LU Factorization

Like diagonalization, LU factorization is another type of matrix factorization. It is based on row reduction and is used to efficiently solve systems of equations.

Example 3.11. Let $A = \begin{bmatrix} 4 & 3 & -5 \\ -4 & -5 & 7 \\ 8 & 6 & -8 \end{bmatrix}$. Solve the equation $A\vec{x} = \vec{b}$ where $\vec{b} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$.

Solution

We can write A as the following product

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 & -5 \\ 0 & -2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

Let $L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ and $U = \begin{bmatrix} 4 & 3 & -5 \\ 0 & -2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$. Note that L is a lower triangle matrix with 1's on the diagonal and U is a row echelon form of A (so upper triangle matrix).

The equation $A\vec{x} = \vec{b}$ is the same as $LU\vec{x} = \vec{b}$. In order to solve this equation, we proceed in two steps: solve $L\vec{y} = \vec{b}$ and then solve $U\vec{x} = \vec{y}$. The equation $L\vec{y} = \vec{b}$ corresponds to

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$$

We can deduce that $y_1 = 2$, $-y_1 + y_2 = -4$ and $2y_1 + y_3 = 6$. It follows that

$$\vec{y} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

Next, the equation $U\vec{x} = \vec{y}$ corresponds to

$$\begin{bmatrix} 4 & 3 & -5 \\ 0 & -2 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

We see that $x_3 = 2$, $-2x_2 + 3x_3 = -2$ and $4x_1 + 3x_3 - 5x_3 = 2$. We conclude that

$$\vec{x} = \begin{bmatrix} \frac{1}{4} \\ 2 \\ 1 \end{bmatrix}$$

.

Definition 3.12. Let A be an $m \times n$ matrix that can be row reduced to echelon form without row interchanges. Then A can be written in the form $A = LU$ where L is an $m \times m$ lower triangle matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A. Such factorization is called an LU factorization of A.

Note 3.13. The LU factorization is motivated by the problem of solving a sequence of equations, all with the same coefficients matrix

$$A\vec{x} = \vec{b}_1, A\vec{x} = \vec{b}_2, \ldots, A\vec{x} = \vec{b}_p$$

Algorithm for an LU Factorization

Let A be an $m \times n$ matrix. Suppose that A can be reduced to row echelon form without row interchanges. Let U be an echelon form of A.

$$A \xrightarrow{R_1} \xrightarrow{R_2} \cdots \xrightarrow{R_p} U$$

Let E_i be the elementary matrix that corresponds to the row operation R_i. That is

$$E_p E_{p-1} \cdots E_2 E_1 A = U \quad \text{(see Sec. 2.2)}$$

Therefore $A = (E_p E_{p-1} \cdots E_2 E_1)^{-1} U$ is an LU factorization of A, with $L = (E_p E_{p-1} \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_p^{-1}$. There is a faster way to compute an LU factorization.

Formula

Let A be an $m \times n$ matrix that can be reduced to row echelon form with row interchanges. Note that the matrix L is an $m \times m$ lower triangle matrix with 1's on the diagonal and U is an echelon form of A. While row reducing A to row echelon form, from left to right, perform the following steps:

(a) Consider the next pivot and divide the entries below the pivot by the this pivot. These entries form the next column of L below the diagonal entry.

(b) Perform row operations using this pivot value and using only row replacement operations (do not use row interchange nor row scaling).

Example 3.14. Let $A = \begin{bmatrix} 1 & -2 & -4 & -3 \\ 2 & -7 & -7 & -6 \\ -1 & 2 & 6 & 4 \\ -4 & -1 & 9 & 8 \end{bmatrix}$. Find an LU factorization of A.

Solution

We row reduce A and keep track of the entries below the pivots. We have

$$A = \begin{bmatrix} 1 & -2 & -4 & -3 \\ \mathbf{2} & -7 & -7 & -6 \\ \mathbf{-1} & 2 & 6 & 4 \\ \mathbf{-4} & -1 & 9 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -4 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & \mathbf{0} & 2 & 1 \\ 0 & \mathbf{-9} & -7 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -4 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & \mathbf{-10} & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -4 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Hence $U = \begin{bmatrix} 1 & -2 & -4 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

For the matrix L, the entries below each pivots are the following

$$\begin{bmatrix} 2 \\ -1 \\ -4 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 9 \end{bmatrix}, \quad \begin{bmatrix} -10 \end{bmatrix}$$

Divide these entries by the corresponding pivot, which are respectively $1, -3$ and 2, to obtain the columns of L. So L is given by

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -4 & 3 & -5 & 1 \end{bmatrix}$$

Check that $A = LU$.

Chap VI: ORTHOGONALITY and LEAST SQUARES

1 Inner Product, Length, and Orthogonality

Definition 1.1. Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$. The **inner product** (or **dot product**) $\vec{u} \cdot \vec{v}$ of $\vec{u}$ and $\vec{v}$ is defined by $\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$. That is

$$\vec{u} \cdot \vec{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Example 1.2. Compute $\vec{u} \cdot \vec{v}$ and $\vec{v} \cdot \vec{u}$ for $\vec{u} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix}$.

Solution

We have

$$\vec{u} \cdot \vec{v} = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix} = (-1)(-2) + (2)(0) + (1)(3) = 5$$

and

$$\vec{v} \cdot \vec{u} = \begin{bmatrix} -2 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = (-2)(-1) + (0)(2) + (3)(1) = 5$$

Theorem 1.3. Let $\vec{u}, \vec{v}$, and $\vec{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

(a) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$

(b) $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$

(c) $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v})$

(d) $\vec{u} \cdot \vec{u} \geq 0$, and $\vec{u} \cdot \vec{u} = 0$ if and only if $\vec{u} = \vec{0}$

(e) $(c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_p\vec{u}_p) \cdot \vec{w} = c_1(\vec{u}_1 \cdot \vec{w}) + c_2(\vec{u}_2 \cdot \vec{w}) + \cdots + c_p(\vec{u}_p \cdot \vec{w})$

Definition 1.4. The **length** (or **norm**) $||\vec{v}||$ of a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is defined by

$$||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \qquad \text{(i.e. } ||\vec{v}||^2 = \vec{v} \cdot \vec{v}\text{)}$$

Note 1.5. In $\mathbb{R}^2$, the definition of the length $||\vec{v}|| = \sqrt{v_1^2 + v_2^2}$ of $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ coincides with the standard definition of length of the line segment from the origin to the point $\vec{v}$. With this definition, the Pythagorean Theorem holds (i.e. the length l of the vector which is the hypotenuse of the right triangle with horizontal length v_1 and vertical height v_2 satisfies $v_1^2 + v_2^2 = l^2$).

Fact 1.6. For any vector $\vec{v}$ and for any scalar c we have

$$||c\vec{v}|| = |c|||\vec{v}||$$

Definition 1.7.

(a) A vector of length 1 is called a **unit vector**.

(b) If $\vec{v} \neq \vec{0}$ then $\frac{1}{||\vec{v}||}\vec{v}$ is a unit vector and is in **the same direction** as $\vec{v}$.

(c) The process in part b) is called **normalizing** $\vec{v}$.

Example 1.8. Find a unit vector $\vec{u}$ in $\mathbb{R}^4$ in the same direction as $\vec{v} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 2 \end{bmatrix}$.

Solution

First, compute the length of $\vec{v}$:

$$|||\vec{v}||^2 = \vec{v} \cdot \vec{v} = (-2)^2 + (0)^2 + (1)^2 + (2)^2 = 9$$

So

$$|||\vec{v}|| = \sqrt{9} = 3$$

Now, multiply $\vec{v}$ by $\frac{1}{||\vec{v}||}$ to obtain the vector $\vec{u}$ (the process is normalizing $\vec{v}$):

$$\vec{u} = \frac{1}{||\vec{v}||}\vec{v} = \frac{1}{3}\vec{v} = \frac{1}{3}\begin{bmatrix} -2 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 0 \\ 1/3 \\ 2/3 \end{bmatrix}$$

Check that $\vec{u}$ is indeed a unit vector (i.e. $||\vec{u}|| = 1$). It suffices to only check that $||\vec{u}||^2 = 1$. We have

$$||\vec{u}||^2 = \vec{u} \cdot \vec{u} = \left(\frac{-2}{3}\right)^2 + (0)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \frac{4}{9} + 0 + \frac{1}{9} + \frac{4}{9} = 1$$

Distance in $\mathbb{R}^n$

Definition 1.9. For $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, the distance between $\vec{u}$ and $\vec{v}$, denoted by $\text{dist}(\vec{u}, \vec{v})$, is the length of the vector $\vec{u} - \vec{v}$. That is

$$\text{dist}(\vec{u}, \vec{v}) = ||\vec{u} - \vec{v}||.$$

Example 1.10. Compute the distance of $\vec{u} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$.

Solution

By definition $\text{dist}(\vec{u}, \vec{v}) = ||\vec{u} - \vec{v}||$. We have

$$\begin{aligned}
\vec{u} - \vec{v} &= \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix} \\
||\vec{u} - \vec{v}|| &= \sqrt{(-3)^2 + (-1)^2 + (1)^2} = \sqrt{11}.
\end{aligned}$$

Hence $\text{dist}(\vec{u}, \vec{v}) = \sqrt{11}$.

Definition 1.11. Two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\vec{u} \cdot \vec{v} = 0$.

Note 1.12. We have

$$\begin{aligned}
||\vec{u} + \vec{v}||^2 &= (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) \\
&= \vec{u} \cdot (\vec{u} + \vec{v}) + \vec{v} \cdot (\vec{u} + \vec{v}) \\
&= \vec{u} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{v} \\
&= ||\vec{u}||^2 + ||\vec{v}||^2 + 2\vec{u} \cdot \vec{v}
\end{aligned}$$

Theorem 1.13 (Pythagorean Theorem)**.** .
Two vectors $\vec{u}$ and $\vec{v}$ are orthogonal if and only if $||\vec{u} + \vec{v}|| = ||\vec{u}||^2 + ||\vec{v}||^2$.

Example 1.14. The vector $\vec{0}$ is orthogonal to every vector $\vec{v}$ in $\mathbb{R}^n$ since $\vec{0} \cdot \vec{v} = 0$.

Definition 1.15. Let W be a subspace of $\mathbb{R}^n$.

(a) A vector $\vec{z}$ in $\mathbb{R}^n$ is said to be orthogonal to W if $\vec{z}$ is orthogonal to every vector in W.

(b) The **orthogonal complement** of W, denoted by $W^\perp$, is the collection of all vectors orthogonal to W.

$$W^\perp = \{\vec{z} \in \mathbb{R}^n \mid \ \vec{z} \cdot \vec{w} = 0 \text{ for every } \vec{w} \in W\}$$

Fact 1.16. Let W be a subspace of $\mathbb{R}^n$. Then

(a) $(W^\perp)^\perp = W$

(b) A vector $\vec{z}$ is in $W^\perp$ if and only if $\vec{z}$ is orthogonal to every vector in a set that spans W.

(c) $W^\perp$ is a subspace of $\mathbb{R}^n$.

Theorem 1.17. Let A be an $m \times n$ matrix. Then

$$(\text{Row}(A))^\perp = \text{Nul}(A) \quad \text{and} \quad (\text{Col}(A))^\perp = \text{Nul}(A^T)$$

Example 1.18. Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ be vectors in $\mathbb{R}^3$, and let $W = \text{Span}(\vec{v}_1, \vec{v}_2)$ be a subspace of $\mathbb{R}^3$. Give an explicit description (a parametric form) of the elements of $W^\perp$.

Solution

By Fact 1.16 (b), a vector $\vec{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$ is in $W^\perp$ if and only if

$$\begin{cases} \vec{z} \cdot \vec{v}_1 = 0 \\ \vec{z} \cdot \vec{v}_2 = 0 \end{cases} \quad \text{which is equivalent to} \quad \begin{cases} z_1 - z_3 = 0 \\ 2z_1 - z_2 + z_3 = 0 \end{cases}$$

So, we solve this linear system to find an explicit description of the vector $\vec{z}$. The augmented matrix of this system is given by

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 2 & -1 & 1 & 0 \end{bmatrix}$$

This matrix is row reduced to

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

So a solution $\vec{z}$ of the system is of the form

$$\vec{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} z_3 \\ 3z_3 \\ z_3 \end{bmatrix} = z_3 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \quad z_3 \in \mathbb{R}.$$