gauss elimination.

6. Find all eigenvalues and eigenvectors of matrix $A = \begin{pmatrix} -2 & 1 \\ 12 & -3 \end{pmatrix}$.

7. Solve $Ax = b$ in the least-squares sense, where $A = \begin{pmatrix} 1 & 3 & 0 \\ -1 & 1 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{pmatrix}$ $\quad b = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 2 \end{pmatrix}$.

Part C: Advanced Questions

8. Consider the following tridiagonal matrix $A = \begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{pmatrix}$

   (a) What is the LU decomposition of A?

   (b) Using the result above, solve the following linear system, $\begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & -1 & 4 & -1 & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 0 \\ 6 \\ 2 \end{pmatrix}$.

9. Let A be an $n \times n$ square matrix such that A^{-1} exists. Prove: if λ is an eigenvalue of A, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

Part D: Solutions

1.

$$\|a\|_1 = 1+2+3 = 6 \quad \|b\|_1 = 2+0+1 = 3 \quad \|c\|_1 = 0+1+4 = 5$$

$$\|a\|_2 = \sqrt{1+4+9} \approx 3.74 \quad \|b\|_2 = \sqrt{4+0+1} \approx 2.24 \quad \|c\|_2 = \sqrt{0+1+16} \approx 4.12$$

$$\|a\|_\infty = \max\{1,2,3\} = 3 \quad \|b\|_\infty = \max\{2,0,1\} = 2 \quad \|c\|_\infty = \max\{0,1,4\} = 4.$$

2.

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \max_\theta \left(\sin^2\theta + 9\cos^2\theta\right)^{1/2} = \max_\theta \left(1 + 8\cos^2\theta\right)^{1/2} = 3.$$

3.

$$\|A\|_1 = \max\{13,13,6\} = 13, \quad \|A\|_\infty = \max\{11,11,10\} = 11.$$

4. Gaussian elimination process is following:

(a)

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 0 & -2 & 3 & 3 \\ 0 & -1 & 1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 0 & 1 & -1.5 & -1.5 \\ 0 & -1 & 1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 0 & 1 & -1.5 & -1.5 \\ 0 & 0 & -0.5 & 0.5 \end{array}\right)$$

The final solution is given by $\begin{cases} x_1 = 4 \\ x_2 = -3 \\ x_3 = -1 \end{cases}$

(b)

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 2 & -3 & 1 & 6 \\ 3 & -5 & 0 & 7 \\ 1 & 0 & 5 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 1 & 3 & 4 \\ 0 & 1 & 3 & 4 \\ 0 & 2 & 6 & 8 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

The final solution is given by $\begin{cases} x_1 = 9 - 5a \\ x_2 = 4 - 3a \\ x_3 = a \end{cases}$, for all $a \in \mathbb{R}$

(c)

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 2 & 6 \\ 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & -4 & -7 & 7 & -5 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 0 & 2 & 6 \\ 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & -3 & 7 & 1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 0 & 2 & 6 \\ 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

The final solution is given by $\begin{cases} x_1 = 2 \\ x_2 = 3 \\ x_3 = -2 \\ x_4 = -1 \end{cases}$

(d)

$$\left(\begin{array}{cccc|c} 1 & -1 & -2 & 3 & -63 \\ 0 & 1 & 10 & 0 & 124 \\ 0 & -1 & -8 & 3 & -126 \\ 0 & 1 & 15 & 1 & 184 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -1 & -2 & 3 & -63 \\ 0 & 1 & 10 & 0 & 124 \\ 0 & 0 & 2 & 3 & -2 \\ 0 & 0 & 5 & 1 & 60 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -1 & -2 & 3 & -63 \\ 0 & 1 & 10 & 0 & 124 \\ 0 & 0 & 2 & 3 & -2 \\ 0 & 0 & 0 & -6.5 & 65 \end{array}\right)$$

The final solution is given by $\begin{cases} x_1 = -21 \\ x_2 = -16 \\ x_3 = 14 \\ x_4 = -10 \end{cases}$

5. (a) Multiplying out LU and setting the answer equal to A gives

$$\begin{pmatrix} U_{11} & U_{12} & U_{13} \\ L_{21}U_{11} & L_{21}U_{12} + U_{22} & L_{21}U_{13} + U_{23} \\ L_{31}U_{11} & L_{31}U_{12} + L_{32}U_{22} & L_{31}U_{13} + L_{32}U_{23} + U_{33} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 8 & 14 \\ 2 & 6 & 13 \end{pmatrix}.$$

From this, we get

$$U_{11} = 1, \quad U_{12} = 2, \quad U_{13} = 4.$$

Now consider the second row, we get

$$L_{21} = 3, \quad U_{22} = 2, \quad U_{23} = 2$$

Solving the rest unknowns, we get $A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 8 & 14 \\ 2 & 6 & 13 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$

(b) With Gauss elimination, we will get the same result.

6.

$$A - \lambda I = \begin{pmatrix} -2 & 1 \\ 12 & -3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2-\lambda & 1 \\ 12 & -3-\lambda \end{pmatrix}$$

Solve the equation

$$\det(A - \lambda I) = 0$$

Thus, the eigenvalues are: $\lambda_1 = 1$ and $\lambda_2 = -6$.
For $\lambda_1 = 1$:

$$(A - I)\mathbf{v} = \begin{pmatrix} -3 & 1 \\ 12 & -4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathbf{0}$$

This gives

$$-3v_1 + v_2 = 0 \Rightarrow \mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

For $\lambda_2 = -6$:

$$(A + 6I)\mathbf{v} = \begin{pmatrix} 4 & 1 \\ 12 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathbf{0}$$

This gives

$$4v_1 + v_2 = 0 \Rightarrow \mathbf{v}_2 = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

Therefore, the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = -6$. The corresponding eigenvectors are $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$.

7. Note that

$$A^\top A = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 18 & -4 \\ 0 & -4 & 2 \end{pmatrix}$$

Then $A^\top Ax = A^\top b$ gives that $x = \begin{pmatrix} \frac{5}{4} \\ \frac{1}{4} \\ -1 \end{pmatrix}$

8. (a) By direct LU decomposition, we have $A = \begin{pmatrix} 1 & 0 & & & \\ l_2 & 1 & & & \\ & l_3 & 1 & & \\ & & \ddots & \ddots & \\ & & & l_n & 1 \end{pmatrix} \begin{pmatrix} v_1 & c_1 & & & \\ & v_2 & c_2 & & \\ & & \ddots & \ddots & \\ & & & v_{n-1} & c_{n-1} \\ & & & & v_n \end{pmatrix}$ where $l_k = \frac{a_k}{v_{k-1}}$ and $v_k = b_k - l_k c_{k-1}$ for $k = 2, \ldots, n$, with $v_1 = b_1$.

(b) For the given system, the LU factorization yields:

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & 1 & 0 & 0 & 0 \\ 0 & -\frac{4}{15} & 1 & 0 & 0 \\ 0 & 0 & -\frac{15}{56} & 1 & 0 \\ 0 & 0 & 0 & -\frac{56}{209} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 4 & -1 & 0 & 0 & 0 \\ 0 & \frac{15}{4} & -1 & 0 & 0 \\ 0 & 0 & \frac{56}{15} & -1 & 0 \\ 0 & 0 & 0 & \frac{209}{56} & -1 \\ 0 & 0 & 0 & 0 & \frac{780}{209} \end{pmatrix}$$

Solving $L\mathbf{y} = \mathbf{b}$ via forward substitution: $\begin{cases} y_1 = 2 \\ y_2 = 6 + \frac{1}{4}y_1 = \frac{13}{2} \\ y_3 = 0 + \frac{4}{15}y_2 = \frac{26}{15} \\ y_4 = 6 + \frac{15}{56}y_3 = \frac{347}{56} \\ y_5 = 2 + \frac{56}{209}y_4 = \frac{780}{209} \end{cases}$ Then solving $U\mathbf{x} = \mathbf{y}$ via backward substitution: $\begin{cases} x_5 = \frac{y_5}{u_{55}} = 1 \\ x_4 = \frac{y_4 + x_5}{u_{44}} = 2 \\ x_3 = \frac{y_3 + x_4}{u_{33}} = 1 \\ x_2 = \frac{y_2 + x_3}{u_{22}} = 2 \\ x_1 = \frac{y_1 + x_2}{u_{11}} = 1 \end{cases}$

9. Since λ is an eigenvalue of A, there is a non-zero $n \times 1$ matrix x satisfying

$$Ax = \lambda x$$

Then,

$$A^{-1}Ax = \lambda A^{-1}x$$

Therefore,

$$x = \lambda A^{-1}x \rightarrow A^{-1}x = \frac{1}{\lambda}x$$

which completes the proof.