CS 3331 Numerical Methods

Lecture 3: Linear Systems

Cherung Lee

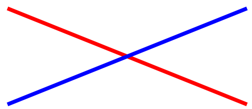
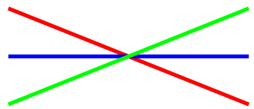
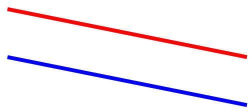
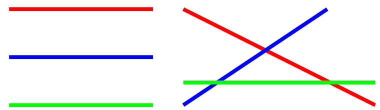
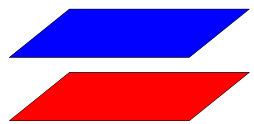
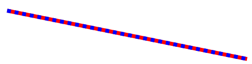
Outline

- Basic ideas from linear algebra
- Triangular system
- Gaussian elimination and LU decomposition
  - Pivoting and stability
- Special matrices
  - Symmetric positive definite matrix: Cholesky decomp
  - Tridiagonal matrix: Thomas algorithm

Basic Ideas from Linear Algebra

Linear system

- $\mathbf{Ax} = \mathbf{b}$: $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^{m}$, $\mathbf{x} \in \mathbb{R}^{n}$ (unknown)

	Square $m = n$	Overdetermined $m > n$	Underdetermined $m < n$
One solution			
No solution			
Inf-many sol			

Vector norm LVF pp.125

- A vector norm $\|.\|$ is a function $C^n \rightarrow \{0\} \cup R^+$ that satisfies
  - $\|\mathbf{x}\| > 0, \forall \mathbf{x} \neq 0$, $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$.
  - $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$ for $c \in C$.
  - $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality)
- Some frequently used vector norms:
  - $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ (1-norm)
  - $\|\mathbf{x}\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$ (2-norm, Euclidean norm)
  - $\|\mathbf{x}\|_\infty = \max_{i=1,..n} |x_i|$ (infinite-norm)

- Example: Let $\mathbf{x} = [3\ -1\ -4\ 1\ -3]^T$
  - $\|\mathbf{x}\|_1 = |3| + |-1| + |-4| + |1| + |-3| = 12$
  - $\|\mathbf{x}\|_2 = (3^2 + (-1)^2 + (-4)^2 + 1^2 + (-3)^2)^{1/2} = 36^{1/2} = 6$
  - $\|\mathbf{x}\|_\infty = \max\{|3|, |-1|, |-4|, |1|, |-3|\} = 4$

Matrix norm LVF pp.125

- Similar definition as a vector norm.
- Consistency property: $\|\mathbf{AB}\| < \|\mathbf{A}\|\|\mathbf{B}\|$
- Subordinate matrix norm: derived from a vector norm

$$\|\mathbf{A}\|_b = \max_{\|\mathbf{x}\|_b=1} \|\mathbf{Ax}\|_b$$

  - $\|\mathbf{A}\|_1 = \max_{j=1..n} \sum_{i=1}^{m} |a_{ij}|$ (1-norm)
  - $\|\mathbf{A}\|_2 =$ largest singular value (2-norm, Euclidean norm)
  - $\|\mathbf{A}\|_\infty = \max_{i=1..m} \sum_{j=1}^{n} |a_{ij}|$ (infinite-norm)

- Forbenius norm: $\|\mathbf{A}\|_F = \left[\sum_{i,j=1}^{m,n} |a_{ij}|^2\right]^{1/2}$

- Example:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -4 \\ 0 & 3 & -3 \\ 4 & -2 & 1 \end{pmatrix} \quad \begin{matrix} 7 \\ 6 \\ 7 \end{matrix}$$

$$\begin{matrix} 5 & 7 & 8 \end{matrix}$$

  - $\|\mathbf{A}\|_1 = \max\{5, 7, 8\} = 8$
  - $\|\mathbf{A}\|_\infty = \max\{7, 6, 7\} = 7$
  - $\begin{aligned} \|\mathbf{A}\|_F &= (1^2 + 0^2 + 4^2 + 2^2 + 3^2 + 2^2 + 4^2 + 3^2 + 1^2)^{1/2} \\ &= \sqrt{60} \end{aligned}$
  - Singular values of $\mathbf{A}$ are $6.45, 4.22, 0.77$. $\|\mathbf{A}\|_2 = 6.45$

Condition number of a linear system LVF pp.125

- How the solution $\mathbf{x}$ changes when input problem $\mathbf{A}$ is perturbed by $\mathbf{E}$.
- Let $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ be the solution of the original system; $\tilde{\mathbf{x}} = (\mathbf{A} + \mathbf{E})^{-1}\mathbf{b}$ be the solution of the perturbed one.
  - $(\mathbf{A} + \mathbf{E})\tilde{\mathbf{x}} = \mathbf{b} = \mathbf{A}\mathbf{x}$. Let $\Delta\mathbf{x} = \mathbf{x} - \tilde{\mathbf{x}}$. $\mathbf{A}\Delta\mathbf{x} = \mathbf{E}\tilde{\mathbf{x}}$.
$$\|\Delta\mathbf{x}\| = \|\mathbf{A}^{-1}\mathbf{E}\tilde{\mathbf{x}}\| \leq \|\mathbf{A}^{-1}\|\|\mathbf{E}\|\|\tilde{\mathbf{x}}\|$$
$$\frac{\|\Delta\mathbf{x}\|}{\|\tilde{\mathbf{x}}\|} \leq \|\mathbf{A}^{-1}\|\|\mathbf{E}\| = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\frac{\|E\|}{\|A\|}$$
  - $\text{cond}(\mathbf{A}) = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|$.

Triangular System

Triangular matrix

- Definition
  - Upper triangular matrix: $\{\mathbf{A} = (a_{ij}) : a_{ij} = 0, \forall i < j\}$
  - Lower triangular matrix: $\{\mathbf{A} = (a_{ij}) : a_{ij} = 0, \forall i > j\}$
  - Unit triangular matrix: $\mathbf{A}$ is triangular and $a_{ii} = 1$.
- Properties:
  - The (unit) inverse of upper (lower) triangular matrix is (unit) upper(lower) triangular.
  - The (unit) product of two upper (lower) triangular matrix is (unit) upper(lower) triangular.

Forward substitution and back substitution

- Forward substitution: solves $\mathbf{Lx} = \mathbf{b}$. $\mathbf{L}$ is lower triangular.

$$\begin{pmatrix} \ell_{11} & 0 \\ \ell_{21} & \ell_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Rightarrow \begin{cases} x_1 & = & b_1/\ell_{11} \\ x_2 & = & (b_2 - \ell_{21}x_1)/\ell_{22} \end{cases}$$

  – General form: $x_i = \left(b_i - \sum_{j=1}^{i-1} \ell_{ij} x_j \right) \ell_{ii}^{-1}$

- Back substitution: solves $\mathbf{Ux} = \mathbf{b}$, $\mathbf{U}$ is upper triangular.

  – General form: $x_i = \left(b_i - \sum_{j=i+1}^{n} u_{ij} x_j \right) u_{ii}^{-1}$

- Computational effort: n^2 flops

Gaussian Elimination

Gaussian Elimination LVF pp.102-110

- Solve $\mathbf{Ax} = \mathbf{b}$ for a general matrix $\mathbf{A}$

  1. Transform $[\mathbf{A}\ \mathbf{b}]$ into an upper triangular matrix $[\mathbf{U}\ \mathbf{c}]$
  2. Solve $\mathbf{Ux} = \mathbf{c}$ by back substitution.

- Step 1 is performed by forward elimination.

1.1 For $i = 1, \cdots, n-1$

1.2 For $j = i+1, \cdots, n$

1.3 Multiply ith row by $-a_{ji}/a_{ii}$ and add to jth row.

- Computational effort of step 1 is about $\frac{2}{3}n^3$.

An example

$$\mathbf{A}^{(1)} = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ x & x & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix}$$

$$\mathbf{A}^{(2)} = \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & x & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{pmatrix}$$

$$\mathbf{A}^{(3)} = \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{pmatrix}$$

LU decomposition LVF pp.144-145, 150-151

- Elementary transformation

$$\mathbf{Mx} = \begin{pmatrix} 1 & 0 \\ -x_2/x_1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

- General form

$$\mathbf{M}_k\mathbf{x} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & & 1 & 0 & & 0 \\ 0 & & m_{k+1,k} & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & m_{n,k} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $m_{i,k} = -x_i/x_k$ for $i = k+1, \cdots, n$.

- Forward elimination can be implemented by a sequence of elementary transformations.

$$\mathbf{M}_{n-1} \cdots \mathbf{M}_1 \mathbf{A} = \mathbf{U}$$

- $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{L}\mathbf{U}$, where $\mathbf{L} = \mathbf{M}_1^{-1} \cdots \mathbf{M}_{n-1}^{-1}$.

- $\mathbf{M}_k^{-1} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & & 1 & 0 & & 0 \\ 0 & & -m_{k+1,k} & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -m_{n,k} & 0 & \cdots & 1 \end{pmatrix}$

LU decomposition for solving linear system

- Algorithm
  1. Decompose $\mathbf{A} = \mathbf{LU}$.
  2. Solve $\mathbf{Ly} = \mathbf{b}$ by forward substitution.
  3. Solve $\mathbf{Ux} = \mathbf{y}$ by back substitution.
- $\mathbf{x} = \mathbf{U}^{-1}\mathbf{y} = \mathbf{U}^{-1}(\mathbf{L}^{-1}\mathbf{b}) = (\mathbf{U}^{-1}\mathbf{L}^{-1})\mathbf{b} = (\mathbf{LU})^{-1}\mathbf{b} = \mathbf{A}^{-1}\mathbf{b}$
- Computational effort
  - Step 1: $2/3n^3$
  - Step 2 and step 3: $2n^2$

Stability and Pivoting

An example LVF pp.28-29

$$\begin{cases} 0.001x_1 & + & x_2 & = & 3 \\ x_1 & + & 2x_2 & = & 5 \end{cases}$$

- Forward elimination: $-998x_2 = -2995$
- Rounding: $x_2 = 3$.
- Back substitution to first equation: $x_1 = 0$

• Solution is $x_1 = -1.002, x_2 = 3.001$.

- Relative error of x_2 is 0.3×10^{-3}.
- Relative error of x_1 is 1.

Alternative formulation

- Exchange the order of two equations

$$\begin{cases} x_1 + 2x_2 = 5 \\ 0.001x_1 + x_2 = 3 \end{cases}$$

  - Forward elimination: $0.998x_2 = 2.995$.
  - Rounding: $x_2 = 3$.
  - Back substitution to first equation: $x_1 = -1$

- Relative error of x_1 becomes 2×10^{-3}.

Growth factor

- Growth factor
$$\gamma = \frac{\text{the largest element during forward elimination}}{\text{the largest element of } \mathbf{A}}.$$

- A large growth factor implies a large backward error

Row pivoting – partial pivoting LVF pp.148-151

- $\mathbf{A} = \begin{pmatrix} 0.001 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1000 & 1 \end{pmatrix} \begin{pmatrix} 0.001 & 1 \\ 0 & -998 \end{pmatrix} = \mathbf{LU}$

- $\bar{\mathbf{A}} = \begin{pmatrix} 1 & 2 \\ 0.001 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.001 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 0.998 \end{pmatrix} = \bar{\mathbf{L}}\bar{\mathbf{U}}$

- Before applying $\mathbf{M}_k$

  1. Find the maximum value of $|\mathbf{A}^{(k)}(i,k)|$, for $i = k, \cdots, n$, and let the row index be j.

  2. Exchange row k and row j. (Use row j to eliminate other rows.)

Permutation matrix

- Permutation matrix:

$$\bar{\mathbf{A}} = \begin{pmatrix} 1 & 2 \\ 0.001 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0.001 & 1 \\ 1 & 2 \end{pmatrix} = \mathbf{P}\mathbf{A},$$

- LU decomposition with row pivoting:

$$\mathbf{M}_{n-1}\mathbf{P}_{n-1} \cdots \mathbf{M}_2\mathbf{P}_2\mathbf{M}_1\mathbf{P}_1\mathbf{A} = \mathbf{U}$$

  which is identical to

$$\mathbf{M}_{n-1} \cdots \mathbf{M}_2\mathbf{M}_1\mathbf{P}_{n-1} \cdots \mathbf{P}_2\mathbf{P}_1\mathbf{A} = \mathbf{U}$$

- The matrix decomposition form: $\mathbf{PA} = \mathbf{LU}$

A counter example

- With partial pivoting, the growth factor can still be large.

$$\mathbf{W}_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 1 \\ -1 & -1 & 1 & 0 & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 \end{pmatrix}$$

- The growth factor of $\mathbf{W}_n$ is 2^{n-1}.

Special Linear Systems

Tridiagonal matrix LVF pp.146

- $\mathbf{A}$ is tridiagonal if $a_{ij} = 0$ for $|i-j| > 1$, $i,j = 1,\cdots,n$.
- Thomas method: solving $\mathbf{Ax} = \mathbf{r}$ where

$$\mathbf{A} = \begin{pmatrix} d_1 & a_1 & & \\ b_2 & d_2 & a_2 & \\ & \ddots & \ddots & \ddots \\ & & b_n & d_n \end{pmatrix}, \mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$$

$$a_1 = \frac{a_1}{d_1} \qquad r_1 = \frac{r_1}{d_1} \qquad x_i = r_i - a_i x_{i+1}$$

$$\hookrightarrow \quad a_i = \frac{a_i}{d_i - b_i a_{i-1}} \qquad r_i = \frac{r_i - b_i r_{i-1}}{d_i - b_i a_{i-1}} \qquad \uparrow$$

$$\hookrightarrow \quad r_n = \frac{r_n - b_n r_{n-1}}{d_n - b_n a_{n-1}} \quad \longrightarrow \quad x_n = r_n$$

- Operation count $O(n)$

Symmetric positive definite matrix LVF pp.11

- $\mathbf{A}$ is symmetric if $\mathbf{A}^T = \mathbf{A}$. ($a_{ij} = a_{ji}$ for $i, j = 1, \cdots, n$.)
- If $\mathbf{A}$ is real and symmetric, the following are equivalent:
  - $\mathbf{A}$ is positive definite.
  - $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for all nonzero vector $\mathbf{x}$.
  - All eigenvalues of $\mathbf{A}$ are positive.
  - All upper left submatrices of $\mathbf{A}$ have positive determinants.
  - There exists a unique lower triangular nonsingular matrix $\mathbf{L}$ such that $\mathbf{A} = \mathbf{L}\mathbf{L}^T$. (Cholesky decomposition)

Cholesky decomposition LVF pp.154

- $\mathbf{A} = \mathbf{L}\mathbf{L}^T = \begin{pmatrix} \ell_{11} & & & \\ \ell_{21} & \ell_{22} & & \\ \vdots & & \ddots & \\ \ell_{n1} & \cdots & \ell_{n,n-1} & \ell_{nn} \end{pmatrix} \begin{pmatrix} \ell_{11} & \ell_{21} & \cdots & \ell_{n1} \\ & \ell_{22} & & \vdots \\ & & \ddots & \ell_{n,n-1} \\ & & & \ell_{nn} \end{pmatrix}$

  - $\ell_{ij} = \frac{1}{\ell_{jj}} \sqrt{a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk}}$

  - $\ell_{ii} = a_{ii} - \sum_{k=1}^{i-1} \ell_{ik}^2$

- Operation count $= \frac{1}{3}n^3$. (Half of LU decomposition.)