

3.6 LU and PLU Factorizations

Gaussian elimination solves one system $\mathbf{Ax} = \mathbf{b}$. In practice the matrix often stays fixed while the right-hand side changes: the same network under different loads, the same model fed different data. The row operations depend only on $\mathbf{A}$, so repeating them for every $\mathbf{b}$ wastes almost all of the work. This section stores the elimination as a factorization $\mathbf{A} = \mathbf{LU}$, a lower triangular matrix that remembers the operations times an upper triangular matrix that is their result. We prove when it exists and when it is unique, and we repair its failures with row swaps.

Throughout, F is a field and all matrices are square, in $M_n(F)$.

3.6.1 Triangular systems are cheap

Some systems need no elimination at all. Take

$$\mathbf{U} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & -2 \end{pmatrix}, \quad \mathbf{U}\mathbf{x} = \begin{pmatrix} -1 \\ -1 \\ -4 \end{pmatrix}.$$

The last equation mentions only x_3 : $-2x_3 = -4$, so $x_3 = 2$. The second mentions x_2 and the known x_3 : $3x_2 + 2 = -1$, so $x_2 = -1$. The first then gives $2x_1 - 1 - 2 = -1$, so $x_1 = 1$. This is **back substitution**: in an upper triangular system with nonzero diagonal entries, each equation introduces exactly one new unknown, read from the bottom up. A lower triangular system is solved the same way from the top down, by **forward substitution**.

Now suppose a matrix $\mathbf{A}$ factors as $\mathbf{A} = \mathbf{LU}$ with $\mathbf{L}$ lower triangular and $\mathbf{U}$ upper triangular. Then $\mathbf{Ax} = \mathbf{b}$ reads $\mathbf{L}(\mathbf{U}\mathbf{x}) = \mathbf{b}$. Naming the inner vector $\mathbf{y} = \mathbf{U}\mathbf{x}$ splits the problem in two:

$$\text{first solve } \mathbf{L}\mathbf{y} = \mathbf{b} \text{ for } \mathbf{y}, \quad \text{then solve } \mathbf{U}\mathbf{x} = \mathbf{y} \text{ for } \mathbf{x}.$$

Both are triangular, so both are cheap. The expensive part, finding $\mathbf{L}$ and $\mathbf{U}$, is done once for all right-hand sides.

An LU factorization is Gaussian elimination written down once, so that it never has to be done again.

3.6.2 LU factorizations

Definition 3.6.1. LU factorization

A matrix $\mathbf{L} \in M_n(F)$ is **unit lower triangular** if it is lower triangular and **every** diagonal entry equals 1. Let $\mathbf{A} \in M_n(F)$. An **LU factorization** of $\mathbf{A}$ is a pair $(\mathbf{L}, \mathbf{U})$ of matrices in $M_n(F)$ such that

$$\mathbf{A} = \mathbf{LU}, \quad \mathbf{L} \text{ is unit lower triangular,} \quad \mathbf{U} \text{ is upper triangular.}$$

In words: $\mathbf{L}$ has 1's on the diagonal and zeros above it, and the entries below the diagonal are unrestricted. $\mathbf{U}$ has zeros strictly below the diagonal, and its diagonal entries are unrestricted; some of them may be 0. The definition says **an** factorization: whether there is only one is a theorem, and it needs a hypothesis.

Example. An LU factorization from elimination

Let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & -1 \\ 4 & 5 & -1 \\ -2 & 8 & 2 \end{pmatrix} \in M_3(\mathbb{Q}).$$

Reduce $\mathbf{A}$ to upper triangular form using only operations “subtract a multiple of a higher row from a lower row”, and record each multiplier. Compare the result with the matrix $\mathbf{U}$ of the previous subsection.

Solution. The first pivot is 2. To clear column 1 below it we subtract $\ell_{21} = 4/2 = 2$ times row 1 from row 2, and $\ell_{31} = -2/2 = -1$ times row 1 from row 3:

$$\begin{pmatrix} 2 & 1 & -1 \\ 4 & 5 & -1 \\ -2 & 8 & 2 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - (-1)R_1}]{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - (-1)R_1}} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 9 & 1 \end{pmatrix}.$$

The second pivot is 3. We subtract $\ell_{32} = 9/3 = 3$ times row 2 from row 3:

$$\begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 9 & 1 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 3R_2} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & -2 \end{pmatrix} = \mathbf{U}.$$

Put the multipliers below the diagonal of a unit lower triangular matrix, each in the position of the entry it cleared:

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix}.$$

Multiplying out, row by row,

$$\mathbf{LU} = \begin{pmatrix} 2 & 1 & -1 \\ 2 \cdot 2 & 2 \cdot 1 + 3 & 2 \cdot (-1) + 1 \\ -2 & -1 + 9 & 1 + 3 - 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 4 & 5 & -1 \\ -2 & 8 & 2 \end{pmatrix} = \mathbf{A}.$$

So $(\mathbf{L}, \mathbf{U})$ is an LU factorization of $\mathbf{A}$, and $\mathbf{U}$ is exactly the triangular matrix of the back substitution above.

The example suggests the general statement: elimination produces $\mathbf{U}$, and the multipliers, **with their signs as subtracted**, fill in $\mathbf{L}$. We prove this below. First, the smallest cases.

- **Upper triangular $\mathbf{A}$.** $\mathbf{A} = \mathbf{I} \cdot \mathbf{A}$, and $\mathbf{I}$ is unit lower triangular. So every upper triangular matrix, including $\mathbf{O}$ and $\mathbf{I}$, has the LU factorization $(\mathbf{I}, \mathbf{A})$.
- 1×1 . Every $(a) \in M_1(F)$ is upper triangular, so $(a) = (1)(a)$. Nothing can go wrong here, and that is exactly why the 2×2 failure below is instructive.

Nonexample by minimal change. The matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ has the LU factorization

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Change the top-left entry from 1 to 0, and everything collapses.

Example. An invertible matrix with no LU factorization

Show that $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(F)$ has **no** LU factorization, over any field F , although $\mathbf{A}$ is invertible.

Solution. $\mathbf{A}$ is invertible, since $\mathbf{A}\mathbf{A} = \mathbf{I}_2$. Suppose, for a contradiction, that $\mathbf{A} = \mathbf{L}\mathbf{U}$ with

$$\mathbf{L} = \begin{pmatrix} 1 & 0 \\ \ell & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} u & v \\ 0 & w \end{pmatrix}.$$

Then

$$\mathbf{L}\mathbf{U} = \begin{pmatrix} u & v \\ \ell u & \ell v + w \end{pmatrix}.$$

Comparing $(1,1)$ -entries with $\mathbf{A}$ gives $u = 0$. Comparing $(2,1)$ -entries gives $\ell u = 1$. But $\ell u = \ell \cdot 0 = 0$ by Theorem 0.7.1, and $0 \neq 1$ in a field. This contradiction shows that no LU factorization exists.

What failed is visible: the $(1,1)$ -entry of any product $\mathbf{L}\mathbf{U}$ is u_{11} alone, so a zero there forces the whole first column of $\mathbf{L}\mathbf{U}$ to vanish. In elimination terms, the first pivot is 0 while the entry below it is not, and the only cure is to **swap** rows, which “subtract a multiple of a higher row from a lower row” cannot do. We return to this with permutation matrices.

Why “unit”. Without the 1’s on the diagonal of $\mathbf{L}$, uniqueness would fail even for invertible matrices. If $\mathbf{A} = \mathbf{L}\mathbf{U}$ and $\mathbf{D}$ is any invertible diagonal matrix, then also $\mathbf{A} = (\mathbf{L}\mathbf{D})(\mathbf{D}^{-1}\mathbf{U})$, where $\mathbf{L}\mathbf{D}$ is still lower triangular and $\mathbf{D}^{-1}\mathbf{U}$ still upper triangular. When $\mathbf{L}$ is invertible and $\mathbf{D} \neq \mathbf{I}$, this is a different pair, and over any field with more than two elements such a $\mathbf{D}$ exists. For instance $\begin{pmatrix} 2 \\ \end{pmatrix} = \begin{pmatrix} 1 \\ \end{pmatrix} \begin{pmatrix} 2 \\ \end{pmatrix} = \begin{pmatrix} 2 \\ \end{pmatrix} \begin{pmatrix} 1 \\ \end{pmatrix} = \begin{pmatrix} 4 \\ \end{pmatrix} \begin{pmatrix} 1/2 \\ \end{pmatrix}$ over $\mathbb{Q}$. Fixing the diagonal of $\mathbf{L}$ removes this freedom and matches what elimination does: it never rescales a row.

Warning. Invertible does not mean “has an LU factorization”. The matrix of [Example \(An invertible matrix with no LU factorization\)](#) is invertible and has none. Conversely, a singular matrix can have one: $\mathbf{O} = \mathbf{I} \cdot \mathbf{O}$. Existence of an LU factorization is a property of the **order of the rows**, not of invertibility.

3.6.3 Unit lower triangular matrices

Elimination builds $\mathbf{L}$ out of elementary matrices, and the uniqueness proof below divides one factorization by another. Both need the same facts: products and inverses of unit lower triangular matrices stay unit lower triangular. We also need to know exactly when a triangular matrix is invertible.

The key observation is a statement about where the nonzero entries of a product can sit.

Lemma 3.6.1. Unit lower triangular matrices are closed

Let $\mathbf{L}, \mathbf{L}' \in M_n(F)$ be unit lower triangular. Then:

- (a) $\mathbf{L}\mathbf{L}'$ is unit lower triangular;
- (b) $\mathbf{L}$ is invertible, and $\mathbf{L}^{-1}$ is unit lower triangular.

Idea. Write $\mathbf{L} = \mathbf{I} + \mathbf{N}$, where $\mathbf{N}$ has nonzero entries only strictly below the diagonal. Multiplying two such matrices pushes the nonzero entries **further** below the diagonal, so $\mathbf{N}^n = \mathbf{O}$. Then the geometric series $1/(1+t) = 1 - t + t^2 - \cdots$ stops after n terms and gives the inverse exactly.

Proof. Say that $\mathbf{X} \in M_n(F)$ is **p -low** if $x_{ij} = 0$ whenever $i < j + p$. So 0-low means lower triangular, and 1-low means zero on and above the diagonal.

Claim. If $\mathbf{X}$ is p -low and $\mathbf{Y}$ is q -low, then $\mathbf{XY}$ is $(p+q)$ -low. A matrix that is n -low is the zero matrix.

Proof. By Definition 0.9.9, $(\mathbf{XY})_{ij} = \sum_{l=1}^n x_{il}y_{lj}$. A term can be nonzero only if $i \geq l + p$ and $l \geq j + q$, which together give $i \geq j + p + q$. So if $i < j + p + q$, every term is 0 and $(\mathbf{XY})_{ij} = 0$. For the second statement, every pair of indices satisfies $i \leq n < 1 + n \leq j + n$, so every entry of an n -low matrix is 0. ■

Write $\mathbf{L} = \mathbf{I} + \mathbf{N}$ and $\mathbf{L}' = \mathbf{I} + \mathbf{N}'$. Since $\mathbf{L}$ and $\mathbf{L}'$ are unit lower triangular, $\mathbf{N}$ and $\mathbf{N}'$ are 1-low.

For (a), $\mathbf{L}\mathbf{L}' = \mathbf{I} + \mathbf{N} + \mathbf{N}' + \mathbf{N}\mathbf{N}'$ by Theorem 0.9.4. Here $\mathbf{N} + \mathbf{N}'$ is 1-low, since a sum of entries that are both 0 is 0, and $\mathbf{N}\mathbf{N}'$ is 2-low, hence 1-low, by the Claim. So $\mathbf{L}\mathbf{L}' - \mathbf{I}$ is 1-low, which says that $\mathbf{L}\mathbf{L}'$ is unit lower triangular.

For (b), let $\mathbf{K} = \mathbf{I} - \mathbf{N} + \mathbf{N}^2 - \cdots + (-1)^{n-1}\mathbf{N}^{n-1}$. Multiplying out, the sum telescopes:

$$(\mathbf{I} + \mathbf{N})\mathbf{K} = \mathbf{I} + (-1)^{n-1}\mathbf{N}^n = \mathbf{I},$$

where $\mathbf{N}^n = \mathbf{O}$ because $\mathbf{N}^n$ is n -low by the Claim applied $n-1$ times. The same computation gives $\mathbf{K}(\mathbf{I} + \mathbf{N}) = \mathbf{I}$, since $\mathbf{N}$ commutes with its own powers. Hence $\mathbf{L}$ is invertible with $\mathbf{L}^{-1} = \mathbf{K}$. Finally $\mathbf{K} - \mathbf{I} = \sum_{p=1}^{n-1} (-1)^p \mathbf{N}^p$ is a sum of 1-low matrices, by the Claim, so $\mathbf{K}$ is unit lower triangular. This proves the lemma. ■

Transposing everything, the same statements hold for **unit upper triangular** matrices (upper triangular with 1's on the diagonal).

The next lemma decides invertibility of a triangular matrix from its diagonal, without any computation beyond back substitution.

Lemma 3.6.2. Invertibility of a triangular matrix

An upper triangular matrix $\mathbf{U} \in M_n(F)$ is invertible if and only if **every** diagonal entry $u_{11}, \dots, u_{nn}$ is nonzero.

Proof. ($\Leftarrow$) Suppose every $u_{kk} \neq 0$, and let $\mathbf{U}\mathbf{x} = \mathbf{0}$. We show $x_k = 0$ for $k = n, n-1, \dots, 1$ in turn. Suppose $x_{k+1} = \dots = x_n = 0$ (nothing is assumed when $k = n$). Row k of $\mathbf{U}\mathbf{x} = \mathbf{0}$ reads $u_{kk}x_k + \sum_{j>k} u_{kj}x_j = 0$, because $u_{kj} = 0$ for $j < k$. The sum vanishes, so $u_{kk}x_k = 0$, and $x_k = 0$ since $u_{kk} \neq 0$. Hence $\mathbf{x} = \mathbf{0}$ is the only solution of $\mathbf{U}\mathbf{x} = \mathbf{0}$, and $\mathbf{U}$ is invertible by Theorem 3.4.4.

($\Rightarrow$) We prove the contrapositive. Suppose $u_{kk} = 0$ for some k . For each $j \leq k$, the entries of column j in rows $i \geq k$ are zero: if $i > j$ this is upper triangularity, and the only remaining case $i = j = k$ is $u_{kk} = 0$. So the first k columns of $\mathbf{U}$ lie in $W = \text{span}(\mathbf{e}_1, \dots, \mathbf{e}_{k-1})$. The list $(\mathbf{e}_1, \dots, \mathbf{e}_{k-1})$ is independent and spans W , so $\dim W = k-1$ (for $k=1$, $W = \{\mathbf{0}\}$ has dimension 0). By Theorem 1.6.3 (a), these k columns are linearly dependent, so the columns of $\mathbf{U}$ are dependent, and $\mathbf{U}$ is not invertible by Theorem 3.4.4. This proves the lemma. ■

3.6.4 Elimination without row swaps

To prove that elimination always produces an LU factorization, we first need to say exactly which elimination we mean. It is Gaussian elimination restricted to one kind of operation: subtracting a multiple of a row from a row **below** it.

Algorithm. Elimination without row swaps. Input: $\mathbf{A} \in M_n(F)$. Put $\mathbf{A}^{(0)} = \mathbf{A}$. For $k = 1, 2, \dots, n-1$, write $\mathbf{A}^{(k-1)} = (a_{ij})$ and do the following.

- If $a_{kk} \neq 0$, set $\ell_{ik} = a_{ik}/a_{kk}$ for $i = k+1, \dots, n$.
- If $a_{kk} = 0$ and $a_{ik} = 0$ for **every** $i > k$, set $\ell_{ik} = 0$ for $i = k+1, \dots, n$.
- If $a_{kk} = 0$ but $a_{ik} \neq 0$ for **some** $i > k$, **stop**: a row swap is needed.

If we did not stop, obtain $\mathbf{A}^{(k)}$ from $\mathbf{A}^{(k-1)}$ by replacing row i with $(\text{row } i) - \ell_{ik} \cdot (\text{row } k)$, for each $i = k+1, \dots, n$.

If the loop finishes without stopping, we say **elimination without row swaps succeeds** on $\mathbf{A}$. The numbers ℓ_{ik} ($i > k$) are its **multipliers**, and $\mathbf{U} = \mathbf{A}^{(n-1)}$ is its **result**.

In both nonstopping cases the new (i, k) -entry is $a_{ik} - \ell_{ik}a_{kk} = 0$: in the first case by the choice of ℓ_{ik} , in the second because both terms are already 0. The multipliers are well defined: row k is not changed during step k , so it does not matter in which order the rows $i > k$ are updated.

The proof will write each step as a matrix. By Definition 3.4.1, the elementary matrix of the operation $R_i \rightarrow R_i + cR_k$ ($i \neq k$) is $\mathbf{I} + c\mathbf{E}_{ik}$, where $\mathbf{E}_{ik}$ is the matrix unit with 1 in entry (i, k) . We will write the matrix unit as a product of a column and a row,

$$\mathbf{E}_{ik} = \mathbf{e}_i \mathbf{e}_k^\top,$$

which holds because $\mathbf{e}_i \mathbf{e}_k^\top$ has (p, q) -entry $(\mathbf{e}_i)_p (\mathbf{e}_k)_q$, equal to 1 if $p = i$ and $q = k$, and 0 otherwise. We will use one more rule constantly: for a vector $\mathbf{v} \in F^n$, the 1×1 matrix $\mathbf{e}_k^\top \mathbf{v}$ is the k -th entry v_k , which we treat as a scalar. In particular $\mathbf{e}_k^\top \mathbf{e}_j = 0$ for $j \neq k$.

For $1 \leq k \leq n$ and a vector $\mathbf{v} \in F^n$ whose first k entries are zero, put

$$\mathbf{G}_k(\mathbf{v}) = \mathbf{I} + \mathbf{v} \mathbf{e}_k^\top.$$

Its only nonidentity entries are v_i in position (i, k) , for $i > k$: it is the identity with the column vector $\mathbf{v}$ written into column k below the diagonal. So $\mathbf{G}_k(\mathbf{v})$ is unit lower triangular.

Lemma 3.6.3. Multiplier matrices

Let $1 \leq k \leq n$, and let $\mathbf{v}, \mathbf{v}_1, \dots, \mathbf{v}_m \in F^n$ be vectors such that the first k entries of $\mathbf{v}$ are zero and, for each j , the first j entries of $\mathbf{v}_j$ are zero ($m \leq n$). Then:

- (a) $\mathbf{G}_k(\mathbf{v})$ is the product, in any order, of the elementary matrices $\mathbf{I} + v_i \mathbf{E}_{ik}$ for $i = k+1, \dots, n$;
- (b) $\mathbf{G}_k(\mathbf{v})\mathbf{G}_k(-\mathbf{v}) = \mathbf{G}_k(-\mathbf{v})\mathbf{G}_k(\mathbf{v}) = \mathbf{I}$;
- (c) $\mathbf{G}_1(\mathbf{v}_1)\mathbf{G}_2(\mathbf{v}_2) \cdots \mathbf{G}_m(\mathbf{v}_m) = \mathbf{I} + \mathbf{v}_1 \mathbf{e}_1^\top + \mathbf{v}_2 \mathbf{e}_2^\top + \cdots + \mathbf{v}_m \mathbf{e}_m^\top$.

Proof. For (a), let $i, j > k$. By Theorem 0.9.4,

$$(\mathbf{I} + a \mathbf{e}_i \mathbf{e}_k^\top)(\mathbf{I} + b \mathbf{e}_j \mathbf{e}_k^\top) = \mathbf{I} + a \mathbf{e}_i \mathbf{e}_k^\top + b \mathbf{e}_j \mathbf{e}_k^\top + ab \mathbf{e}_i (\mathbf{e}_k^\top \mathbf{e}_j) \mathbf{e}_k^\top,$$

and the last term is zero because $\mathbf{e}_k^\top \mathbf{e}_j = 0$ for $j \neq k$. Repeating this, one factor at a time, the product of the $\mathbf{I} + v_i \mathbf{E}_{ik} = \mathbf{I} + v_i \mathbf{e}_i \mathbf{e}_k^\top$ in any order is $\mathbf{I} + \sum_{i>k} v_i \mathbf{e}_i \mathbf{e}_k^\top = \mathbf{I} + \mathbf{v} \mathbf{e}_k^\top$, since $\mathbf{v} = \sum_{i>k} v_i \mathbf{e}_i$.

For (b), $(\mathbf{I} + \mathbf{v} \mathbf{e}_k^\top)(\mathbf{I} - \mathbf{v} \mathbf{e}_k^\top) = \mathbf{I} - \mathbf{v}(\mathbf{e}_k^\top \mathbf{v}) \mathbf{e}_k^\top = \mathbf{I}$, because $\mathbf{e}_k^\top \mathbf{v} = v_k = 0$. Swapping the signs gives the other order.

For (c), we use induction on m ; the case $m = 1$ is the definition. Suppose the formula holds for $m - 1$. Then

$$\left(\mathbf{I} + \sum_{j<m} \mathbf{v}_j \mathbf{e}_j^\top \right) (\mathbf{I} + \mathbf{v}_m \mathbf{e}_m^\top) = \mathbf{I} + \sum_{j<m} \mathbf{v}_j \mathbf{e}_j^\top + \mathbf{v}_m \mathbf{e}_m^\top + \sum_{j<m} \mathbf{v}_j (\mathbf{e}_j^\top \mathbf{v}_m) \mathbf{e}_m^\top.$$

For $j < m$, $\mathbf{e}_j^\top \mathbf{v}_m$ is entry j of $\mathbf{v}_m$, which is zero because the first m entries of $\mathbf{v}_m$ are zero. So the last sum vanishes, and the formula holds for m . This proves the lemma. ■

Part (c) is the heart of the matter: when the factors come in the order $\mathbf{G}_1, \mathbf{G}_2, \dots$, they **do not interfere**, and each vector $\mathbf{v}_j$ lands untouched in column j .

Theorem 3.6.4. LU from elimination

Let $\mathbf{A} \in M_n(F)$, and suppose elimination without row swaps succeeds on $\mathbf{A}$, with multipliers ℓ_{ik} and result $\mathbf{U}$. Let $\mathbf{L}$ be the unit lower triangular matrix whose (i, k) -entry is ℓ_{ik} for every $i > k$. Then $\mathbf{U}$ is upper triangular and

$$\mathbf{A} = \mathbf{L}\mathbf{U}.$$

In particular $\mathbf{A}$ has an LU factorization.

Idea. Each step of elimination is left multiplication by a product of elementary matrices, which collapses to one matrix $\mathbf{G}_k(-\ell_k)$. So $\mathbf{U} = \mathbf{G}_{n-1}(-\ell_{n-1}) \cdots \mathbf{G}_1(-\ell_1) \mathbf{A}$. To get $\mathbf{A}$ back, undo the steps in reverse order. Their inverses $\mathbf{G}_k(\ell_k)$ then appear in the order $\mathbf{G}_1, \mathbf{G}_2, \dots$, which is exactly the order in which, by Lemma 3.6.3 (c), the multipliers drop into place.

Proof. Let $\mathbf{A}^{(0)}, \dots, \mathbf{A}^{(n-1)} = \mathbf{U}$ be the matrices produced by the algorithm. For each k , let $\ell_k = \sum_{i>k} \ell_{ik} \mathbf{e}_i \in F^n$, a vector whose first k entries are zero.

Step 1: $\mathbf{U}$ is upper triangular. We show by induction on k that in $\mathbf{A}^{(k)}$ every entry below the diagonal in columns $1, \dots, k$ is zero. For $k = 0$ there is nothing to prove. Suppose it holds for $\mathbf{A}^{(k-1)}$. Step k changes only rows $i > k$, replacing row i by row i minus a multiple of row k . In columns $j < k$, the entry of row k lies below the diagonal ($k > j$), so it is 0 by the induction hypothesis, and the entries of row i in these columns stay 0. In column k , the new entries below the diagonal are 0, as noted after the algorithm. This completes the induction, and for $k = n - 1$ it says that $\mathbf{U}$ is upper triangular.

Step 2: each step is a matrix. Step k performs the operations “add $-\ell_{ik}$ times row k to row i ” for $i = k + 1, \dots, n$. By Theorem 3.4.1, applying them one after another multiplies $\mathbf{A}^{(k-1)}$ on the left by the corresponding elementary matrices $\mathbf{I} - \ell_{ik} \mathbf{E}_{ik}$. By Lemma 3.6.3 (a), their product is $\mathbf{G}_k(-\ell_k)$. Hence $\mathbf{A}^{(k)} = \mathbf{G}_k(-\ell_k) \mathbf{A}^{(k-1)}$, and therefore

$$\mathbf{U} = \mathbf{G}_{n-1}(-\ell_{n-1}) \cdots \mathbf{G}_2(-\ell_2) \mathbf{G}_1(-\ell_1) \mathbf{A}.$$

Step 3: undo the steps. Multiply both sides on the left by $\mathbf{G}_{n-1}(\ell_{n-1})$, then by $\mathbf{G}_{n-2}(\ell_{n-2})$, and so on down to $\mathbf{G}_1(\ell_1)$. By Lemma 3.6.3 (b), each one cancels the factor next to it, and we get

$$\mathbf{G}_1(\ell_1) \mathbf{G}_2(\ell_2) \cdots \mathbf{G}_{n-1}(\ell_{n-1}) \mathbf{U} = \mathbf{A}.$$

By Lemma 3.6.3 (c), the product on the left of $\mathbf{U}$ is $\mathbf{I} + \sum_{k=1}^{n-1} \ell_k \mathbf{e}_k^\top$. Its (i, k) -entry for $i > k$ is ℓ_{ik} ; its diagonal entries are 1; its entries above the diagonal are 0. That is, it is $\mathbf{L}$. Hence $\mathbf{A} = \mathbf{L}\mathbf{U}$ with $\mathbf{L}$ unit lower triangular and $\mathbf{U}$ upper triangular, as claimed. ■

So the multipliers are not something to be remembered separately: they **are** the lower triangular factor. When a computer performs elimination it stores each ℓ_{ik} in the slot (i, k) that it has just cleared, and at the end the one array holds $\mathbf{L}$ below the diagonal and $\mathbf{U}$ on and above it.

Warning. The matrix that performs elimination, $\mathbf{L}^{-1}$, is not $\mathbf{L}$ with its multipliers negated. In Example (An LU factorization from elimination) the operations multiply $\mathbf{A}$ on the left by

$$\mathbf{G}_2(-\ell_2) \mathbf{G}_1(-\ell_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 7 & -3 & 1 \end{pmatrix},$$

and the entry $7 = 1 + (-3)(-2)$ mixes two multipliers. This matrix is $\mathbf{L}^{-1}$, by Step 3 of the proof. It is $\mathbf{L}$, whose factors come in the order $\mathbf{G}_1, \mathbf{G}_2$, that has the multipliers cleanly in

place.

Quick check. Find the LU factorization of $\mathbf{A} = \begin{pmatrix} 2 & 4 \\ 6 & 5 \end{pmatrix}$ over $\mathbb{Q}$ by elimination, and check it by multiplying.

Solution. The pivot is 2 and the multiplier is $\ell_{21} = 6/2 = 3$. Subtracting 3 times row 1 from row 2 gives $\mathbf{U} = \begin{pmatrix} 2 & 4 \\ 0 & 5 - 12 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 0 & -7 \end{pmatrix}$, and $\mathbf{L} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$. Check: $\mathbf{LU} = \begin{pmatrix} 2 & 4 \\ 6 & 12 - 7 \end{pmatrix} = \mathbf{A}$.

3.6.5 Solving with an LU factorization

Once $\mathbf{A} = \mathbf{LU}$ is known, each system $\mathbf{Ax} = \mathbf{b}$ costs two substitutions.

Example. Two right-hand sides, one factorization

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 & -1 \\ 4 & 5 & -1 \\ -2 & 8 & 2 \end{pmatrix}$ with the factorization $\mathbf{A} = \mathbf{LU}$ of [Example \(An LU factorization from elimination\)](#). Solve $\mathbf{Ax} = \mathbf{b}$ for

$$\mathbf{b} = \begin{pmatrix} -1 \\ -3 \\ -6 \end{pmatrix} \quad \text{and} \quad \mathbf{b}' = \begin{pmatrix} 3 \\ 7 \\ -2 \end{pmatrix}.$$

Solution. First right-hand side. Forward substitution in $\mathbf{Ly} = \mathbf{b}$:

$$\begin{aligned} y_1 &= -1, \\ 2y_1 + y_2 &= -3 \Rightarrow y_2 = -1, \\ -y_1 + 3y_2 + y_3 &= -6 \Rightarrow y_3 = -6 - 1 + 3 = -4. \end{aligned}$$

Back substitution in $\mathbf{Ux} = \mathbf{y}$:

$$\begin{aligned} -2x_3 &= -4 \Rightarrow x_3 = 2, \\ 3x_2 + x_3 &= -1 \Rightarrow x_2 = -1, \\ 2x_1 + x_2 - x_3 &= -1 \Rightarrow x_1 = 1. \end{aligned}$$

So $\mathbf{x} = (1, -1, 2)$. Check: $\mathbf{Ax} = (2 - 1 - 2, 4 - 5 - 2, -2 - 8 + 4) = (-1, -3, -6) = \mathbf{b}$.

Second right-hand side. $\mathbf{Ly} = \mathbf{b}'$: $y_1 = 3$, $y_2 = 7 - 2 \cdot 3 = 1$, $y_3 = -2 + y_1 - 3y_2 = -2 + 3 - 3 = -2$. Then $\mathbf{Ux} = \mathbf{y}$: $-2x_3 = -2$, so $x_3 = 1$; $3x_2 + 1 = 1$, so $x_2 = 0$; $2x_1 + 0 - 1 = 3$, so $x_1 = 2$. So $\mathbf{x}' = (2, 0, 1)$, and indeed $\mathbf{Ax}' = (4 + 0 - 1, 8 + 0 - 1, -4 + 0 + 2) = (3, 7, -2)$.

Since the diagonal of $\mathbf{U}$ is $2, 3, -2$, all nonzero, $\mathbf{U}$ is invertible by [Lemma 3.6.2](#), $\mathbf{L}$ is invertible by [Lemma 3.6.1](#), and so $\mathbf{A} = \mathbf{L}\mathbf{U}$ is invertible. Each solution found is therefore the only one.

How much is saved? Count the arithmetic operations $(+, -, \times, \div)$ for an $n \times n$ matrix. At step k of elimination there are $n - k$ rows below the pivot. Each needs one division for its multiplier, then $n - k$ multiplications and $n - k$ subtractions in the columns to the right of column k . With $j = n - k$, the total is

$$\sum_{j=1}^{n-1} j(2j+1) = \frac{n(n-1)(4n+1)}{6} \approx \frac{2}{3}n^3.$$

Forward substitution with a unit lower triangular $\mathbf{L}$ uses $n(n-1)$ operations, and back substitution n^2 , so a pair of triangular solves costs $2n^2 - n$. For $n = 1000$ that is about 6.7×10^8 operations for the factorization, done once, against about 2×10^6 for each new right-hand side.

3.6.6 Uniqueness

The factorization of [Example \(An LU factorization from elimination\)](#) came from one particular procedure. Could another procedure give a different pair $(\mathbf{L}, \mathbf{U})$? For invertible matrices, no.

Theorem 3.6.5. Uniqueness of the LU factorization

Let $\mathbf{A} \in M_n(F)$ be invertible. If $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1$ and $\mathbf{A} = \mathbf{L}_2\mathbf{U}_2$ are LU factorizations, then $\mathbf{L}_1 = \mathbf{L}_2$ and $\mathbf{U}_1 = \mathbf{U}_2$.

Idea. Take the difference in multiplicative form: from $\mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$ we get $(\mathbf{L}_2^{-1}\mathbf{L}_1)\mathbf{U}_1 = \mathbf{U}_2$. The matrix $\mathbf{M} = \mathbf{L}_2^{-1}\mathbf{L}_1$ is unit lower triangular, and multiplying $\mathbf{U}_1$ by it must leave an upper triangular matrix. Clearing below the diagonal column by column, the nonzero pivots of $\mathbf{U}_1$ force $\mathbf{M}$ to have nothing below its diagonal. Invertibility of $\mathbf{A}$ is what provides those nonzero pivots.

Proof. By [Lemma 3.6.1](#) (b), $\mathbf{L}_1$ and $\mathbf{L}_2$ are invertible. Hence $\mathbf{U}_1 = \mathbf{L}_1^{-1}\mathbf{A}$ is a product of invertible matrices, so it is invertible by Theorem 0.9.7, and by [Lemma 3.6.2](#) its diagonal entries $u_{11}, \dots, u_{nn}$ are all nonzero.

Let $\mathbf{M} = \mathbf{L}_2^{-1}\mathbf{L}_1$. By [Lemma 3.6.1](#), $\mathbf{M} = (m_{ij})$ is unit lower triangular. Multiplying $\mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$ on the left by $\mathbf{L}_2^{-1}$ gives $\mathbf{M}\mathbf{U}_1 = \mathbf{U}_2$, so $\mathbf{M}\mathbf{U}_1$ is upper triangular. We show $m_{ij} = 0$ for all $i > j$, by induction on the column j .

Fix j , and suppose $m_{ik} = 0$ whenever $k < j$ and $i > k$. Let $i > j$. Since $\mathbf{M}\mathbf{U}_1$ is upper triangular and $i > j$,

$$0 = (\mathbf{M}\mathbf{U}_1)_{ij} = \sum_{k=1}^n m_{ik}u_{kj} = \sum_{k \leq j} m_{ik}u_{kj} = m_{ij}u_{jj},$$

where the third equality uses $u_{kj} = 0$ for $k > j$, and the last uses $m_{ik} = 0$ for $k < j$ (here $i > j > k$). Since $u_{jj} \neq 0$, we get $m_{ij} = 0$. This completes the induction, so $\mathbf{M}$ is lower triangular with 1's on the diagonal and nothing below it: $\mathbf{M} = \mathbf{I}$.

Hence $\mathbf{L}_2^{-1}\mathbf{L}_1 = \mathbf{I}$, and multiplying by $\mathbf{L}_2$ gives $\mathbf{L}_1 = \mathbf{L}_2$. Then $\mathbf{U}_1 = \mathbf{L}_1^{-1}\mathbf{A} = \mathbf{L}_2^{-1}\mathbf{A} = \mathbf{U}_2$.

This proves the theorem. ■

Two consequences are worth saying. The LU factorization of an invertible matrix does not depend on how it was found, so it is legitimate to speak of **the** factors **L** and **U**. And if elimination without row swaps succeeds on an invertible **A**, then every pivot is nonzero: row k is not changed after step k , so the pivot of step k is u_{kk} , which is nonzero by the first paragraph of the proof. So the second case of the algorithm (a zero pivot over a zero column) never occurs.

Remark. The hypothesis “invertible” cannot be dropped. For every $c \in F$,

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

since the first column of the right factor is zero. Taking $c = 0$ and $c = 1$ gives two different LU factorizations. The proof breaks exactly at “ $u_{jj} \neq 0$ ”, here for $j = 1$.

3.6.7 Row swaps and the PLU factorization

Example (An invertible matrix with no LU factorization) showed that some matrices, even invertible ones, have no LU factorization, because elimination meets a zero pivot with a nonzero entry below it. The cure in Gaussian elimination is to swap rows. We now show that the swaps can all be collected in front, as a single matrix that rearranges the rows of **A** **before** elimination starts. After that rearrangement, no swaps are needed.

Matrices that rearrange rows were named in Definition 3.4.2: $\mathbf{P}_\sigma$ has j -th column $\mathbf{e}_{\sigma(j)}$, a product of permutation matrices is one, $\mathbf{P}_\sigma^{-1} = \mathbf{P}_\sigma^\top$, and a transposition gives the row-swap elementary matrix (Lemma 3.4.7). One more fact about them is needed here and nowhere else, because it is about the Gauss transforms of this section: a swap that leaves the first k coordinates alone can be pushed through $\mathbf{G}_k$.

Lemma 3.6.6. Pushing a Swap Past a Gauss Transform

Let $1 \leq k < r < s \leq n$, let $\tau = (r \ s)$, and let $\mathbf{v} \in F^n$ have its first k entries zero. Then $\mathbf{P}_\tau \mathbf{v}$ has its first k entries zero, and

$$\mathbf{P}_\tau \mathbf{G}_k(\mathbf{v}) \mathbf{P}_\tau = \mathbf{G}_k(\mathbf{P}_\tau \mathbf{v}).$$

Proof. Write $\mathbf{P} = \mathbf{P}_\tau$. Its effect on a vector is to exchange entries r and s , which are both beyond position k , so $\mathbf{P}\mathbf{v}$ still has its first k entries zero. Since $k \neq r, s$, we have $\mathbf{P}\mathbf{e}_k = \mathbf{e}_{\tau(k)} = \mathbf{e}_k$. By Lemma 3.4.7 (b) and (c), $\mathbf{P}^\top = \mathbf{P}^{-1} = \mathbf{P}$, so $\mathbf{e}_k^\top \mathbf{P} = \mathbf{e}_k^\top \mathbf{P}^\top = (\mathbf{P}\mathbf{e}_k)^\top = \mathbf{e}_k^\top$ by Theorem 0.9.5. Therefore

$$\mathbf{P}(\mathbf{I} + \mathbf{v} \mathbf{e}_k^\top) \mathbf{P} = \mathbf{P}^2 + (\mathbf{P}\mathbf{v})(\mathbf{e}_k^\top \mathbf{P}) = \mathbf{I} + (\mathbf{P}\mathbf{v}) \mathbf{e}_k^\top = \mathbf{G}_k(\mathbf{P}\mathbf{v}),$$

using $\mathbf{P}^2 = \mathbf{I}$ from Lemma 3.4.7 (c). This proves the lemma. ■

Definition 3.6.2. PLU factorization

Let $\mathbf{A} \in M_n(F)$. A **PLU factorization** of $\mathbf{A}$ is a triple $(\mathbf{P}, \mathbf{L}, \mathbf{U})$ of matrices in $M_n(F)$ such that

$$\begin{aligned} \mathbf{PA} &= \mathbf{LU}, & \mathbf{P} &\text{ is a permutation matrix,} \\ \mathbf{L} &\text{ is unit lower triangular,} & \mathbf{U} &\text{ is upper triangular.} \end{aligned}$$

In words: after rearranging the rows of $\mathbf{A}$ by $\mathbf{P}$, the matrix has an LU factorization. By Lemma 3.4.7 (b) this can also be written $\mathbf{A} = \mathbf{P}^\top \mathbf{L} \mathbf{U}$, which is why some books write the factorization as $\mathbf{A} = \mathbf{PLU}$ with their $\mathbf{P}$ our $\mathbf{P}^\top$. For example, $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has the PLU factorization $(\mathbf{A}, \mathbf{I}, \mathbf{I})$, since $\mathbf{AA} = \mathbf{I} = \mathbf{I} \cdot \mathbf{I}$. An LU factorization is the special case $\mathbf{P} = \mathbf{I}$.

To produce $\mathbf{P}$, run elimination and allow a swap whenever a zero pivot has a nonzero entry below it.

Algorithm. Elimination with row swaps. Input: $\mathbf{A} \in M_n(F)$. Put $\mathbf{A}^{(0)} = \mathbf{A}$. For $k = 1, 2, \dots, n-1$, with $\mathbf{A}^{(k-1)} = (a_{ij})$:

- If $a_{ik} = 0$ for **every** $i \geq k$, set $\mathbf{P}_k = \mathbf{I}$ and all $\ell_{ik} = 0$, and let $\mathbf{A}^{(k)} = \mathbf{A}^{(k-1)}$.
- Otherwise choose some $r_k \geq k$ with $a_{r_k k} \neq 0$. Let $\mathbf{P}_k$ be the matrix swapping rows k and r_k (so $\mathbf{P}_k = \mathbf{I}$ if $r_k = k$), and let $\mathbf{B} = \mathbf{P}_k \mathbf{A}^{(k-1)}$. Set $\ell_{ik} = b_{ik}/b_{kk}$ for $i > k$, and obtain $\mathbf{A}^{(k)}$ from $\mathbf{B}$ by replacing row i with (row i) $-\ell_{ik} \cdot$ (row k) for each $i > k$.

This version never stops. Here is the result it delivers.

Theorem 3.6.7. Existence of a PLU factorization

Every $\mathbf{A} \in M_n(F)$ has a PLU factorization.

Idea. Step roadmap. ① As before, each step is a matrix: $\mathbf{A}^{(k)} = \mathbf{G}_k(-\ell_k) \mathbf{P}_k \mathbf{A}^{(k-1)}$, and the final matrix $\mathbf{U}$ is upper triangular. ② The swaps are tangled with the multiplier matrices. A swap of two rows below row k can be pushed to the right past $\mathbf{G}_k$ at the price of swapping the entries of the multiplier vector: this is Lemma 3.4.7 (d). ③ Once all swaps sit next to $\mathbf{A}$, the rest is the LU proof word for word. In practice, ② says: when you swap two rows, also swap the multipliers already stored in those rows.

Proof. Run elimination with row swaps on $\mathbf{A}$, producing $\mathbf{A}^{(0)}, \dots, \mathbf{A}^{(n-1)} = \mathbf{U}$, swap matrices $\mathbf{P}_1, \dots, \mathbf{P}_{n-1}$, and multipliers ℓ_{ik} . Let $\ell_k = \sum_{i>k} \ell_{ik} \mathbf{e}_i$, whose first k entries are zero.

Step 1. Exactly as in Step 1 of the proof of Theorem 3.6.4, every entry of $\mathbf{A}^{(k)}$ below the diagonal in columns $1, \dots, k$ is zero. The only new point is the swap in step k : it exchanges rows k and $r_k \geq k$, whose entries in columns $j < k$ are below the diagonal and hence zero, so the swap does not disturb these columns. After the swap the pivot b_{kk} is nonzero, and the new entries below it are $b_{ik} - \ell_{ik} b_{kk} = 0$. Hence $\mathbf{U}$ is upper triangular. By Lemma 3.4.7

(c), Theorem 3.4.1 and Lemma 3.6.3 (a), as in Step 2 of that proof,

$$\mathbf{A}^{(k)} = \mathbf{G}_k(-\ell_k) \mathbf{P}_k \mathbf{A}^{(k-1)} \quad (k = 1, \dots, n-1).$$

This also holds in the first case of the algorithm, where $\mathbf{P}_k = \mathbf{I}$ and $\mathbf{G}_k(\mathbf{0}) = \mathbf{I}$.

Step 2. For $1 \leq k \leq m \leq n-1$ put $\ell_k^{(m)} = \mathbf{P}_m \mathbf{P}_{m-1} \cdots \mathbf{P}_{k+1} \ell_k$, with $\ell_m^{(m)} = \ell_m$. We claim that the first k entries of $\ell_k^{(m)}$ are zero and that

$$\mathbf{G}_m(-\ell_m) \mathbf{P}_m \cdots \mathbf{G}_1(-\ell_1) \mathbf{P}_1 = \mathbf{G}_m(-\ell_m^{(m)}) \cdots \mathbf{G}_1(-\ell_1^{(m)}) \mathbf{P}_m \cdots \mathbf{P}_1.$$

For $m = 1$ both sides are $\mathbf{G}_1(-\ell_1) \mathbf{P}_1$. Suppose the claim holds for m , and multiply it on the left by $\mathbf{G}_{m+1}(-\ell_{m+1}) \mathbf{P}_{m+1}$. The matrix $\mathbf{P}_{m+1}$ is $\mathbf{I}$ or swaps rows $m+1$ and $r_{m+1} > m+1$, both beyond every $k \leq m$. So by Lemma 3.4.7 (c) and Lemma 3.6.6, for each $k \leq m$ and each vector $\mathbf{v}$ with its first k entries zero,

$$\mathbf{P}_{m+1} \mathbf{G}_k(\mathbf{v}) = \mathbf{P}_{m+1} \mathbf{G}_k(\mathbf{v}) \mathbf{P}_{m+1} \mathbf{P}_{m+1} = \mathbf{G}_k(\mathbf{P}_{m+1} \mathbf{v}) \mathbf{P}_{m+1},$$

and $\mathbf{P}_{m+1} \mathbf{v}$ again has its first k entries zero. Applying this for $k = m, m-1, \dots, 1$ moves $\mathbf{P}_{m+1}$ to the right past every factor $\mathbf{G}_k$, replacing $\ell_k^{(m)}$ by $\mathbf{P}_{m+1} \ell_k^{(m)} = \ell_k^{(m+1)}$ (note $\mathbf{P}_{m+1}(-\mathbf{v}) = -\mathbf{P}_{m+1} \mathbf{v}$). This is the claim for $m+1$.

Step 3. By Step 1, $\mathbf{U} = \mathbf{G}_{n-1}(-\ell_{n-1}) \mathbf{P}_{n-1} \cdots \mathbf{G}_1(-\ell_1) \mathbf{P}_1 \mathbf{A}$. Write $\ell'_k = \ell_k^{(n-1)}$ and $\mathbf{P} = \mathbf{P}_{n-1} \cdots \mathbf{P}_1$. By Step 2,

$$\mathbf{U} = \mathbf{G}_{n-1}(-\ell'_{n-1}) \cdots \mathbf{G}_1(-\ell'_1) \mathbf{P} \mathbf{A}.$$

$\mathbf{P}$ is a permutation matrix by Lemma 3.4.7 (a) and (c). Multiplying on the left by $\mathbf{G}_{n-1}(\ell'_{n-1})$, then $\mathbf{G}_{n-2}(\ell'_{n-2})$, and so on, and using Lemma 3.6.3 (b) and (c) as in Step 3 of the proof of Theorem 3.6.4,

$$\mathbf{P} \mathbf{A} = \mathbf{L} \mathbf{U}, \quad \mathbf{L} = \mathbf{I} + \sum_{k=1}^{n-1} \ell'_k \mathbf{e}_k^\top,$$

where $\mathbf{L}$ is unit lower triangular because the first k entries of each ℓ'_k are zero. So $(\mathbf{P}, \mathbf{L}, \mathbf{U})$ is a PLU factorization of $\mathbf{A}$. This proves the theorem. ■

Example. A swap in the middle

Find a PLU factorization of

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 7 & 2 \end{pmatrix} \in M_3(\mathbb{Q}).$$

Solution. Step 1. The pivot 1 is nonzero, so no swap. The multipliers are $\ell_{21} = 2$ and

$\ell_{31} = 3$:

$$\mathbf{A} \xrightarrow[R_3 \rightarrow R_3 - 3R_1]{R_2 \rightarrow R_2 - 2R_1} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \ell_1 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}.$$

Step 2. The pivot position $(2, 2)$ holds 0, but the entry below it is 1. Elimination without row swaps would stop here. We swap rows 2 and 3, so $\mathbf{P}_2 = \mathbf{P}_{23}$:

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{U}.$$

The entry below the new pivot is already 0, so $\ell_{32} = 0$, and $\mathbf{U}$ is upper triangular.

By Step 2 of the proof, the stored multipliers of step 1 are swapped too: $\ell'_1 = \mathbf{P}_2 \ell_1 = (0, 3, 2)$. Hence

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Check: $\mathbf{PA}$ is $\mathbf{A}$ with rows 2 and 3 swapped, and

$$\mathbf{LU} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 6+1 & 3-1 \\ 2 & 4 & 2+1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 7 & 2 \\ 2 & 4 & 3 \end{pmatrix} = \mathbf{PA}.$$

Notice where the multipliers ended up: 3 in row 2 and 2 in row 3, traveling with the rows they belong to. To solve $\mathbf{Ax} = \mathbf{b}$, multiply by $\mathbf{P}$ and solve $\mathbf{LUx} = \mathbf{Pb}$ by two substitutions; since $\mathbf{P}$ is invertible, this changes no solutions. For $\mathbf{b} = (2, 3, 8)$, $\mathbf{Pb} = (2, 8, 3)$. Forward substitution in $\mathbf{Ly} = \mathbf{Pb}$ gives $y_1 = 2$, $y_2 = 8 - 3 \cdot 2 = 2$, $y_3 = 3 - 2 \cdot 2 = -1$. Back substitution in $\mathbf{Ux} = \mathbf{y}$ gives $x_3 = -1$, $x_2 = 2 + x_3 = 1$, $x_1 = 2 - 2x_2 - x_3 = 1$. Check: $\mathbf{A}(1, 1, -1) = (1 + 2 - 1, 2 + 4 - 3, 3 + 7 - 2) = (2, 3, 8)$.

Remark. Over $\mathbb{R}$, computers do not choose just any nonzero pivot. They choose the entry of **largest absolute value** in the column, which is called **partial pivoting**. The reason is rounding. For $\mathbf{A} = \begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}$ with $\varepsilon = 10^{-20}$, elimination without swaps gives the multiplier 10^{20} and $u_{22} = 1 - 10^{20}$, which in standard double-precision arithmetic is stored as -10^{20} . Multiplying back, the computed $\mathbf{LU}$ has $(2, 2)$ -entry $10^{20} - 10^{20} = 0$ instead of 1. Swapping the rows first gives the harmless multiplier ε . Chapter 24 makes this precise.

3.6.8 Symmetric matrices: the $\mathbf{LDL}^\top$ factorization

When $\mathbf{A}$ is symmetric, the two factors of $\mathbf{A} = \mathbf{LU}$ are secretly related: up to a diagonal matrix, $\mathbf{U}$ is the transpose of $\mathbf{L}$. In [Example \(An LU factorization from elimination\)](#) the factors look unrelated, but that matrix is not symmetric. Try a symmetric one:

$$\begin{pmatrix} 1 & 3 & -2 \\ 3 & 7 & -8 \\ -2 & -8 & 5 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 + 2R_1]{R_2 \rightarrow R_2 - 3R_1} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -2 & -2 \\ 0 & -2 & 1 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -2 & -2 \\ 0 & 0 & 3 \end{pmatrix} = \mathbf{U},$$

with multipliers $\ell_{21} = 3$, $\ell_{31} = -2$, $\ell_{32} = 1$. Divide each row of $\mathbf{U}$ by its diagonal entry: the rows become $(1, 3, -2)$, $(0, 1, 1)$, $(0, 0, 1)$. These are the columns of $\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix}$, read as rows. So $\mathbf{U} = \mathbf{D}\mathbf{L}^\top$ with $\mathbf{D} = \text{diag}(1, -2, 3)$, and $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^\top$.

Theorem 3.6.8. LDL[⊤] factorization

Let $\mathbf{A} \in M_n(F)$ be **symmetric** and **invertible**, and suppose $\mathbf{A} = \mathbf{L}\mathbf{U}$ is an LU factorization. Then the diagonal entries $u_{11}, \dots, u_{nn}$ of $\mathbf{U}$ are nonzero, and with $\mathbf{D} = \text{diag}(u_{11}, \dots, u_{nn})$,

$$\mathbf{U} = \mathbf{D}\mathbf{L}^\top \quad \text{and} \quad \mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^\top.$$

Idea. Pull the diagonal out of $\mathbf{U}$, writing $\mathbf{U} = \mathbf{D}\mathbf{V}$ with $\mathbf{V}$ unit upper triangular, so $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{V}$. Transposing, $\mathbf{A} = \mathbf{A}^\top = \mathbf{V}^\top(\mathbf{D}\mathbf{L}^\top)$, which is a second LU factorization of the same invertible matrix. Uniqueness forces the two to agree, so $\mathbf{L} = \mathbf{V}^\top$.

Proof. As in the proof of Theorem 3.6.5, $\mathbf{U} = \mathbf{L}^{-1}\mathbf{A}$ is invertible, so its diagonal entries are nonzero by Lemma 3.6.2. Hence $\mathbf{D}$ is invertible, with $\mathbf{D}^{-1} = \text{diag}(u_{11}^{-1}, \dots, u_{nn}^{-1})$. Let $\mathbf{V} = \mathbf{D}^{-1}\mathbf{U}$. Row i of $\mathbf{V}$ is u_{ii}^{-1} times row i of $\mathbf{U}$, so $\mathbf{V}$ is upper triangular with $v_{ii} = u_{ii}^{-1}u_{ii} = 1$, and $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{V}$.

Since $\mathbf{A}$ is symmetric and $\mathbf{D}^\top = \mathbf{D}$, Theorem 0.9.5 gives

$$\mathbf{A} = \mathbf{A}^\top = (\mathbf{L}\mathbf{D}\mathbf{V})^\top = \mathbf{V}^\top\mathbf{D}\mathbf{L}^\top = \mathbf{V}^\top(\mathbf{D}\mathbf{L}^\top).$$

Here $\mathbf{V}^\top$ is unit lower triangular, because $\mathbf{V}$ is unit upper triangular. And $\mathbf{D}\mathbf{L}^\top$ is upper triangular: its (i, j) -entry is $u_{ii}\ell_{ji}$, and $\ell_{ji} = 0$ when $j < i$, since $\mathbf{L}$ is lower triangular. So $(\mathbf{V}^\top, \mathbf{D}\mathbf{L}^\top)$ is an LU factorization of the invertible matrix $\mathbf{A}$. By Theorem 3.6.5, $\mathbf{V}^\top = \mathbf{L}$ and $\mathbf{D}\mathbf{L}^\top = \mathbf{U}$. Therefore $\mathbf{A} = \mathbf{L}\mathbf{U} = \mathbf{L}\mathbf{D}\mathbf{L}^\top$, as claimed. ■

The factorization stores about half as much: $\mathbf{L}$ and the diagonal $\mathbf{D}$ determine everything. Over $\mathbb{R}$, when every entry of $\mathbf{D}$ is positive, one can split $\mathbf{D} = \mathbf{D}^{1/2}\mathbf{D}^{1/2}$ and write $\mathbf{A} = (\mathbf{L}\mathbf{D}^{1/2})(\mathbf{L}\mathbf{D}^{1/2})^\top$; this is the **Cholesky factorization**, which Chapter 13 connects with positive definite matrices.

Remark. Symmetry alone does not give an LU factorization: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is symmetric and invertible, and has none by Example (An invertible matrix with no LU factorization). The theorem assumes the factorization exists and then improves it.

3.6.9 Exercises

3.6.9.1 A. Check your understanding

Exercise. A1

- (a) Define an LU factorization of a matrix $\mathbf{A} \in M_n(F)$, and a PLU factorization.
- (b) True or false: every invertible matrix has an LU factorization. Justify your answer.
- (c) Suppose $\mathbf{A} = \mathbf{L}\mathbf{U}$. In which order are the two triangular systems solved to find a solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, and what is each one?
- (d) True or false: the product of two unit lower triangular $n \times n$ matrices is unit lower triangular. Justify your answer.
- (e) True or false: the zero matrix in $M_2(F)$ has more than one LU factorization. Justify your answer.
- (f) What is the inverse of a permutation matrix $\mathbf{P}$?

Solution.

- (a) See [Definition 3.6.1](#): $\mathbf{A} = \mathbf{L}\mathbf{U}$ with $\mathbf{L}$ unit lower triangular and $\mathbf{U}$ upper triangular. See [Definition 3.6.2](#): $\mathbf{PA} = \mathbf{L}\mathbf{U}$ with $\mathbf{P}$ a permutation matrix, $\mathbf{L}$ unit lower triangular, and $\mathbf{U}$ upper triangular.
- (b) False. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is invertible but has no LU factorization, by [Example \(An invertible matrix with no LU factorization\)](#).
- (c) First $\mathbf{L}\mathbf{y} = \mathbf{b}$, by forward substitution; then $\mathbf{U}\mathbf{x} = \mathbf{y}$, by back substitution. Then $\mathbf{A}\mathbf{x} = \mathbf{L}(\mathbf{U}\mathbf{x}) = \mathbf{L}\mathbf{y} = \mathbf{b}$.
- (d) True, by [Lemma 3.6.1](#) (a).
- (e) True. For every $c \in F$, $\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \mathbf{O} = \mathbf{O}$, so $\left(\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}, \mathbf{O} \right)$ is an LU factorization; $c = 0$ and $c = 1$ give two different ones. This does not contradict [Theorem 3.6.5](#), because $\mathbf{O}$ is not invertible.
- (f) $\mathbf{P}^{-1} = \mathbf{P}^\top$, by [Lemma 3.4.7](#) (b).

3.6.9.2 B. Practice**Exercise. B1: An LU factorization**

Find the LU factorization of

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ -3 & -7 & 4 \\ 2 & 6 & -3 \end{pmatrix} \in M_3(\mathbb{Q})$$

by elimination without row swaps. Hence show that $\mathbf{A}$ is invertible.

Solution. Step 1. The pivot is 1, with multipliers $\ell_{21} = -3$ and $\ell_{31} = 2$:

$$\mathbf{A} \xrightarrow[R_3 \rightarrow R_3 - 2R_1]{R_2 \rightarrow R_2 - (-3)R_1} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 4 \\ 0 & 2 & -3 \end{pmatrix}.$$

Step 2. The pivot is -1 , with multiplier $\ell_{32} = 2/(-1) = -2$:

$$\xrightarrow{R_3 \rightarrow R_3 - (-2)R_2} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 4 \\ 0 & 0 & 5 \end{pmatrix} = \mathbf{U}.$$

By [Theorem 3.6.4](#),

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -2 & 1 \end{pmatrix}, \quad \mathbf{A} = \mathbf{L}\mathbf{U}.$$

$$\text{Check: } \mathbf{L}\mathbf{U} = \begin{pmatrix} 1 & 2 & 0 \\ -3 & -6-1 & 4 \\ 2 & 4+2 & -8+5 \end{pmatrix} = \mathbf{A}.$$

The diagonal entries $1, -1, 5$ of $\mathbf{U}$ are nonzero, so $\mathbf{U}$ is invertible by [Lemma 3.6.2](#), and $\mathbf{L}$ is invertible by [Lemma 3.6.1](#). Hence $\mathbf{A} = \mathbf{L}\mathbf{U}$ is invertible by [Theorem 0.9.7](#).

Exercise. B2: Two systems, one factorization

With $\mathbf{A} = \mathbf{L}\mathbf{U}$ as in [Exercise \(B1: An LU factorization\)](#), solve $\mathbf{A}\mathbf{x} = \mathbf{b}$ for $\mathbf{b} = (3, -6, 5)$ and for $\mathbf{b} = (0, 1, -2)$, using only forward and back substitution.

Solution. For $\mathbf{b} = (3, -6, 5)$. $\mathbf{L}\mathbf{y} = \mathbf{b}$: $y_1 = 3$; $-3y_1 + y_2 = -6$, so $y_2 = 3$; $2y_1 - 2y_2 + y_3 = 5$, so $y_3 = 5 - 6 + 6 = 5$. Then $\mathbf{U}\mathbf{x} = \mathbf{y}$: $5x_3 = 5$, so $x_3 = 1$; $-x_2 + 4x_3 = 3$, so $x_2 = 1$; $x_1 + 2x_2 = 3$, so $x_1 = 1$. Hence $\mathbf{x} = (1, 1, 1)$. Check: $\mathbf{A}\mathbf{x} = (1 + 2, -3 - 7 + 4, 2 + 6 - 3) = (3, -6, 5)$.

For $\mathbf{b} = (0, 1, -2)$. $\mathbf{L}\mathbf{y} = \mathbf{b}$: $y_1 = 0$; $y_2 = 1 + 3y_1 = 1$; $y_3 = -2 - 2y_1 + 2y_2 = 0$. Then $\mathbf{U}\mathbf{x} = \mathbf{y}$: $5x_3 = 0$, so $x_3 = 0$; $-x_2 + 0 = 1$, so $x_2 = -1$; $x_1 + 2(-1) = 0$, so $x_1 = 2$. Hence $\mathbf{x} = (2, -1, 0)$. Check: $\mathbf{A}\mathbf{x} = (2 - 2, -6 + 7, 4 - 6) = (0, 1, -2)$.

Since $\mathbf{A}$ is invertible ([Exercise \(B1: An LU factorization\)](#)), each solution is the only one, by [Theorem 0.9.7](#).

Exercise. B3: A PLU factorization

$$\text{Let } \mathbf{A} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 3 & 1 \end{pmatrix} \in M_3(\mathbb{Q}).$$

- Explain why $\mathbf{A}$ has no LU factorization.
- Find a PLU factorization of $\mathbf{A}$.
- Hence solve $\mathbf{A}\mathbf{x} = (0, 2, 7)$.

Solution.

- If $\mathbf{A} = \mathbf{L}\mathbf{U}$, then the $(1, 1)$ -entry gives $u_{11} = a_{11} = 0$, and the $(2, 1)$ -entry gives $a_{21} = \ell_{21}u_{11} = 0$. But $a_{21} = 1 \neq 0$, a contradiction, exactly as in [Example \(An](#)

invertible matrix with no LU factorization).

- (b) The first pivot position holds 0 and the entry below is 1, so swap rows 1 and 2, $\mathbf{P}_1 = \mathbf{P}_{12}$:

$$\begin{aligned} \mathbf{A} &\xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 2R_1]{R_2 \rightarrow R_2 - 0R_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix} \\ &\xrightarrow{R_3 \rightarrow R_3 - 1R_2} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -3 \end{pmatrix} = \mathbf{U}. \end{aligned}$$

The multipliers are $\ell_{21} = 0$, $\ell_{31} = 2$, $\ell_{32} = 1$, and no swap happens after step 1, so they need no rearranging. Hence

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}.$$

$$\text{Check: } \mathbf{PA} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \text{ and } \mathbf{LU} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 2+1 & 2+2-3 \end{pmatrix} = \mathbf{PA}.$$

- (c) $\mathbf{Ax} = \mathbf{b}$ holds if and only if $\mathbf{PAx} = \mathbf{Pb}$, since $\mathbf{P}$ is invertible (Lemma 3.4.7 (b)). Swapping the first two entries of $\mathbf{b} = (0, 2, 7)$, we solve $\mathbf{LUx} = \mathbf{Pb} = (2, 0, 7)$. $\mathbf{Ly} = (2, 0, 7)$: $y_1 = 2$, $y_2 = 0$, $y_3 = 7 - 2 \cdot 2 - 0 = 3$. $\mathbf{Ux} = \mathbf{y}$: $-3x_3 = 3$, so $x_3 = -1$; $x_2 + 2x_3 = 0$, so $x_2 = 2$; $x_1 + x_2 + x_3 = 2$, so $x_1 = 1$. Hence $\mathbf{x} = (1, 2, -1)$. Check: $\mathbf{Ax} = (0 + 2 - 2, 1 + 2 - 1, 2 + 6 - 1) = (0, 2, 7)$.

3.6.9.3 C. Going deeper

Exercise. C1: Leading principal submatrices

For $\mathbf{A} \in M_n(F)$ and $1 \leq k \leq n$, the **leading principal submatrix** $\mathbf{A}_k \in M_k(F)$ is the top-left $k \times k$ corner of $\mathbf{A}$, with entries a_{ij} for $i, j \leq k$.

- Let $\mathbf{L}$ be lower triangular and $\mathbf{U}$ upper triangular in $M_n(F)$. Show that $(\mathbf{LU})_k = \mathbf{L}_k \mathbf{U}_k$ for every k .
- Deduce that if $\mathbf{A}$ has an LU factorization $\mathbf{A} = \mathbf{LU}$ with $\mathbf{U}$ invertible, then every leading principal submatrix $\mathbf{A}_1, \dots, \mathbf{A}_n$ is invertible.
- Prove the converse: if $\mathbf{A}_1, \dots, \mathbf{A}_n$ are all invertible, then $\mathbf{A}$ has an LU factorization with $\mathbf{U}$ invertible.

Hint for (c): induction on n . Write $\mathbf{A}$ with $\mathbf{A}_{n-1}$ in its top-left corner, a column $\mathbf{b}$, a row $\mathbf{c}^\top$, and a corner entry d , and look for $\mathbf{L}$ and $\mathbf{U}$ of the same shape.

Solution.

- Let $i, j \leq k$. By Definition 0.9.9, $(\mathbf{LU})_{ij} = \sum_{l=1}^n \ell_{il} u_{lj}$. If $l > k$, then $l > i$, so $\ell_{il} = 0$ because $\mathbf{L}$ is lower triangular. Hence $(\mathbf{LU})_{ij} = \sum_{l=1}^k \ell_{il} u_{lj} = (\mathbf{L}_k \mathbf{U}_k)_{ij}$, as claimed.
- Suppose $\mathbf{A} = \mathbf{LU}$ with $\mathbf{U}$ invertible. By (a), $\mathbf{A}_k = \mathbf{L}_k \mathbf{U}_k$. The matrix $\mathbf{L}_k$ is unit lower triangular, so it is invertible by Lemma 3.6.1. By Lemma 3.6.2, $u_{11}, \dots, u_{nn}$ are

nonzero; $\mathbf{U}_k$ is upper triangular with diagonal $u_{11}, \dots, u_{kk}$, so $\mathbf{U}_k$ is invertible by the same lemma. Hence $\mathbf{A}_k$ is invertible by Theorem 0.9.7.

- (c) We prove by induction on n that every $\mathbf{A} \in M_n(F)$ whose leading principal submatrices are all invertible has an LU factorization with $\mathbf{U}$ invertible. For $n = 1$, $\mathbf{A} = (a_{11})$ with $a_{11} \neq 0$, and $\mathbf{A} = (1)(a_{11})$ works.

Let $n \geq 2$ and suppose the statement holds for $n - 1$. Write

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{b} \\ \mathbf{c}^\top & d \end{pmatrix}, \quad \mathbf{b}, \mathbf{c} \in F^{n-1}, d \in F.$$

The leading principal submatrices of $\mathbf{A}_{n-1}$ are $\mathbf{A}_1, \dots, \mathbf{A}_{n-1}$, all invertible, so by the induction hypothesis $\mathbf{A}_{n-1} = \mathbf{L}'\mathbf{U}'$ with $\mathbf{L}'$ unit lower triangular and $\mathbf{U}'$ upper triangular and invertible. Put $\mathbf{y} = \mathbf{L}'^{-1}\mathbf{b}$, $\mathbf{x}^\top = \mathbf{c}^\top \mathbf{U}'^{-1}$, and $z = d - \mathbf{x}^\top \mathbf{y}$, and let

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}' & \mathbf{0} \\ \mathbf{x}^\top & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} \mathbf{U}' & \mathbf{y} \\ \mathbf{0}^\top & z \end{pmatrix}.$$

Then $\mathbf{L}$ is unit lower triangular and $\mathbf{U}$ is upper triangular. Computing $\mathbf{L}\mathbf{U}$ entry by entry, splitting each sum $\sum_{l=1}^n$ into $l \leq n - 1$ and $l = n$: the top-left block is $\mathbf{L}'\mathbf{U}' + \mathbf{0}\mathbf{0}^\top = \mathbf{A}_{n-1}$; the top-right column is $\mathbf{L}'\mathbf{y} + \mathbf{0}z = \mathbf{b}$; the bottom-left row is $\mathbf{x}^\top \mathbf{U}' + 1 \cdot \mathbf{0}^\top = \mathbf{c}^\top$; the corner is $\mathbf{x}^\top \mathbf{y} + z = d$. Hence $\mathbf{A} = \mathbf{L}\mathbf{U}$.

Finally $\mathbf{A} = \mathbf{A}_n$ is invertible by hypothesis, and $\mathbf{L}$ is invertible by Lemma 3.6.1, so $\mathbf{U} = \mathbf{L}^{-1}\mathbf{A}$ is invertible by Theorem 0.9.7. This completes the induction.

In the language of elimination: z is the n -th pivot, and (c) shows that invertible leading corners are exactly what keeps every pivot nonzero.

Exercise. C2: An LDL[⊤] factorization

Let $\mathbf{A} = \begin{pmatrix} 4 & -4 & 8 \\ -4 & 5 & -11 \\ 8 & -11 & 27 \end{pmatrix} \in M_3(\mathbb{R})$.

- (a) Find the LU factorization of $\mathbf{A}$, show that $\mathbf{A}$ is invertible, and write $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^\top$.
 (b) Deduce that $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$ for every nonzero $\mathbf{x} \in \mathbb{R}^3$.

Solution.

- (a) Step 1: pivot 4, multipliers $\ell_{21} = -1$, $\ell_{31} = 2$. Row 2 becomes $(-4, 5, -11) + (4, -4, 8) = (0, 1, -3)$, row 3 becomes $(8, -11, 27) - 2(4, -4, 8) = (0, -3, 11)$. Step 2: pivot 1, multiplier $\ell_{32} = -3$; row 3 becomes $(0, -3, 11) + 3(0, 1, -3) = (0, 0, 2)$. Hence

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -3 & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} 4 & -4 & 8 \\ 0 & 1 & -3 \\ 0 & 0 & 2 \end{pmatrix},$$

and $\mathbf{A} = \mathbf{L}\mathbf{U}$ by Theorem 3.6.4. The diagonal 4, 1, 2 of $\mathbf{U}$ is nonzero, so $\mathbf{U}$ is invertible (Lemma 3.6.2), $\mathbf{L}$ is invertible (Lemma 3.6.1), and $\mathbf{A} = \mathbf{L}\mathbf{U}$ is invertible. $\mathbf{A}$ is

symmetric, so by [Theorem 3.6.8](#), with $\mathbf{D} = \text{diag}(4, 1, 2)$,

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^\top, \quad \mathbf{D}\mathbf{L}^\top = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -4 & 8 \\ 0 & 1 & -3 \\ 0 & 0 & 2 \end{pmatrix} = \mathbf{U},$$

which confirms the theorem directly.

- (b) Let $\mathbf{x} \neq \mathbf{0}$ and put $\mathbf{y} = \mathbf{L}^\top \mathbf{x}$. By [Lemma 3.6.1](#), $\mathbf{L}$ is invertible, so $\mathbf{L}^\top$ is invertible by Theorem 0.9.7 (4), so $\mathbf{y} \neq \mathbf{0}$ (otherwise $\mathbf{x} = (\mathbf{L}^\top)^{-1} \mathbf{y} = \mathbf{0}$). By Theorem 0.9.5, $\mathbf{y}^\top = \mathbf{x}^\top \mathbf{L}$, so

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{x}^\top \mathbf{L} \mathbf{D} \mathbf{L}^\top \mathbf{x} = \mathbf{y}^\top \mathbf{D} \mathbf{y} = 4y_1^2 + y_2^2 + 2y_3^2,$$

reading the 1×1 matrix as a real number. Each term is ≥ 0 , and at least one $y_i \neq 0$, which makes its term strictly positive. Hence $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$.

Exercise. C3: Counting matrices without an LU factorization

- (a) Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(F)$. Prove that $\mathbf{A}$ has an LU factorization if and only if $a \neq 0$ or $c = 0$.
- (b) Let q be a prime. How many matrices in $M_2(\mathbb{F}_q)$ have **no** LU factorization? List them for $q = 2$.

Solution.

- (a) ($\Rightarrow$) Suppose $\mathbf{A} = \mathbf{L}\mathbf{U}$ with $\mathbf{L} = \begin{pmatrix} 1 & 0 \\ \ell & 1 \end{pmatrix}$ and $\mathbf{U} = \begin{pmatrix} u & v \\ 0 & w \end{pmatrix}$. As in [Example \(An invertible matrix with no LU factorization\)](#), $\mathbf{L}\mathbf{U} = \begin{pmatrix} u & v \\ \ell u & \ell v + w \end{pmatrix}$, so $a = u$ and $c = \ell u = \ell a$. If $a = 0$, then $c = 0$. Hence $a \neq 0$ or $c = 0$.
- ($\Leftarrow$) If $a \neq 0$, elimination without row swaps succeeds with multiplier $\ell = c/a$, and indeed

$$\begin{pmatrix} 1 & 0 \\ c/a & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d - cb/a \end{pmatrix} = \begin{pmatrix} a & b \\ c & cb/a + d - cb/a \end{pmatrix} = \mathbf{A}.$$

If $c = 0$, then $\mathbf{A}$ is upper triangular and $\mathbf{A} = \mathbf{I}\mathbf{A}$.

- (b) By (a), $\mathbf{A}$ has no LU factorization exactly when $a = 0$ and $c \neq 0$. There are $q - 1$ choices of c , and q choices each of b and d , so there are $(q - 1)q^2$ such matrices. For $q = 2$ this gives 4 of the 16 matrices: $a = 0$, $c = 1$, and $b, d \in \{0, 1\}$, that is,

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$