Last time: 1) Additional examples on the four fundamental subspaces of a matrix

2) Orthogonal vectors

Def: The inner product (dot product) of two vectors \( x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \) and \( y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n \) is defined as

\[
\langle x, y \rangle = x \cdot y = \sum_{j=1}^{n} x_j y_j
\]

Useful properties (of the inner product for vectors in \( \mathbb{R}^n \))

i) \( \langle x, y \rangle = x^T y \)

ii) \( \langle x, y \rangle = \langle y, x \rangle \)

iii) \( \langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle \) for vectors \( x, y, z \in \mathbb{R}^n \) and scalars \( a, b \in \mathbb{R} \)

iv) \( \langle x, Ay \rangle = \langle A^T x, y \rangle \) for \( A \in \mathbb{R}^{m \times n} \), \( x \in \mathbb{R}^n \), \( y \in \mathbb{R}^m \)

v) \( \langle x, x \rangle = \sum_{j=1}^{n} |x_j|^2 = \|x\|^2 \) for \( x \in \mathbb{R}^n \)

vi) \( |\langle x, y \rangle| \leq \|x\|_2 \|y\|_2 \) (Cauchy-Schwarz inequality)

vii) \( \langle x, y \rangle = \|x\|_2 \|y\|_2 \cos \theta \) where \( \theta \) is the angle between \( x \) and \( y \)

Def: We call two vectors \( x, y \) orthogonal if the angle between them is \( \frac{\pi}{2} \)

Fact: Two non-zero vectors \( x, y \) are orthogonal if and only if \( \langle x, y \rangle = 0 \)
Example: Determine if any pair among

\[ u = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}, \quad v = \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}, \quad w = \begin{bmatrix} 2 \\ 9 \\ 6 \end{bmatrix} \]

is orthogonal.

Sol: We have

\[ \langle u, v \rangle = (2)(3) + (-1)(2) + (5)(-4) + (-2)(0) = -16 \neq 0 \]

\[ \Rightarrow u, v \text{ are not orthogonal.} \]

\[ \langle u, w \rangle = 17 \neq 0 \quad \Rightarrow u, w \text{ are not orthogonal} \]

\[ \langle v, w \rangle = 24 - 24 = 0 \quad \Rightarrow v, w \text{ are orthogonal.} \]
Orthogonal subspaces

\[ S_1 : \text{xy-plane in } \mathbb{R}^3 = \text{span} \{ e_1, e_2 \} \]
\[ S_2 : \text{z-plane} = \text{span} e_3 \]

Observe: Let \( u \in S_1 \) be arbitrary and non-zero. Let \( v \in S_2 \) be arbitrary and non-zero.

Then \( \langle u, v \rangle = 0 \), i.e. \( u \perp v \).
Defn. Let $S_1$ and $S_2$ be two subspaces. We say $S_1$ and $S_2$ are orthogonal if and only if

$$\forall u \in S_1, \forall v \in S_2 \text{ we have } \langle u, v \rangle = 0$$

In this case we write $S_1 \perp S_2$.

**How to determine if two subspaces are orthogonal?**

**Thm.** Two subspaces are orthogonal iff their respective basis vectors are orthogonal.

**Remark:** The theorem says the following:

Let $S_1 = \{ b_1, b_2, \ldots, b_k \}$ be a basis of $S_1$ and let $S_2 = \{ c_1, c_2, \ldots, c_k \}$ be a basis of $S_2$.

Then

$$S_1 \perp S_2 \text{ iff } \langle b_j, c_i \rangle = 0 \text{ for } 1 \leq i, j \leq k$$
How to make use of this theorem?

Let $S_1$ and $S_2$ be two subspaces of $\mathbb{R}^n$.

Let $\{b_1, b_2, \ldots, b_k\}$ be a basis of $S_1$, and $\{c_1, c_2, \ldots, c_l\}$ be a basis of $S_2$.

Then

$$B = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix} \in \mathbb{R}^{n \times k} \quad \text{($n \times k$) matrix}$$

$$C = \begin{bmatrix} c_1 & \cdots & c_l \end{bmatrix} \in \mathbb{R}^{n \times l} \quad \text{($n \times l$) matrix}$$

and thus

$$B^T C = \begin{bmatrix} b_1^T c_1 & b_1^T c_2 & \cdots & b_1^T c_l \\ b_2^T c_1 & b_2^T c_2 & \cdots & b_2^T c_l \\ \vdots & \vdots & \ddots & \vdots \\ b_k^T c_1 & b_k^T c_2 & \cdots & b_k^T c_l \end{bmatrix}$$

Therefore

$$S_1 \subseteq S_2 \iff B^T C = 0 \iff \text{($l \times l$ zero matrix)}$$
**Def:** Let $U$ be a subspace of a vector space $W$. We define

$$U^\perp := \{ w \in W : w \perp U \}$$

where $w \perp U$ means that $w \cdot u = 0$ for all $u \in U$.

We call $U^\perp$ the orthogonal complement of $U$ in $W$.

Note that $U^\perp$ is a subspace of $W$ (Homework).

**Example:**

1) Let $S_1 = \text{xy-plane in } \mathbb{R}^3 = \text{span } \{e_1, e_2\}$

$S_2 = \text{z-plane in } \mathbb{R}^3 = \text{span } \{e_3\}$

Then $S_2 = S_1^\perp$ (and $S_1 = S_2^\perp$) in $W = \mathbb{R}^3$

2) $U = \text{span } \{e_1, e_3\}$ in $W = \mathbb{R}^5$.

Then $U^\perp = \text{span } \{e_2, e_4, e_5\}$
Facts: 1) Let $U$ be a subspace of a vector space $W$. Any $x \in W$ can be written (uniquely) as

$$x = x_U + x_{U^\perp}$$

where $x_U \in U$ and $x_{U^\perp} \in U^\perp$.

Note: This is an orthogonal decomposition. Later we will say that $x_U$ is the orthogonal projection of $x$ onto $U$.

2) $(U^\perp)^\perp = U$

3) If $U$ is subspace of $W$ and $\dim(W) = n$, then

$$\dim(U^\perp) = n - \dim(U)$$

4) If $B$ is a basis for a subspace $U$ of $W$ and $B_\perp$ is a basis for $U^\perp$ then $B \cup B_\perp$ is a basis for $W$. 
From these facts we obtain

Let \( U \) and \( U^\perp \) in \( W \) be given, then

\[ W = \{ x+y : x \in U \text{ and } y \in U^\perp \} \]

Note: \( W = U \cup U^\perp \)

Relations of \( W(A), W(A^\top), R(A), R(A^\top) \)

Then: Let \( A \) be a linear transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^m \)

Then

\[ \text{rank}(A) = \text{rank}(A^\top) \]

Proof: A can be represented as a \( m \times n \) matrix which we call again \( A \).

Let the row rank of \( A \) be \( r \), i.e.

\[ \text{rank}(A^\top) = r. \]

and let \( x_1, x_2, \ldots, x_r \) be a basis of the row space of \( A \).

Claim: The vector \( Ax_1, Ax_2, \ldots, Ax_r \) are linear independent.

Proof of claim: Let \( c_1, \ldots, c_r \) be coefficients such that

\[ 0 = c_1 Ax_1 + c_2 Ax_2 + \ldots + c_r Ax_r \]
0 = c_1 A x_1 + c_2 A x_2 + ... + c_r A x_r
= A (c_1 x_1 + c_2 x_2 + ... + c_r x_r)

1. Observation: \( V = c_1 x_1 + c_2 x_2 + ... + c_r x_r \) is a linear combination of vectors in the row space of \( A \), thus \( V \) is an element of \( \text{R}(A^T) \).

2. Observation: \( A V = 0 \), thus \( V \) is orthogonal to every vector of \( A \) and therefore orthogonal to every vector in \( \text{R}(A^T) \).

\[ \Rightarrow \text{ } V \text{ must be orthogonal to itself and thus } V = 0. \]

Since the vectors \( x_1, ..., x_r \) are basis vectors and therefore by definition linearly independent, the only way that \( V = c_1 x_1 + c_2 x_2 + ... + c_r x_r = 0 \) is that \( c_1 = c_2 = ... = c_r = 0 \). Thus \( A x_1, A x_2, ..., A x_r \) are linearly independent. This concludes the proof of the claim.
Both to the proof of the theorem:

The vectors $A x_1, A x_2, \ldots, A x_r$ are obviously in the range of $A$. So $\{A x_1, A x_2, \ldots, A x_r\}$ is a set of linearly independent vectors in $\mathbb{R}(A)$ and hence the rank of $A$ must be at least as big as $r = \text{rank}(A^T)$. Thus

\[ \text{rank}(A^T) = r \leq \text{rank}(A) \]

If we now apply this result to the transpose of $A$ we obtain

\[ \text{rank}((A^T)^T) \leq \text{rank}(A^T) \]

\[ \Rightarrow \text{rank}(A^T) = \text{rank}(A^T) \]

Therefore we have proved: $\text{rank}(A) = \text{rank}(A^T)$
Theorem: Let $A$ be an $m \times n$ matrix. Then

1. $\text{N}(A) = \left[ \text{R}(A^T) \right]^\perp$

2. $\text{W}(A^T) = \left[ \text{R}(A) \right]^\perp$

Remark: By replacing $A$ with $A^T$ we have that (1) $\Rightarrow$ (2)

Proof of (1): To prove (1) we need to prove

a) $\text{W}(A) \subseteq \left[ \text{R}(A^T) \right]^\perp$

b) $\left[ \text{R}(A^T) \right]^\perp \subseteq \text{N}(A)$

c) Let $x \in \text{W}(A)$ arbitrary.

Then $A x = 0 \Rightarrow < y, A x > = 0 \forall y \in \mathbb{R}^m$

$\Leftrightarrow < A^T y, x > = 0 \forall y \in \mathbb{R}^m$

$\Leftrightarrow x \perp A^T y \text{ for all } y \in \mathbb{R}^m$
\[ \Rightarrow x \perp \{ A^T y : y \in \mathbb{R}^m \} \]
\[ \Rightarrow x \perp \mathbb{R}(A^T) \]
\[ \Rightarrow x \in \left[ \mathbb{R}(A^T) \right]^\perp \]

Since \( x \in \mathbb{W}(A) \) was arbitrary we have that \( \mathbb{W}(A) \subseteq \left[ \mathbb{R}(A^T) \right]^\perp \)

b) Let \( x \in \left[ \mathbb{R}(A^T) \right]^\perp \) be arbitrary

\[ x \in \left[ \mathbb{R}(A^T) \right]^\perp \iff x \perp \mathbb{R}(A^T) \]

\[ \iff \langle x, A^T y \rangle = 0 \quad \forall y \in \mathbb{R}^m \]

\[ \iff \langle A^T x, y \rangle = 0 \quad \forall y \in \mathbb{R}^m \]

Note that this holds for all \( y \in \mathbb{R}^m \), so in particular for \( y = A^T x \)

Therefore

\[ \langle A^T x, A^T x \rangle = 0 \]

\[ \iff \| A^T x \|^2 = 0 \implies A^T x = 0 \]

and thus we shown that \( x \in \mathbb{W}(A) \). Since \( x \in \left[ \mathbb{R}(A^T) \right]^\perp \) was arbitrary, this concludes the proof.
Why is it useful to know that $W(A) = [R(A^T)]^+$?

Because it may make your life ("the exam") much easier.

Q1: Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 0 \\ 2 & 5 & 1 \end{bmatrix}$.

Does there exist $x$ such that $Ax = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$?

Q2: (equivalent to Q1)

Is $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ in $R(A)$?

Q3: (equivalent to Q1 and Q2)

Is $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ in $[N(A^T)]^T$?

Q4: (equivalent to Q1, Q2 and Q3)

Is $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ orthogonal to $N(A^T)$?