2/22/2019

LU Factorization with Pivoting

Recall the definition of the LU factorization.

The $\underline{\underline{L}}\underline{\underline{U}}$ Factorization, if it exists, is the product $\underline{\underline{L}}\,\underline{\underline{U}} = \underline{\underline{A}}$, where $\underline{\underline{L}}$ is unit lower triangular and $\underline{\underline{U}}$ is upper triangular.

<u>Remark:</u> In one of our 1st examples,

we reduced a system of equations as:

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

$$\Downarrow$$

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

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

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

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

So we have $\underline{\underline{L}}_1 \underline{\underline{A}} = \underline{\underline{U}}_1$, where

$$\underline{\underline{L}}_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad \underline{\underline{U}}_1 = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$$

$\rightarrow$

Now, we can write $\underline{\underline{A}} = \underline{\underline{L}}_1^{-1}\underline{\underline{U}}_1$, $\underline{\underline{L}}_1^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

$$\underline{\underline{A}} = \underline{\underline{L}}_1^{-1}\underline{\underline{U}}_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}.$$

This is <u>the</u> LU Factorization of $\underline{\underline{A}}$ by our definition, but it is not the only factorization of A into the product of an upper and lower Δ matrices. In particular, we can squeeze in $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{a} \end{pmatrix} = I_{2\times 2}$ for $a \neq 0$.

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

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

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

This is why it is important to require $\underline{\underline{L}}$ to be unit lower Δ for uniqueness.

<u>Last time</u>: We wrote down formules to actually compute the matrices $\underline{\underline{L}} = (l_{ij})$ and $\underline{\underline{U}} = (u_{ij})$.

$$l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right).$$

$i = 2, \ldots, n$ and $j = 1, \ldots, i-1$

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

$i = 1, \ldots, n$ and $j = i, \ldots, n$

When can we do the factorization $\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}$ "safely": We need a condition on $\underline{\underline{A}}$, which is weaker than invertibility.

<u>Definition 2.4</u>: $\underline{\underline{A}} \in \mathbb{R}^{n \times n}$, $n \geq 2$.

$$1 \leq k \leq n.$$

The leading principle submatrix of order k of $\underline{\underline{A}}$ is defined as $\underline{\underline{A}}^{(k)} \in \mathbb{R}^{k \times k}$ whose element in row i and column j corresponds to the (i, j) element of $\underline{\underline{A}}$.

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

Suppose every leading principle submatrix $\underline{\underline{A}}^{(k)}$, with $1 \leq k < n$, is nonsingular. Then we can write $\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}$, where $\underline{\underline{L}}$ is unit lower Δ and $\underline{\underline{U}}$ is upper Δ.

Pf: The proof is by induction.

Base Case: $n = 2$. By assumption $a \neq 0$ if $\underline{\underline{A}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

We want to show the existence of

$$\underline{\underline{L}} = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}, \quad \underline{\underline{U}} = \begin{pmatrix} u & v \\ 0 & \eta \end{pmatrix}$$

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

This is equivalent to the following eqs:

$$a = u, \quad b = v, \quad mu = c, \quad mv + \eta = d.$$

Since $a \neq 0$, we can solve for everything

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

So the base case holds.

Inductive hypothesis: Assume statement is true for matrices of order k, $2 \le k < n$.

Now we try to prove the statement is true for $k+1 = n$. Partition $\underline{\underline{A}} \in \mathbb{R}^{(k+1)\times(k+1)}$ as

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

where $\underline{\underline{A}}^{(k)} \in \mathbb{R}^{k\times k}$. Apply the inductive hypothesis on $\underline{\underline{A}}^{(k)}$, i.e. $\exists\ \underline{\underline{L}}^{(k)}, \underline{\underline{U}}^{(k)}$ so that

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

We guess the form of L and U as:

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

Then $A = LU$ is equivalent to

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

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

Also, note that:

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

Since $\underline{\underline{A}}^{(k)}$ is invertible by assumption, $\det(\underline{\underline{U}}^{(k)}) \neq 0$. Thus:

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

uniquely determines m. The last equation $\underline{m}^T \underline{v} + \eta = d$ allows us to determine η, given m and v. This concludes the proof since

$$\begin{pmatrix} \underline{\underline{L}}^{(k)} & \underline{0} \\ \underline{m}^T & 1 \end{pmatrix} \text{ is unit lower } \Delta$$

$$\text{and } \begin{pmatrix} \underline{\underline{U}}^{(k)} & \underline{v} \\ \underline{0}^T & \eta \end{pmatrix} \text{ is upper } \Delta.$$

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 ~~[illegible]~~ dividing by numbers close to zero when we do Gaussian elimination.

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

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 $\underline{\underline{A}}$ by a permutation matrix $\underline{\underline{P}}$ which rearranges the rows.

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

$$\underline{\underline{P}}\,\underline{\underline{A}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

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

Trivially, we have $\underline{\underline{P}}\,\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}$,

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

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

$$\underline{\underline{P}}\,\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}.$$

Pivoting:

Definition: Suppose $n \geq 2$. A matrix $\underline{\underline{P}} \in \mathbb{R}^{n \times n}$ in which every element is either a 0 or a 1, and whose every row and every column contain exactly 1 nonzero element, is called a permutation matrix.

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

Let's revisit this Theorem:

Theorem 2.3. $n \geq 2$, $\underline{\underline{A}} \in \mathbb{R}^{n \times n}$. There exists a permutation matrix $\underline{\underline{P}}$, a unit lower Δ matrix $\underline{\underline{L}}$, and an upper Δ matrix $\underline{\underline{U}}$, all three in $\mathbb{R}^{n \times n}$, so that

$$\underline{\underline{P}}\,\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}.$$

Pf: By induction: on n.

Base Case: $n=2$. Express $\underline{\underline{A}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

There are three possibilities:

$$\underbrace{a \neq 0}_{①}, \quad \underbrace{a = 0,\ c \neq 0}_{②}, \quad \text{or} \quad \underbrace{a = c = 0}_{③}.$$

In ①, we can procede as usual with Gaussian elimination.

In ②, let $\underline{\underline{P}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Then $$\underline{\underline{P}}\,\underline{\underline{A}} = \begin{pmatrix} c & d \\ 0 & b \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} c & d \\ 0 & b \end{pmatrix}}_{U}.$$

In ③, $$\underline{\underline{A}} = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\underline{\underline{L}}} \underbrace{\begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix}}_{\underline{\underline{U}}}.$$

This proves the Base Case.

Inductive hypothesis: Assume statement is true for $\underline{\underline{A}} \in \mathbb{R}^{k \times k}$, $2 \le k < n$.

Now, we let $\underline{\underline{A}} \in \mathbb{R}^{(k+1) \times (k+1)}$ and prove the statement is true

Find the element in the 1st column of $\underline{\underline{A}}$ with largest absolute value. Say it is α, and it is in row r. Then we can multiply $\underline{\underline{A}}$ by a permutation $\underline{\underline{P}}^{(1r)}$ which swaps the 1st and r^{th} rows to put α in the upper left corner of $\underline{\underline{P}}^{(1r)}\underline{\underline{A}}$. i.e.

$$\underline{\underline{P}}^{(1r)}\underline{\underline{A}} = \begin{pmatrix} \alpha & \underline{w}^T \\ \underline{p} & \underline{\underline{B}} \end{pmatrix}, \quad \text{with}$$

$\underline{p}, \underline{w} \in \mathbb{R}^k$, $\underline{\underline{B}} \in \mathbb{R}^{k\times k}$. Let us

Suppose we can express $\underline{\underline{P}}^{(1r)}\underline{\underline{A}}$ as:

$$\underline{\underline{P}}^{(1r)}\underline{\underline{A}} = \begin{pmatrix} \alpha & \underline{w}^T \\ \underline{p} & \underline{\underline{B}} \end{pmatrix} = \begin{pmatrix} 1 & \underline{0}^T \\ \underline{m} & \underline{\underline{I}}_{k\times k} \end{pmatrix}\begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{C}} \end{pmatrix},$$

where we do not yet know $\underline{m}$, $\underline{\underline{C}}$, and $\underline{v}$. Note that $\underline{\underline{C}} \in \mathbb{R}^{k\times k}$ and $\underline{m}, \underline{v} \in \mathbb{R}^k$. This expression of $\underline{\underline{A}}$ is equivalent to the following equations for $\underline{m}$, $\underline{\underline{C}}$, and $\underline{v}$:

$$\underline{v}^T = \underline{w}^T$$
$$\alpha \underline{m} = \underline{p}$$
$$\underline{\underline{C}} = \underline{\underline{B}} - \underline{m}\,\underline{v}^T.$$

→

Here there are two cases:

$\alpha = 0$: In this case, $\underline{f} = 0$. We can also set $\underline{m} = \underline{0}$, so $\underline{v} = \underline{w}$ and $\underline{\underline{C}} = \underline{\underline{B}}$.

So we have determined $\underline{m}$, $\underline{v}$, $\underline{\underline{C}}$ from our data.

$\alpha \neq 0$: Solving for $\underline{m}$, we get $\underline{m} = \left(\frac{1}{\alpha}\right)\underline{f}$.

Then also $\underline{v} = \underline{w}$ as before and

$$\underline{\underline{C}} = \underline{\underline{B}} - \underline{m}\,\underline{v}^T = \underline{\underline{B}} - \left(\frac{1}{\alpha}\right)\underline{f}\,\underline{w}^T,$$

So $\underline{m}$, $\underline{v}$, $\underline{\underline{C}}$ are determined from our data. Since $\underline{\underline{C}} \in \mathbb{R}^{k \times k}$, we can apply the inductive hyp to conclude

$$\underline{\underline{P}}^* \underline{\underline{C}} = \underline{\underline{L}}^* \underline{\underline{U}}^*.$$

Now, notice that we can write:

$$\begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{C}} \end{pmatrix} = \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{P}}^*\underline{\underline{C}} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{L}}^*\underline{\underline{U}}^* \end{pmatrix} = \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{L}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}.$$

Noting that $\underline{\underline{P}}^* \underline{\underline{P}}^* = \underline{\underline{I}}_{k \times k}$, we can multiply both sides of the previous equation (by $\begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix}$) to get

$$\underbrace{\begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix}}_{= \underline{\underline{I}}_{(k+1) \times (k+1)}} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{C}} \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{C}} \end{pmatrix} = \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{L}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}}_{(\star)}.$$

Now, plug in the equation $(\star)$ for the formula for $\underline{\underline{P}}^{(1r)} \underline{\underline{A}}$:

$$\underline{\underline{P}}^{(1r)} \underline{\underline{A}} = \begin{pmatrix} 1 & \underline{0}^T \\ \underline{m} & \underline{\underline{I}} \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{C}} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \underline{0}^T \\ \underline{m} & \underline{\underline{I}} \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{L}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}.$$

$$= \begin{pmatrix} 1 & \underline{0}^T \\ \underline{m} & \underline{\underline{P}}^* \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{L}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}$$

$\rightsquigarrow$

$$= \begin{pmatrix} 1 & \underline{0}^T \\ \underline{m} & \underline{\underline{P}}^* \underline{\underline{L}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{\underline{P}}^* \underline{m} & \underline{\underline{L}}^* \end{pmatrix} \begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}.$$

where in this last equality, we used again the fact $\underline{\underline{P}}^* \underline{\underline{P}}^* = \underline{\underline{I}}_{k \times k}$.

This string of equalities allows us to conclude:

$$\begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \underline{\underline{P}}^{(1r)} \underline{\underline{A}} = \underbrace{\begin{pmatrix} 1 & \underline{0}^T \\ \underline{\underline{P}}^* \underline{m} & \underline{\underline{L}}^* \end{pmatrix}}_{\underline{\underline{L}}} \underbrace{\begin{pmatrix} \alpha & \underline{v}^T \\ \underline{0} & \underline{\underline{U}}^* \end{pmatrix}}_{\underline{\underline{U}}}$$

with $\begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{\underline{P}}^* \end{pmatrix} \underline{\underline{P}}^{(1r)}$ a permutation matrix, and $\underline{\underline{L}}$ and $\underline{\underline{U}}$ above satisfying the conditions of the theorem.