MATH426: Computational Mathematics — Lecture Notes

Andrew Ngo

1 Computer Numbers

2 (Square) Linear Systems of Equations

2.1 LU Factorization for Solving $Ax = b$

Assume $A = [a_i j]_{i,j=1,n}$ is a $n \times n$ matrix with $det(A) \neq 0$.

Definition 1. *The matrix A has a LU factorization if there exist two $n \times n$ matrices L-lower triangular, U-upper triangular such that $A = LU$. Note that $det(A) = det(L)det(U) \neq 0$ and $det(L) = l_{11}l_{22} \dots l_{nn} \neq 0$ and $det(U) = u_{11}u_{22} \dots u_{nn} \neq 0$. In other words, all diagonal entries of L and U are nonzero.*

Example 1 (Solving $Ax = b$ using $A = LU$)**.** *Assume that $A = LU$ is given and A is an invertible matrix ($det(A) \neq 0$).*

$$Ax = b \longleftrightarrow (LU)x = b \longleftrightarrow L(Ux) = b \tag{1}$$

Denote $Ux = y$. Then,

$$Ax = b \longleftrightarrow L(Ux) = b, y = Ux \longleftrightarrow Ly = b, Ux = y \tag{2}$$

Solving for the solution $x \in \mathbb{R}^n$ reduces to two steps:

1. *Solve $Ly = b$ for y using Forward Solve.*
2. *Solve $Ux = y$ for x using Back Solve.*

Remark 1. *Solving LU factorization is much faster because it involves less operations than Gaussian Elimination (GE). Operations for GE is $O(n^3)$. Operations for Back Solve is $O(n^2)$ and Forward Solve is $O(n^2)$, therefore operations for LU factorization is $O(2n^2) = O(n^2)$.*

Definition 2. *We say that a $n \times n$ lower triangular matrix L is unit lower triangular if $l_{11} = l_{22} = \dots = l_{nn} = 1$.*

Theorem 1. *If A is an invertible $n \times n$ matrix, then there are* ***unique*** *L-unit lower triangular and U-upper triangular matrices such that $A = LU$.*

2.1.1 Gaussian Elimination on A produces L and U such that $A = LU$.

Example 1. *Assume that $n = 3$ and look for a factorization $A = LU$:*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a32 & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \tag{3}$$

The multipliers computed in Gaussian Elimination for A are exactly the entries under the main diagonal of L, while the upper triangular matrix U is exactly the triangular matrix at the end of Gaussian Elimination.

2.1.2 Naive Algorithm for Gaussian Elimination

If A is an invertible matrix, solve for x in the linear system $Ax = b$. The inputs are A, an invertible $n \times n$ matrix, and b, an n-dimensional column vector ($b \in \mathbb{R}^n$). The output is $x \in \mathbb{R}^n$, a column vector.

for $j = 1 : n - 1$

for $k = j + 1 : n$

$m = a_{kj}/a_{jj}$

$A(k, j + 1 : n) = A(k, j + 1 : n) - m * A(j, j + 1 : n)$

$b_k = b_k - m * b_j$

end

end

2.1.3 Algorithm for Forward Solve

2.1.4 Algorithm for Back Solve

2.2 $PA = LU$ factorization for Solving $Ax = b$

This factorization involves Gaussian Elimination with partial pivoting. Dividing by small numbers leads to bad pivots and large errors for row operations, leading to numerical instability.

2.2.1 Permutation Matrices

Definition 1. *A $n \times n$ matrix P is a permutation matrix if P is obtained from the identity matrix I_n by swapping rows. Properties of permutation matrices are:*

1. *Each row or column has exactly one nonzero entry, which is 1.*
2. *For each permutation matrix P, $P^{-1} = P^t$.*
3. *If $P = (R_i \longleftrightarrow R_j)(I_n)$, and A is an $n \times n$ matrix, PA swaps R_i and R_j of A.*

Example 1.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = (R_2 \longleftrightarrow R_1) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Example 2.

$$PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{bmatrix} \tag{5}$$

Theorem 1 ($PA = LU$ Factorization Theorem)**.** *If A is an invertible $n \times n$ matrix and P is a permutation matrix corresponding to pivoting in Gaussian Elimination, there are **unique** L-unit lower triangular and U-unit upper triangular matrices, and a permutation matrix P such that $PA = LU$.*

Example 1 (Solving $Ax = b$ using $PA = LU$)**.** *Assume that $PA = LU$ where A is an invertible matrix and P is a permutation matrix. Use LU factorization (with Gaussian elimination and the multiplier operations) to find L and U where $PA = LU$.*

$$\begin{aligned} Ax &= b \\ \iff PAx &= Pb \\ \iff (LU)x &= Pb \\ \iff L(Ux) &= Pb \end{aligned} \tag{6}$$

Denote $Ux = y$. Then,

$$Ax = b \iff L(Ux) = Pb, y = Ux \iff Ly = Pb, Ux = y \tag{7}$$

To solve for $x \in \mathbb{R}^n$, solve $Ly = Pb$ for y using Forward Solve and $Ux = y$ for x using Back Solve.

Remark 1. *If $P = I_n$, then we have the standard $A = LU$ factorization.*

2.3 Symmetric Positive Definite (SPD) Matrices

Definition 1. *For a general square matrix $A = [a_{ij}]_{i,j=1,n}$ (real entries) and $x = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n$:*

$$x^T A x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{8}$$

is a quadratic form. Also, A is called symmetric positive definite if $A^T = A$ and $x^T A x > 0$, for all nonzero $x \in \mathbb{R}^n$.

Theorem 1. *A symmetric matrix A is positive definite (PD) if and only if all the eigenvalues are real and strictly positive.*

Theorem 2 (Sylvester Criterion of Positive Definiteness)**.** *A symmetric matrix*

$$A = \begin{bmatrix} a_{11} & a12 & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \tag{9}$$

is positive definite if and only if

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0 \tag{10}$$

2.4 Cholesky Factorization of SPD Matrices

2.5 Vector and Matrix Norms

Definition 1. *A norm on $V = \mathbb{R}^n$ is a function $|| \cdot ||$ from $\mathbb{R}^n$ with values in $\mathbb{R}$, satisfying*

1. *$\|x\| \geq 0$, for all $x \in V$.*

2. $\|x\| = 0$, *if and only if* $x = 0$.

3. $\|\alpha x\| = |\alpha|\|x\|$, *for all* $x \in V$, *and* $\alpha \in \mathbb{R}$.

4. $\|x + y\| \leq \|x\| + normy$, *for all* $x, y \in V$.

Definition 2. $\|x\|$ *is a distance from* x *to the zero vector* $\boldsymbol{0}$.

Definition 3. *Assume* $x = [x_1, x_2, \dots, x_n]^T \in V$. *The important vector norms are*

$$\|x\|_2 = \sqrt{\sum_{k=1}^{n} {x_k}^2}$$
$$\|x\|_1 = \sum_{k=1}^{n} |x_k|$$
$$\|x\|_\infty = max\{|x_k| : k = 1, 2, \dots, n\}$$

2.6 Unit Vectors and Limits

Definition 1. *In any norm, a vector* $x \in V$ *satisfying* $\|x\| = 1$ *is a* ***unit vector****. If* $v \in V, v \neq 0$, *then* $x = \frac{v}{\|v\|}$ *is the* ***normalization*** *of* v.

Definition 2. *We say that a sequence (in* $V = \mathbb{C}^n$ *or* $\mathbb{R}^n$*),*

$$\{x^k\}_{k\geq 1} = ({x_1}^k, {x_2}^k, \dots, {x_n}^k)^T \tag{11}$$

converges to

$$x = (x_1, \dots, x_n)^T \tag{12}$$

if

$$\lim_{x\to\infty} \|x^k - x\| = 0 \tag{13}$$

Theorem 1. *If a sequence* ${x^k}_{k\geq 1}$ *in* V *converges to* $x \in V$ *in a norm, then it converges to* x *in any other norm.*

Theorem 2. *A sequence* x^k *in* V *converges to* $x \in V$ *if and only if it converges* ***component wise****.*

2.7 Induced Matrix Norms

Definition 1. *Given a vector norm* $\|.\|_a$ *and a* $m \times n$ *matrix* $A = [a_{ij}]$, *the* ***induced norm*** $\|A\|_a$ *is*

$$\|A\|_a = \max_{\|x\|_a = 1} \|Ax\|_a = \max_{x\neq 0} \frac{\|Ax\|_a}{\|x\|_a} \tag{14}$$

Example 1. *Assume* A *has real entries. Then*

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\lambda_{max}(A^T A)}$$
$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{x\neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$$
$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$$

2.7.1 Finding x such that $\|Ax\| = \|A\|\|x\|$

Assume that $A = [a_{ij}]$ is a $m \times n$ matrix with real entries.

Theorem 1 (For $\|A\|_2$)**.**

Theorem 2 (For $\|A\|_1$)**.**

Theorem 3 (For $\|A\|_\infty$)**.** *e*

2.7.2 Key Properties of Matrix Norms

Theorem 4. $\|AB\| \leq \|A\|\|B\|$

Proof.

$$\begin{aligned} \|ABx\| = \|A(Bx)\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\| \\ max\frac{\|ABx\|}{\|x\|} \leq \|A\|\|B\| \\ \|AB\| \leq \|A\|\|B\| \end{aligned}$$

□

Example 1. $\begin{pmatrix} 1 & -2 & 3 \\ -3 & 4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

a) Find $x \in \mathbb{R}^3$ such that $\|Ax\|_1 = \|A\|_1\|x\|_1$
c) Find $\|A\|_\infty$

$$\|A\|_\infty = 12$$