Chapter 4 - LU Factorization - Part 2

Maggie Myers
Robert A. van de Geijn

The University of Texas at Austin

Practical Linear Algebra – Fall 2009

Back Substitution = Solving an Upper Triangular System

Backward Substitution = Solving an Upper Triangular System

- Let upper triangular matrix $U \in \mathbb{R}^{n \times n}$ and $x, b \in \mathbb{R}^n$.
- Consider the equation $Ux = b$ where U and b are known and x is to be computed.
- Partition

$$U \rightarrow \left(\begin{array}{c|c} \upsilon_{11} & u_{12}^T \\ \hline 0 & U_{22} \end{array}\right), \quad x \rightarrow \left(\begin{array}{c} \chi_1 \\ \hline x_2 \end{array}\right) \quad \text{and} \quad b \rightarrow \left(\begin{array}{c} \beta_1 \\ \hline b_2 \end{array}\right).$$

- $Ux = b$ implies

$$\overbrace{\left(\begin{array}{c} \beta_1 \\ \hline b_2 \end{array}\right)}^{b} = \overbrace{\left(\begin{array}{c|c} \upsilon_{11} & u_{12}^T \\ \hline 0 & U_{22} \end{array}\right)}^{U} \overbrace{\left(\begin{array}{c} \chi_1 \\ \hline x_2 \end{array}\right)}^{x} = \overbrace{\left(\begin{array}{c} \upsilon_{11}\chi_1 + u_{12}^T x_2 \\ \hline U_{22} x_2 \end{array}\right)}^{Ux}$$

-
$$\overbrace{\left(\begin{array}{c}\beta_1\\ \hline b_2\end{array}\right)}^{b} = \overbrace{\left(\begin{array}{c|c}\upsilon_{11} & u_{12}^T\\ \hline 0 & U_{22}\end{array}\right)}^{U}\overbrace{\left(\begin{array}{c}\chi_1\\ \hline x_2\end{array}\right)}^{x} = \overbrace{\left(\begin{array}{c}\upsilon_{11}\chi_1 + u_{12}^T x_2\\ \hline U_{22}x_2\end{array}\right)}^{Ux}$$

- Thus

$$\left(\begin{array}{c}\beta_1 = \upsilon_{11}\chi_1 + u_{12}^T x_2\\ \hline b_2 = U_{22}x_2\end{array}\right) \quad \text{or} \quad \left(\begin{array}{c}\chi_1 = (\beta_1 - u_{12}^T x_2)/\upsilon_{11}\\ \hline U_{22}x_2 = b_2\end{array}\right)$$

- This suggests the following steps for overwriting the vector b with the solution vector x:
  - Partition

$$U \rightarrow \left(\begin{array}{c|c}\upsilon_{11} & u_{12}^T\\ \hline 0 & U_{22}\end{array}\right), \quad \text{and} \quad b \rightarrow \left(\begin{array}{c}\beta_1\\ \hline b_2\end{array}\right)$$

  - Solve $U_{22}x_2 = b_2$ for x_2, overwriting b_2 with the result.
  - Update $\beta_1 = (\beta_1 - u_{12}^T b_2)/\upsilon_{11} (= (\beta_1 - u_{12}^T x_2)/\upsilon_{11})$.

Algorithm: $[b] := \text{UTRSV\_UNB}(U, b)$

Partition $U \rightarrow \left(\begin{array}{c|c} U_{TL} & U_{TR} \\ \hline U_{BL} & U_{BR} \end{array}\right)$, $b \rightarrow \left(\begin{array}{c} b_T \\ \hline b_B \end{array}\right)$

where U_{BR} is 0×0, b_B has 0 rows

while $m(U_{BR}) < m(U)$ **do**

Repartition

$$\left(\begin{array}{c|c} U_{TL} & U_{TR} \\ \hline 0 & U_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} U_{00} & u_{01} & U_{02} \\ \hline 0 & \upsilon_{11} & u_{12}^T \\ \hline 0 & 0 & U_{22} \end{array}\right), \left(\begin{array}{c} b_T \\ \hline b_B \end{array}\right) \rightarrow \left(\begin{array}{c} b_0 \\ \hline \beta_1 \\ \hline b_2 \end{array}\right)$$

where υ_{11} is 1×1 , β_1 has 1 row

$\beta_1 := (\beta_1 - u_{12}^T b_2)/\upsilon_{11}$

Continue with

$$\left(\begin{array}{c|c} U_{TL} & U_{TR} \\ \hline 0 & U_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} U_{00} & u_{01} & U_{02} \\ \hline 0 & \upsilon_{11} & u_{12}^T \\ \hline 0 & 0 & U_{22} \end{array}\right), \left(\begin{array}{c} b_T \\ \hline b_B \end{array}\right) \leftarrow \left(\begin{array}{c} b_0 \\ \hline \beta_1 \\ \hline b_2 \end{array}\right)$$

endwhile

Exercise

Solve the upper triangular linear system

$$\begin{array}{rcrcrcr} -2\chi_0 & - & \chi_1 & + & \chi_2 & = & 6 \\ & & -3\chi_1 & - & 2\chi_2 & = & 9 \\ & & & & \chi_2 & = & 3 \end{array}$$

Exercise

Use `http://www.cs.utexas.edu/users/flame/Spark/` to write a FLAME@lab code for computing the solution of $Ux = b$, overwriting b with the solution and assuming that U is upper triangular.

Solving the Linear System, Again

Solving the Linear System

Let $A = LU$ and assume that $Ax = b$, where A and b are given. Then $(LU)x = b$ or $L(Ux) = b$. Let us introduce a dummy vector $z = Ux$. Then $Lz = b$ and z can be computed as described in the previous section. Once z has been computed, x can be computed by solving $Ux = z$ where now U and z are known.

When LU Factorization Breaks Down

"Does Gaussian elimination always solve a linear system?" Or, equivalently, can an LU factorization always be computed?

- *If* an LU factorization can be computed
- *and* the upper triangular factor U has no zeroes on the diagonal,
- *then* $Ax = b$ can be solved for all right-hand side vectors b.

Are there examples where LU (Gaussian elimination as we have presented it so far) can break down? The answer is *yes*.

Example

- Consider the matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
- In the first step, the algorithm for LU factorization will try to compute the multiplier $1/0$, which will cause an error.
- Now, $Ax = b$ is given by the set of linear equations

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix} = \begin{pmatrix} \chi_1 \\ \chi_0 \end{pmatrix}$$

  so that $Ax = b$ is equivalent to $\begin{pmatrix} \chi_1 \\ \chi_0 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$.
- The solution to $Ax = b$ is given by the vector $x = \begin{pmatrix} \beta_1 \\ \beta_0 \end{pmatrix}$.
- Although Gaussian Elimination (LU factorization) breaks down, the linear system always has a solution.

Example

$$\begin{array}{rcrcrcr} 2\chi_0 & + & 4\chi_1 & + & (-2)\chi_2 & = & -10 \\ 4\chi_0 & + & 8\chi_1 & + & 6\chi_2 & = & 20 \\ 6\chi_0 & + & (-4)\chi_1 & + & 2\chi_2 & = & 18 \end{array}$$

Now,

- Subtract $(4/2) = 2$ times the first row from the second row:

$$\begin{array}{rcrcrcr} 2\chi_0 & + & 4\chi_1 & + & (-2)\chi_2 & = & -10 \\ 0\chi_0 & + & 0\chi_1 & + & 10\chi_2 & = & 40 \\ 6\chi_0 & + & (-4)\chi_1 & + & 2\chi_2 & = & 18 \end{array}$$

- Subtract $(6/2) = 3$ times the first row from the third row:

$$\begin{array}{rcrcrcr} 2\chi_0 & + & 4\chi_1 & + & (-2)\chi_2 & = & -10 \\ 0\chi_0 & + & 0\chi_1 & + & 10\chi_2 & = & 40 \\ 0\chi_0 & + & (-16)\chi_1 & + & 8\chi_2 & = & 48 \end{array}$$

- Now, we've got a problem.

Example (continued)

$$\begin{array}{rrrl} 2\chi_0 + & 4\chi_1 + & (-2)\chi_2 = & -10 \\ 0\chi_0 + & 0\chi_1 + & 10\chi_2 = & 40 \\ 0\chi_0 + & (-16)\chi_1 + & 8\chi_2 = & 48 \end{array}$$

- Swap the second and third row:

$$\begin{array}{rrrl} 2\chi_0 + & 4\chi_1 + & (-2)\chi_2 = & -10 \\ 0\chi_0 + & (-16)\chi_1 + & 8\chi_2 = & 48 \\ 0\chi_0 + & 0\chi_1 + & 10\chi_2 = & 40 \end{array}$$

- at which point we are done transforming our system into an upper triangular system, and the backward substition can commence to solve the problem.

Example

$$\begin{array}{rrrr} 0\chi_0 + & 4\chi_1 + & (-2)\chi_2 = & -10 \\ 4\chi_0 + & 8\chi_1 + & 6\chi_2 = & 20 \\ 6\chi_0 + & (-4)\chi_1 + & 2\chi_2 = & 18 \end{array}$$

Now,

- Subtract $(4/0)$ times the first row from the second row! Yikes!
- We swap the first row with any of the other two rows:

$$\begin{array}{rrrr} 4\chi_0 + & 8\chi_1 + & 6\chi_2 = & 20 \\ 0\chi_0 + & 4\chi_1 + & (-2)\chi_2 = & -10 \\ 6\chi_0 + & (-4)\chi_1 + & 2\chi_2 = & 18 \end{array}$$

- By subtracting $(6/4) = 3/2$ times the first row from the third row, we get

$$\begin{array}{rrrr} 4\chi_0 + & 8\chi_1 + & 6\chi_2 = & 20 \\ 0\chi_0 + & 4\chi_1 + & (-2)\chi_2 = & -10 \\ 0\chi_0 + & (-16)\chi_1 + & (-7)\chi_2 = & -22 \end{array}$$

- Etc.

LU factorization needs to be modified to allow for row exchanges if a zero pivot is encountered.

Permutations

Exercise

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} -2 & 1 & 2 \\ 3 & 2 & 1 \\ -1 & 0 & -3 \end{pmatrix}}_{A} =$$

Answer

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} -2 & 1 & 2 \\ 3 & 2 & 1 \\ -1 & 0 & -3 \end{pmatrix}}_{A} = \begin{pmatrix} 3 & 2 & 1 \\ -1 & 0 & -3 \\ -2 & 1 & 2 \end{pmatrix}.$$

Definition

A vector $p = \left(\begin{array}{c|c|c|c} k_0 & k_1 & \cdots & k_{n-1} \end{array} \right)^T$ is said to be a permutation (vector) if $k_j \in \{0, \ldots, n-1\}$, $0 \leq j < n$, and $k_i = k_j$ implies $i = j$.

We will below write $\left(\begin{array}{c|c|c|c} k_0 & k_1 & \cdots & k_{n-1} \end{array} \right)^T$ to indicate a column vector, for space considerations. This permutation is just a rearrangement of the vector $(0, 1, \ldots, n-1)^T$.

Definition

Let $p = (k_0, \ldots, k_{n-1})^T$ be a permutation. Then

$$P = \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix}$$

is said to be a *permutation matrix*.

Notation

- P is the identity matrix with its rows rearranged as indicated by the n-tuple $(k_0, k_1, \ldots, k_{n-1})$.
- We will donote this matrix by $P(p)$ where p is the permutation vector.

Theorem

Let $p = (k_0, \ldots, k_{n-1})^T$ be a permutation. Consider

$$P = P(p) = \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix}, \quad x = \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix} a_0^T \\ a_1^T \\ \vdots \\ a_{n-1}^T \end{pmatrix}.$$

Then $Px = \begin{pmatrix} \chi_{k_0} \\ \chi_{k_1} \\ \vdots \\ \chi_{k_{n-1}} \end{pmatrix}$, and $PA = \begin{pmatrix} a_{k_0}^T \\ a_{k_1}^T \\ \vdots \\ a_{k_{n-1}}^T \end{pmatrix}$.

In other words, Px and PA rearrange the elements of x and the rows of A in the order indicated by permutation vector p.

Proof

Recall that unit basis vectors have the property that $e_j^T A = \check{a}_j^T$.

$$PA = \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix} A = \begin{pmatrix} e_{k_0}^T A \\ e_{k_1}^T A \\ \vdots \\ e_{k_{n-1}}^T A \end{pmatrix} = \begin{pmatrix} \check{a}_{k_0}^T \\ \check{a}_{k_1}^T \\ \vdots \\ \check{a}_{k_{n-1}}^T \end{pmatrix}.$$

The result for Px can be proved similarly or, alternatively, by viewing x as a matrix with only one column.

Exercise

Let $p = (2,0,1)^T$. Compute

$$P(p)\begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix} \quad \text{and} \quad P(p)\begin{pmatrix} -2 & 1 & 2 \\ 3 & 2 & 1 \\ -1 & 0 & -3 \end{pmatrix}.$$

Corollary

Let $p = \left(\begin{array}{c|c|c|c} k_0 & k_1 & \cdots & k_{n-1} \end{array} \right)^T$ be a permutation and $P = P(p)$. Consider

$$A = \left(\begin{array}{c|c|c|c} a_0 & a_1 & \cdots & a_{n-1} \end{array} \right).$$

Then

$$AP^T = \left(\begin{array}{c|c|c|c} a_{k_0} & a_{k_1} & \cdots & a_{k_{n-1}} \end{array} \right).$$

Proof

Recall that unit basis vectors have the property that $Ae_k = a_k$.

$$AP^T = A \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix}^T = A \left(e_{k_0} \middle| e_{k_1} \middle| \cdots \middle| e_{k_{n-1}} \right)$$

$$= \left(Ae_{k_0} \middle| Ae_{k_1} \middle| \cdots \middle| Ae_{k_{n-1}} \right) = \left(a_{k_0} \middle| a_{k_1} \middle| \cdots \middle| a_{k_{n-1}} \right).$$

Corollary

If P is a permutation matrix, then so is P^T.

Proof

This follows from the observation that if P can be viewed either as a rearrangement of the rows or as a (usually different) rearrangement of the columns.

Corollary

Let P be a permutation matrix. Then $PP^T = P^T P = I$.

Proof

We will first prove that $PP^T = I$ Let $p = \left(\begin{array}{c|c|c|c} k_0 & k_1 & \cdots & k_{n-1} \end{array} \right)^T$ be the permutation that defines P. Then

$$PP^T = \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix} \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix}^T = \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix} \left(e_{k_0} \middle| e_{k_1} \middle| \cdots \middle| e_{k_{n-1}} \right)$$

$$= \begin{pmatrix} e_{k_0}^T e_{k_0} & e_{k_0}^T e_{k_1} & \cdots & e_{k_0}^T e_{k_{n-1}} \\ e_{k_1}^T e_{k_0} & e_{k_1}^T e_{k_1} & \cdots & e_{k_1}^T e_{k_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ e_{k_{n-1}}^T e_{k_0} & e_{k_{n-1}}^T e_{k_1} & \cdots & e_{k_{n-1}}^T e_{k_{n-1}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I.$$

Proof (continued)

We will next prove that $P^T P = I$.

Let $p = \left(\begin{array}{c|c|c|c} k_0 & k_1 & \cdots & k_{n-1} \end{array} \right)^T$ be the permutation that defines P. Then

$$P^T P = \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix}^T \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix} = \left(e_{k_0} \middle| e_{k_1} \middle| \cdots \middle| e_{k_{n-1}} \right) \begin{pmatrix} e_{k_0}^T \\ e_{k_1}^T \\ \vdots \\ e_{k_{n-1}}^T \end{pmatrix}$$

$$= e_{k_0} e_{k_0}^T + e_{k_1} e_{k_1}^T + \cdots + e_{k_{n-1}} e_{k_{n-1}}^T = e_0 e_0^T + e_1 e_1^T + \cdots + e_{n-1} e_{n-1}^T.$$

- Why?
- What does $e_0 e_0^T$ equal?
- What does $e_j e_j^T$ equal?
- What does $e_0 e_0^T + e_1 e_1^T + \cdots + e_{n-1} e_{n-1}^T$ equal?

Definition

Let us call the special permutation matrix of the form

$$\tilde{P}(\pi) = \begin{pmatrix} \boxed{e_\pi^T} \\ e_1^T \\ \vdots \\ e_{\pi-1}^T \\ \boxed{e_0^T} \\ e_{\pi+1}^T \\ \vdots \\ e_{n-1}^T \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

a *pivot matrix*.

Theorem

- When $\tilde{P}(\pi)$ multiplies a matrix from the left, it swaps rows 0 and π.
- When $\tilde{P}(\pi)$ multiplies a matrix from the right, it swaps columns 0 and π.

Back to “When LU Factorization Breaks Down”

Back to "When LU Factorization Breaks Down"

Let us reiterate the algorithmic steps that were exposed for the LU factorization

- Partition
$$A \to \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right).$$
- Update $a_{21} = a_{21}/\alpha_{11} (= l_{21})$.
- Update $A_{22} = A_{22} - a_{21}a_{12}^T (= A_{22} - l_{21}u_{12}^T)$.
- Overwrite A_{22} with L_{22} and U_{22} by continuing recursively with $A = A_{22}$.

A Different Way of Looking at It

- Partition

$$A \rightarrow \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right).$$

- Compute $l_{21} = a_{21}/\alpha_{11}$.
- Update $\left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) := \left(\begin{array}{c|c} 1 & 0 \\ \hline -l_{21} & I \end{array}\right) \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) = \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline 0 & A_{22} - l_{21}a_{12}^T \end{array}\right).$
- Overwrite A_{22} with L_{22} and U_{22} by continuing recursively with $A = A_{22}$.

$[L, A] := \text{LU\_UNB\_VAR5\_ALT}(A)$

Partition $L := I$

$$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right), \; L \rightarrow \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right)$$

where A_{TL} is 0×0

while $m(A_{TL}) < m(A)$ **do**

Repartition

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ \hline A_{20} & a_{21} & A_{22} \end{array}\right), \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} L_{00} & 0 & 0 \\ \hline l_{10}^T & 1 & 0 \\ \hline L_{20} & l_{21} & L_{22} \end{array}\right)$$

where α_{11} is 1×1

$l_{21} := a_{21}/\alpha_{11}$

$$\left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) := \left(\begin{array}{c|c} 1 & 0 \\ \hline -l_{21} & I \end{array}\right) \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) = \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline 0 & A_{22} - l_{21} a_{12}^T \end{array}\right)$$

Continue with

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ \hline A_{20} & a_{21} & A_{22} \end{array}\right), \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} L_{00} & 0 & 0 \\ \hline l_{10}^T & 1 & 0 \\ \hline L_{20} & l_{21} & L_{22} \end{array}\right)$$

endwhile

Example

Step	$\begin{pmatrix} A_{00} & a_{01} & A_{02} \\ a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{pmatrix}$	$\begin{pmatrix} L_{00} & 0 & 0 \\ l_{10}^T & 1 & 0 \\ L_{20} & l_{21} & L_{22} \end{pmatrix}$	$l_{21} := a_{21}/\alpha_{11}$	$($ …
1-2	$\begin{pmatrix} -2 & -1 & 1 \\ 2 & -2 & -3 \\ -4 & 4 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$ $($ …
3	$\begin{pmatrix} -2 & -1 & 1 \\ 0 & -3 & -2 \\ 0 & 6 & 5 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$	(-2)	$\Big(1 \;\; 0 \; \ldots \\ -(-2) \;\; 1 \; \ldots$
	$\begin{pmatrix} -2 & -1 & 1 \\ 0 & -3 & -2 \\ 0 & -2 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -2 & 1 \end{pmatrix}$		

One more time

- Partition

$$A \rightarrow \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline 0 & \alpha_{11} & a_{12}^T \\ \hline 0 & a_{01} & A_{02} \end{array}\right) \text{ and } L \rightarrow \left(\begin{array}{c|c|c} L_{00} & 0 & 0 \\ \hline l_{10}^T & 1 & 0 \\ \hline L_{20} & 0 & I \end{array}\right)$$

- Compute $l_{21} = a_{21}/\alpha_{11}$.
- Update $\left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline 0 & \alpha_{11} & a_{12}^T \\ \hline 0 & 0 & A_{02} \end{array}\right) :=$

$$\left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & -l_{21} & I \end{array}\right) \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline 0 & \alpha_{11} & a_{12}^T \\ \hline 0 & a_{01} & A_{02} \end{array}\right).$$

- Continue by moving the thick line forward one row and column.

$[L, A] := \text{LU\_UNB\_VAR5\_ALT}(A)$

Partition $L := I$

$$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right),\ L \rightarrow \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right)$$

where A_{TL} is 0×0

while $m(A_{TL}) < m(A)$ **do**

Repartition

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ \hline A_{20} & a_{21} & A_{22} \end{array}\right),\ \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c} L_{00} & 0 \\ \hline l_{10}^T & 1 \\ \hline L_{20} & l_{2} \end{array}\right.$$

where α_{11} is 1×1

$l_{21} := a_{21} / \alpha_{11}$

$$\left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ \hline A_{22} & 0 & A_{02} \end{array}\right) := \left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & -l_{21} & I \end{array}\right) \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ \hline A_{20} & a_{01} & A_{02} \end{array}\right)$$

$$= \left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & -l_{21} & I \end{array}\right) \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline 0 & \alpha_{11} & a_{12}^T \\ \hline 0 & a_{01} & A_{02} \end{array}\right) = \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ \hline 0 & \alpha_{11} & a_{12}^T \\ \hline 0 & 0 & A_{02} - l_{21} a \end{array}\right.$$

Continue with

Note

- Upon completion, A is an upper triangular matrix, U.
- The point of this alternative explanation is to show that
  - *if* $\check{L}^{(i)}$ represents the ith Gauss transform, computed during the ith iteration of the algorithms,
  - *then* the final matrix stored in A, the upper triangular matrix U, satisfies $U = \check{L}^{(n-2)}\check{L}^{(n-3)}\cdots\check{L}^{(0)}\hat{A}$,
  - where $\hat{A}$ is the original matrix stored in A.

Example

$$A = A^{(0)} = \begin{pmatrix} -2 & -1 & 1 \\ 2 & -2 & -3 \\ -4 & 4 & 7 \end{pmatrix} \text{ and } L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In the first step,

- Partition

$$\left(\begin{array}{|c|cc}\hline -2 & -1 & 1 \\ \hline 2 & -2 & -3 \\ -4 & 4 & 7\end{array}\right) \text{ and } \left(\begin{array}{|c|cc}\hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1\end{array}\right);$$

- Compute $l_{21} = \begin{pmatrix} 2 \\ -4 \end{pmatrix} / (-2) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

- Update A with

$$A^{(1)} = \underbrace{\left(\begin{array}{|c|cc}\hline 1 & 0 & 0 \\ \hline 1 & 1 & 0 \\ -2 & 0 & 1\end{array}\right)}_{\check{L}^{(0)}} \underbrace{\left(\begin{array}{|c|cc}\hline -2 & -1 & 1 \\ \hline 2 & -2 & -3 \\ -4 & 4 & 7\end{array}\right)}_{A^{(0)}}$$

$$= \left(\begin{array}{|c|cc}\hline -2 & -1 & 1 \\ \hline 0 & -3 & -2 \\ 0 & 6 & 5\end{array}\right).$$

We emphasize that now $A^{(1)} = \check{L}^{(0)} A^{(0)}$.

In the second step,

- Partition

$$\left(\begin{array}{c|c|c} -2 & -1 & 1 \\ \hline 0 & -3 & -2 \\ \hline 0 & 6 & 5 \end{array}\right) \text{ and } \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline -1 & 1 & 0 \\ \hline 2 & 0 & 1 \end{array}\right);$$

- Compute $l_{21} = \left(\ 6\ \right)/(-3) = \left(\ -2\ \right)$.
- Update A with

$$A^{(2)} = \underbrace{\left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 2 & 1 \end{array}\right)}_{\check{L}^{(1)}} \underbrace{\left(\begin{array}{c|c|c} -2 & -1 & 1 \\ \hline 0 & -3 & -2 \\ \hline 0 & 6 & 5 \end{array}\right)}_{A^{(1)}} = \left(\begin{array}{c|c|c} -2 & -1 & 1 \\ \hline 0 & -3 & -2 \\ \hline 0 & 0 & 1 \end{array}\right).$$

Now...

$$A = \underbrace{\begin{pmatrix} -2 & -1 & 1 \\ 0 & -3 & -2 \\ 0 & 0 & 1 \end{pmatrix}}_{A^{(2)}}$$

$$= \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}}_{\check{L}^{(1)}} \underbrace{\begin{pmatrix} -2 & -1 & 1 \\ 0 & -3 & -2 \\ 0 & 6 & 5 \end{pmatrix}}_{A^{(1)}}$$

$$= \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}}_{\check{L}^{(1)}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}}_{\check{L}^{(0)}} \underbrace{\begin{pmatrix} -2 & -1 & 1 \\ 2 & -2 & -3 \\ -4 & 4 & 7 \end{pmatrix}}_{A^{(0)}}$$

The point is...

- LU factorization can be viewed as the computation of a sequence of Gauss transforms so that, upon completion

$$U = \check{L}^{(n-1)} \check{L}^{(n-2)} \check{L}^{(n-3)} \cdots \check{L}^{(0)} A.$$

- Now, let us reconsider the following property of a typical Gauss transform:

$$\underbrace{\left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & -l_{21} & I \end{array}\right)}_{\check{L}^{(i)}} \underbrace{\left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & l_{21} & I \end{array}\right)}_{L^{(i)}} = \underbrace{\left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & I \end{array}\right)}_{I}$$

Example (continued)

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}}_{L^{(0)}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}}_{L^{(1)}} \underbrace{\begin{pmatrix} -2 & -1 & 1 \\ 0 & -3 & -2 \\ 0 & 0 & 1 \end{pmatrix}}_{U}$$

$$= \underbrace{\begin{pmatrix} -2 & -1 & 1 \\ 2 & -2 & -3 \\ -4 & 4 & 7 \end{pmatrix}}_{A}$$

Finally, note that

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}}_{L^{(0)}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}}_{L^{(1)}} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -2 & 1 \end{pmatrix}}_{\check{L}^{(1)}}$$

$[L, A] := \text{LU\_UNB\_VAR5\_PIV}(A)$

Partition $L := I$

$$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right), L \rightarrow \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right), p \rightarrow \left(\begin{array}{c} p_T \\ \hline p_B \end{array}\right)$$

where A_{TL} is 0×0

while $m(A_{TL}) < m(A)$ **do**

Repartition

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \cdots, \left(\begin{array}{c|c} L_{TL} & 0 \\ \hline L_{BL} & L_{BR} \end{array}\right) \rightarrow \cdots, \left(\begin{array}{c} p_T \\ \hline p_B \end{array}\right) \rightarrow \left(\begin{array}{c} p_0 \\ \hline p_1 \\ \hline p_2 \end{array}\right)$$

where α_{11} is 1×1

$$\pi_1 = \text{Pivot}\left(\left(\begin{array}{c} \alpha_{11} \\ \hline a_{21} \end{array}\right)\right)$$

$$\left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) := P(\pi_1) \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right)$$

$$l_{21} := a_{21}/\alpha_{11}$$

$$\left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) := \left(\begin{array}{c|c} 1 & 0 \\ \hline -l_{21} & I \end{array}\right) \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline a_{21} & A_{22} \end{array}\right) = \left(\begin{array}{c|c} \alpha_{11} & a_{12}^T \\ \hline 0 & A_{22} - l_{21}a_{12}^T \end{array}\right)$$

Continue with

Example: Adding Row Swaps (Pivoting)

- Consider again the system of linear equations

$$\begin{array}{rrrrl} 2\chi_0 + & 4\chi_1 + & (-2)\chi_2 & = & -10 \\ 4\chi_0 + & 8\chi_1 + & 6\chi_2 & = & 20 \\ 6\chi_0 + & (-4)\chi_1 + & 2\chi_2 & = & 18 \end{array}$$

- Focus on the matrix of coefficients

$$A = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 8 & 6 \\ 6 & -4 & 2 \end{pmatrix}.$$

Iteration 0

- Apply a pivot to ensure that the diagonal element in the first column is not zero.
- In this example, no pivoting is required, so the first pivot matrix, $\tilde{P}^{(0)} = I$:

$$\underbrace{\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)}_{\tilde{P}^{(0)}} \left(\begin{array}{c|rr} 2 & 4 & -2 \\ \hline 4 & 8 & 6 \\ 6 & -4 & 2 \end{array}\right) = \underbrace{\left(\begin{array}{c|rr} 2 & 4 & -2 \\ \hline 4 & 8 & 6 \\ 6 & -4 & 2 \end{array}\right)}_{\tilde{A}^{(0)}}$$

Iteration 0 (continued)

- Next, a Gauss transform is computed and applied:

$$\underbrace{\left(\begin{array}{|r|rr} \hline 1 & 0 & 0 \\ \hline -2 & 1 & 0 \\ -3 & 0 & 1 \end{array}\right)}_{\check{L}^{(0)}} \underbrace{\left(\begin{array}{|r|rr} \hline 2 & 4 & -2 \\ \hline 4 & 8 & 6 \\ 6 & -4 & 2 \end{array}\right)}_{\tilde{A}^{(0)}} = \underbrace{\left(\begin{array}{|r|rr} \hline 2 & 4 & -2 \\ \hline 0 & 0 & 10 \\ 0 & -16 & 8 \end{array}\right)}_{A^{(1)}}.$$

Iteration 1

- Now the second and third row must be swapped by pivot matrix $\tilde{P}^{(1)}$:

$$\underbrace{\left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline 0 & 1 & 0 \end{array}\right)}_{\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{P}^{(1)} \end{array}\right)} \underbrace{\left(\begin{array}{c|r|r} 2 & 4 & -2 \\ \hline 0 & 0 & 10 \\ \hline 0 & -16 & 8 \end{array}\right)}_{A^{(1)}} = \underbrace{\left(\begin{array}{c|r|r} 2 & 4 & -2 \\ \hline 0 & -16 & 8 \\ \hline 0 & 0 & 10 \end{array}\right)}_{\tilde{A}^{(1)}}$$

Iteration 1 (continued

- A Gauss transform is computed and applied:

$$\underbrace{\left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)}_{\check{L}^{(1)}} \underbrace{\left(\begin{array}{c|r|r} 2 & 4 & -2 \\ \hline 0 & -16 & 8 \\ \hline 0 & 0 & 10 \end{array}\right)}_{\tilde{A}^{(1)}} = \underbrace{\left(\begin{array}{c|r|r} 2 & 4 & -2 \\ \hline 0 & -16 & 8 \\ \hline 0 & 0 & 10 \end{array}\right)}_{A^{(2)}}.$$

Notice

- In each iteration, some permutation matrix is used to swap two rows, after which a Gauss transform is computed and then applied to the resulting (permuted) matrix.
- One can describe this as

$$U = \check{L}^{(n-2)}\check{P}^{(n-2)}\check{L}^{(n-3)}\check{P}^{(n-3)}\cdots\check{L}^{(0)}\check{P}^{(0)}A,$$

  where $P^{(i)}$ represents the permutation applied during iteration i.
- Now, once an LU factorization with pivoting is computed, one can solve $Ax = b$:

$$\begin{aligned} Ux &= \check{L}^{(n-2)}\check{P}^{(n-2)}\check{L}^{(n-3)}\check{P}^{(n-3)}\cdots\check{L}^{(0)}\check{P}^{(0)}Ax \\ &= \check{L}^{(n-2)}\check{P}^{(n-2)}\check{L}^{(n-3)}\check{P}^{(n-3)}\cdots\check{L}^{(0)}\check{P}^{(0)}b. \end{aligned}$$

Note

- If the LU factorization with pivoting completes without encountering a zero pivot, then given any right-hand size b this procedure produces a unique solution x.
- In other words, the procedure computes the net effect of applying A^{-1} to the right-hand side vector b, and therefore A has an inverse.
- If a zero pivot is encountered, then there exists a vector $x \neq 0$ such that $Ax = 0$, and hence the inverse does not exist.