LECTURE 3

I) The Four Fundamental Subspaces of a Matrix

II) Row Rank = Column Rank

III) Pair wise Orthogonality of A

IV) Direct Sum Decomposition of Domain and Target space

V) Orthogonal Complementarity

Discussion Topics
I) THE FOUR FUNDAMENTAL SUBSPACES OF A LINEAR TRANSFORMATION.

Consider an $m \times n$ matrix

$$A : \mathbb{R}^n \rightarrow \mathbb{R}^m; \quad A = \begin{bmatrix} \vdots \end{bmatrix}^T \quad m \text{ or } A = \begin{bmatrix} \vdots \end{bmatrix}^T \quad n$$

and its transpose

$$A^T : \mathbb{R}^m \rightarrow \mathbb{R}^n; \quad A^T = \begin{bmatrix} \vdots \end{bmatrix}^T \quad n \text{ or } A^T = \begin{bmatrix} \vdots \end{bmatrix}^T \quad m$$

Thus one may have

$$A = \begin{bmatrix} \vdots \end{bmatrix}^T \quad m \text{ and } A^T = \begin{bmatrix} \vdots \end{bmatrix}^T \quad n$$

or

$$A = \begin{bmatrix} \vdots \end{bmatrix}^T \quad m \text{ and } A^T = \begin{bmatrix} \vdots \end{bmatrix}^T \quad n$$
Such matrices are characterized by its four fundamental subspaces:

\[ \mathbf{R}(A) = (\text{column space of } A) \subseteq \mathbb{R}^m; \ \dim = r \]

\[ \mathbf{N}(A) = (\text{null space of } A) \subseteq \mathbb{R}^n; \ \dim = n - r \]

\[ \mathbf{R}(A^\top) = (\text{row space of } A) \subseteq \mathbb{R}^m; \ \dim = r \quad (4-?) \]

\[ \mathbf{N}(A^\top) = (\text{left nullspace of } A) \subseteq \mathbb{R}^m; \ \dim = m - r \quad (4-?) \]

II) \ \text{ROW RANK} = \text{COLUMN RANK}

It is blindingly obvious that row space \( \mathbf{R}(A^\top) \) and the column space \( \mathbf{R}(A) \) of \( A \) are in general two entirely different vector spaces. Nevertheless one has the following

**Theorem** (Row rank = column rank)

\[ \dim \mathbf{R}(A) = \dim \mathbf{R}(A^\top) \]

i.e. column rank of \( A \) = row rank of \( A \).
**THEOREM:** Row rank = Column rank

The validity of this theorem is established by means of a 4 step process.

**STEP 1**
Row reduction to echelon form.

\[
A = \begin{bmatrix}
    a_{11} & \cdots & a_{1n} \\
    a_{21} & \cdots & a_{2n} \\
    a_{31} & \cdots & a_{3n} \\
    \vdots & \ddots & \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{bmatrix}
\rightarrow U = \begin{bmatrix}
    d_1 & * & * & * & * & * & * & * & * & * & * & * & * & * & * & * \\
    0 & \cdots & 0 & d_2 & * & * & * & * & * & * & * & * & * & * & * & * \\
    0 & \cdots & 0 & 0 & \cdots & 0 & d_3 & * & * & * & * & * & * & * & * & * \\
    \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & d_k & * & * & * & * & * & * & * \\
    0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
    \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\end{bmatrix}
\]

(1)

The matrix \( A \) is \( m \times n \). The matrix \( U \), which is also \( m \times n \), has \( k \) non-zero rows with non-zero pivot \( d_1, d_2, d_3, \ldots, d_k \) and \( (m - k) \) rows which are zero.

**STEP 2**
The row space of \( A \) and of \( U \) are the same. Row operations guarantee this. Furthermore, being in reduced form, \( U \) has a linearly independent set of \( k \)
row vectors $u_1, u_2, u_3, \cdots, u_k$:

$$U = \begin{bmatrix}
0 \cdots 0 & u_1 \\
0 \cdots 0 & 0 & u_2 \\
\vdots \\
0 \cdots 0 & 0 & 0 & \cdots & 0 & u_k \\
0 \cdots 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0
\end{bmatrix}$$

$$= \begin{bmatrix}
\bar{u}_1 \\
\bar{u}_2 \\
\bar{u}_3 \\
\vdots \\
\bar{u}_k \\
0 \\
\vdots \\
0
\end{bmatrix}$$

These row vectors form a basis for the row space of $U$ and for that of $A$. Consequently,

"row rank" $\equiv \dim \mathcal{R}(A^T) = k$ \hspace{1cm} (2)

**STEP 3**

Row operations are reversible. Consequently, the row reduction process is reversible. Thus the rows of $A$ are linear combinations of the rows of $U$, namely, the $k$ basis elements $\bar{u}_1, \bar{u}_2, \bar{u}_3, \cdots,$ and $\bar{u}_k$:

$$[a_{11} \cdots a_{1n}] = c_{11}\bar{u}_1 + c_{12}\bar{u}_2 + \cdots + c_{1k}\bar{u}_k$$

$$\vdots$$

$$[a_{m1} \cdots a_{mn}] = c_{m1}\bar{u}_1 + c_{m2}\bar{u}_2 + \cdots + c_{mk}\bar{u}_k$$
or

\[
\begin{bmatrix}
    a_{11} & \cdots & a_{1n} \\
    \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{bmatrix}_{m \times n}
= \begin{bmatrix}
    c_{11} & c_{12} & \cdots & c_{1k} & 0 & \cdots & 0 \\
    c_{21} & c_{22} & \cdots & c_{2k} & 0 & \cdots & 0 \\
    \vdots & & & \vdots & & & \vdots \\
    c_{m1} & c_{m2} & \cdots & c_{mk} & 0 & \cdots & 0
\end{bmatrix}_{m \times m}
= \begin{bmatrix}
    \bar{u}_1 \\
    \bar{u}_2 \\
    \vdots \\
    \bar{u}_k \\
    0 \\
    \vdots \\
    0
\end{bmatrix}_{m \text{ rows } \times n \text{ columns}}
\]

or more compactly,

\[ A = CU. \]

The coefficients \( c_{ij} \), which make up the \( m \times m \) matrix \( C \), are uniquely determined by the row reduction process \( A \rightarrow U \).

STEP 4

This step contrasts two different ways of viewing the product of two matrices.

Instead of viewing \( A = CU \) as reconstructing the rows of \( A \) as linear combinations of the rows \( u_1, u_2, \ldots, u_k \) of \( U \), one now views the columns of \( A \) as linear combinations of the columns of \( C \). Note that there are only \( k \) non-zero such columns. Consequently,

\[
\mathcal{R}(A) = \text{Span} \left( \left\{ \begin{bmatrix} c_{i1} \\ \vdots \\ c_{mi} \end{bmatrix} \right\}_{i=1}^{k} \right)
\]

and therefore

\[
\dim \mathcal{R}(A) \leq k = \dim \mathcal{R}(A^T) \quad \text{(using Eq.(2) on page 2).}
\]

Finally, by reversing the roles of \( A \) and \( A^T \) and applying the previous four steps one obtains the analogous inequality

\[
\dim \mathcal{R}(A^T) \leq \dim \mathcal{R}(A).
\]
These last two inequalities imply

$$\dim \mathcal{R}(A^T) = \dim \mathcal{R}(A),$$

i.e. for any matrix $A$, its row rank equals its column rank indeed.
The validity of this theorem is established by means of a 4 step process.

**Step 1** Row reduction to echelon form

\[
A = \begin{bmatrix}
  a_{11} & \cdots & a_{1n} \\
  a_{21} & \cdots & a_{2n} \\
  \vdots & \ddots & \vdots \\
  a_{m1} & \cdots & a_{mn}
\end{bmatrix}
\]

\[
U = \begin{bmatrix}
  d_1 & * & * & \cdots & * \\
  0 & d_2 & * & \cdots & * \\
  0 & 0 & d_3 & \cdots & * \\
  \vdots & \ddots & \ddots & \ddots & \ddots \\
  0 & 0 & 0 & \cdots & d_m
\end{bmatrix}
\]

\[m \times n\] with \(k\) non-zero rows and \(m-k\) rows which are zero.

**Step 2**

Note that, being in echelon form, \(U\) has \(k\) linearly independent rows. Moreover, they span the row space of \(U\). Thus \(\{u_1, u_2, \ldots, u_k\}\) is a basis for the row space of \(U\).

The row reduction process consists of:

(a) interchanging a pair of rows

(b) multiplying a row by a non-zero constant

(c) add one row to another.
STEP 2
The row space of $A$ and of $U$ are the same. Row operation guarantees this.
Furthermore, being in echelon form, $U$ has $k$ linearly independent rows:

$$U = \begin{bmatrix}
-u_1 \\
u_2 \\
u_3 \\
\vdots \\
u_k \\
0 \\
\vdots \\
0
\end{bmatrix} = \begin{bmatrix}
1 & -u_1 \\
0 & u_2 \\
0 & u_3 \\
\vdots \\
0 & -u_k \\
0 & 0 \\
\vdots \\
0
\end{bmatrix}
$$

These row vectors form a basis for the row space of $U$ and for that of $A$, consequently

"row rank" = $\dim R(A^T) = k$. 
STEP 3

Row operations are reversible.

Consequently, the row reduction process is reversible.

Thus, the rows of $A$ are linear combinations of the rows of $U$, namely the $k$ basis elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$:

$$[a_{11} \ldots a_{1n}] = c_{11} \mathbf{u}_1 + c_{12} \mathbf{u}_2 + \ldots + c_{1k} \mathbf{u}_k$$

$$\vdots$$

$$[a_{m1} \ldots a_{mn}] = c_{m1} \mathbf{u}_1 + c_{m2} \mathbf{u}_2 + \ldots + c_{mk} \mathbf{u}_k$$

or

$$\begin{bmatrix}
  a_{11} & \cdots & a_{1n} \\
  \vdots & & \vdots \\
  a_{m1} & \cdots & a_{mn}
\end{bmatrix} =
\begin{bmatrix}
  c_{11} & c_{12} & \cdots & c_{1k} & 0 & \cdots & 0 \\
  c_{21} & c_{22} & \cdots & c_{2k} & 0 & \cdots & 0 \\
  \vdots & & & & \vdots & & \vdots \\
  c_{m1} & c_{m2} & \cdots & c_{mk} & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
  \mathbf{u}_1 \\
  \mathbf{u}_2 \\
  \vdots \\
  \mathbf{u}_k
\end{bmatrix} =
\begin{bmatrix}
  -u_1 \\
  -u_2 \\
  \vdots \\
  -u_k
\end{bmatrix}$$

or more compactly

$$A = CU$$
The constants $c_{ij}$, which make up the $m \times m$ matrix $C$, are uniquely determined by the row reduction process $A \rightarrow U$.

**STEP 4**

This step is the most subtle:

Instead of viewing $CU$ as reconstructing the rows of $A$ in terms of the rows $u_1, ..., u_k$ of $U$, one now views $CU$ as reconstructing the columns of $A$ in terms of the columns of $C$.

Note that there are only $k$ of them.

Consequently,

$$R(A) = \text{span} \left\{ \begin{bmatrix} c_{1,i} \\ \vdots \\ c_{m,i} \end{bmatrix} \right\}_{i=1}^{k}$$
and
\[ \dim \mathcal{R}(A) \leq \mathcal{R} = \dim \mathcal{R}(A^T) \]  \( (*) \)

Now reverse the roles of \( A \) and \( A^T \). The result is
\[ \dim \mathcal{R}(A^T) \leq \dim \mathcal{R}(A) \]  \( (**) \)

It follows that
\[ \dim \mathcal{R}(A^T) = \dim \mathcal{R}(A) \]

i.e.
row rank = column rank.
III) PAIRWISE ORTHOGONALITY

The second important prop of 31.6

The Four Fundamental subspaces and their

pairwise orthogonality,

\[ \text{In } \mathbb{R}^m, \quad R(A) \perp N(A^T) \]

\[ \text{In } \mathbb{R}^m, \quad R(A^T) \perp N(A) \]

This property is based on the

Definition

Two subspaces \( V \) and \( W \) of \( \mathbb{R}^n \) are said to be orthogonal \( (V \perp W) \) if every \( v \in V \) is orthogonal to every \( w \in W \),

and on the

Theorem

For any \( m \times n \) matrix \( A: \mathbb{R}^n \rightarrow \mathbb{R}^m \) on has

(a) \( N(A) \perp R(A^T) \)

(b) \( N(A^T) \perp R(A) \)
The proof is by inspecting the elements of the two nullspaces

(a) \( w \in N(A) \Rightarrow A w = 0 \)

\[
\begin{bmatrix}
\text{row 1} & \vdots & \text{row m}
\end{bmatrix}
\begin{bmatrix}
w_1 \\
\vdots \\
w_m
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
\vdots \\
0
\end{bmatrix}
\]

\( \text{m equations} \)

\( (i^{th} \text{ row of } A) w = (i^{th} \text{ column of } A^T)^T w = 0 \quad i = 1, 2, \ldots, m \)

\( R(A^T) \perp w = 0 \quad \forall w \in N(A) \)

\( R(A^T) \perp N(A) \).

(b) \( y \in N(A^T) \Rightarrow (y^T A = 0 \iff A^T y = 0) \)

\[
\begin{bmatrix}
y_1 \\
\vdots \\
y_m
\end{bmatrix}
\begin{bmatrix}
c_1 & \cdots & c_n
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
\vdots \\
0
\end{bmatrix}
\]

\( \text{n eqns} \)

\( y \perp R(A) \quad \forall y \in N(A^T) \)

\( N(A^T) \perp R(A) \).

**Comment:** These orthogonalities imply that \( A \) induces a unique decomposition of the domain space \( \mathbb{R}^m \) and of the target space \( \mathbb{R}^m \) as follows
a) domain $\mathcal{A} = \mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{R}(A^T)$

b) domain $A^T$ = target $\mathcal{A} = \mathbb{R}^m = \mathcal{N}(A^T) \oplus \mathcal{R}(A)$.

where $\mathcal{N}(A) \cap \mathcal{R}(A^T) = \{0\}$ in $\mathbb{R}^m$

$\mathcal{N}(A^T) \cap \mathcal{R}(A) = \{0\}$ in $\mathbb{R}^m$

Thus $w \in \mathcal{N}(T) \cap \mathcal{R}(A^T)$

\[
\begin{bmatrix}
\text{row}(1) \\
\cdots \\
\text{row}(m)
\end{bmatrix}
\begin{bmatrix}
w_1 \\
\cdots \\
w_m
\end{bmatrix} = 0
\]

$w^T = \frac{1}{m} \sum_{k=1}^{m} x_k \text{row}(k)$

\[
\begin{bmatrix}
w_1 \\
\cdots \\
w_m
\end{bmatrix} = 0
\]

$w^T w = \sum x_k^2 \text{row}(k)^T w_k = 0$

but $w^T w = (w_1)^2 + \cdots + (w_m)^2$
Comment:
The orthogonality of the subspaces of $A$ constitutes the Fundamental Theorem of Linear Algebra.
It is stated most succinctly in terms of orthogonal complements.

Definition:
Given a subspace $W \subseteq \mathbb{R}^n$, the space of vectors orthogonal to $W$ is called the orthogonal complement of $W$ and is denoted by $W^\perp$.

**FUNDAMENTAL THEOREM OF LINEAR ALGEBRA**

\[
N(A) = (R(A^T))^\perp \\
R(A^T) = (N(A))^\perp \\
N(A^T) = (R(A))^\perp \\
R(A) = (N(A^T))^\perp
\]

\[\mathbb{R}^n = N(A) \oplus R(A)\]

\[\mathbb{R}^m = R(A) \oplus N(A^T)\]

Domain space

Target space