2/21/2019

LU Factorization

<u>Last time</u>: We saw by example that we can take $A \in \mathbb{R}^{n \times n}$ and reduce it to an upper triangular matrix U with a sequence of matrices $L_1, L_2, \ldots, L_N$

of the form $I + \mu_{rs} E^{(rs)}$,

where $E^{(rs)}$ is the matrix whose only nonzero entry is in the r^{th} row and s^{th} column:

i.e. if $E^{(rs)} \in \mathbb{R}^{3 \times 3}$, then

$$\mu_{31} E^{(31)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \mu_{31} & 0 & 0 \end{pmatrix}.$$

Also, we saw that if

$$L = I + \mu_{rs} E^{(rs)},$$

then $L^{-1} = I - \mu_{rs} E^{(rs)}$.

<u>Thm</u>: 2.1 These are true for any integer $n \geq 2$.

(i) The product of two lower Δ matrices of order n is lower Δ of order n.

(ii) The product of two <u>unit</u> lower Δ matrices of order n is <u>unit</u> lower Δ of order n.

(iii) a lower Δ matrix is nonsingular $\Leftrightarrow$ all the diagonal elements are nonzero.

(iv) The inverse of a nonsingular lower Δ matrix of order n is lower Δ of order n.

(v) The inverse of a unit lower Δ matrix of order n is unit lower Δ of order n.

<u>Pf of iv</u> (By induction).

Base case: $n = 2$.

let $L = \begin{pmatrix} a & 0 \\ c & b \end{pmatrix}$. Note that we can write down L^{-1} explicitly:

$$L^{-1} = \frac{1}{\det L} \begin{pmatrix} b & 0 \\ -c & a \end{pmatrix}.$$

Since $\det L = ab$, we are implicitly assuming $a \neq 0$ and $b \neq 0$. Thus, L^{-1} is lower Δ.

Inductive hypothesis Assume true for $2 \leq n < k$.

i.e. the inverse of a nonsingular lower Δ matrix of order k is lower Δ of order k.

Prove true for $k+1$. Let L be a lower Δ triangular matrix of order $k+1$.

We can partition L and its inverse L^{-1} as:

$$L = \begin{pmatrix} L_1 & \underline{0} \\ \underline{r}^T & \alpha \end{pmatrix}, \quad L^{-1} = \begin{pmatrix} X & \underline{y} \\ \underline{z}^T & \beta \end{pmatrix}.$$

By definition, $LL^{-1} = I_{k+1}$, so

$$L_1 X = I_k, \quad L_1 \underline{y} = \underline{0}, \quad r^T X + \alpha z^T = \underline{0}^T$$

$$\text{and} \quad \underline{r}^T \underline{y} + \alpha \beta = 1.$$

By induction, since $X = L_1^{-1}$, X is lower Δ. Also, since L_1 is invertible, $\underline{y} = \underline{0}$. This shows L^{-1} is lower Δ.

Computing LU.

How do we actually compute the LU Factorization?

If we assume that A can be factored as $A = LU$, L unit lower Δ and U upper Δ, then express:

$$A = (a_{ij}), \quad L = (l_{ij}), \quad U = (u_{ij}).$$

Then by definition of matrix multiplication

$$a_{ij} = \sum_{k=1}^{n} l_{ik} u_{kj} \quad . \quad 1 \le i, j \le n.$$

Since L and U are lower and upper triangular, the sum only goes to $\min(i,j)$

$$a_{ij} = \sum_{k=1}^{\min(i,j)} l_{ik} u_{kj}.$$

Let's break into cases:

Case 1: $1 \le j < i \le n$

$$a_{ij} = \sum_{k=1}^{j} l_{ik} u_{kj}$$

Case 2: $1 \le i \le j \le n$

$$a_{ij} = \sum_{k=1}^{i} l_{ik} u_{kj}$$

look at case 1:

$$a_{ij} = \sum_{k=1}^{j} l_{ik} u_{kj} = \sum_{k=1}^{j-1} l_{ik} u_{kj} + l_{ij} u_{jj}$$

Solving for l_{ij}, we get

$$\boxed{l_{ij} = \frac{1}{u_{jj}}\left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}\right). \quad \begin{array}{l} i = 2, \ldots, n. \\ j = 1, \ldots, i-1. \end{array}}$$

look at case 2: and use $l_{jj} = 1$, $1 \le j \le n$.

$$a_{ij} = \sum_{k=1}^{i} l_{ik} u_{kj} = \sum_{k=1}^{i-1} l_{ik} u_{kj} + l_{ii} u_{ij}$$

$$\Rightarrow \boxed{u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \quad \begin{array}{l} i = 1, \ldots, n \\ j = i, \ldots, n \end{array}}$$

Example: Consider the matrix:

$$A = \begin{bmatrix} 6 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}.$$

Note that because $L = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$

we can immediately read off the 1st row of U:

$$u_{11} = a_{11} = 6, \quad u_{12} = a_{12} = 2, \quad u_{13} = a_{13} = 3.$$

1st row of L: $\quad l_{11} = 1.$

2nd row of L:

$$l_{21} = \frac{1}{u_{11}} \left(a_{21} - \sum_{k=1}^{0} l_{2k} u_{k1} \right) = \frac{a_{21}}{u_{11}} = \frac{2}{6} = \frac{1}{3}.$$

$$l_{22} = 1.$$

2nd row of U:

$$u_{22} = a_{22} - \sum_{k=1}^{1} l_{2k} u_{k2} = a_{22} - l_{21} u_{12}$$

$$= 2 - \left(\tfrac{1}{3}\right)(2) = 1\tfrac{1}{3}.$$

$$u_{23} = a_{23} - \sum_{k=1}^{1} l_{2k} u_{k3} = a_{23} - l_{21} u_{13} = 0 - \left(\tfrac{1}{3}\right)(3) = -1.$$

3rd row of L:

$$l_{31} = \frac{1}{u_{11}}\left(a_{31} - \sum_{k=1}^{0} \cancel{x}\right) = \frac{a_{31}}{u_{11}} = \frac{3}{6} = \frac{1}{2}.$$

$$l_{32} = \frac{1}{u_{22}}\left(a_{32} - \sum_{k=1}^{1} l_{3k}u_{k2}\right).$$

$$= \frac{-1}{u_{22}} \; l_{31}u_{12} = \frac{-3}{4}\left(\frac{1}{2}\right)(2) = \frac{-3}{4}.$$

$$l_{33} = 1.$$

3rd row of U:

$$u_{33} = a_{33} - \sum_{k=1}^{2} l_{3k}u_{k3}$$

$$= a_{33} - l_{31}u_{13} - l_{32}u_{23}$$

$$= 3 - \left(\frac{1}{2}\right)(3) - \left(-\frac{3}{4}\right)(-1)$$

$$= \frac{3}{2} - \frac{3}{4} = \frac{3}{4}.$$

What about the process for computing LU breaking down? i.e. $u_{jj} = 0$ for some j.

We need a definition:

Thm 2.2: $A \in \mathbb{R}^{n \times n}$, $n \geq 2$.

Suppose every leading principle submatrix $A^{(k)} \in \mathbb{R}^{k \times k}$, with $1 \leq k < n$, is nonsingular. (A is not required to be nonsingular).
Then $A = LU$, where L is unit lower Δ and U is upper Δ.

Definition 2.4: $A \in \mathbb{R}^{n \times n}$, $n \geq 2$.
$1 \leq k \leq n$.
The leading principle submatrix of order k of A is defined as $A^{(k)} \in \mathbb{R}^{k \times k}$ whose element in row i and column j is equal to the element of matrix A in row i and column j, $1 \leq i, j \leq k$.

Pf: (of Thm 2.2)

We do induction on n.

Base case $n = 2$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a \neq 0 \text{ by assumption.}$$

we want to find $L = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}$, $U = \begin{pmatrix} u & v \\ 0 & \eta \end{pmatrix}$

satisfying $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix} \begin{pmatrix} u & v \\ 0 & \eta \end{pmatrix}$.

We have the following equations

$$\begin{cases} a = u. \\ b = v \\ mu = c \\ mv + \eta = d. \end{cases}$$

Since $a \neq 0$, we can solve:

$$m = \frac{c}{u} = \frac{c}{a}.$$

$$\eta = d - mv = d - \left(\frac{c}{a}\right) b.$$

So the base case $n=2$ holds.

$\rightsquigarrow$

Inductive hypothesis:

Statement is true for matrices of order k, $2 \le k < n$.

Prove: Statement is true for matrices of order $k+1$. Let $A \in \mathbb{R}^{k+1 \times k+1}$.

Partition A as follows:

$$A = \begin{pmatrix} A^{(k)} & b \\ c^T & d \end{pmatrix}.$$

We can apply the inductive hypothesis on $A^{(k)}$, i.e. $\exists\, L^{(k)}, U^{(k)}$ so that

$$A^{(k)} = L^{(k)} U^{(k)}.$$

Guess L, U look like

$$L = \begin{pmatrix} L^{(k)} & 0 \\ m^T & 1 \end{pmatrix}, \quad U = \begin{pmatrix} U^{(k)} & v \\ 0^T & \eta \end{pmatrix}.$$

If $A = LU$, we get four equalities:

$$L^{(k)} U^{(k)} = A^{(k)}, \quad L^{(k)} v = b, \quad m^T U^{(k)} = c^T, \quad m^T v + \eta = d.$$

$L^{(k)} v = b$ allows us to uniquely determine v since $L^{(k)}$ is invertible.

Also, note that

$$\det(A^{(k)}) = \det(L^{(k)} U^{(k)})$$

$$= \det(L^{(k)}) \det(U^{(k)})$$

$$= 1 * \det(U^{(k)}).$$

Since $A^{(k)}$ is invertible, $\det(U^{(k)}) \neq 0$, and so $U^{(k)}$ is invertible. Thus

$$m^T U^{(k)} = c^T$$

uniquely determines m.

Then we can use the last equation

$m^T v + \eta = d$ to get η, given m and v.

Some random comments: about "LU-like" Factorizations

<u>Rmk:</u> In our first example, we reduced the system of equations as:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Equivalently, we multiplied A by $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$

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

$$= \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}.$$

So we have $L_1 A = U_1$, where

$$L_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad U_1 = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}.$$

Now, we can write $A = L_1^{-1} U_1$,
with $L_1^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

$$A = L_1^{-1} U_1' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix},$$

This is <u>an</u> LU factorization for A,

but it is not unique! In particular, we can squeeze in $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{a} \end{pmatrix}$
for $a \neq 0$ to get

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{a} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \\ &= \begin{pmatrix} a & 0 \\ a & a \end{pmatrix} \begin{pmatrix} \frac{1}{a} & \frac{1}{a} \\ 0 & -\frac{2}{a} \end{pmatrix}. \end{aligned}$$

<u>Fact</u>: (I think, please check ü). An LU factorization is unique if L is required to be <u>unit</u> lower Δ.

Bose on the nonuniquess of a factorization of A into ~~uppe~~ a product of upper Δ and lower Δ matrizes, we explictly define the (unique) LU Factorization of A as Follows:

<u>Definition</u> The <u>LU Factorization</u> of A,

If it exists, is the product $A = LU$, where L is unit lower Δ and U is upper Δ.

Rmk: let's say we are given an arbitrary matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and we want to transform this to upper Δ, using the techniques just described.
We can multiply the 1^{st} row by $-\frac{c}{a}$ and add to the second row:

$$\begin{pmatrix} 1 & 0 \\ -\frac{c}{a} & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & -\frac{c}{a}(b+d) \end{pmatrix}.$$

But, what if $a \approx 0$. This is not stable. (We will define "not stable" later in the class).

It suffices to say that we need to be careful about dividing by numbers close to zero when we do Gaussian elimination.

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

If we try to use (blindly) our process to convert A to an upper Δ matrix, we are in trouble since $a_{11} = 0$.

But, we can multiply A by a <u>permutation matrix</u> P which rearranges the rows.

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

$$PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

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

Trivially, we have $PA = LU$,

where $L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

<u>Thm</u>: let $n \geq 2$ and $A \in \mathbb{R}^{n \times n}$. There exists a permutation matrix P, a unit lower Δ matrix L, and an upper Δ matrix U, so that $P, L, U \in \mathbb{R}^{n \times n}$ and

$$PA = LU.$$