EE636: Matrix Computations
Lecture 4

Debasattam Pal

EE Department,
Indian Institute of Technology Bombay,
Mumbai, India.

IIT Bombay
January 21, 2021

Recap

- How a computer solves

$$Ax = b, \quad \min \|Ax - b\|, \quad Ax = \lambda x, \quad Av = \sigma u$$

  where $\sigma \geqslant 0, \{u_1, u_2, \ldots, u_m\} \subseteq \mathbb{R}^m, \{v_1, v_2, \ldots, v_n\} \subseteq \mathbb{R}^n$ are orthonormal.
- The effect of round off error.
- When a solution can be trusted.
- FLOPS to measure computation time.
- Counted the FLOPS for basic matrix multiplication: $2mn$ for a matrix of size $m \times n$ and a vector of size n.
- How to solve linear systems of equations: triangular systems.
- The (almost) general case: Gaussian elimination.
- Gaussian elimination (without pivotting) is possible if and only if the leading principal minors of A are non-zero.
- Gaussian elimination without pivotting results in the LU decomposition of A: that is $A = LU$, where U is an upper triangular matrix and L is a unity lower triangular matrix.
- The L and U in the LU decomposition is unique.
- LU requires $(O)(n^3)$ FLOPs.
- Inversion of a matrix that admits LU requires $(O)(n^3)$ FLOPs.
- Symmetric positive definite matrices.

Theorem

$A = A^{\mathrm{T}}$ is PD if and only if every eigenvalue of A is positive.

A $\exists$ Q orthogonal matrix ($Q^TQ = QQ^T = I$)

$Q^T A Q = \Lambda \to$ real entries.

$\Rightarrow A = Q \Lambda Q^T$

$\Rightarrow A$ is PD $\Leftrightarrow \lambda_1, \ldots, \lambda_n > 0$.

PSD $\Leftrightarrow \lambda_1, \ldots, \lambda_n \geq 0$.

||ly for ND/NSD.

PSD: All leading principal minors ≥ 0 is necessary for PSD but NOT sufficient.

$$e_1^T \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{A} e_1 = 0, \quad e_2^T A e_2 = 0$$

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}^T A \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 2 > 0$$

$$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}^T A \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = -2 < 0$$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

All leading principal minor > 0

But A is indefinite.

is indefinite

All principal minor $\geq 0 \iff$ A PSD

Theorem

$A = A^T$ is PD if and only if every leading principal submatrix of of A is PD. ✓

THM I $\Longleftrightarrow$ leading principal minors are (+)ve number.

$$(PD \Rightarrow \det > 0) \equiv (\det \not> 0 \Rightarrow \cancel{PD})$$

$\Rightarrow$ ✓

$\Leftarrow$ use induction.

Lemma

If $A = A^{\mathrm{T}}$ is PD then $\det A > 0$.

$$Q^T A Q = \Lambda$$

$$\det Q = \pm 1$$

$$Q^T Q = I$$

$$\det(Q^T Q) = \det I = 1$$

$$\Rightarrow \det Q^T \cdot \det Q = 1$$

$$\Rightarrow (\det Q)^2 = 1$$

$$\det(Q^T A Q) = \det \Lambda = \prod_{i=1}^{n} \lambda_i > 0$$

$$\Rightarrow (\det Q)^2 \det A > 0$$

$$\Rightarrow \det A > 0$$

Suppose a symmetric matrix is given in block form as

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}.$$

Further, suppose A is invertible. Then M is invertible iff the Schur complement of A is invertible.

$$C - B^T A^{-1} B.$$

The Schur complement

$$\underbrace{\begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix}}_{P^T} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \underbrace{\begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}}_{P} = \begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ B^T & C - B^T A^{-1} B \end{bmatrix}$$

$$\begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^T & y^T \end{bmatrix} P^{-T} S P^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} \xi^T & \eta^T \end{bmatrix} S \begin{bmatrix} \xi \\ \eta \end{bmatrix}$$

$$\begin{bmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{bmatrix}$$

SPECIAL CASE:

$$\det \begin{pmatrix} A & b \\ b^T & c \end{pmatrix} = \det \begin{bmatrix} A & 0 \\ 0 & \boxed{c - b^T A^{-1} b} \end{bmatrix} = (\det A) \cdot (c - b^T A^{-1} b)$$

HW: Using Schur complement, and induction on the size of A prove lemma 1.

Theorem

Let $A = A^{\mathrm{T}}$ be PD, then A admits an LU factorization.

$$A = L\tilde{U} = LDU$$

L: unity lower Δ^r

$\tilde{U}$: upper Δ^r

D: diagonal

U: unity upper Δ^r

L: unity lower Δ^r.

Suppose $A = A^T$ (suppose A admits LU)

$$A = LDU \Rightarrow A^T = U^T D L^T$$

$$\parallel$$

$$A = LDU$$

By uniqueness of LDU decomposition.

$$\boxed{A = LDL^T}$$

Assume A is PD $\Rightarrow$ $A = LDL^T \Rightarrow$ all the diagonal entries in D are > 0.

$$x^T A x = x^T L D \underline{\underline{L^T x}} \qquad \xi := L^T x$$

$$= \xi^T D \xi$$

Since L^T is invertible, $x^T A x > 0 \ \forall x \neq 0$

$$\Leftrightarrow \xi^T D \xi > 0 \ \forall \underline{\xi \neq 0}.$$

$$\Leftrightarrow d_{ii} > 0.$$

$$A = LDL^T = L\begin{bmatrix}\sqrt{d_{11}} & & & \\ & \sqrt{d_{22}} & & \\ & & \ddots & \\ & & & \sqrt{d_{nn}}\end{bmatrix}\left[\sqrt{D}\right]L^T$$

$$\sqrt{D}$$

$$F^T := \sqrt{D}\,L^T$$

$$\Rightarrow \boxed{A = FF^T}$$ where F is Lower Δ? $\in \mathbb{R}^{n\times n}$.

We assume that all leading principal minors > 0, TST A is PD.

Use induction on the size of A.

Base case: $n = 1$ $A \in \mathbb{R}$ $A > 0 \Rightarrow A$ is PD.

Inductive step: Suppose the result holds true for matrices of size $(n-1)$ (or less). TST the result must hold for size n.

$$A = \left[\begin{array}{c|c} A_{11} & a_{12} \\ \hline a_{12}^T & a_{22} \end{array}\right] \qquad A_{11} \in \mathbb{R}^{(n-1)\times(n-1)}$$

$$a_{12} \in \mathbb{R}^{(n-1)}, \; a_{22} \in \mathbb{R}.$$

All leading principal minors of $A > 0 \Rightarrow$

" " " " " $A_{11} > 0$

By inductive hypothesis $\Rightarrow$ $\underline{A_{11}}$ is PD.

$$\underline{\begin{bmatrix} A_{11} & a_{12} \\ a_{12}^T & a_{22} \end{bmatrix}} = \begin{bmatrix} I & 0 \\ a_{12}^T A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \underline{a_{22} - a_{12}^T A_{11}^{-1} a_{12}} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} a_{12} \\ 0 & I \end{bmatrix}$$

$$\det A > 0 \Rightarrow a_{22} - a_{12}^T A_{11}^{-1} a_{12} > 0$$

$$\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix}^T A \begin{pmatrix} x \\ y \end{pmatrix} = \xi^T A_{11} \xi + \eta \left(a_{22} - a_{12}^T A_{11}^{-1} a_{12}\right) \eta > 0$$

$$\therefore \quad 0 \quad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^n \quad \forall \xi \in \mathbb{R}^{n-1}, \; \eta \in \mathbb{R}.$$

$\Rightarrow$ A is PD.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & 0 & \cdots & a_{3n}^{(2)} \\ 0 & 0 & 0 & \cdots & a_{4n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(2)} \end{bmatrix}$$

$\underline{\underline{E}}\,\underline{A} = \underline{\tilde{A}}$

Block upper

$\underline{A}^r$

non invertible

- ~~Orthogonal.~~

If A is invertible and $a^{(k-1)}_{kk}$ is zero

$\Rightarrow \exists\, l = k+1, \ldots, n$ s.t.

$$a^{(k-1)}_{l,k} \neq 0.$$

- ~~Permutaion involving only one swap is idempotent.~~

This idea of pivoting mirrors the LU strategy.

How the permutation matrix pervades through a chain of elementary row operation matrices

$$\pi_{35}\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{32} & 1 & 0 & 0 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{52} & 0 & 0 & 1 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{52} & 0 & 0 & 1 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{32} & 1 & 0 & 0 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix}$$

How the permutation matrix pervades through a chain of elementary row operation matrices

$$\pi_{35}\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{32} & 1 & 0 & 0 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{52} & 0 & 0 & 1 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & {\color{red}-m_{52}} & 0 & 0 & {\color{red}1} & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & {\color{red}-m_{32}} & {\color{red}1} & 0 & 0 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & {\color{red}-m_{52}} & 0 & 0 & {\color{red}1} & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & {\color{red}-m_{32}} & {\color{red}1} & 0 & 0 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & {\color{red}-m_{52}} & {\color{red}1} & 0 & 0 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & {\color{red}-m_{32}} & 0 & 0 & {\color{red}1} & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix}\pi_{35}$$

How the permutation matrix pervades through a chain of elementary row operation matrices

$$\pi_{35}\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{32} & 1 & 0 & 0 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{52} & 0 & 0 & 1 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{52} & 0 & 0 & 1 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{32} & 1 & 0 & 0 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{52} & 0 & 0 & 1 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{32} & 1 & 0 & 0 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & -m_{52} & 1 & 0 & 0 & 0\\ 0 & -m_{42} & 0 & 1 & 0 & 0\\ 0 & -m_{32} & 0 & 0 & 1 & 0\\ 0 & -m_{62} & 0 & 0 & 0 & 1\end{bmatrix}\pi_{35}$$

$$\pi_{35}E_2E_1A = \pi_{35}A^{(2)}$$

$$\widehat{E}_2\pi_{35}E_1A = \widehat{A}^{(2)}$$

$$\widehat{E}_2\widehat{E}_1\pi_{35}A = \widehat{A}^{(2)}$$

Thank you