Learning to Measure Before Iterating: Norms and Error Analysis

Nate DeMaagd, Kurt O'Hearn

MTH 499-02

March 25, 2013

Outline

- Norms: properties, types, examples

Outline

- Norms: properties, types, examples
- Theorems on norms

Outline

- Norms: properties, types, examples
- Theorems on norms
- Relative/absolute error in perturbations, conditioning number

Norm Definition

- Norm is a function $|| \cdot ||$ that maps from vector space V to $\mathbb{Z}^+$. It obeys:

Norm Definition

- Norm is a function $||\cdot||$ that maps from vector space V to $\mathbb{Z}^{+}$. It obeys:
    - $||x|| > 0$ if $x \neq 0,\ x \in V$

Norm Definition

- Norm is a function $||\cdot||$ that maps from vector space V to $\mathbb{Z}^+$. It obeys:
  - $||x|| > 0$ if $x \neq 0,\ x \in V$
  - $||\lambda x|| = |\lambda|\ ||x||$ if $\lambda \in \mathbb{R},\ x \in V$

Norm Definition

- Norm is a function $||\cdot||$ that maps from vector space V to $\mathbb{Z}^+$. It obeys:
  - $||x|| > 0$ if $x \neq 0, \ x \in V$
  - $||\lambda x|| = |\lambda| \ ||x||$ if $\lambda \in \mathbb{R}, \ x \in V$
  - the Triangle Inequality, $||x+y|| \leq ||x|| + ||y||$ if $x, y \in V$.

Norm Definition

- Norm is a function $||\cdot||$ that maps from vector space V to $\mathbb{Z}^+$. It obeys:
  - $||x|| > 0$ if $x \neq 0,\ x \in V$
  - $||\lambda x|| = |\lambda|\ ||x||$ if $\lambda \in \mathbb{R},\ x \in V$
  - the Triangle Inequality, $||x + y|| \leq ||x|| + ||y||$ if $x, y \in V$.
- Generalizes idea of absolute value.

Norm Definition

- Norm is a function $||\cdot||$ that maps from vector space V to $\mathbb{Z}^+$. It obeys:
  - $||x|| > 0$ if $x \neq 0,\ x \in V$
  - $||\lambda x|| = |\lambda|\ ||x||$ if $\lambda \in \mathbb{R},\ x \in V$
  - the Triangle Inequality, $||x + y|| \leq ||x|| + ||y||$ if $x, y \in V$.
- Generalizes idea of absolute value.
- Think of norm as length of vector.

- Defined generally,

$$l_k = ||x||_k = \left(\sum_{i=1}^{n} |x_i|^k \right)^{1/k} \text{ where } k \in \{1, 2, \dots\} \text{ and } x = (x_1, \dots, x_n)^T$$

- Defined generally,

$$l_k = ||x||_k = \left(\sum_{i=1}^{n} |x_i|^k \right)^{1/k} \quad \text{where } k \in \{1, 2, \dots\} \text{ and } x = (x_1, \dots, x_n)^T$$

- Some common norms are

k	l_k
1	$\|\|x\|\|_1 = \sum_{i=1}^{n} \|x_i\|$
2	$\|\|x\|\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$
$\vdots$	$\vdots$
∞	$\|\|x\|\|_\infty = \max_{1 \le i \le n} \|x_i\|$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
x			
y			
z			

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\|\cdot\|\|_1$	$\|\|\cdot\|\|_2$	$\|\|\cdot\|\|_\infty$
x	6		
y			
z			

$$||x||_1 = 2 + 1 + 3 = 6$$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\|\cdot\|\|_1$	$\|\|\cdot\|\|_2$	$\|\|\cdot\|\|_\infty$
x	6	3.74	
y			
z			

$$||x||_2 = \sqrt{2^2 + 1^2 + 3^2} = 3.74$$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\|\cdot\|\|_1$	$\|\|\cdot\|\|_2$	$\|\|\cdot\|\|_\infty$
x	6	3.74	3
y			
z			

$$||x||_\infty = 3$$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\|\cdot\|\|_1$	$\|\|\cdot\|\|_2$	$\|\|\cdot\|\|_\infty$
x	6	3.74	3
y	14	8.83	7
z	13	9.54	9

- Matrix norm associated with vector norm $||\cdot||_k$ is

$$||A||_k = \sup\left\{||Au||_k : u \in \mathbb{R}^n, ||u||_k = 1\right\},$$

where A is an $n \times n$ matrix.

- Matrix norm associated with vector norm $||\cdot||_k$ is

$$||A||_k = \sup\{||Au||_k : u \in \mathbb{R}^n, ||u||_k = 1\},$$

  where A is an $n \times n$ matrix.

- Note matrix norm is *subordinate* to a chosen vector norm.

- Matrix norm associated with vector norm $||\cdot||_k$ is

$$||A||_k = \sup\{||Au||_k : u \in \mathbb{R}^n, ||u||_k = 1\},$$

  where A is an $n \times n$ matrix.
- Note matrix norm is *subordinate* to a chosen vector norm.
- Two common subordinate matrix norms

- Matrix norm associated with vector norm $||\cdot||_k$ is

$$||A||_k = \sup\left\{||Au||_k : u \in \mathbb{R}^n, ||u||_k = 1\right\},$$

  where A is an $n \times n$ matrix.
- Note matrix norm is *subordinate* to a chosen vector norm.
- Two common subordinate matrix norms
  - Subordinate to vector norm $||\cdot||_\infty$ is

$$||A||_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|,$$

    which is the maximum absolute row sum of A.

- Matrix norm associated with vector norm $||\cdot||_k$ is

$$||A||_k = \sup\left\{||Au||_k : u \in \mathbb{R}^n, ||u||_k = 1\right\},$$

  where A is an $n \times n$ matrix.
- Note matrix norm is *subordinate* to a chosen vector norm.
- Two common subordinate matrix norms
  - Subordinate to vector norm $||\cdot||_\infty$ is

$$||A||_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|,$$

    which is the maximum absolute row sum of A.
  - Subordinate to vector norm $||\cdot||_2$ is

$$\begin{aligned} ||A||_2 &= \sup_{||x||_2=1} ||Ax||_2 \\ &= \sqrt{\rho(A^T A)}, \end{aligned}$$

    where $\rho(A^T A)$ is spectral radius of $A^T A$, largest eigenvalue of $A^T A$.

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$
- $(A^T A) = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$
- $(A^T A) = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$
- For $A^T A$, eigenvalues are $\sigma_1^2 = 6.8541$ and $\sigma_2^2 = 0.1459$.

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$
- $(A^T A) = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$
- For $A^T A$, eigenvalues are $\sigma_1^2 = 6.8541$ and $\sigma_2^2 = 0.1459$.
- So, $||A||_2 = \sqrt{6.8541} \approx 2.618$.

- Consider matrix $B = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

- Consider matrix $B = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$
- $||B||_2$ is simply the maximum absolute row sum, which is 9.

Theorem (Subordinate Matrix Norm)

If $\|\cdot\|$ is any norm in $\mathbb{R}^n$, then the equation

$$\|A\| = \sup_{\|u\|=1} \{\|Au\| : u \in \mathbb{R}^n\}$$

defines a norm on the linear space of all n x n matrices.

Theorem (Subordinate Matrix Norm)

If $\|\cdot\|$ is any norm in $\mathbb{R}^n$, then the equation

$$\|A\| = \sup_{\|u\|=1} \{\|Au\| : u \in \mathbb{R}^n\}$$

defines a norm on the linear space of all n x n matrices.

Proof.

TODO $\square$

Theorem (Infinity Matrix Norm)

If the vectorm norm $\|\cdot\|_\infty$ is defined by

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|$$

then its subordinate matrix norm is given by

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|.$$

Theorem (Infinity Matrix Norm)

If the vectorm norm $\|\cdot\|_\infty$ is defined by

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$$

then its subordinate matrix norm is given by

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|.$$

Proof.

TODO $\square$

TODO

The Cholesky Factorization

The Cholesky Factorization

- Given A, we can write A as $A = LL^T$ with L lower triangular

The Cholesky Factorization

- Given A, we can write A as $A = LL^T$ with L lower triangular
- Necessary conditions on A:

The Cholesky Factorization

- Given A, we can write A as $A = LL^T$ with L lower triangular
- Necessary conditions on A:
  - Real-valued entries ($A \in \mathbb{R}^{n \mathrm{x} n}$)
  - Symmetric ($A = A^T$)
  - Positive definite ($x^T A x > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0_n$)

The Cholesky Factorization

- Given A, we can write A as $A = LL^T$ with L lower triangular
- Necessary conditions on A:
  - Real-valued entries ($A \in \mathbb{R}^{n \mathrm{x} n}$)
  - Symmetric ($A = A^T$)
  - Positive definite ($x^T Ax > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0_n$)
- Cholesky: special case of LU-factorization

The Cholesky Factorization

- Given A, we can write A as $A = LL^T$ with L lower triangular
- Necessary conditions on A:
  - Real-valued entries ($A \in \mathbb{R}^{n\text{x}n}$)
  - Symmetric ($A = A^T$)
  - Positive definite ($x^T Ax > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0_n$)
- Cholesky: special case of LU-factorization
- Why are factorizations important? Connections later in pivoting

The Cholesky Factorization

- Given A, we can write A as $A = LL^T$ with L lower triangular
- Necessary conditions on A:
  - Real-valued entries ($A \in \mathbb{R}^{n\text{x}n}$)
  - Symmetric ($A = A^T$)
  - Positive definite ($x^T Ax > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0_n$)
- Cholesky: special case of LU-factorization
- Why are factorizations important? Connections later in pivoting

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite
  $\Rightarrow$ principle minors nonsingular $\Rightarrow A = LU$ (LU factorization exists)

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite
  $\Rightarrow$ principle minors nonsingular $\Rightarrow A = LU$ (LU factorization exists)
- $LU = A = A^T = U^T L^T \Leftrightarrow U(L^T)^{-1} = L^{-1}U^T$

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite
  $\Rightarrow$ principle minors nonsingular $\Rightarrow$ $A = LU$ (LU factorization exists)
- $LU = A = A^T = U^T L^T \Leftrightarrow U(L^T)^{-1} = L^{-1}U^T$
- $U(L^T)^{-1}$ upper triangular, $L^{-1}U^T$ lower triangular
  $\Rightarrow$ both sides diagonal (D)

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite
  $\Rightarrow$ principle minors nonsingular $\Rightarrow$ $A = LU$ (LU factorization exists)
- $LU = A = A^T = U^T L^T \Leftrightarrow U(L^T)^{-1} = L^{-1}U^T$
- $U(L^T)^{-1}$ upper triangular, $L^{-1}U^T$ lower triangular
  $\Rightarrow$ both sides diagonal (D)
- $U = DL^T$

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite $\Rightarrow$ principle minors nonsingular $\Rightarrow A = LU$ (LU factorization exists)
- $LU = A = A^T = U^T L^T \Leftrightarrow U(L^T)^{-1} = L^{-1}U^T$
- $U(L^T)^{-1}$ upper triangular, $L^{-1}U^T$ lower triangular $\Rightarrow$ both sides diagonal (D)
- $U = DL^T$
- D has positive diagonal entries $\Rightarrow D = D^{1/2}D^{1/2}$

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite
  $\Rightarrow$ principle minors nonsingular $\Rightarrow$ $A = LU$ (LU factorization exists)
- $LU = A = A^T = U^T L^T \Leftrightarrow U(L^T)^{-1} = L^{-1}U^T$
- $U(L^T)^{-1}$ upper triangular, $L^{-1}U^T$ lower triangular
  $\Rightarrow$ both sides diagonal (D)
- $U = DL^T$
- D has positive diagonal entries $\Rightarrow$ $D = D^{1/2}D^{1/2}$
- $A = L^*(L^*)^T$, where $L^* = LD^{1/2}$

Theorem

If A is a real, symmetric, and positive definite matrix, then $A = LL^T$ for a unique upper triangular L.

Proof.

Existence:

- A positive definite $\Rightarrow$ A nonsingular
- Leading principle minors of A positive definite
  $\Rightarrow$ principle minors nonsingular $\Rightarrow$ $A = LU$ (LU factorization exists)
- $LU = A = A^T = U^T L^T \Leftrightarrow U(L^T)^{-1} = L^{-1}U^T$
- $U(L^T)^{-1}$ upper triangular, $L^{-1}U^T$ lower triangular
  $\Rightarrow$ both sides diagonal (D)
- $U = DL^T$
- D has positive diagonal entries $\Rightarrow$ $D = D^{1/2}D^{1/2}$
- $A = L^*(L^*)^T$, where $L^* = LD^{1/2}$

Unicity: straightforward (assume two, show same) □

Cholesky Algorithm

function $\textsc{Cholesky}(n, (a_{ij}))$

  for $k = 1$ **to** n **do**

    $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$

    for $i = k+1$ **to** n **do**

      $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$

  return (l_{ij})

where

- $A = (a_{ij})$: factored matrix
- n: order of A
- $L = (l_{ij})$: resulting lower triangular matrix

Example of Cholesky Algorithm

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} & 0 & 0 \\ & & 0 \\ & & \end{bmatrix}$$

function $\textsc{Cholesky}(n, (a_{ij}))$

    for $k = 1$ **to** n **do**

        $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$

        for $i = k+1$ **to** n **do**

            $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$

    return (l_{ij})

Example of Cholesky Algorithm

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} 5 & 0 & 0 \\ & & 0 \\ & & \end{bmatrix}$$

function $\textsc{Cholesky}(n, (a_{ij}))$
  for $k = 1$ **to** n **do**
    $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$
    for $i = k+1$ **to** n **do**
      $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$
  return (l_{ij})

$$k = 1:$$

$$l_{11} = (25 - 0)^{1/2} = 5$$

Example of Cholesky Algorithm

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & & 0 \\ & & \end{bmatrix}$$

function CHOLESKY$(n, (a_{ij}))$
    for $k = 1$ **to** n **do**
        $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$
        for $i = k+1$ **to** n **do**
            $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$
    return (l_{ij})

$$k = 1, \quad i = 2:$$

$$l_{21} = (15 - 0)/5 = 3$$

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & & 0 \\ -1 & & \end{bmatrix}$$

1: **function** CHOLESKY$(n, (a_{ij}))$
2: **for** $k = 1$ **to** n **do**
3: $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$
4: **for** $i = k + 1$ **to** n **do**
5: $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$
6: **return** (l_{ij})

$$k = 1, \quad i = 3:$$

$$l_{31} = (-5 - 0)/5 = -1$$

Example of Cholesky Algorithm

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & & \end{bmatrix}$$

function $\text{CHOLESKY}(n, (a_{ij}))$
  for $k = 1$ **to** n **do**
    $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2\right)^{1/2}$
    for $i = k+1$ **to** n **do**
      $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks}\right) / l_{kk}$
  return (l_{ij})

$$k = 2:$$

$$l_{22} = (18 - 3^2)^{1/2} = 3$$

Example of Cholesky Algorithm

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & \end{bmatrix}$$

function $\textsc{Cholesky}(n, (a_{ij}))$
  for $k = 1$ **to** n **do**
    $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$
    for $i = k+1$ **to** n **do**
      $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$
  return (l_{ij})

$$k = 2, \quad i = 3:$$

$$l_{32} = \left(0 - (-1)(3)\right)/3 = 1$$

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix}$$

function CHOLESKY$(n, (a_{ij}))$
    for $k = 1$ **to** n **do**
        $l_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$
        for $i = k + 1$ **to** n **do**
            $l_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} l_{ks} \right) / l_{kk}$
    return (l_{ij})

$$k = 3 :$$

$$l_{33} = \left(11 - (-1)^2 - (-1)^2\right) = 3$$

Example of Cholesky Algorithm Cont.

- Verify correctness:

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix}$$

- Verify correctness:

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix} \Longrightarrow L^T = \begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

- Verify correctness:

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix} \Longrightarrow L^T = \begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

$$LL^T = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} = A$$

Gaussian Elimination

Gaussian Elimination

- Process which attempts to solve a linear system by reducing the coefficient matrix to triangular form using elementary row operations (the result can then by easily solved using forward/backward substitution)

Gaussian Elimination

- Process which attempts to solve a linear system by reducing the coefficient matrix to triangular form using elementary row operations (the result can then by easily solved using forward/backward substitution)
- Possible row operations:
  - Interchange
  - Scaling
  - Addition

Gaussian Elimination

- Process which attempts to solve a linear system by reducing the coefficient matrix to triangular form using elementary row operations (the result can then by easily solved using forward/backward substitution)
- Possible row operations:
  - Interchange
  - Scaling
  - Addition
- Gaussian elimination only utilizes scaling and addition!

Gaussian Elimination Algorithm

function $\mathrm{GE}(n, (a_{ij}))$

  for $k = 1$ **to** $n - 1$ **do**

    for $i = k + 1$ **to** n **do**

      $z \leftarrow a_{ik} / a_{kk}$

      $a_{ik} \leftarrow 0$

      for $j = k + 1$ **to** n **do**

        $a_{ij} \leftarrow a_{ij} - z a_{kj}$

  return (a_{ij})

- Recall: row operations are equivalent to left multiplication of coefficient matrix by elementary matrices

- Recall: row operations are equivalent to left multiplication of coefficient matrix by elementary matrices
  - Example: scale row 2 of $A \in \mathbb{R}^{3x3}$ by 4

$$EA, \quad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Recall: row operations are equivalent to left multiplication of coefficient matrix by elementary matrices
  - Example: scale row 2 of $A \in \mathbb{R}^{3x3}$ by 4

$$EA, \quad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Define G.E. recursively in terms of multiplication by these matrices:

$$A_{k+1} = E_k A_k$$

- Recall: row operations are equivalent to left multiplication of coefficient matrix by elementary matrices
  - Example: scale row 2 of $A \in \mathbb{R}^{3\text{x}3}$ by 4

$$EA, \quad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Define G.E. recursively in terms of multiplication by these matrices:

$$A_{k+1} = E_k A_k$$

- Note that algorithm terminates after $s = T_n$ operations with $A_{n\text{x}n}$

- Recall: row operations are equivalent to left multiplication of coefficient matrix by elementary matrices
  - Example: scale row 2 of $A \in \mathbb{R}^{3x3}$ by 4

$$EA, \quad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Define G.E. recursively in terms of multiplication by these matrices:

$$A_{k+1} = E_k A_k$$

- Note that algorithm terminates after $s = T_n$ operations with A_{nxn}
- Define $U = A_{s+1}$ and observe that

$$U = E_s \cdots E_1 A \quad \text{or} \quad A = E_1^{-1} \cdots E_s^{-1} U = LU$$

- Recall: row operations are equivalent to left multiplication of coefficient matrix by elementary matrices
  - Example: scale row 2 of $A \in \mathbb{R}^{3\text{x}3}$ by 4

$$EA, \quad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Define G.E. recursively in terms of multiplication by these matrices:

$$A_{k+1} = E_k A_k$$

- Note that algorithm terminates after $s = T_n$ operations with $A_{n\text{x}n}$
- Define $U = A_{s+1}$ and observe that

$$U = E_s \cdots E_1 A \quad \text{or} \quad A = E_1^{-1} \cdots E_s^{-1} U = LU$$

- So, Gaussian elimination constructs an LU factorization (and hence is functionally equivalent to the LU factorization algorithm!)

Pivoting Methods

Pivoting Methods

- Problem: Gaussian elimination breaks down (spectacular examples pg. 167–169)

Pivoting Methods

- Problem: Gaussian elimination breaks down (spectacular examples pg. 167–169)
- Solution: incorporate row interchange operations via permutation matrix (we call this pivoting)

Pivoting Methods

- Problem: Gaussian elimination breaks down (spectacular examples pg. 167–169)
- Solution: incorporate row interchange operations via permutation matrix (we call this pivoting)
- Types of pivoting:
  - Partial/complete: pertains to the amount of the matrix examined to find where to pivot

Pivoting Methods

- Problem: Gaussian elimination breaks down (spectacular examples pg. 167–169)
- Solution: incorporate row interchange operations via permutation matrix (we call this pivoting)
- Types of pivoting:
  - Partial/complete: pertains to the amount of the matrix examined to find where to pivot
    - Partial: analyze one row
    - Complete: analyze entire submatrix

Pivoting Methods

- Problem: Gaussian elimination breaks down (spectacular examples pg. 167–169)
- Solution: incorporate row interchange operations via permutation matrix (we call this pivoting)
- Types of pivoting:
  - Partial/complete: pertains to the amount of the matrix examined to find where to pivot
    - Partial: analyze one row
    - Complete: analyze entire submatrix
  - Unscaled/scaled: pertains to how the potential pivot locations are compared

Pivoting Methods

- Problem: Gaussian elimination breaks down (spectacular examples pg. 167–169)
- Solution: incorporate row interchange operations via permutation matrix (we call this pivoting)
- Types of pivoting:
  - Partial/complete: pertains to the amount of the matrix examined to find where to pivot
    - Partial: analyze one row
    - Complete: analyze entire submatrix
  - Unscaled/scaled: pertains to how the potential pivot locations are compared
    - Unscaled: compare absolute value of entries directly
    - Scaled: compare ratios of pivot entries to maximum entries in magnitude in each row

Pivoting Algorithms

function GE_PP$(n, (a_{ij}), (p_i))$

  for $k = 1$ **to** $n - 1$ **do**

    for $i = k + 1$ **to** n **do**

      $z \leftarrow a_{p_i k} / a_{p_k k}$

      $a_{p_i k} \leftarrow 0$

      for $j = k + 1$ **to** n **do**

        $a_{p_i j} \leftarrow a_{p_i j} - z a_{p_k j}$

  return (a_{ij})

$$A_1 = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 1 & 1 \\ 1 & 3 & 3 \end{bmatrix}$$

$$p = \begin{bmatrix} 1, & 2, & 3 \end{bmatrix}$$

function GE_PP$(n, (a_{ij}), (p_i))$
    for $k = 1$ **to** $n - 1$ **do**
        for $i = k + 1$ **to** n **do**
            $z \leftarrow a_{p_i k} / a_{p_k k}$
            $a_{p_i k} \leftarrow 0$
            for $j = k + 1$ **to** n **do**
                $a_{p_i j} \leftarrow a_{p_i j} - z a_{p_k j}$
    return (a_{ij})

$$\max(|1|, |2|, |1|) = 2$$

Pivoting Examples

$$A_1 = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 1 & 1 \\ 1 & 3 & 3 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

function GE_PP$(n, (a_{ij}), (p_i))$
    for $k = 1$ **to** $n-1$ **do**
        for $i = k+1$ **to** n **do**
            $z \leftarrow a_{p_i k}/a_{p_k k}$
            $a_{p_i k} \leftarrow 0$
            for $j = k+1$ **to** n **do**
                $a_{p_i j} \leftarrow a_{p_i j} - z a_{p_k j}$
    return (a_{ij})

$$\max(|1|, |2|, |1|) = 2$$

$$R_1 \leftarrow R_1 - \frac{1}{2} R_2$$

$$R_3 \leftarrow R_3 - \frac{1}{2} R_2$$

$$R_2 \leftarrow \frac{1}{2} R_2$$

$$A_2 = \begin{bmatrix} 0 & 2.5 & 5.5 \\ 1 & 0.5 & 0.5 \\ 0 & 2.5 & 2.5 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

```
function GE_PP(n, (a_ij), (p_i))
    for k = 1 to n − 1 do
        for i = k + 1 to n do
            z ← a_{p_i k} / a_{p_k k}
            a_{p_i k} ← 0
            for j = k + 1 to n do
                a_{p_i j} ← a_{p_i j} − z a_{p_k j}
    return (a_ij)
```

$$\max(|1|, |2|, |1|) = 2$$

$$R_1 \leftarrow R_1 - \frac{1}{2} R_2$$

$$R_3 \leftarrow R_3 - \frac{1}{2} R_2$$

$$R_2 \leftarrow \frac{1}{2} R_2$$

Pivoting Examples

$$A_2 = \begin{bmatrix} 0 & 2.5 & 5.5 \\ 1 & 0.5 & 0.5 \\ 0 & 2.5 & 2.5 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

```
function GE_PP(n, (a_ij), (p_i))
    for k = 1 to n − 1 do
        for i = k + 1 to n do
            z ← a_{p_i k} / a_{p_k k}
            a_{p_i k} ← 0
            for j = k + 1 to n do
                a_{p_i j} ← a_{p_i j} − z a_{p_k j}
    return (a_ij)
```

$$\max(|2.5|, |2.5|) = 2.5$$

Pivoting Examples

$$A_2 = \begin{bmatrix} 0 & 2.5 & 5.5 \\ 1 & 0.5 & 0.5 \\ 0 & 2.5 & 2.5 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

```
function GE_PP(n, (a_ij), (p_i))
    for k = 1 to n − 1 do
        for i = k + 1 to n do
            z ← a_{p_i k} / a_{p_k k}
            a_{p_i k} ← 0
            for j = k + 1 to n do
                a_{p_i j} ← a_{p_i j} − z a_{p_k j}
    return (a_ij)
```

$$\max(|2.5|, |2.5|) = 2.5$$

$$R_3 \leftarrow R_3 - R_2$$

$$R_1 \leftarrow \frac{2}{5} R_1$$

Pivoting Examples

$$A_3 = \begin{bmatrix} 0 & 1 & 2.2 \\ 1 & 0.5 & -0.6 \\ 0 & 0 & -3 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

function GE_PP$(n, (a_{ij}), (p_i))$
  for $k = 1$ **to** $n - 1$ **do**
    for $i = k + 1$ **to** n **do**
      $z \leftarrow a_{p_i k} / a_{p_k k}$
      $a_{p_i k} \leftarrow 0$
      for $j = k + 1$ **to** n **do**
        $a_{p_i j} \leftarrow a_{p_i j} - z a_{p_k j}$
  return (a_{ij})

$$\max(|2.5|, |2.5|) = 2.5$$

$$R_3 \leftarrow R_3 - R_2$$

$$R_1 \leftarrow \frac{2}{5} R_1$$

Pivoting Examples

$$A_3 = \begin{bmatrix} 0 & 1 & 2.2 \\ 1 & 0.5 & \text{-}0.6 \\ 0 & 0 & \text{-}3 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

function $\mathrm{GE\_PP}(n, (a_{ij}), (p_i))$
    for $k = 1$ **to** $n - 1$ **do**
        for $i = k + 1$ **to** n **do**
            $z \leftarrow a_{p_i k} / a_{p_k k}$
            $a_{p_i k} \leftarrow 0$
            for $j = k + 1$ **to** n **do**
                $a_{p_i j} \leftarrow a_{p_i j} - z a_{p_k j}$
    return (a_{ij})

$$R_3 \leftarrow -\frac{1}{3} R_3$$

Pivoting Examples

$$A_4 = \begin{bmatrix} 0 & 1 & 2.2 \\ 1 & 0.5 & \text{-}0.6 \\ 0 & 0 & 1 \end{bmatrix}$$

$$p = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}$$

function GE_PP($n, (a_{ij}), (p_i)$)
  for $k = 1$ **to** $n - 1$ **do**
    for $i = k + 1$ **to** n **do**
      $z \leftarrow a_{p_i k} / a_{p_k k}$
      $a_{p_i k} \leftarrow 0$
      for $j = k + 1$ **to** n **do**
        $a_{p_i j} \leftarrow a_{p_i j} - z a_{p_k j}$
  return (a_{ij})

$$R_3 \leftarrow -\frac{1}{3} R_3$$