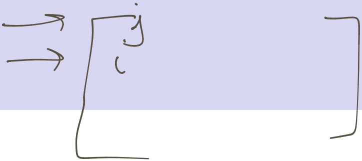
LU Factorization

Definition

If A is an $m \times n$ matrix which can be expressed as

$$A = LU$$

where L is a lower triangular square matrix and U is a matrix in *echelon form*, then this is called an LU-factorization of A.



Assumption on A

It is assumed that A can be reduced to echelon form by using only row replacements of the type $R_i \to R_i + cR_j$, where i is strictly greater than j (in other words, row i is *below* row j.)

Why are we allowed to make this assumption?

Questions about existence and uniqueness of the LU factorization are not strictly a part of the course syllabus, and therefore can be addressed during Office Hours (provided you email me in advance).

$i > j$

If an $m \times n$ matrix A can be reduced to an echelon form using only such row operations, then there exists a sequence of lower triangular matrices $E_1, \ldots, E_p$ such that

$$E_p E_{p-1} \ldots E_1 A = U$$

where U is in echelon form.
Hence

$$A = E_1^{-1} E_2^{-1} \ldots E_p^{-1} U.$$

As the product of lower triangular matrices is lower triangular, the product

$$L = E_1^{-1} E_2^{-1} \ldots E_p^{-1}$$

is a lower triangular matrix.

$E_p \cdots E_3 E_2 E_1 L = I$

This gives us an LU-factorization for the matrix A.

Homework: Prove that the product of 2 lower triangular matrices is lower triangular.

$$R_i \to R_i + cR_j$$
$$i > j$$

Algorithm for an LU factorization

1. Reduce A to an echelon form U by a sequence of row replacement operations of the form above, if possible.
2. Place entries in L such that the same sequence of row operations reduces L to I .

Example

I am not using strict mathematical language

$$A = \begin{bmatrix} 2 & 7 & 1 \\ 3 & -2 & 0 \\ 1 & 5 & 3 \end{bmatrix}$$

$R_2 = R_2 - \frac{3R_1}{2}$

Start from the identity matrix "L" $= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

→ Subtract row 1 multiplied by $\frac{3}{2}$ from row 2 : $R_2 \to R_2 - \frac{3R_1}{2}$.

ok in CS but not acceptable in rigorous math.

$$\to \begin{bmatrix} 2 & 7 & 1 \\ 0 & -\frac{25}{2} & -\frac{3}{2} \\ 1 & 5 & 3 \end{bmatrix}$$

$$\to L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$L = E_1^{-1} E_2^{-1} E_3^{-1}$

Subtract row 1 multiplied by $\frac{1}{2}$ from row 3 : $R_3 \to R_3 - \frac{R_1}{2}$.

$$\begin{bmatrix} 2 & 7 & 1 \\ 0 & -\frac{25}{2} & -\frac{3}{2} \\ 0 & \frac{3}{2} & \frac{5}{2} \end{bmatrix}$$

Write the coefficient $\frac{1}{2}$ in the matrix L at row 3 , column 1 :

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}$$

$I =$

Add row 2 multiplied by $\frac{3}{25}$ to row 3 : $R_3 \to R_3 + \frac{3R_2}{25}$.

$$\begin{bmatrix} 2 & 7 & 1 \\ 0 & -\frac{25}{2} & -\frac{3}{2} \\ 0 & 0 & \frac{58}{25} \end{bmatrix}$$

Write the coefficient $-\frac{3}{25}$ in the matrix L at row 3 , column 2 :

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{3}{25} & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 3/2 & 1 & 0 \\ 1/2 & -3/25 & 1 \end{bmatrix} \begin{bmatrix} 2 & 7 & 1 \\ 0 & -25/2 & -3/2 \\ 0 & 0 & 58/25 \end{bmatrix}$$

Solving a System Using an LU Factorization

We write the equation $A\mathbf{x} = \mathbf{b}$ as

$$LU\mathbf{x} = \mathbf{b}$$

Put $\mathbf{y} = U\mathbf{x}$. We find $\mathbf{x}$ by solving the pair of equations

$$\boxed{L\mathbf{y} = \mathbf{b}}$$
$$U\mathbf{x} = \mathbf{y}$$

in that order.

Let's look at an example.

$$A = \begin{bmatrix} 2 & 7 & 1 \\ 3 & -2 & 0 \\ 1 & 5 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$$

The obvious solution is $(0, 0, 1)$. Let's verify that the LU method yields the same result.

Let us row reduce the matrix $[L \quad \mathbf{b}]$ to reduce echelon form:

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ \frac{3}{2} & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{3}{25} & 1 & 3 \end{bmatrix} \xrightarrow{R_2 \to R_2 - \frac{3}{2}R_1} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{25} & 1 & 3 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 - \frac{1}{2}R_1} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & -\frac{3}{25} & 1 & \frac{5}{2} \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 + \frac{3}{25}R_2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & 1 & \frac{58}{25} \end{bmatrix}$$

So

$$\mathbf{y} = \begin{bmatrix} 1 \\ -\frac{3}{2} \\ \frac{58}{25} \end{bmatrix}$$

Next we solve $U\mathbf{x} = \mathbf{y}$ using back substitution.

$$U \begin{bmatrix} 2 & 7 & 1 \\ 0 & -25/2 & -3/2 \\ 0 & 0 & 58/25 \end{bmatrix} \qquad Ux = \begin{bmatrix} 1 \\ -\frac{3}{2} \\ \frac{58}{25} \end{bmatrix}$$

$$\frac{58}{25} x_3 = \frac{58}{25} \Rightarrow x_3 = 1$$

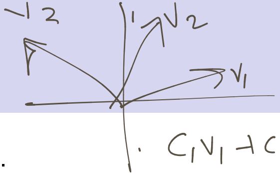
$$-\frac{25}{2} x_2 - \frac{3}{2} x_3 = -\frac{25}{2} x_2 \cancel{-\frac{3}{2}} = \cancel{-\frac{3}{2}}$$

$$\Rightarrow x_2 = 0$$

$$2x_1 + \cancel{7x_2} + x_3 = 2x_1 + \cancel{1} = \cancel{1} \Rightarrow x_1 = 0.$$

→ $\{v_1, v_2, v_3\}$

$\{v_2, v_1, v_3\}$?



$c_1 v_1 + c_2 v_2 = 0$

$c_1 v_2 + c_2 v_1 = 0$

Back to linear independence/dependence ...

Definition

→ An ordered set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} \in \mathbb{R}^n$ is said to be *linearly independent* if the vector equation

$$x_1\mathbf{v}_1 + \ldots + x_p\mathbf{v}_p = 0$$

has only the trivial solution. The (ordered) set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be *linearly dependent* if there exist weights $c_1, \ldots, c_p$, not all zero, such that

at least one is nonzero

$$c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p = 0$$

Why do we care about order? Does reordering a set change independence/dependence?

→ $c_1 v_1 + c_2 v_2 + c_3 v_3 = 0$

→ $c_2 v_2 + c_1 v_1 + c_3 v_3 = 0$

Linear Independence of Matrix Columns

The columns of a matrix A are linearly independent if and only if the equation $A\mathbf{x} = 0$ has only the trivial solution.

Scalar Multiples

A set of two vectors $\{\mathbf{v_1}, \mathbf{v_2}\}$ is linearly dependent if at least one of the vectors is a multiple of the other. The set is linearly independent if and only if neither of the vectors is a multiple of the other.