eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

ZIENTZIA ETA TEKNOLOGIA FAKULTATEA
FACULTAD DE CIENCIA Y TECNOLOGÍA

4. Direct methods for the resolution of linear systems

Jon Asier Bárcena-Petisco (University of the Basque Country UPV/EHU) - Numerical Methods I (2023/24)

Contents on this chapter

- Resolution of linear systems in a systematic way approachable by a computer (Cramer formula is too slow for big systems)
- Resolution of triangular systems
- *LUP* decomposition
- Cholesky decomposition

4.1. Properties of triangular matrices.

Properties already proved

1. The identity matrix I is a diagonal matrix.
2. $D \in \mathbb{R}^{n\times n}$ is a diagonal matrix if and only if it is upper triangular and lower triangular.
3. If $L \in \mathbb{R}^{n\times n}$ is a lower triangular matrix, L' is an upper triangular matrix.
4. If $U \in \mathbb{R}^{n\times n}$ is an upper triangular matrix, U' is a lower triangular matrix.
5. If $L_1, L_2 \in \mathbb{R}^{n\times n}$ are lower triangular matrices, so are $L_1 + L_2$ and $L_1 L_2$.
6. If $U_1, U_2 \in \mathbb{R}^{n\times n}$ are upper triangular matrices, so are $U_1 + U_2$ and $U_1 U_2$.

Characterization of non-singular lower triangular matrix

Exercise

Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix.

1. Show that $\det(L) = L(1,1)\cdots L(n,n)$.
2. Show that L is non-singular if and only if $L(i,i) \neq 0$ for all $i \neq 0$.

Remark

Consequently, the linear system $Lx = b$ for $b \in \mathbb{R}^n$ is well-posed (have a unique solution) if and only if $L(i,i) \neq 0$ for all $i = 1,\ldots,n$.

Additional properties of lower triangular matrices

Exercise

Let $L_1, L_2 \in \mathbb{R}^{n\times n}$ be lower triangular matrices. Show that $(L_1 L_2)(i,i) = L_1(i,i)L_2(i,i)$ for all $i = 1, \ldots, n$.

Exercise

Let $L \in \mathbb{R}^{n\times n}$ be a lower non-singular triangular matrices. Then L^{-1} is also a lower triangular matrix, and $(L^{-1})(i,i) = \frac{1}{L(i,i)}$.

Characterization of non-singular upper triangular matrices

Exercise

Let $U \in \mathbb{R}^{n\times n}$ be an upper triangular matrices

1. Show that $\det(U) = U(1,1)\cdots U(n,n)$.
2. Show that U is non-singular if and only if $U(i,i) \neq 0$ for all $i \neq 0$.

Remark

Consequently, the linear system $Ux = b$ for $b \in \mathbb{R}^n$ is well-posed (have a unique solution) if and only if $U(i,i) \neq 0$ for all $i = 1,\ldots,n$.

Additional properties of upper triangular matrices

Exercise

Let $U_1, U_2 \in \mathbb{R}^{n\times n}$ be upper triangular matrices. Show that $(U_1U_2)(i,i) = U_1(i,i)U_2(i,i)$ for all $i = 1,\dots,n$.

Exercise

Let $U \in \mathbb{R}^{n\times n}$ be an upper non-singular triangular matrix. Then U^{-1} is also an upper triangular matrix, and $(U^{-1})(i,i) = \frac{1}{U(i,i)}$.

4.2. Resolution of linear triangular systems.

Lower triangular system

Let us explain how to solve systems of the form $Lx = b$, for $L \in \mathbb{R}^{n\times n}$ a lower triangular matrix and $b \in \mathbb{R}^n$. That is, let us explain how to solve:

$$\begin{cases} L(1,1)x(1) & = b(1), \\ L(2,1)x(1) + L(2,2)x(2) & = b(2), \\ \qquad\qquad\vdots & \quad\vdots \\ L(i,1)x(1) + \cdots + L(i,i)x(i) & = b(i), \\ \qquad\qquad\vdots & \quad\vdots \\ L(n,1)x(1) + L(n,2)x(2) + \cdots + L(n,n)x(n) & = b(n). \end{cases}$$

Example of resolution of a linear triangular system

Let us consider:

$$\begin{cases} 2x(1) & = 2, \\ -2x(1) + 3x(2) & = 4, \\ 5x(1) + x(2) + 4x(3) & = 8. \end{cases}$$

Then, the solution is given by:

$$\begin{aligned} x(1) &= \frac{2}{2} = 1, \\ x(2) &= \frac{4 + 2x(1)}{3} = \frac{4+2}{3} = 2, \\ x(3) &= \frac{8 - 5x(1) - x(2)}{4} = \frac{8-5-2}{4} = \frac{1}{4}. \end{aligned}$$

Resolution of linear triangular systems

Let us consider a lower triangular system. If $L(i,i) \neq 0$ for all $i = 1, \ldots, n$ (that is, if L is non-singular), the resolution of the previous system is given by:

$$
\begin{aligned}
x(1) &= \frac{b(1)}{L(1,1)}, \\
x(2) &= \frac{b(2) - L(2,1)x(1)}{L(2,2)}, \\
&\vdots \\
x(i) &= \frac{b(i) - L(i,1)x(1) - \cdots - L(i,i-1)x(i-1)}{L(i,i)}, \\
&\vdots \\
x(n) &= \frac{b(n) - L(n,1)x(1) - \cdots - L(n,n-1)x(n-1)}{L(n,n)}.
\end{aligned}
$$

It is clear that any computer can process such formulas, with the help of a for loop.

Upper triangular systems

Let us explain how to solve systems of the form $Ux = b$, for $U \in \mathbb{R}^{n\times n}$ an upper triangular matrix and $b \in \mathbb{R}^n$. That is, let us explain how to solve:

$$\begin{cases} U(1,1)x(1) + U(1,2)x(2) + \cdots + U(1,n-1)x(n-1) + U(1,n)x(n) & = b(1), \\ \qquad\qquad U(2,2)x(2) + \cdots + U(2,n-1)x(n-1) + U(2,n)x(n) & = b(2), \\ \vdots & \vdots \\ U(i,i)x(i) + \cdots U(i,n)x(n) & = b(i) \\ \vdots & \vdots \\ U(n,n)x(n) & = b(n). \end{cases}$$

Example of resolution of a linear triangular system

Let us consider:

$$\begin{cases} 4x(1) + x(2) + 5x(3) & = 8, \\ \qquad\qquad 3x(2) - 2x(3) & = 4, \\ \qquad\qquad\qquad\quad 2x(3) & = 2. \end{cases}$$

Then, the solution is given by:

$$\begin{aligned} x(3) &= \frac{2}{2} = 1, \\ x(2) &= \frac{4 + 2x(3)}{3} = \frac{4+2}{3} = 2, \\ x(1) &= \frac{8 - 5x(3) - x(2)}{4} = \frac{8-5-2}{4} = \frac{1}{4}. \end{aligned}$$

Resolution of linear triangular systems

Let us consider a lower triangular system. If $U(i,i) \neq 0$ for all $i = 1, \ldots, n$ (that is, if U is non-singular), the resolution of the previous system is given by:

$$
\begin{aligned}
x(n) &= \frac{b(n)}{U(n,n)}, \\
x(n-1) &= \frac{b(n-1) - U(n-1,n)x(n)}{U(n-1,n-1)}, \\
&\vdots \\
x(i) &= \frac{b(i) - U(i,n)x(n) - \cdots - U(i,i+1)x(i+1)}{U(i,i)}, \\
&\vdots \\
x(1) &= \frac{b(1) - U(n,n)x(n) - \cdots - U(n,2)x(2)}{U(1,1)}.
\end{aligned}
$$

It is clear that any computer can process such formulas, with the help of a for loop.

4.3. LU decomposition.

Resolution of the linear system $LU = b$

Let $L \in \mathbb{R}^{n\times n}$ a lower triangular non-singular matrix and $U \in \mathbb{R}^{n\times n}$ an upper triangular non-singular matrix. Let $A = LU \in \mathbb{R}^{n\times n}$. Then, the system $Ax = b$ can be solved as follows:

- First, solve $Ly = b$.
- Then, solve $Ux = y$

The computation of both steps was explained in the previous section. The method works because, if x is obtained following those two steps:

$$Ax = LUx = Ly = b.$$

Remark

The objective of this section will be to decompose any non-singular matrix A as LU and identify any potential problem that may arise.

Non-uniqueness of the LU decomposition

It is easy to see that we do not have uniqueness in the LU decomposition. Indeed, if D is any non-singular diagonal matrix:

$$LU = LDD^{-1}U = (LD)(D^{-1}U).$$

As L is lower triangular and D diagonal, LD is lower triangular. Similarly, as D^{-1} is diagonal and U upper triangular, $D^{-1}U$ is upper triangular.

As we are going to see, this will be fixed by requiring that the diagonal elements of L are all 1.

An additional restriction to the LU decomposition (ii)

Proposition

Let $L_1 \in \mathbb{R}^{n\times n}$ be a non-singular lower triangular matrix and $U_1 \in \mathbb{R}^{n\times n}$ be a non-singular upper triangular matrix. Then, there is $L_2 \in \mathbb{R}^{n\times n}$ a non-singular lower triangular matrix such that $L_2(i,i) = 1$ for all $i = 1, \ldots, n$ and U_2 a non-singular lower triangular matrix such that:

$$L_1 U_1 = L_2 U_2.$$

Proof.

Let D the diagonal matrix given by $D(i,i) = \frac{1}{L_1(i,i)}$ for all $i = 1, \ldots, n$. As L_1 is non-singular, $L_1(i,i) \neq 0$ for all $i = 1, \ldots, n$, so D is well-defined. Moreover:

- $L_2 U_2 = L_1 D D^{-1} U_1 = L_1 U_1$.
- $L_2(i,i) = L_1(i,i) D(i,i) = 1$, for all $i = 1, \ldots, n$.

$\square$

An additional restriction to the LU decomposition (iii)

Proposition

Let $L_1, L_2 \in \mathbb{R}^{n\times n}$ be lower triangular matrices such that $L_1(i,i) = L_2(i,i) = 1$ for all $i = 1, \dots, n$ and $U_1, U_2 \in \mathbb{R}^{n\times n}$ be upper triangular non-singular matrices. Then, if $L_1 U_1 = L_2 U_2$ we have that $L_1 = L_2$ and $U_1 = U_2$.

Proof *1st step: getting an equation of diagonal matrices.* Since all the matrices are non-singular:

$$L_2^{-1} L_1 = U_2 U_1^{-1}. \tag{$\dagger$}$$

Since the left hand-side of ($\dagger$) is lower triangular, and the right hand-side of ($\dagger$) is upper triangular, both sides must be diagonal.

An additional restriction to the LU decomposition (iv)

2nd step: obtention of the identity. Since

$$(L_2)^{-1}(i,i) = \frac{1}{L_2(i,i)} = \frac{1}{1} = 1 \ \ \forall i = 1, \ldots, n$$

and the diagonal of the product of lower matrices is given by the products of the diagonal, then the diagonal of $L_2^{-1}L_1$ contains just 1. Moreover, since $L_2^{-1}L_1$ is diagonal:

$$L_2^{-1}L_1 = I. \tag{‡}$$

3rd step: conclusion. By multiplying (‡) by L_2, we obtain that:

$$L_1 = L_2.$$

Finally, combining (†) and (‡), we obtain that:

$$U_2 U_1^{-1} = I,$$

which implies multiplying by U_1 that:

$$U_2 = U_1. \square$$

Non-existence of the LU decomposition

As we are going to see, not every non-singular matrix can be decomposed with a LU decomposition. Let us consider:

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

It is an invertible matrix. However, if

$$\begin{pmatrix} 1 & 0 \\ L(2,1) & 1 \end{pmatrix} \begin{pmatrix} U(1,1) & U(1,2) \\ 0 & U(2,2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

we have the equations:

$$\begin{cases} U(1,1) = 0, \\ U(1,2) = 0, \\ L(2,1)U(1,1) = 1, \\ L(2,1)U(1,2) + U(2,2) = 1, \end{cases}$$

which does not have any solution.

Matrix for whom a LU decomposition exist

Exercise

Let $A \in \mathbb{R}^{n\times n}$ be a non-singular matrix. Then, there is an LU decomposition if and only if $\det(A(1{:}k, 1{:}k)) \neq 0$ for all $k = 1, \dots, n$.

Exercise

Let $A \in \mathbb{R}^{n\times n}$ for $n \geq 2$ be a matrix such that $|A(i,i)| > \sum_{j\neq i} |A(i,j)|$ for all $i = 1, \dots, n$ (a matrix with a *strictly dominant diagonal*). Show that A admits an LU decomposition in which $L(i,i) = 1$ for all $i = 1, \dots, n$.

4.4. LUP decomposition.

Revisit of permutation matrices

Definition

A *permutation matrix* is any matrix $P \in \mathbb{R}^{n\times n}$ such that there is a bijection $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ such that:

$$P(i,j) = \begin{cases} 1 & j = \sigma(i) \\ 0 & j \neq \sigma(i). \end{cases}$$

In that case we say that P is *induced* by σ.

Remark

Recall that if P is a permutation matrix, P is non-singular and $P^{-1} = P'$. Also recall that the product of permutation matrices is a permutation matrix.

Permutation of columns (i)

Let us see that multiplying by a permutation matrix in the left permutes rows:

Proposition

Let $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ be a bijection and P be its induced matrix. Then, for all $A \in \mathbb{R}^{n \times n}$:

$$(PA)(i, j) = A(\sigma(i), j).$$

Proof.

We have that:

$$(PA)(i, j) = \sum_{k=1}^{n} P(i, k)A(k, j) = A(\sigma(i), j),$$

as $P(i, \sigma(i)) = 1$ and $P(i, k) = 0$ if $k \neq \sigma(i)$. □

Resolution of linear systems with permutation matrices

Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular non-singular matrix, $U \in \mathbb{R}^{n\times n}$ be an upper triangular non-singular matrix, and P be a permutation matrix. Let $A = P'LU$; that is, $A \in \mathbb{R}^{n\times n}$ satisfying $PA = LU$. Then, the system $Ax = b$ can be solved as follows:

- First, solve $Ly = Pb$.
- Then, solve $Ux = y$

The computation of both steps was explained in Section 4.2. The method woks because, if x is obtained following the two steps, as $P' = P^{-1}$:

$$Ax = P^{-1}LUx = P^{-1}Ly = P^{-1}Pb = b.$$

Main result of the section

Theorem (LUP Decomposition Theorem)

Let $A \in \mathbb{R}^{n\times n}$ be a non-singular matrix. Then, there are $L \in \mathbb{R}^{n\times n}$ a lower triangular matrix whose diagonal terms are all 1, $U \in \mathbb{R}^{n\times n}$ an upper triangular non-singular matrix, and $P \in \mathbb{R}^{n\times n}$ a permutation matrix such that $LU = PA$.

We are going to give a constructive proof, with an algorithm that can be implemented in a computer.

Remark

The LUP decomposition is also known in the literature as LU decomposition with pivoting.

Rephrasing the main result in an inductive way

Proposition

Let $A \in \mathbb{R}^{n\times n}$ be a non-singular matrix. Then, for all $k = 1, \dots, n$ there are:

- $L_k \in \mathbb{R}^{n\times n}$ *a lower triangular matrix such that* $L_k(i,i) = 1$ *for all* $i = 1, \dots, k$,
- $U_k \in \mathbb{R}^{n\times n}$ *an upper triangular matrix such that* $U_k(i,i) \neq 0$,
- $P_k \in \mathbb{R}^{n\times n}$ *a permutation matrix,*

such that:

$$(L_k U_k)(:, 1{:}k) = (P_k A)(:, 1{:}k).$$

Once this proposition is proved, the LUP Decomposition Theorem follows with $k = n$. Indeed, if $L_n U_n = PA$ and A is non-singular, necessarily $0 \neq \det(L_n U_n) = \det(L_n)\det(U_n)$, so L_n and U_n are also non-singular.

Step 1: the base case: definition of P_1

Let us show that the previous proposition holds for $k = 1$. Let us fix $i_1 \in \{1, \ldots, n\}$ such that:

$$|A(i_1, 1)| = \max_{i=1,\ldots,n} \{|A(i, 1)|\}.$$

(if there are more than one possibility, we choose the smallest index). Then, we fix P_1 the permutation matrix induced by:

$$\sigma_1(i) = \begin{cases} i_1 & i = 1, \\ 1 & i = i_1, \\ i & \text{otherwise.} \end{cases}$$

Step 2: the base case: definition of U_1

We fix U_1 as follows:

$$U_1(i,j) = \begin{cases} (P_1A)(1,1) & (i,j) = (1,1), \\ 0 & \text{otherwise.} \end{cases}$$

Since

$$U_1(1,1) = (P_1A)(1,1) = A_1(\sigma_1(1),1) = A_1(i_1,1),$$

$U_1(1,1) = 0$ if and only if $A(:,1) = 0$, which is absurd because A is non-singular. Thus, $U_1(1,1) \neq 0$.

Step 3: the base case: definition of L_1

Next, we define L_1 as follows:

$$L_1(i,j) = \begin{cases} \frac{(P_1A)(i,1)}{U_1(1,1)} & i = 1, \dots, n; \quad j = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Fist, because of the definition of $U(1,1)$ we have that:

$$L_1(1,1) = \frac{(P_1A)(1,1)}{U(1,1)} = \frac{(P_1A)(1,1)}{(P_1A)(1,1)} = 1.$$

Moreover, if $i = 1, \dots, n$, using that $U_1(k,1) = 0$ for all $k \geq 2$:

$$(L_1U_1)(i,1) = \sum_{k=1}^{n} L_1(i,k)U_1(k,1) = L_1(i,1)U_1(1,1) = (P_1A)(i,1),$$

Thus, the base case is proved.

Step 4: the inductive case: definition of U_{k+1}

Let us suppose the result is true for k and prove it for $k+1$ (here $k \in \{1, \dots, n-1\}$). The inductive hypothesis gives us the matrices (L_k, U_k, P_k). We fix:

$$\tilde{U}_{k+1}(i,j) = \begin{cases} U_k(i,j) & j = 1, \dots, k \\ (P_k A)(1, k+1) & (i,j) = (1, k+1) \\ (P_k A)(2, k+1) - L_k(2,1)\tilde{U}_{k+1}(1, k+1) & (i,j) = (2, k+1) \text{ and } k \leq 2 \\ (P_k A)(3, k+1) - L_k(3, 1{:}2) \cdot \tilde{U}_{k+1}(1{:}2, k+1) & (i,j) = (3, k+1) \text{ and } k \leq 3 \\ \vdots & \vdots \\ (P_k A)(k, k+1) - L_k(k, 1{:}k-1) \cdot \tilde{U}_{k+1}(1{:}k-1, k+1) & (i,j) = (k, k+1) \\ 0 & \text{otherwise} \end{cases}$$

Note that with this definition, we have that:

$$(L_k \tilde{U}_{k+1})(1{:}k, k+1) = (P_k A)(1{:}k, k+1).$$

Step 5: the inductive case: definition of P_{k+1}

Let us fix $i_{k+1} \in \{k+1, \dots, n\}$ such that:

$$
\begin{aligned}
&|(P_k A)(i_{k+1}, k+1) - L_k(i_{k+1}, 1{:}k) \cdot \tilde{U}_{k+1}(1{:}k, k+1)| \\
&= \max_{i=k+1,\dots,n} \{|(P_k A)(i, k+1) - L_k(i, 1{:}k) \cdot \tilde{U}_{k+1}(1{:}k, k+1)|\}.
\end{aligned}
$$

Then, we fix $\tilde{P}_{k+1}$ the permutation matrix induced by:

$$
\sigma_{k+1}(i) = \begin{cases} i_{k+1} & i = k+1, \\ k+1 & i = i_{k+1}, \\ i & \text{otherwise.} \end{cases}
$$

With this we define $P_{k+1} = \tilde{P}_{k+1} P_k$.

Step 6: the inductive case: defining U_{k+1} (i)

We fix U_{k+1} as follows:

$$U_{k+1}(i,j) = \begin{cases} (P_{k+1}A)(k+1,k+1) - L_k(i_{k+1},1{:}k)\cdot \tilde{U}_{k+1}(1{:}k,k+1) & (i,j)=(k+1,k+1), \\ \tilde{U}_{k+1}(i,j) & \text{otherwise.} \end{cases}$$

First, we remark that:

$$U_{k+1}(i,i) = \tilde{U}_{k+1}(i,i) = U_k(i,i) \neq 0 \;\; \forall i = 1,\ldots,k$$

by the inductive hypothesis. Moreover, $U_{k+1}(k+1,k+1) \neq 0$. Let us suppose for the sake of contradiction that $U_{k+1}(k+1,k+1) = 0$. In that case, we have that:

$$L_k\left[\tilde{U}_{k+1}(:,k+1)\right] = (P_kA)(:,k+1),$$

which implies together with the inductive hypothesis $(L_kU_k)(:,1{:}k) = (P_kA_k)(:,1{:}k)$ the equation:

$$L_k\left[\tilde{U}_{k+1}(:,1{:}k+1)\right] = (P_kA)(:,1{:}k+1)$$

Step 6: the inductive case: defining U_{k+1} (ii)

Since A is non-singular, so is $\tilde{A}_k = P_k A$ (all permutation matrices are non-singular), so the rank of $\tilde{A}_k(:, 1{:}k+1)$ is $k+1$, as all the columns of $\tilde{A}_k$ are linearly independent. However, the rank of $\tilde{U}_{k+1}(:, 1{:}k+1)$ is k, because $U_{k+1}(k+1{:}n, 1{:}k+1) = 0$. Consequently, we get an absurd (when we multiply matrices the rank cannot increase), and thus $U_{k+1}(k+1, k+1) \neq 0$.

Step 7: the inductive case: definition of L_{k+1}

Next, we define L_{k+1} as follows:

$$L_{k+1}(i,j) = \begin{cases} (\tilde{P}_{k+1}L_k)(i,j) & j = 1,\dots,k \\ \frac{(P_{k+1}A)(i,j) - L_{k+1}(i,1{:}k)\cdot U_{k+1}(1{:}k,k+1)}{U_{k+1}(k+1,k+1)} & i = k+1,\dots,n; \quad j = k+1 \\ 0 & \text{otherwise.} \end{cases}$$

Let us check that the term of the diagonal are 1. Since $\tilde{P}_{k+1}$ permutes the rows with an index larger than $k+1$:

$$L_{k+1}(i,i) = L_k(i,i) = 1, \quad i = 1,\dots,k.$$

Moreover, considering the definition of $U_{k+1}(k+1,k+1)$:

$$L_{k+1}(k+1,k+1) = \frac{(P_{k+1}A)(k+1,k+1) - L_{k+1}(k+1,1{:}k)\cdot U_{k+1}(1{:}k,k+1)}{U_{k+1}(k+1,k+1)} = 1.$$

Indeed, $L_{k+1}(k+1,1{:}k) = L_k(i_{k+1},1{:}k)$ by the definition of $\tilde{P}_{k+1}$.

Step 8: the inductive case: check the decomposition

To conclude the proof, we must show that:

$$(L_{k+1}U_{k+1})(:, 1{:}k+1) = (P_{k+1}A)(:, 1{:}k+1).$$

Let us start with the first k column. First, considering that U_{k+1} is upper triangular:

$$\begin{aligned}
(L_{k+1}U_{k+1})(:, 1{:}k) &= L_{k+1}(:, 1{:}k)U_{k+1}(1{:}k, 1{:}k) \\
&= \tilde{P}_{k+1}L_k(:, 1{:}k)U_k(1{:}k, 1{:}k) \\
&= \tilde{P}_{k+1}L_kU_k(:, 1{:}k) \\
&= \tilde{P}_{k+1}(P_kA)(:, 1{:}k) = (P_{k+1}A)(:, 1{:}k).
\end{aligned}$$

Finally, the last column is an easy consequence of applying matrix multiplication as in the base case.

Simplifications in the final iteration

In the final iteration, we just have one row. Thus, $i_n = n$ and there is no need to compute. This implies that $\tilde{P}_n = I$, so $P_n = P_{n-1}$. Also, we clearly have:

$$L_n(i,j) = \begin{cases} L_{n-1}(i,j) & j \leq n-1, \\ 0 & i \leq n-1, \quad j = n, \\ 1 & (i,j) = (n,n). \end{cases}$$

Thus, in the final step we can spare some computations and directly compute U_n (with the formulas for $\tilde{U}_n$ and U_n).

All this simplifications can be done when computing the LUP decomposition by hand.

Main steps of the LUP decomposition

The main steps are the following:

1. Obtain P_1, P_1A, U_1 and L_1.
2. Obtain recursively the following matrices $\tilde{U}_{k+1}$, P_{k+1}, $P_{k+1}A$, U_{k+1}, L_{k+1} $(k = 1, \dots, n-2)$-
3. Obtain P_nA, L_n and U_n considering the simplifications on the previous slide.

For that, we have to use the formulas obtained in the proof.

I recommend to take a look at the solved example written in a separate document.

4.5. Cholesky decomposition.

Importance of symmetric matrices

Symmetric positive-definite matrices arises in the following scenarios:

- Least Square Problem.
- Non-linear optimizations.
- Monte Carlo simulations.
- Kalman filters.
- Matrix inversions.

Cholesky ask himself whether LUP could be improved in these cases.

Important properties of symmetric matrices

Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is *positive definite* if $v' A v > 0$ for all $v \in \mathbb{R}^n \setminus \{0\}$.

Definition

A matrix $P \in \mathbb{R}^{n \times n}$ is orthogonal if P is non-singular and $P^{-1} = P'$.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ a symmetric matrix. Then, all the eigenvalues of A are real, and A can be diagonalized with an orthonormal basis with respect to the canonical scalar product. In particular, there is an orthogonal matrix $P \in \mathbb{R}^{n \times n}$ such that PAP' is diagonal.

Theorem (Sylvester criterion)

Let $A \in \mathbb{R}^{n \times n}$ a symmetric matrix. Then, A is positively definite if and only if $\det(A(1{:}k, 1{:}k)) > 0$ *for all* $k = 1, \ldots, n$.

The proofs will be omitted.

LL' decomposition (i)

Cholesky studied decompositions of the type LL', where L is a lower triangular matrix (with a diagonal that is not necessarily formed of ones). Note that the LL' decomposition is a special case of LU decomposition, in which $U = L'$. This implies that, once we have obtained this decomposition, we can solve linear systems in the same way as with the LU decomposition.

Proposition

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix and $A = LL'$. Then, A is a symmetric matrix.

Proof.

$A' = (L')'L' = LL' = A$. □

LL' decomposition (ii)

Proposition

Let $L \in \mathbb{R}^{n\times n}$ a non-singular lower triangular matrix. Let $A = LL'$. Then, A is positive definite.

Proof.

To show this result, by Sylvester criterion: as A is symmetric, it suffices to show that $\det(A(1{:}k, 1{:}k)) > 0$ for all $k = 1, \ldots, n$. Since the determinant of a triangular matrix is the product of the terms in the diagonal:

$$\begin{aligned}
\det(A(1{:}k, 1{:}k)) &= \det((LL')(1{:}k, 1{:}k)) \\
&= \det(L(1{:}k, 1{:}k)) \det(L'(1{:}k, 1{:}k)) \\
&= [L(1,1) \cdots L(k,k)][L(1,1) \cdots L(k,k)] \\
&= [L(1,1)]^2 \cdots [L(k,k)]^2 > 0.
\end{aligned}$$

We have used that $L(i,i) \neq 0$ for all $i = 1, \ldots, k$ as L is non-singular. □

Cholesky decomposition

Theorem (Cholesky Decomposition Theorem)

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. Then, there exists a unique lower triangular matrix L with strictly positive diagonal entries such that:

$$A = LL'.$$

The entries of L can be computed iteratively as follows: $L(1,1) = \sqrt{a_{1,1}}$, and, for $i = 2, \ldots, n$:

$$L(i,j) = \frac{A(i,j) - \sum_{k=1}^{j-1} L(i,k)L(j,k)}{L(j,j)}, \quad j = 1, \ldots, i-1$$

and

$$L(i,i) = \sqrt{A(i,i) - \sum_{k=1}^{i-1} (L(i,k))^2}.$$

Step 1: the entries of L (i)

Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix. If $LL' = A$, then, using that $L(i,k) = 0$ if $k > i$:

$$A(1,1) = \sum_{k=1}^{n} L(1,k)L'(k,1) = L(1,1)L'(1,1) = (L(1,1))^2,$$

and for $i = 2, \ldots, n$:

$$A(i,j) = \sum_{k=1}^{n} L(i,k)L'(k,j) = \sum_{k=1}^{j-1} L(i,k)L(j,k) + L(i,j)L(j,j) \quad j = 1, \ldots, i-1$$

and:

$$A(i,i) = \sum_{k=1}^{n} L(i,k)L'(k,i) = \sum_{k=1}^{i} L(i,k)L(i,k) = \sum_{k=1}^{i}(L(i,k))^2.$$

Step 1: the entries of L (ii)

If

$$A(i,i) - \sum_{k=1}^{i-1} (L(i,k))^2 > 0 \quad \text{for all } i = 1, \dots, n, \qquad (\dagger)$$

we see from the previous equations that there is a unique L with positive entries in the diagonal such that $LL'(i,j) = A(i,j)$ if $i \geq j$.This, by the symmetry of A and of LL' implies that there is a unique lower triangular L with positive entries in the diagonal such that $LL' = A$. Thus, to conclude, we need to prove $(\dagger)$.

Step 2: positivity of $L(i, i)$

(†) is proved by induction on i. The base case, $L(1,1)$ is a direct consequence of Sylvester criterion, which implies $A(1,1) > 0$. Let us now suppose that the result holds for i and prove it for $i+1$ ($i = 1, \ldots, n-1$). If it holds for i; then with the previous formulas there is $L \in \mathbb{C}^{(i+1)\times(i+1)}$ such that $L(1{:}i, 1{:}i) \in \mathbb{R}^{i\times i}$ and $LL' = A(1{:}i+1, 1{:}i+1)$. Thus, by Sylvester criterion:

$$0 < \det(A(1{:}i+1, 1{:}i+1)) = (\det(L))^2$$
$$= (L(1,1))^2 \cdots (L(i,i))^2 \left[A(i+1, i+1) - \sum_{k=1}^{i}(L(i+1,k))^2\right],$$

which implies that:

$$A(i,i) - \sum_{k=1}^{i-1}(L(i,k))^2 > 0.$$

Thus:

$$\sqrt{A(i,i) - \sum_{k=1}^{i-1}(L(i,k))^2} > 0.\square$$

Cholesky VS LUP: final conclusions

- Cholesky has around half the operations than the LUP decomposition. Moreover, in Cholesky decomposition we do not have to deal with permutation matrices and it is easy to program.
- LUP does not require to compute square roots.
- Both rely on solving triangular systems after the composition is made.
- In both decompositions, the number of operations is around $O(n^3)$, so it is not the fastest way of solving linear systems (see Levinson recursion). However, they are both clearly faster than Cramer method, which, if not programmed properly, have $O(n!)$ operations.