# Numerical Analysis

## Chapter Three: Solution of Nonlinear Equations

主讲人 朱亚菲
幻灯片制作 朱亚菲

中国海洋大学 信息科学与工程学院

2013 年 10 月 27 日

# Contents

# Contents

# Contents

# 4.4.1 Vector Norms

- On a vector space $V$, a **norm** is a function $\|\cdot\|$ from $V$ to the set of nonnegtive reals that obeys these postulates:

$$\| x \| > 0 \ \mathit{if}\ x \neq 0, x \in V \tag{1}$$

$$\| \lambda x \| = |\lambda| \, \| x \| \ \ \mathit{if}\ \lambda \in R,\ \ x \in V \tag{2}$$

$$\| x + y \| \leqslant \| x \| + \| y \| \ \ \mathit{if}\ x, y \in V \tag{3}$$

# 4.4.1 Vector Norms

- **Euclidean $l_2$-norm**

$$\| x \|_2 = \left( \sum_{i=1}^{n} x_i^2 \right)^{1/2} \qquad where \quad x = (x_1, x_2, \ldots, x_n)^T$$

- **$l_\infty$-norm**

$$\| x \|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|$$

- **$l_1$-norm**

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

# Contents

# 4.4.2 Matrix Norms

**THEOREM 1 (Theorem on Subordinate Matrix Norm)**

*If $\| \cdot \|$ is any norm on $R^n$, then the equation*

$$\| A \| = \sup_{\|u\|=1} \{ \| Au \| : u \in R^n \}$$

*defines a norm on the linear space of all $n \times n$ matrices.*

# 4.4.2 Matrix Norms

## Proof.

We shall verify the **three** axioms for a norm.

- **First**, if $A \neq 0$, then $A$ has at least one nonzero column, say, $A^{(j)} \neq 0$.
  Consider the vector $x$ having 1 as its $j$th component and 0's elsewhere; that is , $x = (0, \ldots, 0, 1, 0, \ldots, 0)^T$.
  Obviously, $x \neq 0$ and the vector $v = x/ \parallel x \parallel$ is of norm 1.
  Hence, $\parallel A \parallel \geqslant \parallel Av \parallel = \dfrac{\parallel Ax \parallel}{\parallel x \parallel} = \dfrac{\parallel A^{(j)} \parallel}{\parallel x \parallel} > 0$

$\square$

# 4.4.2 Matrix Norms

**Proof.**

- **Next**, from Property (2) of the vector norm, we have

$$\| \lambda A \| = sup\{\| \lambda A u \| : \| u \| = 1\}$$

$$= |\lambda| sup\{\| Au \| : \| u \| = 1\} = |\lambda| \| A \|$$

$\square$

**Proof.**

- **Finally**, for the triangle inequality

$$\begin{aligned} \| A + B \| &= sup\{\| (A+B)u \| : \| u \| = 1\} \\ &\leqslant sup\{\| Au \| + \| Bu \| : \| u \| = 1\} \\ &\leqslant sup\{\| Au \| : \| u \| = 1\} + sup\{\| Bu \| : \| u \| = 1\} = \| A \| + \| B \| \end{aligned}$$

$\square$

# 4.4.2 Matrix Norms

**$\| Ax \| \leqslant \| A \| \| x \| \quad (x \in R^n)$**

**Proof** It is true for $x = 0$.
If $x \neq 0$, then the vector $v = x / \| x \|$ is of norm 1.
Hence,

$$\| A \| \geqslant \| Av \| = \frac{\| Ax \|}{\| x \|}$$

# 4.4.2 Matrix Norms

**THEOREM 2 (Theorem on Infinity Matrix Norm)**

*If the vector norm* $\| \cdot \|_\infty$ *is defined by*

$$\| x \|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|$$

*then its subordinate matrix norm is given by*

$$\| A \|_\infty = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

# 4.4.2 Matrix Norms

**Proof.**

$Au$ is a $n \times 1$ matrix.

$$\| A \|_\infty = \sup_{\|u\|_\infty = 1} \| Au \|_\infty$$

$$= \sup_{\|u\|_\infty = 1} \left\{ \max_{1 \leqslant i \leqslant n} |(Au)_i| \right\} = \max_{1 \leqslant i \leqslant n} \left\{ \sup_{\|u\|_\infty = 1} |(Au)_i| \right\}$$

$$= \max_{1 \leqslant i \leqslant n} \left\{ \sup_{\|u\|_\infty = 1} |\sum_{j=1}^{n} a_{ij} u_j| \right\} = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

$\square$

# Contents

# 4.4.3 Condition Number

## EXAMPLE 1

Consider an equation $Ax = b$, where $A$ is an $n \times n$ matrix. Suppose that $A$ is invertible. Suppose that the vector $b$ is perturbed to obtain a vector $\widetilde{b}$. If $x$ and $\widetilde{x}$ satisfy $Ax = b$ and $A\widetilde{x} = \widetilde{b}$, by how much do $x$ and $\widetilde{x}$ differ, in absolute and relative terms?

## Solution

$$\| x - \widetilde{x} \| = \| A^{-1}b - A^{-1}\widetilde{b} \| = \| A^{-1}(b - \widetilde{b}) \| \leqslant \| A^{-1} \| \| b - \widetilde{b} \|$$

$$= \| A^{-1} \| \| Ax \| \frac{\| b - \widetilde{b} \|}{\| b \|} \leqslant \| A^{-1} \| \| A \| \| x \| \frac{\| b - \widetilde{b} \|}{\| b \|}$$

Hence, $\dfrac{\| x - \widetilde{x} \|}{\| x \|} \leqslant \kappa(A) \dfrac{\| b - \widetilde{b} \|}{\| b \|}$ where $\kappa(A) \equiv \| A \| \cdot \| A^{-1} \|$

The number $\kappa(A)$ is called a **condition number** of the matrix $A$.

# 4.4.3 Condition Number

- If we solve a system of equations $Ax = b$ numerically, we obtain not the exact solution but an approximate solution $\widetilde{x}$.
- One can test $\widetilde{x}$ by forming $A\widetilde{x}$ to see whether it is close to $b$.
- Thus, we obtain the **residual vector** $r = b - A\widetilde{x}$
- The difference between the exact solution $x$ and the approximate solution $\widetilde{x}$ is called the **error vector** $e = x - \widetilde{x}$
- The following relationship $Ae = r$ between the error vector and the residual vector is of fundamental importance.

# 4.4.3 Condition Number

**THEOREM 3 (Theorem on Bounds Involving Condition Number)**

*In solving systems of equations $Ax = b$, the condition number $\kappa(A)$, the residual vector $r$, and the error vector $e$ satisfy the following inequality:*

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

**Proof.**

**Since** $\| e \| \| b \| = \| A^{-1} r \| \| Ax \| \leqslant \| A^{-1} \| \| r \| \| A \| \| x \|$

**Thus** $\frac{\|e\|}{\|x\|} \leqslant \kappa(A) \frac{\|r\|}{\|b\|}$

**Similarly** $\| r \| \| x \| = \| Ae \| \| A^{-1} b \| \leqslant \| A \| \| e \| \| A^{-1} \| \| b \|$

**And thus** $\frac{1}{\kappa(A)} \frac{\|r\|}{\|b\|} \leqslant \frac{\|e\|}{\|x\|}$ □

# 4.4.3 Condition Number

- A matrix with a large condition number is said to be **ill conditioned**. For an illconditioned matrix $A$, there will be cases in which the solution of a system $Ax = b$ will be very sensitive to small changes in the vector $b$.
- In other words, to attain a certain precision in the determination of $x$, we shall require significantly higher precision in $b$. If the condition number of $A$ is of moderate size, the matrix is said to be **well conditioned**.

# Contents

# 4.5 Neumann Series and Iterative Refinement

## The concept of convergence in a vector space

If a vector space $V$ is assigned a norm $\| \cdot \|$, then the pair $(V, \| \cdot \|)$ is a **normed linear space**. The notion of a convergence sequence of vectoes $v^{(1)}, v^{(2)}, \ldots$ is then defined as follows: We say that the given sequence **converges** to a vector $v$ if

$$\lim_{k \to \infty} \| v^{(k)} - v \| = 0$$

# 4.5 Neumann Series and Iterative Refinement

**THEOREM 1 (Theorem on Neumann Series)**

*If $A$ is an $n \times n$ matrix such that $\| A \| < 1$, then $I - A$ is invertible, and*

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k \tag{4}$$

**Proof.**

**First**, we shall show that $I - A$ is invertible. If it is not invertible, then it is singular, and there exists a vector $x$ satisfying $\| x \| = 1$ and $(I - A)x = 0$

Then $1 = \| x \| = \| Ax \| \leqslant \| A \| \| x \| = \| A \|$

which contradicts the hypothesis that $\| A \| < 1$. □

# 4.5 Neumann Series and Iterative Refinement

## Proof.

**Next**, we shall show that the patial sums of the Neumann series converge to $(I - A)^{-1}$: $\sum_{k=0}^{m} A^k \rightarrow (I - A)^{-1} \quad (as\, m \rightarrow \infty)$
It will suffice to prove that

$$(I - A) \sum_{k=0}^{m} A^k \rightarrow I \quad as\, m \rightarrow \infty \tag{5}$$

The left-hand side can be written as

$$(I - A) \sum_{k=0}^{m} A^k = \sum_{k=0}^{m} (A^k - A^{k+1}) = A^0 - A^{m+1} = I - A^{m+1}$$

Since $\| A^{m+1} \| \leqslant \| A \|^{m+1} \rightarrow 0$ as $m \rightarrow \infty$, this establishes (5). □

# 4.5 Neumann Series and Iterative Refinement

## THEOREM 2 (Theorem on Invertible Matrices)

*If $A$ and $B$ are $n \times n$ matrices that $\| I - AB \| < 1$, then $A$ and $B$ are invertible. Furthermore, we have*

$$A^{-1} = B\sum_{k=0}^{\infty}(I - AB)^K \quad and \quad B^{-1} = \sum_{k=0}^{\infty}(I - AB)^k A$$

## Proof.

$$(AB)^{-1} = \sum_{k=0}^{\infty}(I - AB)^k$$

**Hence**, $A^{-1} = BB^{-1}A^{-1} = B(AB)^{-1} = B\sum_{k=0}^{\infty}(I - AB)^k$
$B^{-1} = B^{-1}A^{-1}A = (AB)^(-1)A = \sum_{k=0}^{\infty}(I - AB)^k A$ $\square$

# Contents

# 4.5.1 Iterative Refinement

- If $x^{(0)}$ is an approximate solution of the equation $Ax = b$ then the precise solution $x$ is given by

$$x = x^{(0)} + A^{-1}(b - Ax^{(0)}) = x^{(0)} + e^{(0)}$$

where $e^{(0)} = A^{-1}(b - Ax^{(0)})$ is called the **error vector**.

- The **residual vector** corresponding to the approximate solution $x^{(0)}$ is $r^{(0)} = b - Ax^{(0)}$. It is computable. Of course, we do not want to compute $A^{-1}$, but the vector $e^{(0)} = A^{-1}r^{(0)}$ can be obtained by solving the equation

$$Ae^{(0)} = r^{(0)}$$

These remarks lead to a numerical procedure called **iterative improvement** or **iterative refinement**.

# 4.5.1 Iterative Refinement

- To analyze this algorithm theoretically, we adopt the point of view that our solution $x^{(0)}$ is obtained by the formula $x^{(0)} = Bb$
- where $B$ is an approximate inverse of $A$. The iterative process then can be written

$$x^{(k+1)} = x^{(k)} + B(b - Ax^{(k)}) \quad (k \geqslant 0) \tag{6}$$

- We interpret the loose expression "$B$ is an approximate inverse of $A$" to mean that $\| I - AB \| < 1$. By Theorem 2, $A^{-1}$ is given by

$$A^{-1} = B \sum_{k=0}^{\infty} (I - AB)^k \tag{7}$$

- Thus, the exact solution of the equation $Ax = b$ is

$$x = B \sum_{k=0}^{\infty} (I - AB)^k b \tag{8}$$

# 4.5.1 Iterative Refinement

**THEOREM 3 (Theorem on Iterative Improvement)**

*If $\| I - AB \| < 1$, then the method of iterative improvement given by Equation(6) produces the sequence of vectors*

$$x^{(m)} = B\sum_{k=0}^{m}(I - AB)^k b \quad (m \geqslant 0)$$

*These are the partial sums in Equation(8) and therefore converge to $x$.*

# 4.5.1 Iterative Refinement

**Proof.**

We use induction. Since $x^{(0)} = Bb$, the case $m = 0$ is true. If the $m$th case is assumed true, then the $(m+1)st$ case is true since

$$x^{(m+1)} = x^{(m)} + B(b - Ax^{(m)})$$

$$= B\sum_{k=0}^{m}(I - AB)^{K}b + Bb - BAB\sum_{k=0}^{m}(I - AB)^{k}b$$

$$= B\left\{b + (I - AB)\sum_{k=0}^{m}(I - AB)^{k}b\right\} = B\sum_{k=0}^{m+1}(I - AB)^{k}b$$

□