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.)

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 $E_1, \dots, E_p$ such that

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

where U is in echelon form.
Hence

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

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

$$\underline{L = I\, E_1^{-1} E_2^{-1} \dots E_p^{-1}}$$

is a lower triangular matrix.

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

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. (Alternatively perform flipped column operations on I in the same sequence.)

The "flipped" column operation corresponding to the operation $R_i \to R_i + cR_j$ is $C_j \to C_j - cC_i$.

Example

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

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

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

Column operation on L: $C_1 \to C_1 + \frac{3}{2} C_2$

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

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

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

Column operation on L: $C_1 \to C_1 + \frac{1}{2}C_3$

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

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

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

Column operation on L: $C_2 \to C_2 - \frac{3}{25}C_3$

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

$n \times n$ $\qquad\qquad\qquad$ $A : n \times n$

$$\begin{array}{c} \\ \\ \\ i\text{th row} \rightarrow \\ \\ \\ \end{array}\begin{bmatrix} 1 & & & & & 0 \\ & 1 & & & & \\ & & 1 & & & \\ & c & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ \vdots \\ R_n \end{bmatrix} \begin{array}{c} \\ \\ \leftarrow i\text{th row} \\ \\ \\ \end{array} = \begin{bmatrix} R_1 \\ \vdots \\ R_i + cR_j \\ \vdots \\ R_n \end{bmatrix}$$

$\uparrow$

jth column

$$\begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix}\begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & \\ \vdots & 0 & & \\ \vdots & c & \ddots & \\ 0 & \vdots & & 1 \end{bmatrix} \begin{array}{l} \\ \leftarrow j\text{th entry} \\ \\ \leftarrow i\text{th row} \\ \\ \end{array}$$

j-th column

$$= \begin{bmatrix} c_1 & \cdots & c_j + c\,c_i & \cdots & c_n \end{bmatrix}$$

Definition

An ordered set of vectors $\{v_1, \ldots, v_p\} \in \mathbb{R}^n$ is said to be *linearly independent* if the vector equation

$$x_1 v_1 + \ldots + x_p v_p = 0$$

has only the trivial solution. The (ordered) set $\{v_1, \ldots, v_p\}$ is said to be *linearly dependent* if there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1 v_1 + \ldots + c_p v_p = 0$$

An ordered set of vectors remains linearly independent (or dependent) even if the order is changed. Therefore it is acceptable to say that a set of vectors is linearly independent (or dependent) without actually mentioning the order, unless that particular ordering is actually used elsewhere.

$\rightarrow \underline{Ax = 0}$ $A = [a_1 \ \cdots \ a_n]$

$\rightarrow \underline{x_1 a_1 + \cdots + x_n a_n = 0}$ ← vector equation

Linear Independence of Matrix Columns

The columns of a matrix A are linearly independent if and only if the equation $Ax = 0$ has only the trivial solution.

to be proved

Scalar Multiples

A set of two vectors $\{v_1, v_2\}$ 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.

Proof: At least one of the vectors is a multiple of the other.

Without loss of generality, we may assume that $\exists$ a constant $c \in \mathbb{R}$ such that

there exists.

$$v_1 = c\,v_2$$

$$\Rightarrow v_1 - c\,v_2 = 0$$

$c_1 = 1$

$c_2 = c$

$\Rightarrow$ $\{v_1, v_2\}$ is a linearly dependent set.

A set of two vectors is linearly independent $\iff$ neither of the vectors is a multiple of the other.

We will prove:

A set of two vectors is linearly dependent $\iff$ at least one of the vectors is a multiple of the other

$A \iff B$ logical equivalent to not $A \iff$ not B

Remaining statement:

A set of two vectors is linearly dependent $\Rightarrow$ at least one of the vectors is a scalar multiple of the other.

Proof: $\{v_1, v_2\}$ is linearly dependent.
$\exists\ c_1, c_2 \in \mathbb{R}$, where at least one of c_1, c_2 is nonzero.

such that

$$c_1 v_1 + c_2 v_2 = 0 .$$

Without loss of generality, we may assume that $c_1 \neq 0$.

multiplying both sides by $\frac{1}{c_1}$, we obtain

$$v_1 + \frac{c_2}{c_1} v_2 = 0$$

$$\Rightarrow v_1 = -\frac{c_2}{c_1} v_2 . \quad \square$$

same as ordered (linearly ordered)...

Characterization of Linearly Dependent Sets

An indexed set $S = \{v_1, \ldots, v_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $v_1 \neq 0$, then some v_j (with $j > 1$) is a linear combination of the preceding vectors, $v_1, \ldots, v_{j-1}$.

($\Rightarrow$): Assume that $\{v_1, \ldots, v_p\}$ is linearly dependent.

Since ordering does not change linear dependence, we may assume without loss of generality that $v_1 \neq 0$.

As $\{v_1, \ldots, v_p\}$ is a linearly dependent, $\exists\, c_1, \ldots, c_p \in \mathbb{R}$, not all zero, such that

$$c_1 v_1 + c_2 v_2 + \cdots + c_p v_p = 0.$$

Let $j \in \{1, \ldots, p\}$ be the largest index for which $c_j \neq 0$.

$\Rightarrow$ $c_{j+1}, \ldots, c_p$ are all zero.

$\Rightarrow$ $c_1 v_1 + \cdots + c_j v_j = 0$

Multiplying both sides by $\frac{1}{c_j}$,

we obtain

$$\frac{c_1}{c_j} v_1 + \cdots + \frac{c_{j-1}}{c_j} v_{j-1} + v_j = 0$$

$$\Rightarrow v_j = -\frac{c_1}{c_j} v_1 - \frac{c_2}{c_j} v_2 \cdots - \frac{c_{j-1}}{c_j} v_{j-1}$$

Claim: $j > 1$.

Suppose if possible that $j = 1$.

$\Rightarrow c_2 = c_3 \cdots\cdots = c_p = 0$.

$\Rightarrow c_1 v_1 = 0$.

As $v_1 \neq 0 \Rightarrow c_1 = 0$, $\times$

Hence the claim.

$(\Leftarrow)$: At least one of the vectors in S is a linear combination of the other vectors.

$\Rightarrow \exists\, j \in \{1, \ldots, p\}$

and constants $c_1, \ldots, c_{j-1}, c_{j+1}, \ldots c_p$

such that $\underline{v_j} = c_1 v_1 + \ldots + c_{j-1} v_{j-1} + c_{j+1} v_{j+1} \ldots$

$\ldots + c_p v_p$

$$\Rightarrow c_1 v_1 + \cdots + c_{j-1} v_{j-1} - v_j$$

$$+ c_{j+1} v_{j+1} + \cdots + c_p v_p = 0.$$

$\therefore$ $v_1, \ldots, v_p$ is linearly dependent.

put $c_j = 1$.