2.6.4 Rank and Nullity of Matrices

Let $A : \mathcal{U} \to \mathcal{V}$ be an LT, with $\dim \{\mathcal{U}\} = n$, $\dim \{\mathcal{V}\} = m$.

This implies that $A$ has an $m \times n$ matrix representation.

Definition: Rank and Nullity

The rank and the nullity of $A$ are defined by:

\[
\text{rank} \{A\} := \dim \{\mathcal{R}\{A\}\} \leq m \\
\text{nullity} \{A\} := \dim \{\mathcal{N}\{A\}\} \leq n
\]

Theorem:

Let $A$ be defined as above. Then:

\[
\text{rank} \{A\} + \text{nullity} \{A\} = n
\]

Proof: This is an application of the second theorem in L6.

Proposition:

Let $A$ be as defined above. Then:

\[
\text{rank} \{A\} = \text{maximum number of linearly independent columns of } A.
\]

Proof:

Columns of $A$ represent images under $A$ of the basis vectors in $\mathcal{U}$. Therefore, columns of $A$ span $\mathcal{R}\{A\}$. An independent set of such column vectors must form a basis of $\mathcal{R}\{A\}$, and their number is equal to $\dim \{\mathcal{R}\{A\}\}$, which is the rank.

2.6.5 Rank and Determinants

Recall that determinants are well-defined for square matrices $A$ only. The following result is just a special case of a previous theorem on inverses of LT mapping a space into itself.
Proposition:
Let $A$ be a square matrix. Then, $A^{-1}$ exists $\iff$ $A$ is one-to-one $\iff$ $A$ is onto.

Theorem:
Let $A$ be a square matrix. Then,
\[ \det \{ A \} \neq 0 \iff \text{columns of } A \text{ are independent} \iff \text{rows of } A \text{ are independent.} \]

Proof:
\[
\det \{ A \} \neq 0 \iff A^{-1} \text{ doesn't exist (Cramer's rule)} \\
\iff A \text{ is not onto} \\
\iff \dim \{ \mathcal{R} \{ A \} \} < n \\
\iff A \text{ has less than } n \text{ independent columns}
\]

(rows) The rows of $A$ are the columns of $A^T$, and $\det \{ A \} = \det \{ A^T \}$. So the conclusion for rows follows from that of columns.

Theorem:
Consider a rectangular matrix $A \in \mathcal{F}^{m \times n}$. Then:

(a) $\rank \{ A \} =$ maximum number of linearly independent columns of $A$,

(b) $\rank \{ A \} =$ maximum number of linearly independent rows of $A$,

(c) $\rank \{ A \} =$ order of the largest invertible submatrix of $A$.

Proof:
(a) Follows from the results already obtained showing that columns represent vectors spanning $\mathcal{R} \{ A \}$.

(b) Follows from the property that the rank of a matrix remains unchanged after pre- and post-multiplications with elementary matrices (reduction to echelon form seen later.)

(c) Is postponed.
2.6.6 Equivalence Transformations

Definition:

Equivalence transformations are square, invertible, transformations which have the property that multiplication by them preserves rank and nullity, as follows:

Let \( A : U \to V \) be an LT represented by matrix \( A \). If matrix \( P^{-1} \) of \( P : V \to V \) exists, then \( \mathcal{N}\{PA\} = \mathcal{N}\{A\} \). Since nullspace, domain and codomain are unchanged by such postmultiplications, it must be true that rank and nullity of \( PA \) must be the same as for \( A \). However, the range of \( PA \) in general differs from the range of \( A \).

Similarly, pre-multiplications by an invertible matrix \( Q \) to obtain \( AQ \) leaves the range, rank and nullity invariant, but in general changes the nullspace.

Example:

Invariance of nullspace under an equivalence transformation: \( Au = 0 \Rightarrow PAu = 0 \Rightarrow P^{-1}PAu = 0 \).

2.6.7 Elementary Row Operations on Matrix A

The so-called elementary row/column operations (transformations) are in practice, perhaps the most important techniques in matrix calculus, e.g., Gaussian elimination. In our case, we need them to determine the rank, nullity, determinants, and solutions to \( Ax = y \).

Definition: Elementary Row Operations

There are three kinds of elementary row operations for a matrix \( A : \mathbb{R}^n \to \mathbb{R}^m \), e.g., post multiplications by a matrix \( E \), namely,

(RI): \( E_{p,q} = \) Interchange of \( p^{th} \) and \( q^{th} \) row

(RM): \( E_p(\alpha) = \) Multiplication of \( p^{th} \) row by constant \( \alpha \)

(RA): \( E_{p,q}(\alpha) = \) Multiply \( p^{th} \) row by \( \alpha \) and add to \( q^{th} \) row.

Examples:

Let \( A := \begin{bmatrix} a & b & c & d \\ p & q & r & s \\ w & x & y & z \end{bmatrix} \).

(RI): Row interchange. \( E_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \), \( E_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \), \( E_{2,3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \), e.g.
\[ E_{i,j}A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c & d \\ p & q & r & s \\ w & x & y & z \end{bmatrix} = \begin{bmatrix} p & q & r & s \\ a & b & c & d \\ w & x & y & z \end{bmatrix} \]

(RM) : Row multiplication. \( E_{2}(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \), \( E_{2}(\alpha)A = \begin{bmatrix} a & b & c & d \\ \alpha p & \alpha q & \alpha r & \alpha s \\ w & x & y & z \end{bmatrix} \)

(RA) : Multiply and add. \( E_{3,1}(\alpha) = \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \), \( E_{3,2}(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{bmatrix} \), etc.

\[ E_{3,1}(\alpha)A = \begin{bmatrix} a + \alpha w & b + \alpha x & c + \alpha y & d + \alpha z \\ p & q & r & s \\ w & x & y & z \end{bmatrix} \]

Proposition:

The row operations are invertible, and their inverses are of the same type, e.g.,

(RI) : \( E_{p,q} = E_{p,q}^{T} = E_{p,q}^{-1} \)

(RM) : \( E_{p}(\alpha)^{-1} = E_{p}(\alpha^{-1}) \)

(RA) : \( E_{p,q}(\alpha)^{-1} = E_{p,q}(\alpha^{-1}) \)

Note that elementary column operations are similar but involve pre-multiplications by matrices.

<table>
<thead>
<tr>
<th>Elementary matrix</th>
<th>To obtain from identity matrix</th>
<th>Effect on ( E ) (or ( \tilde{E} )) by multiplication on left (right)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{i,j} = RI )</td>
<td>Interchange rows (columns) ( i, j ) of ( I )</td>
<td>( E_{i,j}A = ) interchange ( i^{th} ) and ( j^{th} ) row of ( A ) ( AE_{i,j} = ) interchange ( i^{th} ) and ( j^{th} ) column of ( A )</td>
</tr>
<tr>
<td>( E_{i}(\alpha) = RM )</td>
<td>Multiply the ( i^{th} ) row (column) of ( I ) by ( \alpha )</td>
<td>( E_{i}(\alpha)A = ) Multiplication of row ( i ) of ( A ) by ( \alpha ) ( AE_{i}(\alpha) = ) Multiplication of column ( i ) of ( A ) by ( \alpha )</td>
</tr>
<tr>
<td>( E_{i,j}(\alpha) = RA )</td>
<td>Multiply the ( i^{th} ) row (( j^{th} ) column) by ( \alpha ) and adding it to the ( j^{th} ) row (( i^{th} ) column) of ( I )</td>
<td>( E_{i,j}(\alpha)A = ) Multiplication of row ( i ) of ( A ) by ( \alpha ) and adding it to row ( j ). ( AE_{i,j}(\alpha) = ) Multiplication of column ( j ) of ( A ) by ( \alpha ) and adding it to column ( i ).</td>
</tr>
</tbody>
</table>
Definition: Row/Column Equivalent Matrices

The matrices $A$ and $B$ are row equivalent if $\exists k$ such that $B = E_1 E_2 \cdots E_k A$, where $\{E_j\}_{j=1}^k$ are elementary row operations, and column equivalent if $\exists k$ such that $B = AE_1 E_2 \cdots E_k^\prime$, where $\{E_j^\prime\}_{j=1}^k$ are elementary column operations.

Note: Following the definitions above,

(1) Row operations preserve $\mathcal{N}\{A\}$, rank, nullity but not $\mathcal{R}\{A\}$;

(2) Column operations preserve $\mathcal{R}\{A\}$, rank, nullity, but not $\mathcal{N}\{A\}$.

2.7 Methods of Simplifying Equations

In order to solve $Ax = y$ ($A$, $y$ given), one might try to use Cramer's rule:

$$x = A^{-1}y, \quad A^{-1} = \frac{[\text{cof}_{i,j}\{A\}]}{\det A}, \quad \det\{A\} \neq 0,$$

where the cofactors are $\text{cof}_{i,j}\{A\} = (-1)^{i+j} \det A(i \mid j)$, and satisfy:

$$[\text{cof}_{i,j}\{A\}]^T A = \begin{bmatrix} \det\{A\} & 0 \\ 0 & \det\{A\} \end{bmatrix}.$$

For an $n \times n$ matrix $A$, the computation of $\det\{A\}$ involves multiplications of the order $n!$. For instance, $70! > 10^{100}$, so this method is not practical for anything other than small $n$.

Methods based on elementary operations, such as Gaussian elimination, involve multiplications of the order $n^3$, a number which is close to the optimal. These methods are based on making the matrix triangular.

2.7.1 Triangular Matrices

Next, we show that any matrix $A$ can be reduced to triangular form by elementary operations. Once this is shown, equations of the form $Ax = y$ will be solved.

An $m \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ is upper triangular if $a_{ij} = 0$ whenever $i > j$:
Suppose we are given the equation \( y = Ax \) with \( A \) upper diagonal. Then, the lower part of the matrix equation can be decoupled from the rest and solved separately.

\[
\begin{bmatrix}
  a_{11} & a_{12} & a_{13} & \cdots & a_{1m} & \cdots & a_{1n} \\
  0 & a_{22} & a_{23} & \cdots & a_{2m} & \cdots & a_{2n} \\
  0 & 0 & a_{33} & \cdots & a_{3m} & \cdots & a_{3n} \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\
  0 & 0 & 0 & \cdots & a_{mn} & \cdots & a_{mn}
\end{bmatrix}
\]

That is, \( y_r = A_r x_r \), and we can solve a much simpler equation for \( x_r \). This is particularly easy if the matrix is of the form

\[
\begin{bmatrix}
  1 & 2 & -3 & -1 & 5 \\
  0 & 0 & 5 & 1 & 1 \\
  0 & 0 & 0 & 2
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4 \\
  x_5
\end{bmatrix}
= \begin{bmatrix}
  3 \\
  8 \\
  0
\end{bmatrix}
\]

**Definition: Echelon Matrix**

A matrix \( A \) is an echelon matrix iff the leading nonzero entry of each row after the first is 1, and is to the right of the corresponding entry for the previous row (a staircase with steps of various widths). For example,

\[
\begin{bmatrix}
  0 & 1 & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\
  0 & 0 & 0 & 1 & a_{25} & a_{26} & a_{27} \\
  0 & 0 & 0 & 0 & 1 & a_{36} & a_{37} \\
  0 & 0 & 0 & 0 & 0 & 0 & 1 \\
  0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
2.7.1.1 Properties of Upper-Triangular Echelon Matrices

1) The nonzero rows are independent (in any subset, the leading elements of the top cannot be linear combination of the others as they are zero; row eliminate the top row and repeat the argument for the remaining rows; continue by induction.)

2) Similarly, columns containing leading elements are independent.

3) Columns containing leading elements form a basis spanning $\mathcal{R}\{A\}$.

4) $\text{Rank}\{A\} = \text{number of leading elements}$

5) Equations $y = \tilde{A}x$, where $\tilde{A}$ is an upper-triangular echelon matrix, can be solved sequentially backwards starting with the last row, with low complexity.

2.7.1.2 Row Reduction to Upper-Triangular Echelon Form

Any matrix $A$ can be reduced to upper triangular echelon form $\tilde{A}$ by row operations, which leave $\mathcal{N}\{A\}$ invariant.

Example:

\[
A = \begin{bmatrix}
0 & 0 & 1 & 1 \\
2 & 3 & 4 & 5 \\
2 & 3 & 3 & 2 \\
4 & 6 & 8 & 8
\end{bmatrix}
\quad \text{use row 1 to change row 2 by 0.5}
\quad \text{use row 2 to change rows 3,4}
\quad \text{use row 3 to change 4 rows}
\quad \text{interchange rows 2,3,4}
\rightarrow
\begin{bmatrix}
0 & 0 & 1 & 1 \\
2 & 3 & 0 & 1 \\
2 & 3 & 0 & -1 \\
4 & 6 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 1 & 1 \\
1 & 3/2 & 0 & 1/2 \\
1 & 3/2 & 0 & 1/2 \\
0 & 0 & 0 & -2
\end{bmatrix}
\rightarrow
\begin{bmatrix}
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
= \tilde{A}
\]

Hence,

\[
\text{rank}\{A\} = \text{rank}\{\tilde{A}\} = 3,
\]

\[
\text{nullity}\{A\} = \text{nullity}\{\tilde{A}\} = 1
\]

but $\mathcal{R}\{A\} \neq \mathcal{R}\{\tilde{A}\}$ because of the matrix multiplications on the left.

To find $P$ such that $PA = \tilde{A}$ (product of all elementary matrices), we can append the identity matrix $I_m$ to $A$ to form an extended matrix, and row reduce the extended matrix to upper triangular form. The identity matrix "records" the postmultiplication of the elementary matrices.
\[ A \rightarrow \tilde{A} \rightarrow P \]

Examples:

(a) \[ A = \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \text{ Hence,} \]

\[ E_1 E_2 E_3 A = I \Rightarrow A^\dagger = P = E_4 E_2 E_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/3 & 2/3 \end{bmatrix} \]

(b) \[ A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & -6 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & -6 \end{bmatrix}. \text{ Thus,} \]

\[ P = E_4 E_2 E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \]

(c) \[ A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \end{bmatrix}. \text{ Construct:} \]

\[ \begin{bmatrix} 2 & 3 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & 1 \\ 2 & 3 & 1 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & 1 \\ 0 & -1 & -1 & 1 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & 1 \\ 0 & 1 & 1 & -1 & 2 \end{bmatrix} \]

Hence, \( \tilde{A} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \) and \( P = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \).

### 2.7.2 Finding \( \mathcal{R}(A) \)

Note that \( \mathcal{R}(A) \neq \mathcal{R}(\tilde{A}) \). Instead, let \( K \) be the index set of the columns containing the leading entries of \( \tilde{A} \).

**Proposition:**

Columns of the (untransformed) matrix \( A \) with indices in \( K \) form a basis for \( \mathcal{R}(A) \).

**Example:**
In a previous example, we found \( A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 2 & 3 & 4 & 5 \\ 2 & 3 & 3 & 2 \\ 4 & 6 & 8 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3/2 & 0 & 1/2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \tilde{A} \). The index set is \( K = \{1, 3, 4\} \), hence the columns \( \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 2 \\ 8 \end{bmatrix} \) span \( \mathcal{R}(A) \).

This is true because:

1) row operations preserve the dependence (and independence) of any set of columns because it amounts to a multiplication on the left by an invertible matrix, i.e.,

\[
\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \neq 0 \iff P \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \neq 0,
\]

2) as we have shown, the largest set of independent columns of \( A \) represents a basis for \( \mathcal{R}(A) \).

Alternatively, one could use column operations to map \( A \) into an upper triangular echelon form \( \hat{A} \). The columns of \( \hat{A} \) containing nonzero leading entries would form a basis for \( \mathcal{R}(A) \).