Numerical Analysis

Mathematics of Scientific Computing

主讲人 邱欣欣
幻灯片制作 邱欣欣

中国海洋大学 信息科学与工程学院

2013 年 10 月 11 日

Solving Systems of Linear Equations

Contents

Introduction

- Constructing a general-purpose algorithm for solving the problem $Ax = b$.
- Analyzing the errors that are associated with the computer solution and study methods for controlling and reducing the error.
- Introducing the important topic of iterative algorithm for this problem.

Introduction

- The overall objective of this chapter is to discuss the numerical aspects of solving systems of linear equations having the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n = b_3 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n = b_n \end{cases}$$

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

Introduction

- Matrix are useful devices for representing systems of equations. The above system of linear equations can be written as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

  We can denote these matrices by A, x, and b, so that the equation becomes simply $Ax = b$.

Contents

Matrix Algebra

- A matrix is a rectangular array of numbers such as

$$\begin{bmatrix} 3.0 & 1.1 & -0.12 \\ 6.2 & 0.0 & 0.15 \\ 0.6 & -4.0 & 1.3 \\ 9.3 & 2.1 & 8.2 \end{bmatrix} \begin{bmatrix} 3 & 6 & \frac{11}{7} & -17 \end{bmatrix} \begin{bmatrix} 3.2 \\ -4.7 \\ 0.11 \end{bmatrix}$$

  These are, a 4×3 matrix, a 1×4 matrix and a 3×1 matrix.

- A $1 \times n$ matrix is called a row vector. A $m \times 1$ matrix is called a column vector or just a vector.

- If A is a matrix, the notations a_{ij}, $(A)_{ij}$, or $A(i,j)$ is used to denote the element at the intersection of the ith row and jth column.

Matrix Algebra

- The **transpose** of a matrix is denoted by A^T and is the matrix defined by $(A^T)_{ij} = a_{ji}$. For example, if A denotes

$$\begin{bmatrix} 3.0 & 1.1 & -0.12 \\ 6.2 & 0.0 & 0.15 \\ 0.6 & -4.0 & 1.3 \\ 9.3 & 2.1 & 8.2 \end{bmatrix}$$

we have

$$A^T = \begin{bmatrix} 3.0 & 6.2 & 0.6 & 9.3 \\ 1.1 & 0.0 & -4.0 & 2.1 \\ -0.12 & 0.15 & 1.3 & 8.2 \end{bmatrix}$$

If a matrix A has the property $A^T = A$, we say that A is **symmetric**.

Matrix Algebra

- If A is matrix and λ is a scalar (that is, a real number in this context), then λA is defined by $(\lambda A)_{ij} = \lambda a_{ij}$.
- If $A = (a_{ij})$ and $B = (b_{ij})$ are $m \times n$ matrices, then $A + B$ is defined by $(A + B)_{ij} = a_{ij} + b_{ij}$.
- If A is an $m \times p$ matrix and B is a $p \times n$ matrix, then AB is an $m \times n$ matrix defined by

$$(AB)_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \qquad (1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n)$$

Matrix Algebra

- Here are some examples of the algebraic operations:

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 4 & -4 \end{bmatrix} + \begin{bmatrix} 6 & 0 \\ 3 & -7 \\ 8 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 3 \\ 5 & -8 \\ 12 & -2 \end{bmatrix}$$

$$3 \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 4 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 9 \\ 6 & -3 \\ 12 & -12 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & -6 \\ 2 & 1 & 5 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -5 & 4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 13 \\ -24 & 2 \\ -3 & 19 \\ -5 & -2 \end{bmatrix}$$

Matrix Algebra

- In the systems of linear equations, a concept of equivalence is important.
- Let us consider two systems, each of them consisting of n equation with n unknowns expressed by

$$Ax = b \qquad Bx = d$$

  If the two systems have precisely the same solutions of equations, we call them **equivalent systems**.
- To solve a system of equations, we can instead solve any equivalent system; no solutions are lost and no new ones appear.

Matrix Algebra

These **elementary operations** are of the following three types. (Here, ε_i denotes the ith equation in the system.)

1. Interchanging two equations in the system: $\varepsilon_i \leftrightarrow \varepsilon_j$
2. Multiplying an equation by a nonzero number: $\lambda\varepsilon_i \to \varepsilon_i$
3. Adding to an equation a multiple of some other equation: $\varepsilon_i + \lambda\varepsilon_j \to \varepsilon_i$

Theorem on Equivalent Systems

THEOREM 1 (Theorem on Equivalent Systems)

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

Proof.

Suppose that an elementary operation transforms the system $Ax = b$ into the system $Bx = d$.

- For the Type 1, then the two systems consist of the same equations. If x solves the first system, then it solves the second, and vice versa.

□

Theorem on Equivalent Systems

Proof.

For the Type 2, then suppose that the ith equation has been multiplied by a scalar λ, with $\lambda \neq 0$. The ith and jth equations in $Ax = b$ are

$$\begin{aligned} a_{i1}x_1 + \ldots + a_{in}x_n &= b_i \qquad (1) \\ a_{j1}x_1 + \ldots + a_{jn}x_n &= b_j \qquad (2) \end{aligned}$$

and the ith equation in $Bx = d$ is

$$\lambda a_{i1}x_1 + \ldots + \lambda a_{in}x_n = \lambda b_i \qquad (3)$$

Any vector x that satisfies Equation (1) satisfies Equation (3), and vice versa, because $\lambda \neq 0$. □

Theorem on Equivalent Systems

Proof.

The ith and jth equations in $Ax = b$ are

$$
\begin{aligned}
a_{i1}x_1 + \ldots + a_{in}x_n &= b_i \qquad (4)\\
a_{j1}x_1 + \ldots + a_{jn}x_n &= b_j \qquad (5)
\end{aligned}
$$

For the Type 3, assume that λ times the jth equation has been added to the ith. Then the ith equation in $Bx = d$ is

$$(a_{i1} + \lambda a_{j1})x_1 + \ldots + (a_{in} + \lambda a_{jn})x_n = b_i + \lambda b_j \qquad (6)$$

If $Ax = b$, then Equation (4) and (5) are true. Hence, (6) is true.
If we suppose that x solves $Bx = d$, then Equation (6) and (5) are true.
If λ times Equation (5) is subtracted from Equation (6), the result is Equation (4). Hence, $Ax = b$. □

Matrix Properties

- The $n \times n$ matrix

$$I = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

is called an **identity matrix**. It has the property that $IA = A = AI$ for any matrix A of size $n \times n$.

Matrix Properties

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \alpha & \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We see that if a matrix has a right inverse, then the latter is not necessarily unique.

Matrix Properties

THEOREM 2 (Theorem on Right Inverse)

A square matrix can possess at most one right inverse.

Proof.

Let $AB = I$, where A, B, and I are all $n \times n$ matrices. Denote by $A^{(j)}$ the jth column of A and by $I^{(k)}$ the kth column of I. The equation $AB = I$ means that

$$\sum_{j=1}^{n} b_{jk} A^{(j)} = I^{(k)} \qquad (1 \leqslant k \leqslant n) \tag{7}$$

Each column of I is a linear combination of the columns of A. Since the columns of I span $\mathbb{R}^n$, the same is true of the columns of A. Hence, the columns of A form a basis for $\mathbb{R}^n$, consequently, the coefficients b_{jk} in Equation (7) are uniquely determined. □

Matrix Properties

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

$$b_{11} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + b_{21} \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \ldots + b_{n1} \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\sum_{j=1}^{n} b_{j1} A^{(j)} = I^{(1)} \qquad (1 \leqslant k \leqslant n)$$

Matrix Properties

THEOREM 3 (Theorem on Matrix Inverse)

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

Proof.

Let $C = BA - I + B$. Then

$$AC = ABA - AI + AB = A - A + I = I$$

Thus, C (as well as B) is a right inverse of A. By Theorem 2, $B = C$; hence, $BA = I$. $\square$

Matrix Properties

- If a square matrix A has a right inverse B, then B is unique and $BA = AB = I$. We call B the **inverse** of A and say that A is **invertible** or **nonsingular**. We write $B = A^{-1}$ and $A = B^{-1}$.
- If A is invertible, the system of equations $Ax = b$ has the solution $x = A^{-1}b$.

Matrix Properties

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.

The elementary operations expressed in terms of the rows of a matrix A, are

1. The interchange of two rows in A: $A_s \leftrightarrow A_t$
2. Multiplying one row by a nonzero constant: $\lambda A_s \rightarrow A_s$
3. Adding to one row a multiple of another: $A_s + \lambda A_t \rightarrow A_s$

Matrix Properties

- Each elementary row operation on A can be accomplished by multiplying A on the left by an elementary matrix.

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

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

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

Matrix Properties

- Introducing elementary matrices E_1, E_2, ..., E_m. If a matrix is invertible, a sequence of elementary row operations can be applied to A, reducing it to I. Thus, we have

$$E_m E_{m-1} \dots E_2 E_1 A = I$$

  From this it follows that $A^{-1} = E_m E_{m-1} \dots E_2 E_1$.
- Consequently, A^{-1} can be obtained by subjecting I to the same sequence of elementary row operations.

Matrix Properties

$$
\begin{aligned}
A &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 2 & 4 & 7 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} &= I \\
E_1 &= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} &= E_2 \\
E_1 A &= \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 2 & 4 & 7 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} &= E_1 I \\
E_2 E_1 A &= \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} &= E_2 E_1 I
\end{aligned}
$$

Matrix Properties

$$
\begin{aligned}
E_3 &= \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_4 \\
E_3 E_2 E_1 A &= \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = E_3 E_2 E_1 I \\
E_4 E_3 E_2 E_1 A &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 9 & -2 & -3 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = E_4 E_3 E_2 E_1 I = A^{-1}
\end{aligned}
$$

Matrix Properties

THEOREM 4 (Theorem on Nonsingular Matrix Properties)

For an $n \times n$ 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 subjective (onto).*
7. *The equation $Ax = 0$ implies $x = 0$.*
8. *For each $b \in \mathbb{R}^n$, there is exactly one $x \in \mathbb{R}^n$ such that $Ax = b$.*
9. *A is a product of elementary matrices.*
10. *0 is not an eigenvalue of A.*

Matrix Properties

- An important fundamental concept is the **positive definiteness** of a matrix. A matrix A is **positive definite** if $x^T Ax > 0$ for every nonzero vector x. For example, the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

is positive definite since

$$x^T Ax = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1 + x_2)^2 + x_1^2 + x_2^2 > 0$$

for all x_1 and x_2 except $x_1 = x_2 = 0$. Here, $x^T Ax$ is called a **quadratic form**.

Partitioned matrices

- It is convenient to partition matrices into submatrices and compute products as if the submatrices were numbers.

$$\begin{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} & \begin{bmatrix} 1 & -1 & 0 & 1 \end{bmatrix} \\ \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 \end{bmatrix} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix} & \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 1 \\ -1 & 1 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} 1 & 2 & 7 \end{bmatrix} & \begin{bmatrix} 2 & 5 \end{bmatrix} \\ \begin{bmatrix} -3 & 1 & 3 \\ -3 & 3 & 2 \\ 4 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix} & \begin{bmatrix} 0 & 2 \\ -2 & 1 \\ 0 & 1 \\ 1 & 6 \end{bmatrix} \end{bmatrix}$$

Matrix Properties

- If the submatrices are denoted by single letters, we have a product of the form

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

  One can verify that $C_{ij} = \sum_{s=1}^{2} A_{is}B_{sj}$.
- For example, $C_{11} = A_{11}B_{11} + A_{12}B_{21}$.

Matrix Properties

THEOREM 5 (Theorem on Multiplication of Partitioned Matrices)

If each product $A_{is}B_{sj}$ can be formed, and if $C_{ij} = \sum_{s=1}^{n} A_{is}B_{sj}$, then $C = AB$.

Proof.

Let the dimensions of A_{ij} be $m_i \times n_j$. Let the dimensions of B_{ij} be $\widehat{m}_i \times \widehat{n}_j$. Since $A_{is}B_{sj}$ exists, we must have $n_s = \widehat{m}_s$ for all s. Then C_{ij} will have dimension $m_i \times \widehat{n}_j$. Now select an arbitrary element c_{ij} in the matrix C. Suppose that c_{ij} lie in the block C_{rs} and is in the pth row and qth column of C_{rs}. Then we must have

$$\begin{aligned} i &= m_1 + m_2 + \ldots + m_{r-1} + p \\ j &= \widehat{n}_1 + \widehat{n}_2 + \ldots + \widehat{n}_{s-1} + q \end{aligned}$$

$\square$

Matrix Properties

Proof.

Then we have

$$c_{ij} = (C_{rs})_{pq} = (\sum_{t=1}^{n} A_{rt}B_{ts})_{pq} = \sum_{t=1}^{n}(A_{rs}B_{ts})_{pq} = \sum_{t=1}^{n}\sum_{\alpha=1}^{n_t}(A_{rt})_{p\alpha}(B_{ts})_{\alpha q}$$

The elements $(A_{rt})_{p\alpha}$ lie in row i of A. These elements fill out the entire row i of A since $1 \leqslant t \leqslant n$ and $1 \leqslant \alpha \leqslant n_t$. The elements $(B_{ts})_{\alpha q}$ lie in column j of B. Also, the entire column j of B is present and appears in its natural order. Hence,

$$c_{ij} = \sum_{\beta=1}^{n}(A)_{i\beta}(B)_{\beta j} = (AB)_{ij}$$

□