MATH 204 C03 – PLU-DECOMPOSITION

JON FICKENSCHER

LU-decomposition is a useful computational tool, but this does not work for every matrix. Consider even the simple example

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Exercise. *Prove that no unipotent lower trangular* $\mathbf{L}$ *and upper triangular* $\mathbf{U}$ *exist such that* $\mathbf{LU} = \mathbf{A}$ *in this case.*

In general, we need to perform moves of type 2 ($R_k \to R_k + cR_\ell$) AND moves of type 1 ($R_k \leftrightarrow R_\ell$) when reducing $\mathbf{A}$ to echelon form by Guassian Elimination. This is the motivation behind **PLU**-decomposition. Here $\mathbf{P}$ is a permutation matrix. This may be interpretted in the following way:

Remark. Instead of permuting rows, eliminating entries by addition of rows, permuting rows again, eliminating by addition of rows, permuting rows ... we may instead permute the rows **once** and then reduce our matrix only by type 2 moves.

We first point out some things about permutation matrices and how they interact with other moves (see Section 1). We then sketch the idea of the proof of **PLU**-decomposition (see Section 2).

1. Permutations and Their Matrices

A permutation σ is nothing more than a bijection on the set $\{1, \ldots, n\}$. We denote this as

$$\sigma = \{\sigma(1), \sigma(2), \ldots, \sigma(n)\}$$

For example, the permutation $\sigma = \{2, 4, 1, 3\}$ satisfies $\sigma(1) = 2$, $\sigma(2) = 4$, $\sigma(3) = 1$ and $\sigma(4) = 3$. A permutation σ on $\{1, \ldots, n\}$ has an associated permutation matrix

$$\mathbf{M}_\sigma = [m_{ij}], \text{ where } m_{ij} = \begin{cases} 1, & j = \sigma(i) \\ 0 & \text{otherwise} \end{cases}$$

So if we look at our $\sigma = \{2, 4, 1, 3\}$, we get matrix

$$\mathbf{M}_\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and see that the associated linear transformation $T_\sigma : \mathbb{R}^4 \to \mathbb{R}^4$ is defined by

$$T_\sigma\left(\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^t\right) = \begin{bmatrix} x_2 & x_4 & x_1 & x_3 \end{bmatrix}^t$$

(compare to see that this is consistent with the text).

Date: March 2, 2012.

Now suppose σ and ρ are both permutations on $\{1,\dots,n\}$. Then we may see that

- $\sigma \circ \rho$ is also a permutation.
- We now show that

$$\mathbf{M}_\rho \mathbf{M}_\sigma = \mathbf{M}_{\sigma\circ\rho}$$

Remark. Yes, the order on the right hand side is intentional. If this is confusing, please at least remember that the product of two permutation matrices is itself a permutation matrix and the new permutation is related to the original two.

Proof. Let m_{ij} denote the entries of $\mathbf{M}_\rho$, m'_{ij} denote the entries of $\mathbf{M}_\sigma$ and c_{ij} denote the entris of $\mathbf{M}_\rho\mathbf{M}_\sigma$. Then for each $i, j \in \{1,\dots,n\}$,

$$c_{ij} = \sum_{\ell=1}^{n} m_{i\ell} m'_{\ell j}$$

By definition $m_{i\ell} = 1$ if and only if $\ell = \rho(i)$. Likewise $m'_{\ell j} = 1$ if and only if $j = \sigma(\ell)$. So $m_{i\ell}m'_{\ell j} = 1$ (and not zero) if and only if $j = \sigma(\ell) = \sigma(\rho(i))$.

So c_{ij} will contain many zero elements in the sum and one '1' element if and only if $j = \sigma(\rho(i))$. Therefore the product is the matrix for permutation $\sigma \circ \rho$. □

- The elementary matrix associated to the elementary operation of switching rows is a permutation matrix. Therefore, performing a series of row switches may be represented as a permutation matrix, since it is a product of permutation matrices.

We end with one more result. This concerns the relationship between moves of type 2 and permuting of rows.

Proposition 1. *Let $\mathbf{E} \in M(m,m)$ be the elementary matrix that represents the action $R_k \to R_k + cR_\ell$ and let $\mathbf{M}_\sigma$ be the permutation matrix for σ on $\{1,\dots,m\}$. Then*

$$\mathbf{E}\mathbf{M}_\sigma = \mathbf{M}_\sigma \mathbf{E}'$$

where $\mathbf{E}'$ is the elementary matrix that represents the action

$$R_{\sigma(k)} \to R_{\sigma(k)} + cR_{\sigma(\ell)}.$$

Proof. Let $\mathbf{E}'\mathbf{M}_\sigma = \mathbf{A} = [a_{ij}]$ and $\mathbf{M}_\sigma\mathbf{E} = \mathbf{B} = [b_{ij}]$. Each row of $\mathbf{A}$ is identical to each row of $\mathbf{M}_\sigma$ except for the $\sigma^{-1}(k)^{\text{th}}$ row. This has a '1' in column $\sigma(\sigma^{-1}(k)) = k$, a '$c$' in column $\sigma(\sigma^{-1}(\ell)) = \ell$and '0' in every other place. Likewise, each column of $\mathbf{B}$ is the same as the column of $\mathbf{M}_\sigma$ except the ℓ column. This column has a '1' in row $\sigma^{-1}(\ell)$, a 'c' in row $\sigma^{-1}(k)$ and each other entry is '0.' It is left to the reader to confirm that indeed $\mathbf{A} = \mathbf{B}$. □

Corollary 2. *The same equation holds if $\mathbf{E}'$ represents the action $R_k \to R_k + cR_\ell$ and $\mathbf{E}$ represents the action $R_{\sigma^{-1}(k)} \to R_{\sigma^{-1}(k)} + cR_{\sigma^{-1}(\ell)}$.*

Proof. Exercise. If this is more than one sentence (or two), the proof is too long. *Note that you should directly use the previous proposition.* □

2. PLU-Decomposition

We show this by induction on m, where $\mathbf{A} \in M(m,n)$. The case $m = 1$ is trivial.

Exercise. *Show that there exists a* **PLU***-decomposition for any matrix in the case* $m = 2$.

Now we assume that any matrix with $m \geq 2$ rows has such a decomposition and use this fact to prove that every matrix with $m+1$ rows also has one. Let n_0 denote the first non-zero column of $\mathbf{A} \in M(m+1,n)$. If no such column exists, we are done because $\mathbf{A} = \mathbf{0}$ has the decomposition $\mathbf{I} \cdot \mathbf{I} \cdot \mathbf{A}$. Then we may permute our rows by matrix $\mathbf{Q} \in M(m+1, m+1)$ so that the $(1, n_0)$-entry in

$$\mathbf{QA}$$

is non-zero. We then act by subtracting the first row from the ones below so that each (i, n_0)-entry is zero for $i > 1$. These operations may be represented by a matrix $\mathbf{E} = [e_{ij}]$ which satisfies

$$e_{ij} = \begin{cases} 1, & i = j \\ -c_i, & i > 1, j = 1 \\ 0, & \text{otherwise} \end{cases}$$

where c_i are the constants the correspond to the actions $R_i \to R_i - c_i R_1$ (recall that these actions commute when they all use the same row). Call the resulting matrix $\mathbf{B} = \mathbf{EQA}$ and note that

- $\mathbf{A} = \mathbf{Q}^{-1}\mathbf{E}^{-1}\mathbf{B}$.
- $\tilde{\mathbf{Q}} := \mathbf{Q}^{-1}$ is a permutation matrix and $\tilde{\mathbf{L}} := \mathbf{E}^{-1}$ is unipotent lower triangular.
- If $n_0 = n$, then $\mathbf{B} = \mathbf{U}$ is in echelon form (hence upper triangular), So we have decomposed $\mathbf{A}$, with $\mathbf{P} = \tilde{\mathbf{Q}}$ and $\mathbf{L} = \tilde{\mathbf{L}}$.
- If $n_0 < n$ then we may partition $\mathbf{B}$ into
$$\mathbf{B} = \begin{bmatrix} 0 \dots 0 & b_{1,n_0} & b_{1,n_0+1} \dots b_{1,n} \\ 0 \quad \dots & 0 & \\ \vdots & \vdots & \mathbf{A}' \\ 0 \quad \dots & 0 & \end{bmatrix}.$$
Here, $\mathbf{A}' \in M(m, n-n_0)$, so $\mathbf{A}' = \mathbf{P}'\mathbf{L}'\mathbf{U}'$ for $\mathbf{P}', \mathbf{L}' \in M(m,m)$ by our inductive hypothesis.

Exercise. *We may now express* $\mathbf{B}$ *as*

$$\mathbf{B} = \underbrace{\begin{bmatrix} 1 & 0 \dots 0 \\ 0 & \\ \vdots & \mathbf{P}' \\ 0 & \end{bmatrix}}_{\mathbf{P}''} \underbrace{\begin{bmatrix} 1 & 0 \dots 0 \\ 0 & \\ \vdots & \mathbf{L}' \\ 0 & \end{bmatrix}}_{\mathbf{L}''} \underbrace{\begin{bmatrix} 0 \dots 0 & b_{1,n_0} & b_{1,n_0+1} \dots b_{1,n} \\ 0 \quad \dots & 0 & \\ \vdots & \vdots & \mathbf{U}' \\ 0 \quad \dots & 0 & \end{bmatrix}}_{\mathbf{U}},$$

where $\mathbf{P}''$ *is a permutation matrix,* $\mathbf{L}''$ *is unipotent lower triangular and* $\mathbf{U}$ *is upper triangular.*

So

$$\mathbf{A} = \tilde{\mathbf{Q}}\tilde{\mathbf{L}}\mathbf{P}''\mathbf{L}''\mathbf{U}$$

Because of Proposition 1,

$$\tilde{\mathbf{L}}\mathbf{P}'' = \mathbf{P}''\hat{\mathbf{L}}$$

for some $\hat{\mathbf{L}}$, which has the same form as $\tilde{\mathbf{L}}$ with the constants reordered (the actions of type $R_k \to R_k + cR_1$ commute, and the permutation associated to $\mathbf{P}''$ fixes 1). Hence, if we let $\mathbf{P} = \tilde{\mathbf{Q}}\mathbf{P}''$ (this is a product of permutation matrices) and $\mathbf{L} = \hat{\mathbf{L}}\mathbf{L}''$ (a product of lower triangular unipotent matrices), we finally conclude that

$$\mathbf{A} = \tilde{\mathbf{Q}}\tilde{\mathbf{L}}\mathbf{P}''\mathbf{L}''\mathbf{U} = \tilde{\mathbf{Q}}\mathbf{P}''\hat{\mathbf{L}}\mathbf{L}''\mathbf{U} = \mathbf{PLU}.$$

$$* * *$$

$$* *$$

$$*$$

$$\boxed{\text{Ta-da!}}$$