Special Matrices

MATH 375 *Numerical Analysis*

J Robert Buchanan

Department of Mathematics

Spring 2022

Objective

In this lesson we will learn to identify special types of matrices for which Gaussian elimination (perhaps with row interchanges) can be used to solve linear systems.

Diagonally Dominant Matrices

Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is **strictly diagonally dominant** when

$$|a_{ii}| > \sum_{j=1, j \neq i}^{n} |a_{ij}|$$

for $i = 1, 2, \ldots, n$. In other words a matrix is strictly diagonally dominant if the magnitude of each element on the diagonal exceeds the sum of the magnitudes of all other elements on that row.

Diagonally Dominant Matrices

Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is **strictly diagonally dominant** when

$$|a_{ii}| > \sum_{j=1, j \neq i}^{n} |a_{ij}|$$

for $i = 1, 2, \ldots, n$. In other words a matrix is strictly diagonally dominant if the magnitude of each element on the diagonal exceeds the sum of the magnitudes of all other elements on that row.

Example

$$A = \begin{bmatrix} -7 & 2 & 0 \\ 3 & -5 & 1 \\ 0 & 5 & 6 \end{bmatrix}$$

Diagonal Dominance and Gaussian Elimination

Theorem

A strictly diagonally dominant matrix A is non-singular. Gaussian elimination can be performed on the linear system $A\mathbf{x} = \mathbf{b}$ without row interchanges.

Proof of Non-singularity (1 of 2)

- For the purposes of contradiction, assume A is singular, *i.e.* there exists $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$.

Proof of Non-singularity (1 of 2)

- For the purposes of contradiction, assume A is singular, *i.e.* there exists $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$.
- Let $k \in \{1, 2, \ldots, n\}$ be such that

$$|x_k| = \max_{1 \leq j \leq n} \{|x_j|\} > 0.$$

Proof of Non-singularity (1 of 2)

- For the purposes of contradiction, assume A is singular, *i.e.* there exists $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$.
- Let $k \in \{1, 2, \ldots, n\}$ be such that

$$|x_k| = \max_{1 \leq j \leq n} \{|x_j|\} > 0.$$

- Since $A\mathbf{x} = \mathbf{0}$

$$\begin{aligned} \sum_{j=1}^{n} a_{kj} x_j &= 0 \\ a_{kk} x_k &= -\sum_{j=1, j \neq k}^{n} a_{kj} x_j. \end{aligned}$$

Proof of Non-singularity (2 of 2)

- Taking absolute values of both sides of the equation yields:

$$
\begin{aligned}
|a_{kk}x_k| &= \left| \sum_{j=1, j\neq k}^{n} a_{kj}x_j \right| \\
|a_{kk}|\,|x_k| &\leq \sum_{j=1, j\neq k}^{n} |a_{kj}|\,|x_j| \\
|a_{kk}| &\leq \sum_{j=1, j\neq k}^{n} |a_{kj}| \underbrace{\frac{|x_j|}{|x_k|}}_{\leq 1} \\
&\leq \sum_{j=1, j\neq k}^{n} |a_{kj}|
\end{aligned}
$$

Proof of Non-singularity (2 of 2)

- Taking absolute values of both sides of the equation yields:

$$
\begin{aligned}
|a_{kk} x_k| &= \left| \sum_{j=1, j \neq k}^{n} a_{kj} x_j \right| \\
|a_{kk}| \, |x_k| &\leq \sum_{j=1, j \neq k}^{n} |a_{kj}| \, |x_j| \\
|a_{kk}| &\leq \sum_{j=1, j \neq k}^{n} |a_{kj}| \underbrace{\frac{|x_j|}{|x_k|}}_{\leq 1} \\
&\leq \sum_{j=1, j \neq k}^{n} |a_{kj}|
\end{aligned}
$$

- This contradicts the strictly diagonally dominant assumption.

Proof No Row Interchanges Necessary (1 of 4)

- Since A is strictly diagonally dominant, $a_{11} \neq 0$ and we can define $M^{(1)}$. Let $A^{(2)} = M^{(1)} A$.

Proof No Row Interchanges Necessary (1 of 4)

- Since A is strictly diagonally dominant, $a_{11} \neq 0$ and we can define $M^{(1)}$. Let $A^{(2)} = M^{(1)} A$.
- For $i = 2, 3, \ldots, n$

$$a_{ij}^{(2)} = a_{ij} - \frac{a_{1j} a_{i1}}{a_{11}} \quad \text{for } 2 \leq j \leq n.$$

and $a_{i1}^{(2)} = 0$.

Proof No Row Interchanges Necessary (2 of 4)

Take absolute values and sum over j.

$$\begin{aligned} \left| a_{ij}^{(2)} \right| &= \left| a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}} \right| \\ \sum_{j=2, j\neq i}^{n} \left| a_{ij}^{(2)} \right| &= \sum_{j=2, j\neq i}^{n} \left| a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}} \right| \end{aligned}$$

Proof No Row Interchanges Necessary (2 of 4)

Take absolute values and sum over j.

$$
\begin{aligned}
\left|a_{ij}^{(2)}\right| &= \left|a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}}\right| \\
\sum_{j=2,j\neq i}^{n} \left|a_{ij}^{(2)}\right| &= \sum_{j=2,j\neq i}^{n} \left|a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}}\right| \\
&\leq \sum_{j=2,j\neq i}^{n} |a_{ij}| + \sum_{j=2,j\neq i}^{n} \left|\frac{a_{1j}a_{i1}}{a_{11}}\right|
\end{aligned}
$$

Proof No Row Interchanges Necessary (2 of 4)

Take absolute values and sum over j.

$$
\begin{aligned}
\left|a_{ij}^{(2)}\right| &= \left|a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}}\right| \\
\sum_{j=2, j\neq i}^{n} \left|a_{ij}^{(2)}\right| &= \sum_{j=2, j\neq i}^{n} \left|a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}}\right| \\
&\leq \sum_{j=2, j\neq i}^{n} |a_{ij}| + \sum_{j=2, j\neq i}^{n} \left|\frac{a_{1j}a_{i1}}{a_{11}}\right| \\
&< |a_{ii}| - |a_{i1}| + \frac{|a_{i1}|}{|a_{11}|} \sum_{j=2, j\neq i}^{n} |a_{1j}|
\end{aligned}
$$

(by the assumption of strict diagonal dominance)

Proof No Row Interchanges Necessary (2 of 4)

Take absolute values and sum over j.

$$\begin{aligned}
\left|a_{ij}^{(2)}\right| &= \left|a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}}\right| \\
\sum_{j=2, j\neq i}^{n} \left|a_{ij}^{(2)}\right| &= \sum_{j=2, j\neq i}^{n} \left|a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}}\right| \\
&\leq \sum_{j=2, j\neq i}^{n} |a_{ij}| + \sum_{j=2, j\neq i}^{n} \left|\frac{a_{1j}a_{i1}}{a_{11}}\right| \\
&< |a_{ii}| - |a_{i1}| + \frac{|a_{i1}|}{|a_{11}|} \sum_{j=2, j\neq i}^{n} |a_{1j}| \\
&< |a_{ii}| - |a_{i1}| + \frac{|a_{i1}|}{|a_{11}|} (|a_{11}| - |a_{1i}|)
\end{aligned}$$

(by the assumption of strict diagonal dominance)

Proof No Row Interchanges Necessary (3 of 4)

- Thus we have

$$\begin{aligned}\sum_{j=2, j\neq i}^{n} \left|a_{ij}^{(2)}\right| &< |a_{ii}| - |a_{i1}| + |a_{i1}| - \frac{|a_{i1}|\,|a_{1i}|}{|a_{11}|} \\ &= |a_{ii}| - \frac{|a_{i1}|\,|a_{1i}|}{|a_{11}|} \\ &\leq \left|a_{ii} - \frac{a_{i1}\,a_{1i}}{a_{11}}\right| \\ &= \left|a_{ii}^{(2)}\right|\end{aligned}$$

for $i = 2, 3, \ldots, n$.

Proof No Row Interchanges Necessary (3 of 4)

- Thus we have

$$\begin{aligned}\sum_{j=2, j\neq i}^{n} \left|a_{ij}^{(2)}\right| &< |a_{ii}| - |a_{i1}| + |a_{i1}| - \frac{|a_{i1}|\,|a_{1i}|}{|a_{11}|} \\ &= |a_{ii}| - \frac{|a_{i1}|\,|a_{1i}|}{|a_{11}|} \\ &\leq \left| a_{ii} - \frac{a_{i1}\,a_{1i}}{a_{11}} \right| \\ &= \left|a_{ii}^{(2)}\right|\end{aligned}$$

  for $i = 2, 3, \ldots, n$.
- This implies row i of $A^{(2)}$ is strictly diagonally dominant for $i = 2, 3, \ldots, n$.

Proof No Row Interchanges Necessary (4 of 4)

- Since row 1 of $A^{(2)}$ is the same as row 1 of A, then $A^{(2)}$ is strictly diagonally dominant.

Proof No Row Interchanges Necessary (4 of 4)

- Since row 1 of $A^{(2)}$ is the same as row 1 of A, then $A^{(2)}$ is strictly diagonally dominant.
- By induction we can show $A^{(n)}$ (the upper triangular form of A) is strictly diagonally dominant, *i.e.* A can be row reduced without row interchanges.

Positive Definite Matrices

Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is **positive definite** if A is symmetric and if $\mathbf{x}^t A \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.

Example

Let $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$, and show A is positive definite.

Example

Let $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$, and show A is positive definite.

If $\mathbf{x} = \langle x_1, x_2, x_3 \rangle \neq \mathbf{0}$ then

$$\mathbf{x}^t A \mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Example

Let $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$, and show A is positive definite.

If $\mathbf{x} = \langle x_1, x_2, x_3 \rangle \neq \mathbf{0}$ then

$$\begin{aligned} \mathbf{x}^t A \mathbf{x} &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4x_1 + 2x_2 + 2x_3 \\ 2x_1 + 4x_2 + 2x_3 \\ 2x_1 + 2x_2 + 4x_3 \end{bmatrix} \end{aligned}$$

Example

Let $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$, and show A is positive definite.

If $\mathbf{x} = \langle x_1, x_2, x_3 \rangle \neq \mathbf{0}$ then

$$
\begin{aligned}
\mathbf{x}^t A \mathbf{x} &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4x_1 + 2x_2 + 2x_3 \\ 2x_1 + 4x_2 + 2x_3 \\ 2x_1 + 2x_2 + 4x_3 \end{bmatrix} \\
&= 4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_1 x_2 + 4x_1 x_3 + 4x_2 x_3 \\
&= 2(x_1 + x_2)^2 + 2(x_1 + x_3)^2 + 2(x_2 + x_3)^2 \\
&> 0.
\end{aligned}
$$

Properties of Positive Definite Matrices

Theorem

If $A \in \mathbb{R}^{n \times n}$ is a positive definite matrix, then

1. *A is non-singular.*
2. *$a_{ii} > 0$ for $i = 1, 2, \ldots, n$.*
3. $\max\limits_{1 \leq k,j \leq n} |a_{kj}| \leq \max\limits_{1 \leq i \leq n} |a_{ii}|$
4. *$a_{ij}^2 < a_{ii} a_{jj}$ for $i \neq j$.*

Proof: A is non-singular

If $\mathbf{x} \neq \mathbf{0}$ and $A\mathbf{x} = \mathbf{0}$ then

$$\mathbf{x}^t A \mathbf{x} = \mathbf{x}^t \mathbf{0} = 0$$

which contradicts the assumption that A is positive definite.

Proof: $a_{ii} > 0$ for $i = 1, 2, \ldots, n$

For $i = 1, 2, \ldots, n$ let $\mathbf{e}_i$ be the ith standard basis vector.

$$0 < \mathbf{e}_i^t A \mathbf{e}_i = a_{ii}$$

Proof: $\max_{1 \le k,j \le n} |a_{kj}| \le \max_{1 \le i \le n} |a_{ii}|$

- Suppose $k \neq j$ and define $\mathbf{x}$ as

$$x_i = \begin{cases} 0 & \text{if } i \neq j \text{ and } i \neq k \\ 1 & \text{if } i = j \\ -1 & \text{if } i = k \end{cases}$$

Proof: $\max_{1 \le k,j \le n} |a_{kj}| \le \max_{1 \le i \le n} |a_{ii}|$

- Suppose $k \neq j$ and define $\mathbf{x}$ as

$$x_i = \begin{cases} 0 & \text{if } i \neq j \text{ and } i \neq k \\ 1 & \text{if } i = j \\ -1 & \text{if } i = k \end{cases}$$

- Since $\mathbf{x} \neq \mathbf{0}$ then

$$\begin{aligned} 0 &< \mathbf{x}^t A \mathbf{x} = a_{jj} + a_{kk} - a_{jk} - a_{kj} \\ 2a_{kj} &< a_{jj} + a_{kk}. \text{ (since } A = A^t\text{)} \end{aligned}$$

Proof: $\max_{1\leq k,j\leq n} |a_{kj}| \leq \max_{1\leq i\leq n} |a_{ii}|$

- Define $\mathbf{z}$ by

$$z_i = \begin{cases} 0 & \text{if } i \neq j \text{ and } i \neq k \\ 1 & \text{if } i = j \text{ or } i = k \end{cases}$$

Proof: $\max_{1 \le k,j \le n} |a_{kj}| \le \max_{1 \le i \le n} |a_{ii}|$

- Define $\mathbf{z}$ by
$$z_i = \begin{cases} 0 & \text{if } i \neq j \text{ and } i \neq k \\ 1 & \text{if } i = j \text{ or } i = k \end{cases}$$
- Since $\mathbf{z} \neq \mathbf{0}$ then
$$\begin{aligned} 0 &< \mathbf{z}^t A \mathbf{z} = a_{jj} + a_{kk} + a_{jk} + a_{kj} \\ -2a_{kj} &< a_{jj} + a_{kk}. \text{ (since } A = A^t) \end{aligned}$$

Proof: $\max_{1 \le k,j \le n} |a_{kj}| \le \max_{1 \le i \le n} |a_{ii}|$

- Define $\mathbf{z}$ by

$$z_i = \begin{cases} 0 & \text{if } i \neq j \text{ and } i \neq k \\ 1 & \text{if } i = j \text{ or } i = k \end{cases}$$

- Since $\mathbf{z} \neq \mathbf{0}$ then

$$\begin{aligned} 0 &< \mathbf{z}^t A \mathbf{z} = a_{jj} + a_{kk} + a_{jk} + a_{kj} \\ -2a_{kj} &< a_{jj} + a_{kk}. \text{ (since } A = A^t\text{)} \end{aligned}$$

- Thus if $k \neq j$

$$\begin{aligned} |a_{kj}| &< \frac{a_{jj} + a_{kk}}{2} \le \max_{1 \le i \le n} |a_{ii}| \\ \max_{1 \le k,j \le n} |a_{kj}| &\le \max_{1 \le i \le n} |a_{ii}| \end{aligned}$$

Proof: $a_{ij}^2 < a_{ii}a_{jj}$ for $i \neq j$

- Suppose $i \neq j$ and define $\mathbf{x}$ as

$$x_k = \begin{cases} 0 & \text{if } k \neq j \text{ and } k \neq i \\ \alpha & \text{if } k = i \\ 1 & \text{if } k = j \end{cases}$$

where $\alpha \in \mathbb{R}$.

Proof: $a_{ij}^2 < a_{ii}a_{jj}$ for $i \neq j$

- Suppose $i \neq j$ and define $\mathbf{x}$ as

$$x_k = \begin{cases} 0 & \text{if } k \neq j \text{ and } k \neq i \\ \alpha & \text{if } k = i \\ 1 & \text{if } k = j \end{cases}$$

  where $\alpha \in \mathbb{R}$.
- Since $\mathbf{x} \neq \mathbf{0}$ then

$$0 < \mathbf{x}^t A \mathbf{x} = \alpha^2 a_{ii} + 2\alpha a_{ij} + a_{jj}$$

Proof: $a_{ij}^2 < a_{ii}a_{jj}$ for $i \neq j$

- Suppose $i \neq j$ and define $\mathbf{x}$ as

$$x_k = \begin{cases} 0 & \text{if } k \neq j \text{ and } k \neq i \\ \alpha & \text{if } k = i \\ 1 & \text{if } k = j \end{cases}$$

  where $\alpha \in \mathbb{R}$.
- Since $\mathbf{x} \neq \mathbf{0}$ then

$$0 < \mathbf{x}^t A \mathbf{x} = \alpha^2 a_{ii} + 2\alpha a_{ij} + a_{jj}$$

- Note the quadratic polynomial in α above has no real roots. Therefore

$$\begin{aligned} 4a_{ij}^2 - 4a_{ii}a_{jj} &< 0 \\ a_{ij}^2 &< a_{ii}a_{jj} \end{aligned}$$

Principal Submatrices

Definition

A **leading principal submatrix** of matrix $A \in \mathbb{R}^{n \times n}$ is a matrix of the form

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

for some $1 \leq k \leq n$.

Principal Submatrices

Definition

A **leading principal submatrix** of matrix $A \in \mathbb{R}^{n \times n}$ is a matrix of the form

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

for some $1 \leq k \leq n$.

Theorem

A symmetric matrix A is positive definite if and only if each of its leading principal submatrices has a positive determinant.

Example

Show that matrix $A = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{bmatrix}$ is positive definite.

Example

Show that matrix $A = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{bmatrix}$ is positive definite.

$$\begin{aligned} |2| &= 2 > 0 \\ \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} &= 3 > 0 \\ \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} &= 2\begin{vmatrix} 2 & 4 \\ 4 & 9 \end{vmatrix} + \begin{vmatrix} -1 & 4 \\ -3 & 9 \end{vmatrix} - 3\begin{vmatrix} -1 & 2 \\ -3 & 4 \end{vmatrix} = 1 > 0 \end{aligned}$$

Significant Results

Theorem

The symmetric matrix A is positive definite if and only if Gaussian elimination without row interchanges can be performed on the linear system $A\mathbf{x} = \mathbf{b}$ with all pivot elements positive. Furthermore, the computations are stable with respect to the growth of round-off errors.

Significant Results

Theorem

The symmetric matrix A is positive definite if and only if Gaussian elimination without row interchanges can be performed on the linear system $A\mathbf{x} = \mathbf{b}$ with all pivot elements positive. Furthermore, the computations are stable with respect to the growth of round-off errors.

Corollary

Matrix A is positive definite if and only if A can be factored as $L\,D\,L^t$ where L is a lower triangular matrix with 1's on its diagonal and D is a diagonal matrix with positive diagonal entries.

Proof (1 of 5)

- We can factor $A = L\,U$ where L is lower triangular with 1's on the diagonal and U is upper triangular.
- Define D to be the diagonal matrix whose entries are the diagonal entries of U.

Proof (1 of 5)

- We can factor $A = L\,U$ where L is lower triangular with 1's on the diagonal and U is upper triangular.
- Define D to be the diagonal matrix whose entries are the diagonal entries of U.
- Since A is non-singular, D is non-singular.

Proof (1 of 5)

- We can factor $A = L\,U$ where L is lower triangular with 1's on the diagonal and U is upper triangular.
- Define D to be the diagonal matrix whose entries are the diagonal entries of U.
- Since A is non-singular, D is non-singular.
- Define $M = U^t \left[D^{-1}\right]^t$. Since D is diagonal then $D^t = D$ and

$$M = U^t \left[D^t\right]^{-1} = U^t D^{-1} \quad \Longrightarrow \quad M^t = D^{-1}\,U.$$

Proof (2 of 5)

$$
\begin{aligned}
A &= L\,U \\
A &= L\left(D\,D^{-1}\right)U \\
A &= L\,D\left(D^{-1}\,U\right) \\
A &= L\,D\,M^t \\
A\left[M^t\right]^{-1} &= L\,D \\
M^{-1}A\left[M^t\right]^{-1} &= M^{-1}L\,D
\end{aligned}
$$

Proof (2 of 5)

$$
\begin{aligned}
A &= L\,U \\
A &= L\left(D\,D^{-1}\right)U \\
A &= L\,D\left(D^{-1}\,U\right) \\
A &= L\,D\,M^t \\
A\left[M^t\right]^{-1} &= L\,D \\
M^{-1}A\left[M^t\right]^{-1} &= M^{-1}L\,D
\end{aligned}
$$

Note:

$$\left(M^{-1}A\left[M^t\right]^{-1}\right)^t = \left(\left[M^t\right]^{-1}\right)^t A^t\left[M^{-1}\right]^t = M^{-1}A\left[M^t\right]^{-1}$$

which implies $M^{-1}A\left[M^t\right]^{-1}$ is a symmetric matrix.

Proof (3 of 5)

- $L\,D$ is lower triangular $\implies A\left[M^t\right]^{-1}$ is lower triangular.

Proof (3 of 5)

- $L\,D$ is lower triangular $\implies A\left[M^t\right]^{-1}$ is lower triangular.
- M is lower triangular $\implies M^{-1}$ is lower triangular.

Proof (3 of 5)

- $L\,D$ is lower triangular $\implies A\left[M^t\right]^{-1}$ is lower triangular.
- M is lower triangular $\implies M^{-1}$ is lower triangular.
- Thus $M^{-1}A\left[M^t\right]^{-1}$ is lower triangular and symmetric, *i.e.*, diagonal $\implies M^{-1}L\,D$ is diagonal $\implies M^{-1}L$ is diagonal.

Proof (3 of 5)

- $L\,D$ is lower triangular $\implies A\left[M^t\right]^{-1}$ is lower triangular.
- M is lower triangular $\implies M^{-1}$ is lower triangular.
- Thus $M^{-1}A\left[M^t\right]^{-1}$ is lower triangular and symmetric, *i.e.*, diagonal $\implies M^{-1}L\,D$ is diagonal $\implies M^{-1}L$ is diagonal.
- What is $M^{-1}L$?

Proof (4 of 5)

$$
\begin{aligned}
M^{-1}L &= D\left[U^t\right]^{-1}L \\
&= \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix} \begin{bmatrix} d_{11}^{-1} & 0 & \cdots & 0 \\ * & d_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & d_{nn}^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix}
\end{aligned}
$$

Proof (4 of 5)

$$
\begin{aligned}
M^{-1}L &= D\left[U^t\right]^{-1}L \\
&= \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix} \begin{bmatrix} d_{11}^{-1} & 0 & \cdots & 0 \\ * & d_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & d_{nn}^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix} \\
M^{-1}L &= I_n \\
M &= L
\end{aligned}
$$

Proof (5 of 5)

Since $M = L = U^t \left[D^{-1}\right]^t$ then transposing both sides yields

$$\begin{aligned} L^t &= D^{-1} U \\ D\,L^t &= U \\ A &= L\,U \\ A &= L\,D\,L^t \end{aligned}$$

Square Root of a Matrix

Corollary

Matrix A is positive definite if and only if A can be factored as $L\,L^t$ where L is a lower triangular matrix with non-zero diagonal entries.

Remark: this is known as the **Cholesky factorization** of a positive definite matrix.

Algorithm: Cholesky

INPUT positive definite matrix $A \in \mathbb{R}^{n \times n}$

STEP 1 Set $l_{11} = \sqrt{a_{11}}$.

STEP 2 For $j = 2, 3, \ldots, n$ set $l_{j1} = \dfrac{a_{j1}}{l_{11}}$.

STEP 3 For $i = 2, 3, \ldots, n-1$ do STEPS 4–5.

STEP 4 Set $l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}$.

STEP 5 For $j = i+1, i+2, \ldots, n$ set

$$l_{ji} = \frac{a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik}}{l_{ii}}.$$

STEP 6 Set $l_{nn} = \sqrt{a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2}$.

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

$$l_{11} = \sqrt{4} = 2$$

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

$$
\begin{aligned}
l_{11} &= \sqrt{4} = 2 \\
l_{21} &= -\frac{1}{2} \\
l_{31} &= \frac{1}{2}
\end{aligned}
$$

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

$$\begin{aligned}
l_{11} &= \sqrt{4} = 2 \\
l_{21} &= -\frac{1}{2} \\
l_{31} &= \frac{1}{2} \\
l_{22} &= \sqrt{3 - \sum_{k=1}^{1} l_{2k}^2} = \sqrt{3 - \frac{1}{4}} = \frac{\sqrt{11}}{2} \\
l_{32} &= \frac{0 - \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)}{\sqrt{11}/2} = \frac{1}{2\sqrt{11}}
\end{aligned}$$

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

$$
\begin{aligned}
l_{11} &= \sqrt{4} = 2 \\
l_{21} &= -\frac{1}{2} \\
l_{31} &= \frac{1}{2} \\
l_{22} &= \sqrt{3 - \sum_{k=1}^{1} l_{2k}^2} = \sqrt{3 - \frac{1}{4}} = \frac{\sqrt{11}}{2} \\
l_{32} &= \frac{0 - \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)}{\sqrt{11}/2} = \frac{1}{2\sqrt{11}} \\
l_{33} &= \sqrt{2 - \sum_{k=1}^{2} l_{3k}^2} = \sqrt{2 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2\sqrt{11}}\right)^2} = \sqrt{\frac{19}{11}}
\end{aligned}
$$

Result

$$L = \begin{bmatrix} 2 & 0 & 0 \\ -\frac{1}{2} & \frac{\sqrt{11}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{11}} & \sqrt{\frac{19}{11}} \end{bmatrix}$$

$$L\,L^t = \begin{bmatrix} 2 & 0 & 0 \\ -\frac{1}{2} & \frac{\sqrt{11}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{11}} & \sqrt{\frac{19}{11}} \end{bmatrix} \begin{bmatrix} 2 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{11}}{2} & \frac{1}{2\sqrt{11}} \\ 0 & 0 & \sqrt{\frac{19}{11}} \end{bmatrix}$$

$$A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

Banded Matrices

Definition

An $n \times n$ matrix is called a **band matrix** if there exist integers p and q with $1 \leq p, q \leq n$ with the property that $a_{ij} = 0$ whenever $p \leq j - i$ or $q \leq i - j$. The **band width** of a band matrix is defined as $w = p + q - 1$.

Remarks:

- Band matrices have their non-zero entries concentrated near the diagonal.
- The frequently occurring case of $p = q = 2$ is sometimes called a **tridiagonal** matrix.

Example

Determine the band width of the following matrix.

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

Homework

- Read Section 6.6.
- Exercises: 1, 5, 9a, 17, 19, 21