數值分析 MA-3021

Chapter 5. Direct and Iterative Methods for Solving Linear Systems

§5.1 Introduction - Review on Linear Algebra

We are interested in solving systems of linear equations of the form:

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & b_n \end{array}\right.$$

This is a system of n equations in the n unknowns, $x_1, x_2, \cdots, x_n$. The elements a_{ij} and b_i are assumed to be prescribed real numbers.

We can rewrite this system of linear equations in a matrix form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a\cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

We can denote these matrices by A, $\boldsymbol{x}$, and $\boldsymbol{b}$, giving the simpler equation: $A\boldsymbol{x} = \boldsymbol{b}$.

§5.1 Introduction - Review on Linear Algebra

We are interested in solving systems of linear equations of the form:

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & b_n \end{array}\right.$$

This is a system of n equations in the n unknowns, $x_1, x_2, \cdots, x_n$. The elements a_{ij} and b_i are assumed to be prescribed real numbers. We can rewrite this system of linear equations in a matrix form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a\cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

We can denote these matrices by A, $\boldsymbol{x}$, and $\boldsymbol{b}$, giving the simpler equation: $A\boldsymbol{x} = \boldsymbol{b}$.

§5.1 Introduction - Review on Linear Algebra

Notation:

1. Let A be a $m \times n$ matrix. Then
   - The (i,j) entry of A is denoted by A_{ij}, a_{ij} or $A(i,j)$.
   - The j-th row of A is denoted by $A(j,:)$.
   - The j-th column of A is denoted by $A(:,j)$.
2. The $n \times n$ identity matrix is denoted by I_n or $I_{n\times n}$. When the dimension n is clear, $n \times n$ we sometimes also use I to denote the identity matrix.

§5.1 Introduction - Review on Linear Algebra

If A and B are two matrices such that $AB = I$, then we say that B is a right inverse of A and that A is a left inverse of B. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \alpha & \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times 2}, \qquad \forall \alpha, \beta \in \mathbb{R}.$$

$$\begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times 2}, \qquad \forall \alpha, \beta \in \mathbb{R}.$$

Notice that right inverse and left inverse may not unique.

§5.1 Introduction - Review on Linear Algebra

Theorem

A square matrix can possess at most one right inverse.

Proof.

Let $AB = I$. Then $\sum_{j=1}^{n} b_{jk} A(:, j) = I(:, k)$ for all $1 \leqslant k \leqslant n$. So, the columns of A form a basis for $\mathbb{R}^n$. Therefore, the coefficients b_{jk} above are uniquely determined. □

Theorem

If A and B are square matrices such that $AB = I$, then $BA = I$.

Proof.

Let $C = BA - I + B$. Then $AC = ABA - AI + AB = A - A + I = I$. Since right inverse for square matrix is at most one, $B = C$. Hence, $C = BA - I + B = BA - I + C$; that is, $BA = I$. □

§5.1 Introduction - Review on Linear Algebra

Theorem

A square matrix can possess at most one right inverse.

Proof.

Let $AB = I$. Then $\sum_{j=1}^{n} b_{jk} A(:, j) = I(:, k)$ for all $1 \leqslant k \leqslant n$. So, the columns of A form a basis for $\mathbb{R}^n$. Therefore, the coefficients b_{jk} above are uniquely determined. □

Theorem

If A and B are square matrices such that $AB = I$, then $BA = I$.

Proof.

Let $C = BA - I + B$. Then $AC = ABA - AI + AB = A - A + I = I$.
Since right inverse for square matrix is at most one, $B = C$.
Hence, $C = BA - I + B = BA - I + C$; that is, $BA = I$. □

§5.1 Introduction - Review on Linear Algebra

1. If a square matrix A has a right inverse B, then B is unique and $BA = AB = I$. We then call B the inverse of A and say that A is invertible or nonsingular. We denote $B = A^{-1}$.
2. If A is invertible, then the system of equations $A\boldsymbol{x} = \boldsymbol{b}$ has the solution $\boldsymbol{x} = A^{-1}b$. If A^{-1} is not available, then in general, A^{-1} should not be computed solely for the purpose of obtaining x.
3. How do we get this A^{-1}?

§5.1 Introduction - Review on Linear Algebra

1. Let two linear systems be given, each consisting of n equations with n unknowns:

$$A\mathbf{x} = \mathbf{b} \quad \text{and} \quad B\mathbf{x} = \mathbf{d}.$$

   If the two systems have precisely the same solutions, we call them equivalent systems.

2. Note that A and B can be very different.

3. Thus, to solve a linear system of equations, we can instead solve any equivalent system. This simple idea is at the heart of our numerical procedures.

§5.1 Introduction - Review on Linear Algebra

Let $\mathcal{E}_i$ denote the i-th equation in the system $A\boldsymbol{x} = \boldsymbol{b}$. The following are the **elementary operations** which can be performed:

- Interchanging two equations in the system: $\mathcal{E}_i \leftrightarrow \mathcal{E}_j$;
- Multiplying an equation by a nonzero number: $\lambda\mathcal{E}_i \to \mathcal{E}_i$;
- Adding to an equation a multiple of some other equation: $\mathcal{E}_i + \lambda\mathcal{E}_j \to \mathcal{E}_i$.

Theorem

If one system of equations is obtained from another by a finite sequence of elementary operations, then the two systems are equivalent.

§5.1 Introduction - Review on Linear Algebra

1. An elementary matrix is defined to be an $n \times n$ matrix that arises when an elementary operation is applied to the $n \times n$ identity matrix.
2. The elementary operations expressed in terms of the rows of matrix A are:
   - The interchange of two rows in A: $A(i,:) \leftrightarrow A(j,:)$;
   - Multiplying one row by a nonzero constant: $\lambda A(i,:) \rightarrow A(:,i)$;
   - Adding to one row a multiple of another:
   $$A(i,:) + \lambda A(j,:) \rightarrow A(i,:).$$
3. Each elementary row operation on A can be accomplished by multiplying A on the left by an elementary matrix.

§5.1 Introduction - Review on Linear Algebra

Example

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \lambda a_{21} + a_{31} & \lambda a_{22} + a_{32} & \lambda a_{23} + a_{33} \end{bmatrix}.$$

§5.1 Introduction - Review on Linear Algebra

1. If matrix A is invertible, then there exists a sequence of elementary row operations can be applied to A, reducing it to the identity matrix I,
$$E_m E_{m-1} \cdots E_2 E_1 A = I.$$
2. This gives us an equation for computing the inverse of a matrix:
$$A^{-1} = E_m E_{m-1} \cdots E_2 E_1 = E_m E_{m-1} \cdots E_2 E_1 I.$$

Remark: This is not a practical method to compute A^{-1}.

§5.1 Introduction - Review on Linear Algebra

Let $A \in \mathbb{C}^{n\times n}$ be a square matrix. If there exists a nonzero vector $\boldsymbol{x} \in \mathbb{C}^n$ and a scalar $\lambda \in \mathbb{C}$ such that

$$A\boldsymbol{x} = \lambda\boldsymbol{x},$$

then λ is called an **eigenvalue** of A and $\boldsymbol{x}$ is called the corresponding **eigenvector** of A.

Remark: Computing λ and $\boldsymbol{x}$ is a major task in numerical linear algebra.

§5.1 Introduction - Review on Linear Algebra

For an $n \times n$ real matrix A, the following properties are equivalent:

1. The inverse of A exists; that is, A is nonsingular;
2. The determinant of A is nonzero;
3. The rows of A form a basis for $\mathbb{R}^n$;
4. The columns of A form a basis for $\mathbb{R}^n$;
5. As a map from $\mathbb{R}^n$ to $\mathbb{R}^n$, A is injective (one to one);
6. As a map from $\mathbb{R}^n$ to $\mathbb{R}^n$, A is surjective (onto);
7. The equation $A\boldsymbol{x} = 0$ implies $\boldsymbol{x} = \mathbf{0}$;
8. For each $\boldsymbol{b} \in \mathbb{R}^n$, there is exactly one $x \in \mathbb{R}^n$ such that $A\boldsymbol{x} = \boldsymbol{b}$;
9. A is a product of elementary matrices;
10. 0 is not an eigenvalue of A.

§5.1 Introduction - Review on Linear Algebra

There are some easy-to-solve systems:

1. **Diagonal Structure**

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}.$$

The solution is: (provided $a_{ii} + 0$ for all $i = 1, 2, \cdots, n$)

$$\boldsymbol{x} = \left(\frac{b_1}{a_{11}}, \frac{b_2}{a_{22}}, \frac{b_3}{a_{33}}, \cdots, \frac{b_n}{a_{nn}} \right)^\top.$$

- If $a_{ii} = 0$ for some index i, and if $b_i = 0$ also, then x_i can be any real number. The number of solutions is infinity.
- If $a_{ii} = 0$ and $b_i + 0$, no solution of the system exists.
- What is the complexity of the method? n divisions.

§5.1 Introduction - Review on Linear Algebra

There are some easy-to-solve systems:

2. **Lower Triangular Systems**

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}.$$

Some simple observations:

- If $a_{11} \neq 0$, then we have $x_1 = b_1/a_{11}$.
- Once we have x_1, we can simplify the second equation, $x_2 = (b_2 - a_{21}x_1)/a_{22}$, provided that $a_{22} \neq 0$.

  Similarly, $x_3 = (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}$, provided that $a_{33} \neq 0$.

  In general, to find the solution to this system, we use forward substitution (assume that $a_{ii} \neq 0$ for all i).

§5.1 Introduction - Review on Linear Algebra

There are some easy-to-solve systems:

2. **Lower Triangular Systems (cont'd)**
   - Algorithm of forward substitution:

     input n, (a_{ij}), $b = (b_1, b_2, \cdots, b_n)^\top$
     for $i = 1$ **to** n **do**
     $$x_i \leftarrow \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j\right) / a_{ii}$$
     end do
     output $x = (x_1, x_2, \cdots, x_n)^\top$
   - Complexity of forward substitution:
     - n divisions.
     - the number of multiplications: 0 for x_1, 1 for x_2, 2 for x_3, $\cdots$ total $= 0 + 1 + 2 + \cdots + (n-1) \approx (n+1)n/2 = \mathcal{O}(n^2)$.
     - the number of subtractions: same as the number of multiplications $= \mathcal{O}(n^2)$.

     Forward substitution is an $\mathcal{O}(n^2)$ algorithm.
   - **Remark:** forward substitution is a sequential algorithm (not parallel at all).

§5.1 Introduction - Review on Linear Algebra

There are some easy-to-solve systems:

3. **Upper Triangular Systems**

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}.$$

- The formal algorithm to solve for x is called backward substitution. It is also an $\mathcal{O}(n^2)$ algorithm.
- Assume that $a_{ii} \neq 0$ for all i. Algorithm:

input n, (a_{ij}), $b = (b_1, b_2, \cdots, b_n)^\top$
for $i = n : -1 : 1$ **do**

$$x_i \leftarrow \Big(b_i - \sum_{j=i+1}^{n} a_{ij}x_j\Big)/a_{ii}$$

end do
output $x = (x_1, x_2, \cdots, x_n)^\top$

§5.2 LU Decomposition

LU decomposition (factorization):

Suppose that A can be factored into the product of a lower triangular matrix L and an upper triangular matrix U:

$$A = LU.$$

Then, $A\mathbf{x} = LU\mathbf{x} = L(U\mathbf{x})$. Thus, to solve the system of equations $A\mathbf{x} = \mathbf{b}$, it is enough to solve this problem in two stages:

$$\begin{aligned} L\mathbf{z} &= \mathbf{b} \qquad \textit{solve for } \ \mathbf{z}, \\ U\mathbf{x} &= \mathbf{z} \qquad \textit{solve for } \ \mathbf{x}. \end{aligned}$$

§5.2 LU Decomposition

Example (Basic Gaussian elimination)

Let $A^{(1)} = (a_{ij}^{(1)}) = A = (a_{ij})$ and $b^{(1)} = b$. Consider the following linear system $A\boldsymbol{x} = \boldsymbol{b}$:

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 34 \\ 27 \\ -38 \end{bmatrix}.$$

pivot row $=$ row1; pivot element: $a_{ij}^{(1)} = 6$.

row2 $-$ $(12/6)\times$row1 $\to$ row2.

row3 $-$ $(3/6)\times$row1 $\to$ row3.

row4 $-$ $(-6/6)\times$row1 $\to$ row4.

$$\Longrightarrow \begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & -12 & 8 & 1 \\ 0 & 2 & 3 & -14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \\ 21 \\ -26 \end{bmatrix}.$$

multipliers: $12/6, \quad 3/6, \quad (-6)/6$

§5.2 *LU* Decomposition

Example (Basic Gaussian elimination - cont'd)

We have the following equivalent system $A^{(2)}\boldsymbol{x} = \boldsymbol{b}^{(2)}$:

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & -12 & 8 & 1 \\ 0 & 2 & 3 & -14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \\ 21 \\ -26 \end{bmatrix}.$$

pivot row $=$ row2; pivot element $a_{22}^{(2)} = -4$.

row3 $-$ $(-12/-4)\times$row2 $\rightarrow$ row3.

row4 $-$ $(2/-4)\times$row2 $\rightarrow$ row4.

$$\Longrightarrow \begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 4 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \\ -9 \\ -21 \end{bmatrix}.$$

multiplier: $(-12)/(-4), \quad 2/(-4)$

§5.2 *LU* Decomposition

Example (Basic Gaussian elimination - cont'd)

We have the following equivalent system $A^{(3)}\boldsymbol{x} = \boldsymbol{b}^{(3)}$:

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 4 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \\ -9 \\ -21 \end{bmatrix}.$$

pivot row = row3; pivot element $a_{33}^{(3)} = 2$.

row4 − $(4/2)\times$row3 → row4.

$$\Longrightarrow \begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \\ -9 \\ -3 \end{bmatrix}.$$

multiplier: $4/2$

§5.2 LU Decomposition

Example (Basic Gaussian elimination - cont'd)

Finally, we have the following equivalent upper triangular system $A^{(4)}\boldsymbol{x} = \boldsymbol{b}^{(4)}$:

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 10 \\ -9 \\ -3 \end{bmatrix}.$$

Using the backward substitution, we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ -2 \\ 1 \end{bmatrix}.$$

§5.2 LU Decomposition

Example (Basic Gaussian elimination - cont'd)

Display the multipliers in an unit lower triangular matrix $L = (\ell_{ij})$:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ \frac{1}{2} & 3 & 1 & 0 \\ -1 & -\frac{1}{2} & 2 & 1 \end{bmatrix}.$$

Let $U = (u_{ij})$ be the final upper triangular matrix $A^{(4)}$. Then we have

$$U = \begin{bmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

and one can check that $A = LU$ (the Doolittle Decomposition).

§5.2 LU Decomposition

Remark:

1. The entire elimination process will break down if any of the pivot elements are 0.
2. The total number of arithmetic operations:
   - multiplication and division $= \dfrac{n^3}{3} - \dfrac{n}{3} \qquad \left(\sum_{k=1}^{n-1} k(k+1) \right)$;
   - addition and subtraction $= \dfrac{n^3}{3} - \dfrac{n^2}{2} + \dfrac{n}{6} \qquad \left(\sum_{k=1}^{n-1} k^2 \right)$.

   Therefore, the Gauss Elimination is an $\mathcal{O}(n^3)$ algorithm.

§5.3 Norms on Vectors and Matrices

Definition

A ***nomed vector space*** $(\mathcal{V}, \|\cdot\|)$ is a vector space $\mathcal{V}$ over field $\mathbb{F}$ associated with a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ such that

1. $\|\boldsymbol{x}\| \geqslant 0$ for all $\boldsymbol{x} \in \mathcal{V}$.
2. $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \mathbf{0}$.
3. $\|\lambda \cdot \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$ for all $\lambda \in \mathbb{F}$ and $\boldsymbol{x} \in \mathcal{V}$.
4. $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$.

A function $\|\cdot\|$ satisfies ①–④ is called a ***norm*** on $\mathcal{V}$.

Remark: The norm of a vector can be viewed as the length of that vector. Moreover, the norm induces the concept of distance on the vector space: the distance between two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in a normed vector space $(\mathcal{V}, \|\cdot\|)$ is defined by $d(\boldsymbol{x}, \boldsymbol{y}) \equiv \|\boldsymbol{x} - \boldsymbol{y}\|$.

§5.3 Norms on Vectors and Matrices

Definition

A ***nomed vector space*** $(\mathcal{V}, \|\cdot\|)$ is a vector space $\mathcal{V}$ over field $\mathbb{F}$ associated with a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ such that

1. $\|\boldsymbol{x}\| \geqslant 0$ for all $\boldsymbol{x} \in \mathcal{V}$.
2. $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \mathbf{0}$.
3. $\|\lambda \cdot \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$ for all $\lambda \in \mathbb{F}$ and $\boldsymbol{x} \in \mathcal{V}$.
4. $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$.

A function $\|\cdot\|$ satisfies ①–④ is called a ***norm*** on $\mathcal{V}$.

Remark: The norm of a vector can be viewed as the length of that vector. Moreover, the norm induces the concept of distance on the vector space: the distance between two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in a normed vector space $(\mathcal{V}, \|\cdot\|)$ is defined by $d(\boldsymbol{x}, \boldsymbol{y}) \equiv \|\boldsymbol{x} - \boldsymbol{y}\|$.

§5.3 Norms on Vectors and Matrices

Example

1. Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^\top \in \mathbb{R}^n$:
   - The 2-norm (Euclidean norm, or ℓ^2 norm): $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$
   - The infinity norm (ℓ^∞-norm): $\|\boldsymbol{x}\|_\infty = \max_{1\leqslant i\leqslant n} |x_i|$
   - The 1-norm (ℓ^1-norm): $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$
   - The p-norm (ℓ^p-norm), $1 \leqslant p < \infty$, is $\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$
2. Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^\top$, $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^\top \in \mathbb{R}^n$. Then
   - $\|\boldsymbol{x} - \boldsymbol{y}\|_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$
   - $\|\boldsymbol{x} - \boldsymbol{y}\|_\infty = \max_{1\leqslant i\leqslant n} |x_i - y_i|$
   - $\|\boldsymbol{x} - \boldsymbol{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|$
   - $\|\boldsymbol{x} - \boldsymbol{y}\|_p = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{\frac{1}{p}}$

§5.3 Norms on Vectors and Matrices

Example

$(\mathbb{R}^2, \|\cdot\|_p)$ is a normed vector space. Consider the ball centered at $\boldsymbol{x}_0 = \mathbf{0}$ with radius 1 and $p = 1$, $p = 2$ and $p = \infty$ respectively.

1. $p = 1$: $\|\boldsymbol{x} - \boldsymbol{x}_0\|_1 = |x_1| + |x_2|$.
2. $p = 2$: $\|\boldsymbol{x} - \boldsymbol{x}_0\|_2 = \sqrt{x_1^2 + x_2^2}$.
3. $p = \infty$: $\|\boldsymbol{x} - \boldsymbol{x}_0\|_\infty = \max\left\{|x_1|, |x_2|\right\}$.

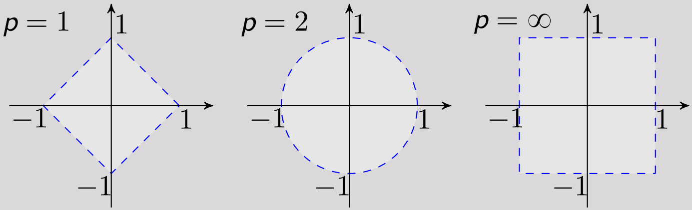

Figure 1: The 1-ball about 0 in $\mathbb{R}^2$ with different p

§5.3 Norms on Vectors and Matrices

Example

Let A be an invertible $n \times n$ matrix. For a given norm $\|\cdot\|_{\mathbb{R}^n}$ on $\mathbb{R}^n$, define a map $|||\cdot||| : \mathbb{R}^n \to \mathbb{R}$ by

$$|||\boldsymbol{x}||| = \|A\boldsymbol{x}\|_{\mathbb{R}^n} .$$

Then $|||\cdot|||$ is a norm on $\mathbb{R}^n$.

Definition

Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $\{\boldsymbol{x}^{(n)}\}_{n=1}^{\infty}$ be a sequence in $\mathcal{V}$. Then $\{\boldsymbol{x}^{(n)}\}_{n=1}^{\infty}$ is said to converge to a vector $\boldsymbol{x} \in \mathcal{V}$, denoted by $\lim\limits_{n\to\infty} \boldsymbol{x}^{(n)} = \boldsymbol{x}$, if for every $\varepsilon > 0$, there exists $N > 0$ such that

$$\left\|\boldsymbol{x}^{(n)} - \boldsymbol{x}\right\| < \varepsilon \qquad \text{whenever} \quad n \geqslant N.$$

Sequence $\{\boldsymbol{x}^{(n)}\}_{n=1}^{\infty}$ in $\mathcal{V}$ is said to be convergent if there exists $\boldsymbol{x} \in \mathcal{V}$ such that $\lim\limits_{n\to\infty} \boldsymbol{x}^{(n)} = \boldsymbol{x}$.

§5.3 Norms on Vectors and Matrices

Example

Let A be an invertible $n \times n$ matrix. For a given norm $\|\cdot\|_{\mathbb{R}^n}$ on $\mathbb{R}^n$, define a map $|||\cdot||| : \mathbb{R}^n \to \mathbb{R}$ by

$$|||\boldsymbol{x}||| = \|A\boldsymbol{x}\|_{\mathbb{R}^n} .$$

Then $|||\cdot|||$ is a norm on $\mathbb{R}^n$.

Definition

Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $\{\boldsymbol{x}^{(n)}\}_{n=1}^{\infty}$ be a sequence in $\mathcal{V}$. Then $\{\boldsymbol{x}^{(n)}\}_{n=1}^{\infty}$ is said to converge to a vector $\boldsymbol{x} \in \mathcal{V}$, denoted by $\lim\limits_{n\to\infty} \boldsymbol{x}^{(n)} = \boldsymbol{x}$, if for every $\varepsilon > 0$, there exists $N > 0$ such that

$$\left\|\boldsymbol{x}^{(n)} - \boldsymbol{x}\right\| < \varepsilon \qquad \text{whenever} \quad n \geqslant N.$$

Sequence $\{\boldsymbol{x}^{(n)}\}_{n=1}^{\infty}$ in $\mathcal{V}$ is said to be convergent if there exists $\boldsymbol{x} \in \mathcal{V}$ such that $\lim\limits_{n\to\infty} \boldsymbol{x}^{(n)} = \boldsymbol{x}$.

§5.3 Norms on Vectors and Matrices

Definition

Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be normed vector spaces, $A \subseteq \mathcal{V}$, and $f : A \to \mathcal{W}$ be a $\mathcal{W}$-valued function. f is said to be continuous at $a \in A$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|f(\boldsymbol{x}) - f(\boldsymbol{a})\|_{\mathcal{W}} < \varepsilon \qquad \text{whenever} \quad \|\boldsymbol{x} - \boldsymbol{a}\|_{\mathcal{V}} < \delta \text{ and } \boldsymbol{x} \in A\,.$$

Definition

Two norms $\|\cdot\|$ and $|||\cdot|||$ on a vector space $\mathcal{V}$ are called equivalent if there are positive constants C_1 and C_2 such that

$$C_1\|\boldsymbol{x}\| \leqslant |||\boldsymbol{x}||| \leqslant C_2\|\boldsymbol{x}\| \qquad \forall\, \boldsymbol{x} \in \mathcal{V}\,.$$

Remark: Equivalent norms induce the same concept of convergence of sequences, continuity of functions, and so on. For example, if $\{x^{(k)}\}_{k=1}^{\infty}$ is a convergent sequence in $(\mathcal{V}, \|\cdot\|_1)$ and $\|\cdot\|_2$ is an equivalent norm of $\|\cdot\|_1$, then $\{x^{(k)}\}_{k=1}^{\infty}$ is convergent in $(\mathcal{V}, \|\cdot\|_2)$.

§5.3 Norms on Vectors and Matrices

Definition

Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be normed vector spaces, $A \subseteq \mathcal{V}$, and $f : A \to \mathcal{W}$ be a $\mathcal{W}$-valued function. f is said to be continuous at $a \in A$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|f(\boldsymbol{x}) - f(\boldsymbol{a})\|_{\mathcal{W}} < \varepsilon \qquad \text{whenever} \quad \|\boldsymbol{x} - \boldsymbol{a}\|_{\mathcal{V}} < \delta \text{ and } \boldsymbol{x} \in A\,.$$

Definition

Two norms $\|\cdot\|$ and $|||\cdot|||$ on a vector space $\mathcal{V}$ are called equivalent if there are positive constants C_1 and C_2 such that

$$C_1\|\boldsymbol{x}\| \leqslant |||\boldsymbol{x}||| \leqslant C_2\|\boldsymbol{x}\| \qquad \forall\, \boldsymbol{x} \in \mathcal{V}\,.$$

Remark: Equivalent norms induce the same concept of convergence of sequences, continuity of functions, and so on. For example, if $\{x^{(k)}\}_{k=1}^{\infty}$ is a convergent sequence in $(\mathcal{V}, \|\cdot\|_1)$ and $\|\cdot\|_2$ is an equivalent norm of $\|\cdot\|_1$, then $\{x^{(k)}\}_{k=1}^{\infty}$ is convergent in $(\mathcal{V}, \|\cdot\|_2)$.

§5.3 Norms on Vectors and Matrices

Definition

Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be normed vector spaces, $A \subseteq \mathcal{V}$, and $f : A \to \mathcal{W}$ be a $\mathcal{W}$-valued function. f is said to be continuous at $a \in A$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|f(\boldsymbol{x}) - f(\boldsymbol{a})\|_{\mathcal{W}} < \varepsilon \qquad \text{whenever} \quad \|\boldsymbol{x} - \boldsymbol{a}\|_{\mathcal{V}} < \delta \text{ and } \boldsymbol{x} \in A\,.$$

Definition

Two norms $\|\cdot\|$ and $|||\cdot|||$ on a vector space $\mathcal{V}$ are called equivalent if there are positive constants C_1 and C_2 such that

$$C_1\|\boldsymbol{x}\| \leqslant |||\boldsymbol{x}||| \leqslant C_2\|\boldsymbol{x}\| \qquad \forall\, \boldsymbol{x} \in \mathcal{V}\,.$$

Remark: Equivalent norms induce the same concept of convergence of sequences, continuity of functions, and so on. For example, if $\{x^{(k)}\}_{k=1}^{\infty}$ is a convergent sequence in $(\mathcal{V}, \|\cdot\|_1)$ and $\|\cdot\|_2$ is an equivalent norm of $\|\cdot\|_1$, then $\{x^{(k)}\}_{k=1}^{\infty}$ is convergent in $(\mathcal{V}, \|\cdot\|_2)$.

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof.

Let $\{\boldsymbol{e}_k\}_{k=1}^N$ be a basis of $\mathcal{V}$. For each $\boldsymbol{x} \in \mathcal{V}$, we write $\boldsymbol{x} = \sum_{k=1}^{N} x_k \boldsymbol{e}_k$ and define a function $\|\cdot\|_2 : \mathcal{V} \to \mathbb{R}$ by

$$\|\boldsymbol{x}\|_2 = \Big(\sum_{k=1}^{N} |x_k|^2\Big)^{\frac{1}{2}}.$$

Then

1. $\|\boldsymbol{x}\|_2 \geqslant 0$ for all $\boldsymbol{x} \in \mathcal{V}$, and $\|\boldsymbol{x}\|_2 = 0$ if and only if $\boldsymbol{x} = \mathbf{0}$.
2. $\|\lambda \boldsymbol{x}\|_2 = |\lambda| \|\boldsymbol{x}\|_2$ for all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) and $\boldsymbol{x} \in \mathcal{V}$.
3. $\|\boldsymbol{x} + \boldsymbol{y}\|_2 \leqslant \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$ becuase of the Cauchy -Schwarz inequality. □

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof (cont'd).

Therefore, $\|\cdot\|_2$ is a normed on $\mathcal{V}$. It then suffices to shows that any norm $\|\cdot\|$ on $\mathcal{V}$ is equivalent to $\|\cdot\|_2$:

$$\text{if}\quad C_1\|\boldsymbol{x}\| \leqslant \|\boldsymbol{x}\|_2 \leqslant C_2\|\boldsymbol{x}\| \quad\text{and}\quad C_3|||\boldsymbol{x}||| \leqslant \|\boldsymbol{x}\|_2 \leqslant C_4|||\boldsymbol{x}|||,$$

$$\text{then}\quad \frac{C_1}{C_4}\|\boldsymbol{x}\| \leqslant |||\boldsymbol{x}||| \leqslant \frac{C_2}{C_3}\|\boldsymbol{x}\|.$$

By the definition of norms and the Cauchy-Schwarz inequality,

$$\|\boldsymbol{x}\| \leqslant \sum_{k=1}^{N} |x_k|\|\boldsymbol{e}_k\| \leqslant \|\boldsymbol{x}\|_2 \Big(\sum_{k=1}^{N} \|\boldsymbol{e}_k\|^2\Big)^{\frac{1}{2}};$$

thus letting $C_2 = \Big(\sum_{k=1}^{N} \|\boldsymbol{e}_k\|^2\Big)^{\frac{1}{2}}$ we have $\|\boldsymbol{x}\| \leqslant C_2\|\boldsymbol{x}\|_2$. □

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof (cont'd).

Therefore, $\|\cdot\|_2$ is a normed on $\mathcal{V}$. It then suffices to shows that any norm $\|\cdot\|$ on $\mathcal{V}$ is equivalent to $\|\cdot\|_2$:

$$\text{if} \quad C_1\|\boldsymbol{x}\| \leqslant \|\boldsymbol{x}\|_2 \leqslant C_2\|\boldsymbol{x}\| \quad \text{and} \quad C_3|||\boldsymbol{x}||| \leqslant \|\boldsymbol{x}\|_2 \leqslant C_4|||\boldsymbol{x}|||,$$

$$\text{then} \quad \frac{C_1}{C_4}\|\boldsymbol{x}\| \leqslant |||\boldsymbol{x}||| \leqslant \frac{C_2}{C_3}\|\boldsymbol{x}\|.$$

By the definition of norms and the Cauchy-Schwarz inequality,

$$\|\boldsymbol{x}\| \leqslant \sum_{k=1}^{N} |x_k|\|\boldsymbol{e}_k\| \leqslant \|\boldsymbol{x}\|_2 \Big(\sum_{k=1}^{N} \|\boldsymbol{e}_k\|^2\Big)^{\frac{1}{2}};$$

thus letting $C_2 = \Big(\sum_{k=1}^{N} \|\boldsymbol{e}_k\|^2\Big)^{\frac{1}{2}}$ we have $\|\boldsymbol{x}\| \leqslant C_2\|\boldsymbol{x}\|_2$. □

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof (cont'd).

Define $f\colon (\mathcal{V}, \|\cdot\|_2) \to \mathbb{R}$ by $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$. Then

$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| = \big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant C_2\|\boldsymbol{x} - \boldsymbol{y}\|_2$$

which implies that f is continuous. Let $\mathbb{S}^{n-1}$ be the unit sphere $\big\{\boldsymbol{x} \in \mathcal{V} \,\big|\, \|\boldsymbol{x}\|_2 = 1\big\}$. Then $\mathbb{S}^{n-1}$ is (sequentially) compact in $(\mathcal{V}, \|\cdot\|_2)$, so f attains its minimum on $\mathbb{S}^{n-1}$. Suppose that $\min\limits_{\boldsymbol{x}\in\mathbb{S}^{n-1}} f(\boldsymbol{x}) = f(\boldsymbol{a})$ for some $\boldsymbol{a} \in \mathbb{S}^{n-1}$. Then $f(\boldsymbol{a}) > 0$ (for otherwise $\boldsymbol{a} = \mathbf{0}$), and

$$\Big\|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big\| = f\Big(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big) \geqslant f(\boldsymbol{a}) \qquad \forall\, \boldsymbol{x} \in \mathcal{V}$$

which implies that $\|\boldsymbol{x}\| \geqslant C_1\|\boldsymbol{x}\|_2$ for $C_1 = f(\boldsymbol{a})$. □

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof (cont'd).

Define $f\colon (\mathcal{V}, \|\cdot\|_2) \to \mathbb{R}$ by $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$. Then

$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| = \big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant C_2\|\boldsymbol{x} - \boldsymbol{y}\|_2$$

which implies that f is continuous. Let $\mathbb{S}^{n-1}$ be the unit sphere $\big\{\boldsymbol{x} \in \mathcal{V} \,\big|\, \|\boldsymbol{x}\|_2 = 1\big\}$. Then $\mathbb{S}^{n-1}$ is (sequentially) compact in $(\mathcal{V}, \|\cdot\|_2)$, so f attains its minimum on $\mathbb{S}^{n-1}$. Suppose that $\min\limits_{\boldsymbol{x}\in\mathbb{S}^{n-1}} f(\boldsymbol{x}) = f(\boldsymbol{a})$ for some $\boldsymbol{a} \in \mathbb{S}^{n-1}$. Then $f(\boldsymbol{a}) > 0$ (for otherwise $\boldsymbol{a} = \boldsymbol{0}$), and

$$\left\|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right\| = f\Big(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big) \geqslant f(\boldsymbol{a}) \qquad \forall\, \boldsymbol{x} \in \mathcal{V}$$

which implies that $\|\boldsymbol{x}\| \geqslant C_1\|\boldsymbol{x}\|_2$ for $C_1 = f(\boldsymbol{a})$. $\square$

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof (cont'd).

Define $f\colon (\mathcal{V}, \|\cdot\|_2) \to \mathbb{R}$ by $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$. Then

$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| = \big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant C_2\|\boldsymbol{x} - \boldsymbol{y}\|_2$$

which implies that f is continuous. Let $\mathbb{S}^{n-1}$ be the unit sphere $\big\{\boldsymbol{x} \in \mathcal{V} \,\big|\, \|\boldsymbol{x}\|_2 = 1\big\}$. Then $\mathbb{S}^{n-1}$ is (sequentially) compact in $(\mathcal{V}, \|\cdot\|_2)$, so f attains its minimum on $\mathbb{S}^{n-1}$. Suppose that $\min\limits_{\boldsymbol{x}\in\mathbb{S}^{n-1}} f(\boldsymbol{x}) = f(\boldsymbol{a})$ for some $\boldsymbol{a} \in \mathbb{S}^{n-1}$. Then $f(\boldsymbol{a}) > 0$ (for otherwise $\boldsymbol{a} = \mathbf{0}$), and

$$\Big\|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big\| = f\Big(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big) \geqslant f(\boldsymbol{a}) \qquad \forall\, \boldsymbol{x} \in \mathcal{V}$$

which implies that $\|\boldsymbol{x}\| \geqslant C_1\|\boldsymbol{x}\|_2$ for $C_1 = f(\boldsymbol{a})$. □

§5.3 Norms on Vectors and Matrices

Theorem

Any two norms on a finite dimensional real (or complex) normed vector space $\mathcal{V}$ are equivalent.

Proof (cont'd).

Define $f \colon (\mathcal{V}, \|\cdot\|_2) \to \mathbb{R}$ by $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$. Then

$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| = \big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant C_2\|\boldsymbol{x} - \boldsymbol{y}\|_2$$

which implies that f is continuous. Let $\mathbb{S}^{n-1}$ be the unit sphere $\big\{\boldsymbol{x} \in \mathcal{V} \,\big|\, \|\boldsymbol{x}\|_2 = 1\big\}$. Then $\mathbb{S}^{n-1}$ is (sequentially) compact in $(\mathcal{V}, \|\cdot\|_2)$, so f attains its minimum on $\mathbb{S}^{n-1}$. Suppose that $\min\limits_{\boldsymbol{x} \in \mathbb{S}^{n-1}} f(\boldsymbol{x}) = f(\boldsymbol{a})$ for some $\boldsymbol{a} \in \mathbb{S}^{n-1}$. Then $f(\boldsymbol{a}) > 0$ (for otherwise $\boldsymbol{a} = \mathbf{0}$), and

$$\Big\|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big\| = f\Big(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big) \geqslant f(\boldsymbol{a}) \qquad \forall\, \boldsymbol{x} \in \mathcal{V}$$

which implies that $\|\boldsymbol{x}\| \geqslant C_1\|\boldsymbol{x}\|_2$ for $C_1 = f(\boldsymbol{a})$. $\square$

§5.3 Norms on Vectors and Matrices

Theorem

Let $\|\cdot\|_{\mathbb{R}^n}$ be a norm on $\mathbb{R}^n$ and $\|\cdot\|_{\mathbb{R}^m}$ be a norm on $\mathbb{R}^m$. Then

$$\|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\|A\boldsymbol{x}\|_{\mathbb{R}^m} : \boldsymbol{x} \in \mathbb{R}^n, \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}$$

defines a norm on the vector space of all $m \times n$ real matrices.

Proof.

1. Clearly $\|A\|_{\mathbb{R}^n,\mathbb{R}^m} \geqslant 0$, and $\|A\| = 0$ if and only if $A = 0$.
2. $$\begin{aligned}\|\lambda A\|_{\mathbb{R}^n,\mathbb{R}^m} &= \max\left\{\|\lambda A\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}\\ &= \max\left\{|\lambda|\|A\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}\\ &= |\lambda| \max\left\{\|A\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\} = |\lambda|\|A\|_{\mathbb{R}^n,\mathbb{R}^m}.\end{aligned}$$
3. $$\begin{aligned}\|A+B\|_{\mathbb{R}^n,\mathbb{R}^m} &= \max\left\{\|(A+B)\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}\\ &\leqslant \max\left\{\|A\boldsymbol{x}\|_{\mathbb{R}^m} + \|B\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}\\ &\leqslant \max\left\{\|A\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\} + \max\left\{\|B\boldsymbol{x}\|_{\mathbb{R}^m} : \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}\\ &= \|A\|_{\mathbb{R}^n,\mathbb{R}^m} + \|B\|_{\mathbb{R}^n,\mathbb{R}^m}.\end{aligned}$$

$\square$

§5.3 Norms on Vectors and Matrices

Remark:

1. $\|\cdot\|_{\mathbb{R}^n,\mathbb{R}^m}$ is called the matrix norm induced by vector norms $\|\cdot\|_{\mathbb{R}^n}$ and $\|\cdot\|_{\mathbb{R}^m}$. Moreover,
$$\|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\|A\mathbf{x}\|_{\mathbb{R}^m} : \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_{\mathbb{R}^n} = 1\right\}$$
$$\Leftrightarrow \|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\frac{\|A\mathbf{x}\|_{\mathbb{R}^m}}{\|\mathbf{x}\|_{\mathbb{R}^n}} : \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}\right\}$$
2. If $\|\cdot\|_{\mathbb{R}^n} = \|\cdot\|_p$ and $\|\cdot\|_{\mathbb{R}^m} = \|\cdot\|_q$, then $\|A\|_{\mathbb{R}^n,\mathbb{R}^m}$ is simply denoted by $\|A\|_{p,q}$. If in addition $p = q$, then $\|A\|_{p,q}$ is simply denoted by $\|A\|_p$.

Theorem (Additional properties of matrix norms)

Let A be a $m \times n$ matrix, and B be a $n \times k$ matrix.

1. $\|A\mathbf{x}\|_{\mathbb{R}^m} \leqslant \|A\|_{\mathbb{R}^n,\mathbb{R}^m}\|\mathbf{x}\|_{\mathbb{R}^n}$ *for all* $\mathbf{x} \in \mathbb{R}^n$ (*sub-ordinance*)
2. $\|AB\|_{\mathbb{R}^k,\mathbb{R}^m} \leqslant \|A\|_{\mathbb{R}^n,\mathbb{R}^m}\|B\|_{\mathbb{R}^k,\mathbb{R}^n}$ (*sub-multiplicativity*)
3. $\|I_{n\times n}\|_p = 1$ *for all* $p \in [1,\infty]$.

§5.3 Norms on Vectors and Matrices

Remark:

1. $\|\cdot\|_{\mathbb{R}^n,\mathbb{R}^m}$ is called the matrix norm induced by vector norms $\|\cdot\|_{\mathbb{R}^n}$ and $\|\cdot\|_{\mathbb{R}^m}$. Moreover,
$$\|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\|A\boldsymbol{x}\|_{\mathbb{R}^m} : \boldsymbol{x} \in \mathbb{R}^n, \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}$$
$$\Leftrightarrow \|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\frac{\|A\boldsymbol{x}\|_{\mathbb{R}^m}}{\|\boldsymbol{x}\|_{\mathbb{R}^n}} : \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{0}\right\}$$
2. If $\|\cdot\|_{\mathbb{R}^n} = \|\cdot\|_p$ and $\|\cdot\|_{\mathbb{R}^m} = \|\cdot\|_q$, then $\|A\|_{\mathbb{R}^n,\mathbb{R}^m}$ is simply denoted by $\|A\|_{p,q}$. If in addition $p = q$, then $\|A\|_{p,q}$ is simply denoted by $\|A\|_p$.

Theorem (Additional properties of matrix norms)

Let A be a $m \times n$ matrix, and B be a $n \times k$ matrix.

1. $\|A\boldsymbol{x}\|_{\mathbb{R}^m} \leqslant \|A\|_{\mathbb{R}^n,\mathbb{R}^m}\|\boldsymbol{x}\|_{\mathbb{R}^n}$ *for all* $\boldsymbol{x} \in \mathbb{R}^n$ (*sub-ordinance*)
2. $\|AB\|_{\mathbb{R}^k,\mathbb{R}^m} \leqslant \|A\|_{\mathbb{R}^n,\mathbb{R}^m}\|B\|_{\mathbb{R}^k,\mathbb{R}^n}$ (*sub-multiplicativity*)
3. $\|I_{n\times n}\|_p = 1$ *for all* $p \in [1,\infty]$.

§5.3 Norms on Vectors and Matrices

Remark:

1. $\|\cdot\|_{\mathbb{R}^n,\mathbb{R}^m}$ is called the matrix norm induced by vector norms $\|\cdot\|_{\mathbb{R}^n}$ and $\|\cdot\|_{\mathbb{R}^m}$. Moreover,
$$\|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\|A\boldsymbol{x}\|_{\mathbb{R}^m} : \boldsymbol{x}\in\mathbb{R}^n, \|\boldsymbol{x}\|_{\mathbb{R}^n} = 1\right\}$$
$$\Leftrightarrow \|A\|_{\mathbb{R}^n,\mathbb{R}^m} \equiv \max\left\{\frac{\|A\boldsymbol{x}\|_{\mathbb{R}^m}}{\|\boldsymbol{x}\|_{\mathbb{R}^n}} : \boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{x}\neq\mathbf{0}\right\}$$
2. If $\|\cdot\|_{\mathbb{R}^n} = \|\cdot\|_p$ and $\|\cdot\|_{\mathbb{R}^m} = \|\cdot\|_q$, then $\|A\|_{\mathbb{R}^n,\mathbb{R}^m}$ is simply denoted by $\|A\|_{p,q}$. If in addition $p=q$, then $\|A\|_{p,q}$ is simply denoted by $\|A\|_p$.

Theorem (Additional properties of matrix norms)

Let A be a $m\times n$ matrix, and B be a $n\times k$ matrix.

1. $\|A\boldsymbol{x}\|_{\mathbb{R}^m} \leqslant \|A\|_{\mathbb{R}^n,\mathbb{R}^m}\|\boldsymbol{x}\|_{\mathbb{R}^n}$ *for all* $\boldsymbol{x}\in\mathbb{R}^n$ $\quad(\|A\boldsymbol{x}\| \leqslant \|A\|\|\boldsymbol{x}\|)$
2. $\|AB\|_{\mathbb{R}^k,\mathbb{R}^m} \leqslant \|A\|_{\mathbb{R}^n,\mathbb{R}^m}\|B\|_{\mathbb{R}^k,\mathbb{R}^n}$ $\quad(\|AB\| \leqslant \|A\|\|B\|)$
3. $\|I_{n\times n}\|_p = 1$ *for all* $p\in[1,\infty]$.

§5.3 Norms on Vectors and Matrices

Example ($\|A\|_\infty$)

Let $A = \left[a_{ij}\right]_{m\times n}$ and $\max\limits_{1\leqslant i\leqslant m} \sum\limits_{j=1}^{n} |a_{ij}| = \sum\limits_{j=1}^{n} |a_{kj}|$ for some $1 \leqslant k \leqslant m$.

1. Let $\boldsymbol{x} = \big(\operatorname{sgn}(a_{k1}), \operatorname{sgn}(a_{k2}), \cdots, \operatorname{sgn}(a_{kn})\big)$. Then $\|\boldsymbol{x}\|_\infty = 1$, and $\|A\boldsymbol{x}\|_\infty = \sum\limits_{j=1}^{n} |a_{kj}|$.
2. Let $\boldsymbol{x} = (x_1, \cdots, x_n)$. If $\|\boldsymbol{x}\|_\infty = 1$, then $|x_j| \leqslant 1$ for all $1 \leqslant j \leqslant n$; thus
$$|a_{i1}x_1 + a_{i2}x_2 + \cdots a_{in}x_n| \leqslant \sum_{j=1}^{n} |a_{ij}| \leqslant \sum_{j=1}^{n} |a_{kj}| .$$

By the definition of matrix norms, ① implies that $\|A\|_\infty \geqslant \sum\limits_{j=1}^{n} |a_{kj}|$ while ② implies that $\|A\|_\infty \leqslant \sum\limits_{j=1}^{n} |a_{kj}|$. Therefore,
$$\|A\|_\infty = \max\left\{\sum_{j=1}^{n} |a_{1j}|, \sum_{j=1}^{n} |a_{2j}|, \cdots, \sum_{j=1}^{n} |a_{nj}|\right\};$$
that is, $\|A\|_\infty$ is the largest sum of the absolute value of row entries.

§5.3 Norms on Vectors and Matrices

Theorem

For each $\boldsymbol{x} \in \mathbb{R}^n$,

$$\|\boldsymbol{x}\|_1 = \max\left\{\boldsymbol{y}^\top \boldsymbol{x} : \|\boldsymbol{y}\|_\infty = 1\right\}, \quad \|\boldsymbol{x}\|_\infty = \max\left\{\boldsymbol{y}^\top \boldsymbol{x} : \|\boldsymbol{y}\|_1 = 1\right\}.$$

Example ($\|A\|_1$)

By the theorem above,

$$\begin{aligned} \|A\|_1 &= \max_{\|\boldsymbol{x}\|_1=1} \|A\boldsymbol{x}\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \max_{\|\boldsymbol{y}\|_\infty=1} \boldsymbol{y}^\top A\boldsymbol{x} \\ &= \max_{\|\boldsymbol{y}\|_\infty=1} \max_{\|\boldsymbol{x}\|_1=1} \boldsymbol{y}^\top A\boldsymbol{x} = \max_{\|\boldsymbol{y}\|_\infty=1} \max_{\|\boldsymbol{x}\|_1=1} \boldsymbol{x}^\top A^\top \boldsymbol{y} \\ &= \max_{\|\boldsymbol{y}\|_\infty=1} \|A^\top \boldsymbol{y}\|_\infty = \|A^\top\|_\infty \,; \end{aligned}$$

thus

$$\|A\|_1 = \max\left\{\sum_{i=1}^m |a_{i1}|, \sum_{i=1}^m |a_{i2}|, \cdots, \sum_{i=1}^m |a_{in}|\right\};$$

that is, $\|A\|_1$ is the largest sum of the absolute value of column entries.

§5.3 Norms on Vectors and Matrices

Example ($\|A\|_2$)

Let A be an $m \times n$ matrix. Then by the definition of the 2-norm,

$$\|A\|_2^2 = \max\left\{\|A\boldsymbol{x}\|_2^2 : \|\boldsymbol{x}\|_2 = 1\right\} = \max\left\{\boldsymbol{x}^\top A^\top A\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 1\right\}.$$

Since $A^\top A$ is an $n \times n$ symmetric matrix, $A^\top A$ has n real eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ and corresponding **orthogonal unit** eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$. Then each $\boldsymbol{x} \in \mathbb{R}^n$ can be expressed as

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots x_n\boldsymbol{v}_n \qquad (\star)$$

and the condition $\|\boldsymbol{x}\|_2 = 1$ is translated into $\sum_{i=1}^{n} x_i^2 = 1$. Using $(\star)$,

$$\boldsymbol{x}^\top A^\top A\boldsymbol{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

whose maximum, under the constraint $\sum_{i=1}^{n} x_i^2 = 1$, is λ_n. Therefore,

$$\|A\|_2 = \text{the square root of the maximum eigenvalue of } A^\top A.$$

§5.3 Norms on Vectors and Matrices

Example ($\|A\|_2$)

Let A be an $m \times n$ matrix. Then by the definition of the 2-norm,

$$\|A\|_2^2 = \max\left\{\|A\boldsymbol{x}\|_2^2 : \|\boldsymbol{x}\|_2 = 1\right\} = \max\left\{\boldsymbol{x}^\top A^\top A\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 1\right\}.$$

Since $A^\top A$ is an $n \times n$ symmetric matrix, $A^\top A$ has n real eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ and corresponding **orthogonal unit** eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$. Then each $\boldsymbol{x} \in \mathbb{R}^n$ can be expressed as

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots x_n\boldsymbol{v}_n \qquad (\star)$$

and the condition $\|\boldsymbol{x}\|_2 = 1$ is translated into $\sum_{i=1}^{n} x_i^2 = 1$. Using $(\star)$,

$$\boldsymbol{x}^\top A^\top A\boldsymbol{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

whose maximum, under the constraint $\sum_{i=1}^{n} x_i^2 = 1$, is λ_n. Therefore,

$\|A\|_2 =$ the square root of the maximum eigenvalue of $A^\top A$.

§5.3 Norms on Vectors and Matrices

Example ($\|A\|_2$)

Let A be an $m \times n$ matrix. Then by the definition of the 2-norm,

$$\|A\|_2^2 = \max\left\{\|A\boldsymbol{x}\|_2^2 : \|\boldsymbol{x}\|_2 = 1\right\} = \max\left\{\boldsymbol{x}^\top A^\top A\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 1\right\}.$$

Since $A^\top A$ is an $n \times n$ symmetric matrix, $A^\top A$ has n real eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ and corresponding **orthogonal unit** eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$. Then each $\boldsymbol{x} \in \mathbb{R}^n$ can be expressed as

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots x_n\boldsymbol{v}_n \qquad (\star)$$

and the condition $\|\boldsymbol{x}\|_2 = 1$ is translated into $\sum_{i=1}^{n} x_i^2 = 1$. Using $(\star)$,

$$\boldsymbol{x}^\top A^\top A\boldsymbol{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

whose maximum, under the constraint $\sum_{i=1}^{n} x_i^2 = 1$, is λ_n. Therefore,

$$\|A\|_2 = \text{the square root of the maximum eigenvalue of } A^\top A.$$

§5.3 Norms on Vectors and Matrices

Example ($\|A\|_2$)

Let A be an $m \times n$ matrix. Then by the definition of the 2-norm,

$$\|A\|_2^2 = \max\left\{\|A\boldsymbol{x}\|_2^2 : \|\boldsymbol{x}\|_2 = 1\right\} = \max\left\{\boldsymbol{x}^\top A^\top A\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 1\right\}.$$

Since $A^\top A$ is an $n \times n$ symmetric matrix, $A^\top A$ has n real eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ and corresponding **orthogonal unit** eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$. Then each $\boldsymbol{x} \in \mathbb{R}^n$ can be expressed as

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots x_n\boldsymbol{v}_n \tag{$\star$}$$

and the condition $\|\boldsymbol{x}\|_2 = 1$ is translated into $\sum\limits_{i=1}^n x_i^2 = 1$. Using ($\star$),

$$\boldsymbol{x}^\top A^\top A\boldsymbol{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

whose maximum, under the constraint $\sum\limits_{i=1}^n x_i^2 = 1$, is λ_n. Therefore,

$$\|A\|_2 = \text{the square root of the maximum eigenvalue of } A^\top A.$$

§5.3 Norms on Vectors and Matrices

Example ($\|A\|_2$)

Let A be an $m \times n$ matrix. Then by the definition of the 2-norm,

$$\|A\|_2^2 = \max\left\{\|A\boldsymbol{x}\|_2^2 : \|\boldsymbol{x}\|_2 = 1\right\} = \max\left\{\boldsymbol{x}^\top A^\top A\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 1\right\}.$$

Since $A^\top A$ is an $n \times n$ symmetric matrix, $A^\top A$ has n real eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ and corresponding **orthogonal unit** eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$. Then each $\boldsymbol{x} \in \mathbb{R}^n$ can be expressed as

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots x_n\boldsymbol{v}_n \tag{$\star$}$$

and the condition $\|\boldsymbol{x}\|_2 = 1$ is translated into $\sum\limits_{i=1}^{n} x_i^2 = 1$. Using ($\star$),

$$\boldsymbol{x}^\top A^\top A\boldsymbol{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

whose maximum, under the constraint $\sum\limits_{i=1}^{n} x_i^2 = 1$, is λ_n. Therefore,

$\|A\|_2 =$ the square root of the maximum eigenvalue of $A^\top A$.

§5.3 Norms on Vectors and Matrices

Example

Consider the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$. Then

$$A^\top A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & 4 \\ -1 & 4 & 5 \end{bmatrix}$$

which implies that the characteristic equation of $A^\top A$ is

$$\det(A^\top A - \lambda I) = \begin{vmatrix} 3-\lambda & 2 & -1 \\ 2 & 6-\lambda & 4 \\ -1 & 4 & 5-\lambda \end{vmatrix} = -\lambda(\lambda^2 - 14\lambda + 42) = 0\,.$$

Therefore, the eigenvalues of $A^\top A$ are $\lambda = 0, 7+\sqrt{7}, 7-\sqrt{7}$; thus

$$\|A\|_2 = \sqrt{\rho(A^\top A)} = \sqrt{7+\sqrt{7}} \approx 3.106\,.$$

§5.3 Norms on Vectors and Matrices

Example (Frobenius Norm)

Not every norm on the space of $m \times n$ real matrices is of the form $\|\cdot\|_{\mathbb{R}^n,\mathbb{R}^m}$ (called the **natural norm**). For example, the **Frobenius norm**, sometimes also called the **Euclidean norm** (a term unfortunately also used for the vector ℓ^2-norm), is matrix norm of an $m \times n$ matrix A defined as the square root of the sum of the absolute squares of its elements; that is,

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

This is clear a norm because this is to identify the space of real $m \times n$ matrices as the space $\mathbb{R}^{mn}$ with ℓ^2-norm. The Frobenius norm can also be computed by

$$\|A\|_F = \sqrt{\mathrm{Tr}(AA^\top)}\,,$$

where $\mathrm{Tr}(M)$ is the trace of (a square matrix) M.

§5.3 Norms on Vectors and Matrices

Example (Frobenius Norm)

Not every norm on the space of $m \times n$ real matrices is of the form $\|\cdot\|_{\mathbb{R}^n,\mathbb{R}^m}$ (called the **natural norm**). For example, the **Frobenius norm**, sometimes also called the **Euclidean norm** (a term unfortunately also used for the vector ℓ^2-norm), is matrix norm of an $m\times n$ matrix A defined as the square root of the sum of the absolute squares of its elements; that is,

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

This is clear a norm because this is to identify the space of real $m\times n$ matrices as the space $\mathbb{R}^{mn}$ with ℓ^2-norm. The Frobenius norm can also be computed by

$$\|A\|_F = \sqrt{\mathrm{Tr}(AA^\top)},$$

where $\mathrm{Tr}(M)$ is the trace of (a square matrix) M.

§5.3 Norms on Vectors and Matrices

Example (Frobenius Norm)

Not every norm on the space of $m \times n$ real matrices is of the form $\|\cdot\|_{\mathbb{R}^n,\mathbb{R}^m}$ (called the **natural norm**). For example, the **Frobenius norm**, sometimes also called the **Euclidean norm** (a term unfortunately also used for the vector ℓ^2-norm), is matrix norm of an $m \times n$ matrix A defined as the square root of the sum of the absolute squares of its elements; that is,

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

This is clear a norm because this is to identify the space of real $m \times n$ matrices as the space $\mathbb{R}^{mn}$ with ℓ^2-norm. The Frobenius norm can also be computed by

$$\|A\|_F = \sqrt{\mathrm{Tr}(AA^\top)}\,,$$

where $\mathrm{Tr}(M)$ is the trace of (a square matrix) M.

§5.3 Norms on Vectors and Matrices

Definition

The spectral radius of a square matrix is the largest absolute value of its eigenvalues. The spectral radius of A is denoted by $\rho(A)$.

Theorem

Let A be an $m \times n$ real matrix. Then $\|A\|_2 = \sqrt{\rho(A^\top A)}$.

Remark: The ℓ^2-matrix norm is also called the spectral norm.

Corollary

If A is a real symmetric matrix, then $\|A\|_2 = \rho(A)$.

Theorem

$\rho(A) \leqslant \|A\|$ for any real square matrix A and natural norm $\|\cdot\|$.

§5.3 Norms on Vectors and Matrices

Theorem

Let A be a real square matrix. Then for every $\varepsilon > 0$ there exists a (subordinate) matrix norm $\|\cdot\|$ such that $\|A\| \leqslant \rho(A) + \varepsilon$.

Proof.

Let A be an $n \times n$ real matrix. The Jordan canonical form of A is

$$A = S \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_k}(\lambda_k) \end{bmatrix} S^{-1},$$

where S is an invertible matrix, $\lambda_1, \lambda_2, \cdots, \lambda_k$ are (complex) eigenvalues of A, $n_1 + n_2 + \cdots + n_k = n$, and $J_{n_j}(\lambda_j)$ are Jordan blocks of size $n_j \times n_j$. □

§5.3 Norms on Vectors and Matrices

Proof (cont'd).

For each $m \in \mathbb{N}$ and $\eta > 0$, define

$$D(\eta) = \begin{bmatrix} D_{n_1}(\eta) & 0 & \cdots & 0 \\ 0 & D_{n_2}(\eta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_{n_k}(\eta) \end{bmatrix}, \text{ where } D_m(\eta) = \begin{bmatrix} \eta & 0 & \cdots & 0 \\ 0 & \eta^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \eta^m \end{bmatrix}.$$

Then the norm defined by

$$|||M||| \equiv \left\| D\left(\frac{1}{\varepsilon}\right) S^{-1} M S D(\varepsilon) \right\|_1$$

has the property that $|||A||| \leqslant \rho(A) + \varepsilon$. Define a norm on $\mathbb{R}^n$ by $\|\mathbf{x}\|_{\mathbb{R}^n} = \left\| D\left(\frac{1}{\varepsilon}\right) S^{-1} \mathbf{x} \right\|_1$. Then

$$\begin{aligned} \|M\mathbf{x}\|_{\mathbb{R}^n} &= \left\| D\left(\frac{1}{\varepsilon}\right) S^{-1} M \mathbf{x} \right\|_1 = \left\| D\left(\frac{1}{\varepsilon}\right) S^{-1} M S D(\varepsilon) D\left(\frac{1}{\varepsilon}\right) S^{-1} \mathbf{x} \right\|_1 \\ &\leqslant \left\| D\left(\frac{1}{\varepsilon}\right) S^{-1} M S D(\varepsilon) \right\|_1 \left\| D\left(\frac{1}{\varepsilon}\right) S^{-1} \mathbf{x} \right\|_1 = |||M||| \, \|\mathbf{x}\|_{\mathbb{R}^n} \end{aligned}$$

which implies that $|||\cdot|||$ is an subordinate norm. □

§5.3 Norms on Vectors and Matrices

Proof (cont'd).

For each $m \in \mathbb{N}$ and $\eta > 0$, define

$$D(\eta) = \begin{bmatrix} D_{n_1}(\eta) & 0 & \cdots & 0 \\ 0 & D_{n_2}(\eta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_{n_k}(\eta) \end{bmatrix}, \text{ where } D_m(\eta) = \begin{bmatrix} \eta & 0 & \cdots & 0 \\ 0 & \eta^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \eta^m \end{bmatrix}.$$

Then the norm defined by

$$|||M||| \equiv \left\| D\left(\tfrac{1}{\varepsilon}\right) S^{-1} M S D(\varepsilon) \right\|_1$$

has the property that $|||A||| \leqslant \rho(A) + \varepsilon$. Define a norm on $\mathbb{R}^n$ by $\|\mathbf{x}\|_{\mathbb{R}^n} = \left\| D\left(\tfrac{1}{\varepsilon}\right) S^{-1} \mathbf{x} \right\|_1$. Then

$$\begin{aligned} \|M\mathbf{x}\|_{\mathbb{R}^n} &= \left\| D\left(\tfrac{1}{\varepsilon}\right) S^{-1} M \mathbf{x} \right\|_1 = \left\| D\left(\tfrac{1}{\varepsilon}\right) S^{-1} M S D(\varepsilon) D\left(\tfrac{1}{\varepsilon}\right) S^{-1} \mathbf{x} \right\|_1 \\ &\leqslant \left\| D\left(\tfrac{1}{\varepsilon}\right) S^{-1} M S D(\varepsilon) \right\|_1 \left\| D\left(\tfrac{1}{\varepsilon}\right) S^{-1} \mathbf{x} \right\|_1 = |||M||| \, \|\mathbf{x}\|_{\mathbb{R}^n} \end{aligned}$$

which implies that $|||\cdot|||$ is an subordinate norm. □

§5.3 Norms on Vectors and Matrices

Definition

A square matrix A is said to be convergent (to zero matrix) if for all $1 \leqslant i, j \leqslant n$ the (i, j)-entry of A^n converges to 0 as $n \to \infty$.

Example

$$A = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \Rightarrow A^2 = \begin{bmatrix} \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix} \Rightarrow A^3 = \begin{bmatrix} \frac{1}{8} & 0 \\ \frac{3}{16} & \frac{1}{8} \end{bmatrix} \Rightarrow \cdots$$

By induction, one can show that

$$A^k = \begin{bmatrix} (\frac{1}{2})^k & 0 \\ \frac{k}{2^{k+1}} & (\frac{1}{2})^k \end{bmatrix}.$$

Since $\lim\limits_{k\to\infty} (\frac{1}{2})^k = 0$ and $\lim\limits_{k\to\infty} \frac{k}{2^{k+1}} = 0$, A is a convergent matrix.

§5.3 Norms on Vectors and Matrices

Theorem

The following statements are equivalent:

1. *A is a convergent matrix;*
2. $\lim_{n\to\infty} \|A^n\| = 0$ *for some matrix norm;*
3. $\lim_{n\to\infty} \|A^n\| = 0$ *for all matrix norms;*
4. $\rho(A) < 1$;
5. $\lim_{n\to\infty} A^n \mathbf{x} = 0$ *for all* $\mathbf{x}$.

Remark:

- ② ⇔ ③ because all norms on a finite dimensional real vector space are equivalent.
- ① ⇔ ④ ⇔ ⑤ by writing A into Jordan canonical form.
- ① ⇔ ② by using the Frobenius norm.

§5.3 Norms on Vectors and Matrices

Lemma

Let A be a square matrix. If $\rho(A) < 1$, then $(I - A)^{-1}$ exists and

$$(I - A)^{-1} = I + A + A^2 + \cdots \left(:= \sum_{n=0}^{\infty} A^n \right).$$

Proof.

Since $\rho(A) < 1$, 1 is not an eigenvalue of A; thus $(I - A)\boldsymbol{x} = \mathbf{0}$ has only trivial solution. Moreover, if A is $m \times m$, the for all $\boldsymbol{x} \in \mathbb{R}^m$,

$$\begin{aligned}
(I - A)\sum_{n=0}^{\infty} A^n \boldsymbol{x} = (I - A) \lim_{N\to\infty} \sum_{n=0}^{N} A^n \boldsymbol{x} &= \lim_{N\to\infty} (I - A) \sum_{n=0}^{N} A^n \boldsymbol{x} \\
&= \lim_{N\to\infty} \left(\sum_{n=0}^{N} A^n \boldsymbol{x} - \sum_{n=0}^{N} A^{n+1} \boldsymbol{x} \right) \\
&= \lim_{N\to\infty} \left(\boldsymbol{x} - A^{N+1} \boldsymbol{x} \right) = \boldsymbol{x};
\end{aligned}$$

thus $(I - A)^{-1}\boldsymbol{x} = \sum_{n=0}^{\infty} A^n \boldsymbol{x}$. $\left(\text{Does } \sum_{n=0}^{\infty} A^n \boldsymbol{x} \text{ converges for all } \boldsymbol{x}?\right)$ □

§5.3 Norms on Vectors and Matrices

Lemma

Let A be a square matrix. If $\rho(A) < 1$, then $(I - A)^{-1}$ exists and

$$(I - A)^{-1} = I + A + A^2 + \cdots \Big(:= \sum_{n=0}^{\infty} A^n \Big).$$

Proof.

Since $\rho(A) < 1$, 1 is not an eigenvalue of A; thus $(I - A)\mathbf{x} = \mathbf{0}$ has only trivial solution. Moreover, if A is $m \times m$, the for all $\mathbf{x} \in \mathbb{R}^m$,

$$\begin{aligned}(I - A)\sum_{n=0}^{\infty} A^n\mathbf{x} &= (I - A)\lim_{N\to\infty}\sum_{n=0}^{N} A^n\mathbf{x} = \lim_{N\to\infty}(I - A)\sum_{n=0}^{N} A^n\mathbf{x} \\ &= \lim_{N\to\infty}\Big(\sum_{n=0}^{N} A^n\mathbf{x} - \sum_{n=0}^{N} A^{n+1}\mathbf{x}\Big) \\ &= \lim_{N\to\infty}\big(\mathbf{x} - A^{N+1}\mathbf{x}\big) = \mathbf{x};\end{aligned}$$

thus $(I - A)^{-1}\mathbf{x} = \sum_{n=0}^{\infty} A^n\mathbf{x}$. (Does $\sum_{n=0}^{\infty} A^n\mathbf{x}$ converges for all $\mathbf{x}$?) □

§5.3 Norms on Vectors and Matrices

Lemma

Let A be a square matrix. If $\rho(A) < 1$, then $(I - A)^{-1}$ exists and

$$(I - A)^{-1} = I + A + A^2 + \cdots \Big(:= \sum_{n=0}^{\infty} A^n \Big).$$

Proof.

Since $\rho(A) < 1$, 1 is not an eigenvalue of A; thus $(I - A)\boldsymbol{x} = \mathbf{0}$ has only trivial solution. Moreover, if A is $m \times m$, the for all $\boldsymbol{x} \in \mathbb{R}^m$,

$$\begin{aligned}(I - A)\sum_{n=0}^{\infty} A^n \boldsymbol{x} &= (I - A) \lim_{N\to\infty} \sum_{n=0}^{N} A^n \boldsymbol{x} = \lim_{N\to\infty} (I - A) \sum_{n=0}^{N} A^n \boldsymbol{x} \\ &= \lim_{N\to\infty} \Big(\sum_{n=0}^{N} A^n \boldsymbol{x} - \sum_{n=0}^{N} A^{n+1} \boldsymbol{x} \Big) \\ &= \lim_{N\to\infty} \big(\boldsymbol{x} - A^{N+1} \boldsymbol{x} \big) = \boldsymbol{x};\end{aligned}$$

thus $(I - A)^{-1}\boldsymbol{x} = \sum_{n=0}^{\infty} A^n \boldsymbol{x}$. $\Big($Does $\sum_{n=0}^{\infty} A^n \boldsymbol{x}$ converges for all $\boldsymbol{x}$?$\Big)$ □

§5.4 Iterative Methods

Recall that in Chapter 3 to solve a nonlinear equation $f(x) = 0$ we introduce iterative method

$$x^{(k+1)} = g\big(x^{(k)}\big) \quad \text{for } k \in \mathbb{N} \cup \{0\} \text{ with } x^{(0)} \text{ given},$$

where $f(x) = 0 \Leftrightarrow x = g(x)$, and the fixed-point of g is a solution of f.

The idea of solving $A\boldsymbol{x} = \boldsymbol{b}$ using the iterative method is based on the same concept:

1. $A\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow \boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$ for some fixed matrix T and vector c.
2. Given $\boldsymbol{x}^{(0)}$, $\boldsymbol{x}^{(k+1)} := T\boldsymbol{x}^{(k)} + \boldsymbol{c}$ for $k = 0, 1, 2, \cdots$

§5.4 Iterative Methods

Let $A\boldsymbol{x} = \boldsymbol{b}$ be a linear system of n equations, where $A = [a_{ij}]_{n\times n}$ and $\boldsymbol{b} \in \mathbb{R}^n$. Then A can be decomposed into a diagonal component D, a lower triangular part L and an upper triangular part U:

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}, \quad D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}.$$

$$L = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1} & \cdots & a_{n(n-1)} & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{(n-1)n} \\ 0 & \cdots & 0 & 0 \end{bmatrix}$$

1. Jacobi method: $A\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow D\boldsymbol{x} = -(L+U)\boldsymbol{x} + \boldsymbol{b}$.
2. Gauss-Seidel method: $A\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow (D+L)\boldsymbol{x} = -U\boldsymbol{x} + \boldsymbol{b}$.

§5.4 Iterative Methods

Let $A\boldsymbol{x} = \boldsymbol{b}$ be a linear system of n equations, where $A = [a_{ij}]_{n\times n}$ and $\boldsymbol{b} \in \mathbb{R}^n$. Then A can be decomposed into a diagonal component D, a lower triangular part L and an upper triangular part U:

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}, \quad D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}.$$

$$L = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1} & \cdots & a_{n(n-1)} & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{(n-1)n} \\ 0 & \cdots & 0 & 0 \end{bmatrix}$$

1. Jacobi method: $A\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow D\boldsymbol{x} = -(L+U)\boldsymbol{x} + \boldsymbol{b}$.
2. Gauss-Seidel method: $A\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow (D+L)\boldsymbol{x} = -U\boldsymbol{x} + \boldsymbol{b}$.

§5.4 Iterative Methods

1. The Jacobi method of solving $A\boldsymbol{x} = \boldsymbol{b}$ is the iterative method
$$\boldsymbol{x}^{(k+1)} = D^{-1}\left[\boldsymbol{b} - (L+U)\boldsymbol{x}^{(k)}\right] = -D^{-1}(L+U)\boldsymbol{x}^{(k)} + D^{-1}\boldsymbol{b},$$
and the element-based formula is thus
$$x_i^{(k+1)} = \frac{-\sum\limits_{j=1, j\neq i}^{n} a_{ij}x_j^{(k)} + b_i}{a_{ii}} \quad \forall\, k \in \mathbb{N} \cup \{0\}.$$

2. The Gauss-Seidel method of solving $A\boldsymbol{x} = \boldsymbol{b}$ is the iterative method
$$\boldsymbol{x}^{(k+1)} = (D+L)^{-1}\left[\boldsymbol{b} - U\boldsymbol{x}^{(k)}\right] = -(D+L)^{-1}U\boldsymbol{x}^{(k)} + (D+L)^{-1}\boldsymbol{b},$$
and the element-based formula is thus
$$x_i^{(k+1)} = \frac{-\sum\limits_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum\limits_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i}{a_{ii}} \quad \forall\, k \in \mathbb{N} \cup \{0\}.$$

§5.4 Iterative Methods

1. The Jacobi method of solving $A\boldsymbol{x}=\boldsymbol{b}$ is the iterative method
$$\boldsymbol{x}^{(k+1)}=D^{-1}\left[\boldsymbol{b}-(L+U)\boldsymbol{x}^{(k)}\right]=-D^{-1}(L+U)\boldsymbol{x}^{(k)}+D^{-1}\boldsymbol{b},$$
and the element-based formula is thus
$$x_i^{(k+1)}=\frac{-\sum\limits_{j=1,\,j\neq i}^{n} a_{ij}x_j^{(k)}+b_i}{a_{ii}}\quad\forall\, k\in\mathbb{N}\cup\{0\}.$$
2. The Gauss-Seidel method of solving $A\boldsymbol{x}=\boldsymbol{b}$ is the iterative method
$$\boldsymbol{x}^{(k+1)}=(D+L)^{-1}\left[\boldsymbol{b}-U\boldsymbol{x}^{(k)}\right]=-(D+L)^{-1}U\boldsymbol{x}^{(k)}+(D+L)^{-1}\boldsymbol{b},$$
and the element-based formula is thus
$$x_i^{(k+1)}=\frac{-\sum\limits_{j=1}^{i-1} a_{ij}x_j^{(k+1)}-\sum\limits_{j=i+1}^{n} a_{ij}x_j^{(k)}+b_i}{a_{ii}}\quad\forall\, k\in\mathbb{N}\cup\{0\}.$$

§5.4 Iterative Methods

Example (Solving $A\boldsymbol{x} = \boldsymbol{b}$ using Jacobi and Gauss-Seidel methods)

Consider a linear system:

$$\begin{cases} 10x_1 - 1x_2 + 2x_3 + 0x_4 = 6 \\ -x_1 + 11x_2 - 1x_3 + 3x_4 = 25 \\ 2x_1 - 1x_2 + 10x_3 - 1x_4 = -11 \\ 0x_1 + 3x_2 - 1x_3 + 8x_4 = 15 \end{cases}$$

or equivalently,

$$\begin{bmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 25 \\ -11 \\ 15 \end{bmatrix}.$$

Exact unique solution: $\boldsymbol{x} = (1, 2, -1, 1)^\top$.

§5.4 Iterative Methods

Example (cont'd)

We first rewrite the linear system as

$$\begin{aligned}
x_1 &= 0 + \frac{1}{10}x_2 - \frac{2}{10}x_3 + 0 + \frac{6}{10} \\
x_2 &= \frac{1}{11}x_1 + 0 + \frac{1}{11}x_3 - \frac{3}{11}x_4 + \frac{25}{11} \\
x_3 &= -\frac{2}{10}x_1 + \frac{1}{10}x_2 + 0 + \frac{1}{10}x_4 - \frac{11}{10} \\
x_4 &= 0 - \frac{3}{8}x_2 + \frac{1}{8}x_3 + 0 + \frac{15}{8}
\end{aligned}$$

which, written in matrix form, is

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = T\boldsymbol{x} + \boldsymbol{c} \equiv \begin{bmatrix} 0 & \frac{1}{10} & -\frac{2}{10} & 0 \\ \frac{1}{11} & 0 & \frac{1}{11} & -\frac{3}{11} \\ -\frac{2}{10} & \frac{1}{10} & 0 & \frac{1}{10} \\ 0 & -\frac{3}{8} & \frac{1}{8} & 0 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} \frac{6}{10} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{bmatrix}.$$

§5.4 Iterative Methods

Example (cont'd)

If $\boldsymbol{x}^{(0)} = (0,0,0,0)^\top$, then the Jacobi method provides

$$\boldsymbol{x}^{(1)} = T\boldsymbol{x}^{(0)} + \boldsymbol{c} = \begin{bmatrix} \frac{6}{10} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{bmatrix} = \begin{bmatrix} 0.6000 \\ 2.2727 \\ -1.1000 \\ 1.8750 \end{bmatrix}.$$

$$\Rightarrow \boldsymbol{x}^{(2)} = T\boldsymbol{x}^{(1)} + \boldsymbol{c} \Rightarrow \cdots$$

$$\Rightarrow \frac{\|\boldsymbol{x}^{(10)} - \boldsymbol{x}^{(9)}\|_\infty}{\|\boldsymbol{x}^{(10)}\|_\infty} \approx \frac{8.0 \times 10^{-4}}{1.9998} < 10^{-3} \quad \text{stop!}$$ (Stopping criteria)

$$\Rightarrow \boldsymbol{x} \approx \boldsymbol{x}^{(10)} \approx \begin{bmatrix} 1.00011860 \\ 1.99976795 \\ -0.99982814 \\ 0.99978598 \end{bmatrix}.$$

§5.4 Iterative Methods

Example (cont'd)

For the Gauss-Seidel method, we let $x^{(0)} = (0,0,0,0)^\top$ and for $k = 0,1,2,\cdots$ define

$$\begin{aligned}
x_1^{(k+1)} &= 0 + \frac{1}{10}x_2^{(k)} - \frac{2}{10}x_3^{(k)} + 0 + \frac{6}{10}\\
x_2^{(k+1)} &= \frac{1}{11}x_1^{(k+1)} + 0 + \frac{1}{11}x_3^{(k)} - \frac{3}{11}x_4^{(k)} + \frac{25}{11}\\
x_3^{(k+1)} &= -\frac{2}{10}x_1^{(k+1)} + \frac{1}{10}x_2^{(k+1)} + 0 + \frac{1}{10}x_4^{(k)} - \frac{11}{10}\\
x_4^{(k+1)} &= 0 - \frac{3}{8}x_2^{(k+1)} + \frac{1}{8}x_3^{(k+1)} + 0 + \frac{15}{8}
\end{aligned}$$

- Need to proceed from the top line to the bottom line: Solving for $x_1^{(k+1)}$ from the first equation, and then use this solution to solve $x_2^{(k+1)}$ from the second equation, and so on.

$$\Rightarrow \frac{\|\boldsymbol{x}^{(5)} - \boldsymbol{x}^{(4)}\|_\infty}{\|\boldsymbol{x}^{(5)}\|_\infty} = 4.0 \times 10^{-4} < 10^{-3} \quad \text{stop!} \quad \boldsymbol{x} \approx \boldsymbol{x}^{(5)}.$$

§5.4 Iterative Methods

Theorem

Let T be an $n \times n$ real matrix. For any $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$, the sequence $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ defined by

$$\boldsymbol{x}^{(k+1)} := T\boldsymbol{x}^{(k)} + \boldsymbol{c}, \quad k \in \mathbb{N} \cup \{0\},$$

converges to the unique solution of $\boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$ if and only if $\rho(T) < 1$.

Proof.

($\Leftarrow$) Since $\rho(T) < 1$, $(I - T)^{-1}$ exists; thus $\boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$ has a unique solution. Moreover, there exists a subordinate matrix norm $|||\cdot|||$ and a norm $\|\cdot\|$ on $\mathbb{R}^n$ such that $|||T||| < 1$ and $\|T\boldsymbol{x}\| \leqslant |||T||| \|\boldsymbol{x}\|$ for all $x \in \mathbb{R}^n$. Therefore, the mapping $\boldsymbol{x} \mapsto T\boldsymbol{x} + \boldsymbol{c}$ is a contraction mapping, and the contraction mapping principle implies that the sequence $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ defined by $\boldsymbol{x}^{(k+1)} = T\boldsymbol{x}^{(k)} + \boldsymbol{c}$ converges (to the solution of $\boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$). □

§5.4 Iterative Methods

Theorem

Let T be an $n \times n$ real matrix. For any $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$, the sequence $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ defined by

$$\boldsymbol{x}^{(k+1)} := T\boldsymbol{x}^{(k)} + \boldsymbol{c}, \quad k \in \mathbb{N} \cup \{0\},$$

converges to the unique solution of $\boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$ if and only if $\rho(T) < 1$.

Proof.

($\Leftarrow$) Since $\rho(T) < 1$, $(I - T)^{-1}$ exists; thus $\boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$ has a unique solution. Moreover, there exists a subordinate matrix norm $|||\cdot|||$ and a norm $\|\cdot\|$ on $\mathbb{R}^n$ such that $|||T||| < 1$ and $\|T\boldsymbol{x}\| \leqslant |||T||| \|\boldsymbol{x}\|$ for all $x \in \mathbb{R}^n$. Therefore, the mapping $\boldsymbol{x} \mapsto T\boldsymbol{x} + \boldsymbol{c}$ is a contraction mapping, and the contraction mapping principle implies that the sequence $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ defined by $\boldsymbol{x}^{(k+1)} = T\boldsymbol{x}^{(k)} + \boldsymbol{c}$ converges (to the solution of $\boldsymbol{x} = T\boldsymbol{x} + \boldsymbol{c}$). □

§5.4 Iterative Methods

Theorem

Let T be an $n \times n$ real matrix. For any $\mathbf{x}^{(0)} \in \mathbb{R}^n$, the sequence $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ defined by

$$\mathbf{x}^{(k+1)} := T\mathbf{x}^{(k)} + \mathbf{c}, \quad k \in \mathbb{N} \cup \{0\},$$

converges to the unique solution of $\mathbf{x} = T\mathbf{x} + \mathbf{c}$ if and only if $\rho(T) < 1$.

Proof.

($\Leftarrow$) Since $\rho(T) < 1$, $(I - T)^{-1}$ exists; thus $\mathbf{x} = T\mathbf{x} + \mathbf{c}$ has a unique solution. Moreover, there exists a subordinate matrix norm $|||\cdot|||$ and a norm $\|\cdot\|$ on $\mathbb{R}^n$ such that $|||T||| < 1$ and $\|T\mathbf{x}\| \leqslant |||T||| \|\mathbf{x}\|$ for all $x \in \mathbb{R}^n$. Therefore, the mapping $\mathbf{x} \mapsto T\mathbf{x} + \mathbf{c}$ is a contraction mapping, and the contraction mapping principle implies that the sequence $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ defined by $\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{c}$ converges (to the solution of $\mathbf{x} = T\mathbf{x} + \mathbf{c}$). $\square$

§5.4 Iterative Methods

Theorem

Let T be an $n \times n$ real matrix. For any $\mathbf{x}^{(0)} \in \mathbb{R}^n$, the sequence $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ defined by

$$\mathbf{x}^{(k+1)} := T\mathbf{x}^{(k)} + \mathbf{c}, \quad k \in \mathbb{N} \cup \{0\},$$

converges to the unique solution of $\mathbf{x} = T\mathbf{x} + \mathbf{c}$ if and only if $\rho(T) < 1$.

Proof.

($\Rightarrow$) Let $\mathbf{z} \in \mathbb{R}^n$ be given, and $\mathbf{x}$ be the unique solution to $\mathbf{x} = T\mathbf{x} + \mathbf{c}$. Define $\mathbf{x}^{(0)} = \mathbf{x} - \mathbf{z}$. Then

$$\mathbf{x}^{(1)} = T\mathbf{x}^{(0)} + \mathbf{c} = T\mathbf{x} - T\mathbf{z} + \mathbf{c} = \mathbf{x} - T\mathbf{z}$$

which further implies

$$\mathbf{x}^{(2)} = T\mathbf{x}^{(1)} + \mathbf{c} = T\mathbf{x} - T^2\mathbf{z} + \mathbf{c} = \mathbf{x} - T^2\mathbf{z}.$$

By induction, $\mathbf{x}^{(k)} = \mathbf{x} - T^k\mathbf{z}$. Since $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$, we have $\lim_{k\to\infty} T^k\mathbf{z} = \mathbf{0}$. Then $\rho(T) < 1$ due to the previous theorem. □

§5.4 Iterative Methods

Corollary

1. *Let* $\mathbf{x}^{(0)} \in \mathbb{R}^n$, *and* $\left\{\mathbf{x}^{(k)}\right\}_{k=1}^{\infty}$ *be a sequence defined by* $\mathbf{x}^{(k+1)} := T\mathbf{x}^{(k)} + \mathbf{c}$, $k \geqslant 0$. *If* $\|T\| < 1$ *for some natural matrix norm, then* $\left\{\mathbf{x}^{(k)}\right\}_{k=1}^{\infty}$ *converges to the unique solution of* $\mathbf{x} = T\mathbf{x} + \mathbf{c}$ *and*
   - $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leqslant \|T\|^k \|\mathbf{x} - \mathbf{x}^{(0)}\|$.
   - $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leqslant \dfrac{\|T\|^k}{1 - \|T\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|$.
2. *If A is strictly diagonally dominant, then for any* $\mathbf{x}^{(0)} \in \mathbb{R}^n$, *both the Jacobi and Gauss-Seidel methods give sequences* $\left\{\mathbf{x}^{(k)}\right\}_{k=1}^{\infty}$ *that converge to the unique solution of* $A\mathbf{x} = \mathbf{b}$.

§5.4 Iterative Methods

Successive Over Relaxation (SOR):

1. **The Gauss-Seidel method:**

$$x_i^{(k)} = \frac{1}{a_{ii}}\left[-\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)} + b_i\right].$$

2. **Successive over-relaxation:** for $\omega > 0$,

$$x_i^{(k)} = (1-\omega)x_i^{(k-1)} + \frac{\omega}{a_{ii}}\left[-\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)} + b_i\right].$$

$$\Leftrightarrow a_{ii}x_i^{(k)} + \omega\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} = (1-\omega)a_{ii}x_i^{(k-1)} - \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)} + \omega b_i$$

$$\Leftrightarrow (D+\omega L)\mathbf{x}^{(k)} = \Big[(1-\omega)D-\omega U\Big]\mathbf{x}^{(k-1)} + \omega\mathbf{b}$$

$$\Leftrightarrow \mathbf{x}^{(k)} = (D+\omega L)^{-1}\Big[(1-\omega)D-\omega U\Big]\mathbf{x}^{(k-1)} + \omega(D+\omega L)^{-1}\mathbf{b}$$

$$\Leftrightarrow \mathbf{x}^{(k)} = T_\omega\mathbf{x}^{(k-1)} + \mathbf{c}_\omega$$

§5.4 Iterative Methods

Gauss-Seidel:

$$\mathbf{x}^{(k+1)} = -(D+L)^{-1}U\mathbf{x}^{(k)} + (D+L)^{-1}\mathbf{b}\,.$$

SOR:

$$\mathbf{x}^{(k)} = (D+\omega L)^{-1}\Big[(1-\omega)D - \omega U\Big]\mathbf{x}^{(k-1)} + \omega(D+\omega L)^{-1}\mathbf{b}\,.$$

Different parameter ω can be chosen according to the need. In general,

- $\omega = 1$: the Gauss-Seidel method.
- $0 < \omega < 1$: when Gauss-Seidel diverges.
- $\omega > 1$: when Gauss-Seidel converges.

§5.4 Iterative Methods

Example

Consider a linear system

$$\begin{cases} 4x_1 + 3x_2 + 0 &= 24 \\ 3x_1 + 4x_2 - x_3 &= 30 \\ 0 - x_2 + 4x_3 &= -24 \end{cases}$$

Exact unique solution: $x = (3, 4, -5)^\top$.

1. Let $x^{(0)} = (1, 1, 1)^\top$. The Gauss-Seidel method:

$$\begin{cases} x_1^{(k)} = -0.75x_2^{(k-1)} + 6 \\ x_2^{(k)} = -0.75x_1^{(k)} + 0.25x_3^{(k-1)} + 7.5 \\ x_3^{(k)} = 0.25x_2^{(k)} - 6 \end{cases}$$

2. Let $x^{(0)} = (1, 1, 1)^\top$. The SOR with $\omega = 1.25$:

$$\begin{cases} x_1^{(k)} = -0.25x_1^{(k-1)} - 0.9375x_2^{(k-1)} + 7.5 \\ x_2^{(k)} = -0.9375x_1^{(k)} - 0.25x_2^{(k-1)} + 0.3125x_3^{(k-1)} + 9.375 \\ x_3^{(k)} = 0.3125x_2^{(k)} - 0.25x_3^{(k-1)} - 7.5 \end{cases}$$

§5.4 Iterative Methods

Theorem

1. *If* $a_{ii} \neq 0$ *for all* $i = 1, 2, \cdots, n$, *then* $\rho(T_\omega) \geqslant |\omega - 1|$. *This implies the SOR method can converge* ***only if*** $0 < \omega < 2$.
2. *If* A *is symmetric positive definite and* $0 < \omega < 2$, *then the SOR method converges for any* $\mathbf{x}^{(0)}$.

§5.5 Absolute Error, Relative Error and Condition Number

1. Suppose that we want to solve the linear system $A\boldsymbol{x} = \boldsymbol{b}$, but $\boldsymbol{b}$ is somehow perturbed to $\widetilde{\boldsymbol{b}}$ (this may happen when we convert a real $\boldsymbol{b}$ to a floating-point b).
2. Then actual solution would satisfy a slightly different linear system
$$A\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{b}}.$$
3. **Question**: Is $\widetilde{\boldsymbol{x}}$ very different from the desired solution $\boldsymbol{x}$ of the original system?
4. Of course, the answer should depend on how good the matrix A is.
5. Let $\|\cdot\|$ be a vector norm, we consider two types of errors:
   - **absolute error**: $\|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\|$
   - **relative error**: $\|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\| / \|\boldsymbol{x}\|$

§5.5 Absolute Error, Relative Error and Condition Number

- For the absolute error, we have

$$\|\boldsymbol{x}-\tilde{\boldsymbol{x}}\| = \|A^{-1}\boldsymbol{b}-A^{-1}\tilde{\boldsymbol{b}}\| = \|A^{-1}(\boldsymbol{b}-\tilde{\boldsymbol{b}})\| \leqslant \|A^{-1}\|\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|.$$

Therefore, the absolute error of $\boldsymbol{x}$ depends on two factors: **the absolute error of $\boldsymbol{b}$** and **the matrix norm of A^{-1}**.

- For the relative error, we have

$$\begin{aligned}\|\boldsymbol{x}-\tilde{\boldsymbol{x}}\| &= \|A^{-1}\boldsymbol{b}-A^{-1}\tilde{\boldsymbol{b}}\| = \|A^{-1}(\boldsymbol{b}-\tilde{\boldsymbol{b}})\| \\ &\leqslant \|A^{-1}\|\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\| = \|A^{-1}\|\|A\boldsymbol{x}\|\frac{\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|}{\|\boldsymbol{b}\|} \\ &\leqslant \|A^{-1}\|\|A\|\|\boldsymbol{x}\|\frac{\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|}{\|\boldsymbol{b}\|};\end{aligned}$$

that is

$$\frac{\|\boldsymbol{x}-\tilde{\boldsymbol{x}}\|}{\|\boldsymbol{x}\|} \leqslant \|A^{-1}\|\|A\|\frac{\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|}{\|\boldsymbol{b}\|}.$$

Therefore, the relative error of $\boldsymbol{x}$ depends on two factors: **the relative error of $\boldsymbol{b}$** and $\|A\|\|A^{-1}\|$.

§5.5 Absolute Error, Relative Error and Condition Number

- For the absolute error, we have
$$\|\boldsymbol{x}-\tilde{\boldsymbol{x}}\| = \|A^{-1}\boldsymbol{b}-A^{-1}\tilde{\boldsymbol{b}}\| = \|A^{-1}(\boldsymbol{b}-\tilde{\boldsymbol{b}})\| \leqslant \|A^{-1}\|\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|.$$
Therefore, the absolute error of $\boldsymbol{x}$ depends on two factors: **the absolute error of $\boldsymbol{b}$** and **the matrix norm of A^{-1}**.

- For the relative error, we have
$$\begin{aligned}\|\boldsymbol{x}-\tilde{\boldsymbol{x}}\| &= \|A^{-1}\boldsymbol{b}-A^{-1}\tilde{\boldsymbol{b}}\| = \|A^{-1}(\boldsymbol{b}-\tilde{\boldsymbol{b}})\| \\ &\leqslant \|A^{-1}\|\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\| = \|A^{-1}\|\|A\boldsymbol{x}\|\frac{\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|}{\|\boldsymbol{b}\|} \\ &\leqslant \|A^{-1}\|\|A\|\|\boldsymbol{x}\|\frac{\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|}{\|\boldsymbol{b}\|};\end{aligned}$$
that is
$$\frac{\|\boldsymbol{x}-\tilde{\boldsymbol{x}}\|}{\|\boldsymbol{x}\|} \leqslant \|A^{-1}\|\|A\|\frac{\|\boldsymbol{b}-\tilde{\boldsymbol{b}}\|}{\|\boldsymbol{b}\|}.$$
Therefore, the relative error of $\boldsymbol{x}$ depends on two factors: **the relative error of $\boldsymbol{b}$** and $\|A\|\|A^{-1}\|$.

§5.5 Absolute Error, Relative Error and Condition Number

Definition

For a given subordinate matrix norm $\|\cdot\|$, the condition number of the matrix A is the number

$$\kappa(A) := \|A\|\|A^{-1}\|.$$

$\kappa(A)$ measures how good the matrix A is.

Example

Let $\varepsilon > 0$ and

$$A = \begin{bmatrix} 1 & 1+\varepsilon \\ 1-\varepsilon & 1 \end{bmatrix} \Rightarrow A^{-1} = \varepsilon^{-2} \begin{bmatrix} 1 & -1-\varepsilon \\ -1+\varepsilon & 1 \end{bmatrix}.$$

Then $\|A\|_\infty = 2+\varepsilon$, $\|A^{-1}\|_\infty = \varepsilon^{-2}(2+\varepsilon)$, and

$$\kappa(A) = \left(\frac{2+\varepsilon}{\varepsilon}\right)^2 \geqslant \frac{4}{\varepsilon^2}.$$

§5.5 Absolute Error, Relative Error and Condition Number

1. For example, if $\varepsilon = 0.01$, then $\kappa(A) \geqslant 40000$.
2. What does this mean?

   It means that the relative error in $\boldsymbol{x}$ can be 40000 times greater than the relative error in b.
3. If $\kappa(A)$ is large, we say that A is **ill-conditioned,** otherwise A is **well-conditioned.**
4. In the ill-conditioned case, the solution is very sensitive to the small changes in the right-hand vector b (higher precision in b may be needed).

§5.5 Absolute Error, Relative Error and Condition Number

Consider the linear system $A\mathbf{x} = \mathbf{b}$. Let $\widetilde{\mathbf{x}}$ be a computed solution (which is an approximation to $\mathbf{x}$). We define

1. Residual vector: $\mathbf{r} = \mathbf{b} - A\widetilde{\mathbf{x}}$.
2. Error vector: $\mathbf{e} = \mathbf{x} - \widetilde{\mathbf{x}}$.

Then $A\mathbf{e} = A\mathbf{x} - A\widetilde{\mathbf{x}} = \mathbf{b} - A\widetilde{\mathbf{x}} = \mathbf{r}$.

Theorem (bounds involving condition number)

Let A be a square matrix, $\mathbf{x}$ be the solution of $A\mathbf{x} = \mathbf{b}$, and $\mathbf{r}$, $\mathbf{e}$ are the residual vector and the error vector associated with a computed solution $\widetilde{\mathbf{x}}$, respectively. Then

$$\frac{1}{\kappa(A)}\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \leqslant \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \leqslant \kappa(A)\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}.$$

§5.5 Absolute Error, Relative Error and Condition Number

Theorem (bounds involving condition number)

Let A be a square matrix, $\boldsymbol{x}$ be the solution of $A\boldsymbol{x} = \boldsymbol{b}$, and $\boldsymbol{r}$, $\boldsymbol{e}$ are the residual vector and the error vector associated with a computed solution $\tilde{\mathbf{x}}$, respectively. Then

$$\frac{1}{\kappa(A)}\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} \leqslant \frac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|} \leqslant \kappa(A)\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}.$$

Proof.

Since $A\boldsymbol{e} = \boldsymbol{r}$, $\boldsymbol{e} = A^{-1}\boldsymbol{r}$; thus

$$\|\boldsymbol{e}\|\|\boldsymbol{b}\| = \|A^{-1}\boldsymbol{r}\|\|A\boldsymbol{x}\| \leqslant \|A^{-1}\|\|\boldsymbol{r}\|\|A\|\|\boldsymbol{x}\| = \kappa(A)\|\boldsymbol{r}\|\|\boldsymbol{x}\|$$

which further implies that $\dfrac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|} \leqslant \kappa(A)\dfrac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$.

On the other hand, we have

$$\|\boldsymbol{r}\|\|\boldsymbol{x}\| = \|A\boldsymbol{e}\|\|A^{-1}\boldsymbol{b}\| \leqslant \|A\|\|\boldsymbol{e}\|\|A^{-1}\|\|\boldsymbol{b}\| = \kappa(A)\|\boldsymbol{e}\|\|\boldsymbol{b}\|$$

which shows that $\dfrac{1}{\kappa(A)}\dfrac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} \leqslant \dfrac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|}$. □

§5.5 Absolute Error, Relative Error and Condition Number

Theorem (bounds involving condition number)

Let A be a square matrix, $\boldsymbol{x}$ be the solution of $A\boldsymbol{x} = \boldsymbol{b}$, and $\boldsymbol{r}$, $\boldsymbol{e}$ are the residual vector and the error vector associated with a computed solution $\tilde{\boldsymbol{x}}$, respectively. Then

$$\frac{1}{\kappa(A)}\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} \leqslant \frac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|} \leqslant \kappa(A)\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}.$$

Proof.

Since $A\boldsymbol{e} = \boldsymbol{r}$, $\boldsymbol{e} = A^{-1}\boldsymbol{r}$; thus

$$\|\boldsymbol{e}\|\|\boldsymbol{b}\| = \|A^{-1}\boldsymbol{r}\|\|A\boldsymbol{x}\| \leqslant \|A^{-1}\|\|\boldsymbol{r}\|\|A\|\|\boldsymbol{x}\| = \kappa(A)\|\boldsymbol{r}\|\|\boldsymbol{x}\|$$

which further implies that $\dfrac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|} \leqslant \kappa(A)\dfrac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$.

On the other hand, we have

$$\|\boldsymbol{r}\|\|\boldsymbol{x}\| = \|A\boldsymbol{e}\|\|A^{-1}\boldsymbol{b}\| \leqslant \|A\|\|\boldsymbol{e}\|\|A^{-1}\|\|\boldsymbol{b}\| = \kappa(A)\|\boldsymbol{e}\|\|\boldsymbol{b}\|$$

which shows that $\dfrac{1}{\kappa(A)}\dfrac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} \leqslant \dfrac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|}$. □