AMS526: Numerical Analysis I  
(Numerical Linear Algebra for Computational and Data Sciences)  
Lecture 9: Positive-Definite Systems; Cholesky Factorization  

Xiangmin Jiao  

Stony Brook University
Outline

1. Positive-Definite Systems (MC§4.2)

2. Cholesky Factorization (NLA§23)
Symmetric Positive-Definite Matrices

- Symmetric matrix $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite (SPD) if $x^T Ax > 0$ for $x \in \mathbb{R}^n \setminus \{0\}$

- Hermitian matrix $A \in \mathbb{C}^{n \times n}$ is Hermitian positive definite (HPD) if $x^* Ax > 0$ for $x \in \mathbb{C}^n \setminus \{0\}$

- SPD matrices have positive real eigenvalues and orthogonal eigenvectors

- Note: A positive-definite matrix does not need to be symmetric or Hermitian! A real matrix $A$ is positive definite iff $A + A^T$ is SPD

- If $x^T Ax \geq 0$ for $x \in \mathbb{R}^n \setminus \{0\}$, then $A$ is said to be positive semidefinite
Properties of Symmetric Positive-Definite Matrices

- SPD matrix often arises as Hessian matrix of some convex functional
  - E.g., least squares problems; partial differential equations
- If $A$ is SPD, then $A$ is nonsingular
- Let $X$ be any $n \times m$ matrix with full rank and $n \geq m$. Then
  - $X^TX$ is symmetric positive definite, and
  - $XX^T$ is symmetric positive semidefinite
- If $A$ is $n \times n$ SPD and $X \in \mathbb{R}^{n \times m}$ has full rank and $n \geq m$, then $X^TAX$ is SPD
- Any principal submatrix (picking some rows and corresponding columns) of $A$ is SPD and $a_{ii} > 0$
Outline

1. Positive-Definite Systems (MC§4.2)

2. Cholesky Factorization (NLA§23)
Cholesky Factorization

- If $A$ is symmetric positive definite, then there is factorization of $A$

$$A = R^T R$$

where $R$ is upper triangular, and all its diagonal entries are positive.

- Key idea: take advantage and preserve symmetry and positive-definiteness during factorization.

- Eliminate below diagonal and to the right of diagonal.

$$A = \begin{bmatrix} a_{11} & b^T \\ b & K \end{bmatrix} = \begin{bmatrix} r_{11} & 0 \\ b/r_{11} & I \end{bmatrix} \begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & K - bb^T/a_{11} \end{bmatrix}$$

$$= \begin{bmatrix} r_{11} & 0 \\ b/r_{11} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & K - bb^T/a_{11} \end{bmatrix} \begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & I \end{bmatrix} = R_1^T A_1 R_1$$

where $r_{11} = \sqrt{a_{11}}$, where $a_{11} > 0$.

- $K - bb^T/a_{11}$ is principal submatrix of SPD $A_1 = R_1^{-T} A R_1^{-1}$ and therefore is SPD, with positive diagonal entries.
Cholesky Factorization

- Apply recursively to obtain

\[ A = \begin{pmatrix} R_1^T & R_2^T & \cdots & R_n^T \end{pmatrix} (R_n \cdots R_2 R_1) = R^T R, \quad r_{jj} > 0 \]

which is known as *Cholesky factorization*

- How to obtain \( R \) from \( R_n, \ldots, R_2, R_1 \)? Recursively:

\[
A = \begin{bmatrix} r_{11} & 0 \\ s & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} r_{11} & s^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} r_{11} & 0 \\ s & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{R}^T \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{R} \end{bmatrix} \begin{bmatrix} r_{11} & s^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} r_{11} & 0 \\ s & \tilde{R}^T \end{bmatrix} \begin{bmatrix} r_{11} & s^T \\ 0 & \tilde{R} \end{bmatrix} = R^T R
\]

- \( R \) is “union” of \( k \)th rows of \( R_k \) (\( R^T \) is “union” of columns of \( R_k^T \))
- Matrix \( A_1 \) is called the *Schur complement* of \( a_{11} \) in \( A \)
Existence and Uniqueness

- Every SPD matrix has a unique Cholesky factorization
  - It exists because algorithm for Cholesky factorization always works for SPD matrices
  - Unique because once $\alpha = \sqrt{a_{11}}$ is determined at each step, entire column $w/\alpha$ is determined

- Question: How to check whether a symmetric matrix is positive definite?

- Answer: Run Cholesky factorization and it succeeds iff the matrix is positive definite.
Algorithm of Cholesky Factorization

- Factorize SPD matrix \( A \in \mathbb{R}^{n \times n} \) into \( A = R^T R \)

```
Algorithm: Cholesky factorization

\[ R = A \]

```
```
\textbf{for} \( k = 1 : n \)
```
```
\textbf{for} \( j = k + 1 : n \)
```
```
\[ r_{j,j:n} \leftarrow r_{j,j:n} - \left( \frac{r_{kj}}{r_{kk}} \right) r_{k,j:n} \]
```
```
\[ r_{k,k:n} \leftarrow r_{k,k:n}/\sqrt{r_{kk}} \]
```

- Note: \( r_{j,j:n} \) denotes subvector of \( j \)th row with columns \( j, j + 1, \ldots, n \)

- Operation count

\[
\sum_{k=1}^{n} \sum_{j=k+1}^{n} 2(n - j) \approx 2 \sum_{k=1}^{n} \sum_{j=1}^{k} j \approx \sum_{k=1}^{n} k^2 \approx \frac{n^3}{3}
\]

- In practice, \( R \) overwrites \( A \), and only upper-triangular part is stored.
Notes on Cholesky Factorization

- Stability of Cholesky factorization
  - Cholesky factorization is backward stable
  - This is because $\| R \|_2^2 = \| A \|_2$, so entries in $R$ are well bounded

- Cholesky factorization $A = R^* R$ exists for HPD matrices, where $R$ is upper-triangular and its diagonal entries are positive real values

- Implementations
  - Different versions of Cholesky factorization can all use block-matrix operators to achieve better performance, and actual performance depends on sizes of blocks
  - Different versions may have different amount of parallelism
$LDL^T$ Factorization

- What happens if $A$ is symmetric but not positive definite?
- Cholesky factorization is sometimes given by $A = LDL^T$ where $D$ is diagonal matrix and $L$ is unit lower triangular matrix.
- This avoids computing square roots.
- Symmetric indefinite systems can be factorized with $PAP^T = LDL^T$, where
  - $P$ is a permutation matrix
  - $D$ is diagonal (if $A$ is complex, $D$ is block diagonal with $1 \times 1$ and $2 \times 2$ blocks)
  - its cost is similar to Cholesky factorization
A matrix $A$ is \textit{banded} if there is a narrow band around the main diagonal such that all of the entries of $A$ outside of the band are zero.

If $A$ is $n \times n$, and there is an $s \ll n$ such that $a_{ij} = 0$ whenever $|i - j| > s$, then we say $A$ is banded with bandwidth $2s + 1$.

For symmetric matrices, only half of band is stored. We say that $A$ has semi-bandwidth $s$.

\textbf{Theorem}

\textit{Let $A$ be a banded, symmetric positive definite matrix with semi-bandwidth $s$. Then its Cholesky factor $R$ also has semi-bandwidth $s$.}

- This is easy to prove using bordered form of Cholesky factorization.
- Total flop count of Cholesky factorization is only $\sim ns^2$.
- However, $A^{-1}$ of a banded matrix may be dense, so it is not economical to compute $A^{-1}$.