Chapter 4 - LU Factorization - Part 3

Maggie Myers
Robert A. van de Geijn
The University of Texas at Austin

Practical Linear Algebra – Fall 2009
Inverse of a Matrix
Definition

Given $A \in \mathbb{R}^{n \times n}$, a matrix $B$ that has the property that $BA = I$, the identity, is called the inverse of matrix $A$ and is denoted by $A^{-1}$.

- Not every square matrix has an inverse!
- The inverse of a nonsquare matrix is not defined.
- Indeed, we will periodically relate other properties of a matrix to the matrix having an inverse as these notes unfold.
Notice that $A^{-1}$ is the matrix that “undoes” the transformation $A$: $A^{-1}(Ax) = x$. It acts as the inverse function of the linear transformation $L(x) = Ax$. 
Examples

- $I^{-1} =$
  \[
  \begin{pmatrix}
  \delta_0 & 0 & \cdots & 0 \\
  0 & \delta_1 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \delta_{n-1}
  \end{pmatrix}
  \]

- $I^{-1} =$
  \[
  \begin{pmatrix}
  \delta_0 & 0 & \cdots & 0 \\
  0 & \delta_1 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \delta_{n-1}
  \end{pmatrix}
  \]

- If $D = \delta I$ then $D^{-1} =$

- **Zero matrix**: Let $O$ denote the $n \times n$ matrix of all zeroes.
  \[
  O^{-1} =
  \]
Examples

- $I^{-1} = I$.
- 
  \[
  \begin{pmatrix}
  \delta_0 & 0 & \cdots & 0 \\
  0 & \delta_1 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \delta_{n-1}
  \end{pmatrix}
  \]

  \[
  = \begin{pmatrix}
  \frac{1}{\delta_0} & 0 & \cdots & 0 \\
  0 & \frac{1}{\delta_1} & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \frac{1}{\delta_{n-1}}
  \end{pmatrix}.
  \]

- If $D = \delta I$ then $D^{-1} = 1/\delta I$.

- **Zero matrix:** Let $O$ denote the $n \times n$ matrix of all zeroes. This matrix does not have an inverse. Why? Pick any vector $x \neq 0$ (not equal to the zero vector). Then $O^{-1}(Ox) = 0$ and, if $O^{-1}$ existed, $(O^{-1}O)x = Ix = x$, which is a contradiction.
**Theorem**

Let $Ax = b$ and assume that $A$ has an inverse, $A^{-1}$. Then $x = A^{-1}b$.

**Proof**

If $Ax = b$ then $A^{-1}Ax = A^{-1}b$ and hence $Ix = x = A^{-1}b$. 
Corollary

Assume that $A$ has an inverse, $A^{-1}$. Then $Ax = 0$ implies that $x = 0$.

Proof

If $A$ has an inverse and $Ax = 0$, then

$$x =Ix = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}0 = 0.$$
Theorem
If $A$ has an inverse $A^{-1}$, then $AA^{-1} = A^{-1}A = I$.

Theorem: Uniqueness of the inverse
If $A$ has an inverse, then that inverse is unique.

Proof
Assume that $AB = BA = I$ and $AC = CA = I$. Then by associativity of matrix multiplication
$C = CI = C(AB) = (CA)B = B$.  

http://z.cs.utexas.edu/wiki/pla.wiki/
Computing the inverse

- Assume $A$ has an inverse and that $C = A^{-1}$.
- By definition matrix $C$ satisfies $AC = I$.

$$
A \begin{pmatrix}
c_0 & c_1 & \cdots & c_{n-1}
\end{pmatrix}_C = \begin{pmatrix}
Ac_0 & Ac_1 & \cdots & Ac_{n-1}
\end{pmatrix}
= \begin{pmatrix}
e_0 & e_1 & \cdots & e_{n-1}
\end{pmatrix}_I,
$$

- Thus, the $j$th column of $C$, $c_j$, must solve $Ac_j = e_j$. 
If Gaussian elimination without row swapping works then you can solve $Ax = b$ by applying Gaussian elimination to the augmented system $(A | b)$, leaving the result as $(U | z)$ (where we later saw that $z$ solves $Lz = b$), after which backward substitution could be used to solve the upper triangular system $Ux = z$. 
Computing the inverse, continued

So, this means that we should do this for each of the equations
\( Ac_j = e_j \):
- Append \(( A \mid e_j )\), leaving the result as \(( U \mid z_j )\).
- Perform back substitution to solve \( Uc_j = z_j \).

We will perfect this, later.
Example

Consider the $2 \times 2$ matrix \[
\begin{pmatrix}
2 & -1 \\
1 & 1
\end{pmatrix}.
\]

Let $A^{-1} = \begin{pmatrix}
\beta_{0,0} & \beta_{0,1} \\
\beta_{1,0} & \beta_{1,1}
\end{pmatrix}$.

Then \[
\begin{pmatrix}
2 & -1 \\
1 & 1
\end{pmatrix}\begin{pmatrix}
\beta_{0,0} & \beta_{0,1} \\
\beta_{1,0} & \beta_{1,1}
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}.
\]

This yields two linear systems:

\[
\begin{pmatrix}
2 & -1 \\
1 & 1
\end{pmatrix}\begin{pmatrix}
\beta_{0,0} \\
\beta_{1,0}
\end{pmatrix} = \begin{pmatrix}
1 \\
0
\end{pmatrix}
\]

and

\[
\begin{pmatrix}
2 & -1 \\
1 & 1
\end{pmatrix}\begin{pmatrix}
\beta_{0,1} \\
\beta_{1,1}
\end{pmatrix} = \begin{pmatrix}
0 \\
1
\end{pmatrix}.
\]

Solving these yields $A^{-1} = \begin{pmatrix}
\frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & \frac{2}{3}
\end{pmatrix}$.  

http://z.cs.utexas.edu/wiki/pla.wiki/
Exercise

Check that \[
\begin{pmatrix}
2 & -1 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
\frac{1}{3} & \frac{1}{3} \\
-\frac{1}{3} & \frac{2}{3}
\end{pmatrix}
= \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}.
\]
Inverting a $2 \times 2$ matrix

- Consider

$$
\begin{pmatrix}
\alpha_{0,0} & \alpha_{0,1} \\
\alpha_{1,0} & \alpha_{1,1}
\end{pmatrix}
\begin{pmatrix}
\beta_{0,0} & \beta_{0,1} \\
\beta_{1,0} & \beta_{1,1}
\end{pmatrix} = 
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}.
$$

- This yields two linear systems:

$$
\begin{pmatrix}
\alpha_{0,0} & \alpha_{0,1} \\
\alpha_{1,0} & \alpha_{1,1}
\end{pmatrix}
\begin{pmatrix}
\beta_{0,0} \\
\beta_{1,0}
\end{pmatrix} = 
\begin{pmatrix}
1 \\
0
\end{pmatrix}
$$

and

$$
\begin{pmatrix}
\alpha_{0,0} & \alpha_{0,1} \\
\alpha_{1,0} & \alpha_{1,1}
\end{pmatrix}
\begin{pmatrix}
\beta_{0,1} \\
\beta_{1,1}
\end{pmatrix} = 
\begin{pmatrix}
0 \\
1
\end{pmatrix}.
$$

- Solving these yields

$$
\begin{pmatrix}
\alpha_{0,0} & \alpha_{0,1} \\
\alpha_{1,0} & \alpha_{1,1}
\end{pmatrix}^{-1} = \frac{1}{\left(\alpha_{0,0}\alpha_{1,1} - \alpha_{0,1}\alpha_{1,0}\right)}
\begin{pmatrix}
\alpha_{1,1} & -\alpha_{0,1} \\
-\alpha_{1,0} & \alpha_{0,0}
\end{pmatrix}.
$$

http://z.cs.utexas.edu/wiki/pla.wiki/
The expression

\[
\frac{1}{(\alpha_{0,0}\alpha_{1,1} - \alpha_{0,1}\alpha_{1,0})}
\begin{pmatrix}
\alpha_{1,1} & -\alpha_{0,1} \\
-\alpha_{1,0} & \alpha_{0,0}
\end{pmatrix}
\]

is known as the determinant.

The inverse of the $2 \times 2$ matrix exists if and only if this expression is not equal to zero.
Generalizing

- A determinant can be defined for any $n \times n$ matrix $A$.
- Kramer’s rule for solving linear equations (taught in high school algebra classes) requires computation the determinants of various matrices.
- But this method is completely impractical and therefore does not deserve any of our time.
- We will revisit determinants when we discuss eigenvalues.
Gauss-Jordan method for inverting a matrix

Consider the matrix

\[ A = \begin{pmatrix} 2 & 4 & -2 \\ 4 & -2 & 6 \\ 6 & -4 & 2 \end{pmatrix}. \]

Compute the first column of the inverse matrix by applying Gaussian elimination to the augmented systems

\[
\begin{pmatrix} 2 & 4 & -2 & | & 1 \\ 4 & -2 & 6 & | & 0 \\ 6 & -4 & 2 & | & 0 \end{pmatrix},
\]

\[
\begin{pmatrix} 2 & 4 & -2 & | & 0 \\ 4 & -2 & 6 & | & 1 \\ 6 & -4 & 2 & | & 0 \end{pmatrix},
\]

\[
\begin{pmatrix} 2 & 4 & -2 & | & 0 \\ 4 & -2 & 6 & | & 0 \\ 6 & -4 & 2 & | & 1 \end{pmatrix}.
\]
Better

Apply them all at once:

$$
\begin{pmatrix}
2 & 4 & -2 & 1 & 0 & 0 \\
4 & -2 & 6 & 0 & 1 & 0 \\
6 & -4 & 2 & 0 & 0 & 1 \\
\end{pmatrix}.
$$
Then, proceeding with Gaussian elimination:

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

\[
\begin{pmatrix}
2 & 4 & -2 \\
0 & -10 & 10 \\
0 & -16 & 8
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 \\
-2 & 1 & 0 \\
-3 & 0 & 1
\end{pmatrix}
\]

- By subtracting \(((-16)/(-10)) = 1.6\) times the second row from the third row, we get

\[
\begin{pmatrix}
2 & 4 & -2 \\
0 & -10 & 10 \\
0 & 0 & -8
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 \\
-2 & 1 & 0 \\
0.2 & -1.6 & 1
\end{pmatrix}
\]
Simultaneous backsubstitution

\[
\begin{pmatrix}
2 & 4 & -2 & | & 1 & 0 & 0 \\
0 & -10 & 10 & | & -2 & 1 & 0 \\
0 & 0 & -8 & | & 0.2 & -1.6 & 1 \\
\end{pmatrix}
\]

Look at the “10” and the “-8” in last column on the left of the $|$. Subtract $(10)/(-8)$ times the last row from the second row, producing

\[
\begin{pmatrix}
2 & 4 & -2 & | & 1 & 0 & 0 \\
0 & -10 & 0 & | & -1.75 & -1 & 1.25 \\
0 & 0 & -8 & | & 0.2 & -1.6 & 1 \\
\end{pmatrix}
\]
Now take the “-2” and the “-8” in last column on the left of the \( | \) and subtract \((-2)/(-8)\) times the last row from the first row, producing

\[
\begin{pmatrix}
2 & 4 & -2 & & 1 & 0 & 0 \\
0 & -10 & 0 & & -1.75 & -1 & 1.25 \\
0 & 0 & -8 & & 0.2 & -1.6 & 1 \\
\end{pmatrix}
\]
Finally take the “4” and the “-10” in second column on the left of the | and subtract \((4)/(-10)\) times the second row from the first row, producing

\[
\begin{pmatrix}
2 & 0 & 0 & | & 0.25 & 0 & 0.25 \\
0 & -10 & 0 & | & -1.75 & -1 & 1.25 \\
0 & 0 & -8 & | & 0.2 & -1.6 & 1
\end{pmatrix}
\]

Finally, divide the first, second, and third row by the diagonal elements on the left, respectively, yielding

\[
\begin{pmatrix}
1 & 0 & 0 & | & 0.125 & 0 & 0.125 \\
0 & 1 & 0 & | & 0.175 & 0.1 & -0.125 \\
0 & 0 & 1 & | & -0.025 & 0.2 & -0.125
\end{pmatrix}
\]
Lo and behold, the matrix on the right is the inverse of the original matrix:

\[
\begin{pmatrix}
2 & 4 & -2 \\
4 & -2 & 6 \\
6 & -4 & 2
\end{pmatrix}
\begin{pmatrix}
0.125 & 0 & 0.125 \\
0.175 & 0.1 & -0.125 \\
-0.025 & 0.2 & -0.125
\end{pmatrix}
= 
\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

Notice that this procedure works only if no divide by zero is encountered.
Note

Just like Gaussian elimination and LU factorization could be fixed if a zero pivot were encountered by swapping rows (pivoting), Gauss-Jordan can be fixed similarly. It is only if in the end swapping of rows does not yield a nonzero pivot that the process fully breaks down. More on this, later.
Although we don’t state the above remark as a formal theorem, let us sketch a proof anyway:

We previous showed that LU factorization with pivoting could be viewed as computing a sequence of Gauss transforms and pivoting matrices that together transform $n \times n$ matrix $A$ to an upper triangular matrix:

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

where $\tilde{L}^{(i)}$ and $\tilde{P}^{(i)}$ represents the permutation applied during iteration $i$. 
Now, if $U$ has no zeroes on the diagonal (no zero pivots were encountered during LU with pivoting) then it has an inverse. So,

$$U^{-1} \tilde{L}^{(n-2)} \tilde{P}^{(n-2)} \tilde{L}^{(n-3)} \tilde{P}^{(n-3)} \ldots \tilde{L}^{(0)} \tilde{P}^{(0)} A = I,$$

which means that $A$ has an inverse:

$$A^{-1} = U^{-1} \tilde{L}^{(n-2)} \tilde{P}^{(n-2)} \tilde{L}^{(n-3)} \tilde{P}^{(n-3)} \ldots \tilde{L}^{(0)} \tilde{P}^{(0)}.$$
What the first stage of Gauss-Jordan process does is to compute

$$U = \left( \tilde{L}^{(n-2)} \left( \tilde{P}^{(n-2)} \left( \tilde{L}^{(n-3)} \left( \tilde{P}^{(n-3)} \ldots \left( \left( \tilde{L}^{(0)} \left( \tilde{P}^{(0)} A \right) \right) \right) \right) \right) \right) \ldots \right),$$

applying the computed transformations also to the identity matrix:

$$B = \left( \tilde{L}^{(n-2)} \left( \tilde{P}^{(n-2)} \left( \tilde{L}^{(n-3)} \left( \tilde{P}^{(n-3)} \ldots \left( \left( \tilde{L}^{(0)} \left( \tilde{P}^{(0)} I \right) \right) \right) \right) \right) \right) \ldots \right).$$

The second stage of Gauss-Jordan (where the elements of $A$ above the diagonal are eliminated) is equivalent to applying $U^{-1}$ from the left to both $U$ and $B$.

By viewing the problems as the appended (augmented) system

$$\left( \begin{array}{c|c} A & I \end{array} \right)$$

is just a convenient way for writing all the intermediate results, applying each transformation to both $A$ and $I$. 
Inverting a matrix using the LU factorization

An alternative to the Gauss-Jordan method illustrated above:

- Compute the LU factorization of matrix $A$:

  \[ A = LU. \]

  Solve each $Ab_j = e_j$ via
  - $Lz_j = e_j$,
  - $Ub_j = z_j$.

Notice that, like for the Gauss-Jordan procedure, this approach works only if no zero pivot is encountered.
Theorem

Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix with (all) nonzero diagonal elements. Then its inverse $L^{-1}$ exists and is lower triangular.
Proof

Proof by Induction. Let $L$ be a lower triangular matrix with (all) nonzero diagonal elements.

- **Base case:** Let $L = \begin{pmatrix} \lambda_{11} \end{pmatrix}$. Then, since $\lambda_{11} \neq 0$, we let $L^{-1} = \begin{pmatrix} 1/\lambda_{11} \end{pmatrix}$, which is lower triangular and well-defined.

- **Inductive step:** Inductive hypothesis: Assume that for a given $k \geq 0$ the inverse of all $k \times k$ lower triangular matrices with nonzero diagonal elements exist and are lower triangular. We will show that the inverse of a $(k + 1) \times (k + 1)$ lower triangular matrix with (all) nonzero diagonal elements exists and is lower triangular.
Let \((k + 1) \times (k + 1)\) matrix \(L\) be lower triangular with (all) nonzero diagonal elements. We will construct a matrix that is its inverse and is lower triangular. Partition

\[
L \rightarrow \begin{pmatrix}
\lambda_{11} & 0 \\
l_{21} & L_{22}
\end{pmatrix}
\]

where \(L_{22}\) is lower triangular (why?) and has (all) nonzero diagonal elements (why?), and \(\lambda_{11} \neq 0\) (why?). Then

\[
\begin{pmatrix}
\lambda_{11} & 0 \\
l_{21} & L_{22}
\end{pmatrix} \begin{pmatrix}
1/\lambda_{11} & 0 \\
-L_{22}^{-1} l_{21} / \lambda_{11} & L_{22}^{-1}
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & I
\end{pmatrix}!
\]

The desired matrix exists because \(\lambda_{11} \neq 0\) and (by I.H.) \(L_{22}^{-1}\) exists.
By the Principle of Mathematical Induction the result holds for all lower triangular matrices with nonzero diagonal elements.
Corollary

Let $L \in \mathbb{R}^{n \times n}$ be a unit lower triangular matrix with (all) nonzero diagonal elements. Then its inverse $L^{-1}$ exists and is unit lower triangular.

Corollary

Let $U \in \mathbb{R}^{n \times n}$ be an upper triangular matrix with (all) nonzero diagonal elements. Then its inverse $U^{-1}$ exists and is upper triangular.

Let $U \in \mathbb{R}^{n \times n}$ be a unit upper triangular matrix. Then its inverse $U^{-1}$ exists and is unit upper triangular.
Exercise

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with (all) nonzero diagonal elements. Partition $L \rightarrow \begin{pmatrix} L_{TL} & 0 \\ L_{BL} & L_{BR} \end{pmatrix}$, where $L_{TL}$ is $k \times k$. Show that $L^{-1} = \begin{pmatrix} L_{TL}^{-1} & 0 \\ -L_{BR}^{-1}L_{BL}L_{TL}^{-1} & L_{BR}^{-1} \end{pmatrix}$.
Inverting the LU factorization

Yet another way to compute $A^{-1}$ is to compute its LU factorization, $A = LU$ and to then note that $A^{-1} = (LU)^{-1} = U^{-1}L^{-1}$. But that requires us to discuss algorithms for inverting a triangular matrix, which is also beyond the scope of this document. This is actually (closer to) how matrices are inverted in practice. Again, this approach works only if no zero pivot is encountered.
In practice, do not use inverted matrices!

Inverses of matrices are a wonderful theoretical tool. They are not a practical tool.

We noted that if one wishes to solve $Ax = b$, and $A$ has an inverse, then $x = A^{-1}b$. Does this mean we should compute the inverse of a matrix in order to compute the solution of $Ax = b$? The answer is a resounding “no”.

Note

If anyone ever indicates they invert a matrix in order to solve a linear system of equations, they either are

1. very naive and need to be corrected; or
2. they really mean that they are just solving the linear system and don’t really mean that they invert the matrix.
Theorem
Let $A, B \in \mathbb{R}^{n \times n}$ assume that $A^{-1}$ and $B^{-1}$ exist. Then $(AB)^{-1}$ exists and equals $B^{-1}A^{-1}$.

Proof
Let $C = AB$. It suffices to find a matrix $D$ such that $CD = I$ since then $C^{-1} = D$. Now,

$$C(B^{-1}A^{-1}) = (AB)(B^{-1}A^{-1}) = A \underbrace{(BB^{-1})}_{I} A^{-1} = AA^{-1} = I$$

and thus $D = B^{-1}A^{-1}$ has the desired property.
Theorem
Let $A \in \mathbb{R}^{n \times n}$ and assume that $A^{-1}$ exists. Then $(A^T)^{-1}$ exists and equals $(A^{-1})^T$.

Proof
We need to show that $A^T(A^{-1})^T = I$ or, equivalently, that $(A^T(A^{-1})^T)^T = I^T = I$. But

$$(A^T(A^{-1})^T)^T = ((A^{-1})^T)^T(A^T)^T = A^{-1}A = I,$$

which proves the desired result.
Theorem

Let $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$, and assume that $A$ has an inverse. Then $Ax = 0$ if and only if $x = 0$.

Proof

- Assume that $A^{-1}$ exists. If $Ax = 0$ then
  \[ x = Ix = A^{-1}Ax = A^{-1}0 = 0. \]
- Let $x = 0$. Then clearly $Ax = 0$. 
Theorem

Let $A \in \mathbb{R}^{n \times n}$. Then $A$ has an inverse if and only if Gaussian elimination with row pivoting does not encounter a zero pivot.

Proof

 Assume Gaussian elimination with row pivoting does not encounter a zero pivot. We will show that $A$ then has an inverse.

Let $\tilde{L}_{n-1}P_{n-1} \cdots \tilde{L}_0P_0A = U$, where $P_k$ and $\tilde{L}_k$ are the permutation matrix that swaps the rows and the Gauss transform computed and applied during the $k$-th iteration of Gaussian elimination, and $U$ is the resulting upper triangular matrix. The fact that Gaussian elimination does not encounter a zero pivot means that all these permutation matrices and Gauss transforms exist and it means that $U$ has only nonzeroes on the diagonal. We have already seen that the inverse of a permutation matrix is its transpose and that the inverse of each $\tilde{L}_k$ exists (let us call it $L_k$). We also have