

Lecture 9: 3.1 Orthogonal vectors and subspaces.

The **inner product** or **dot product** between two vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = [x_1 \ \cdots \ x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + \cdots + x_n y_n.$$

The dot product satisfy $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ and $\mathbf{x} \cdot \mathbf{x} > 0$, if $\mathbf{x} \neq \mathbf{0}$.

The **length** of the vector \mathbf{x} is $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$

We know this is the Euclidean length in two dimensions by the Phytagorean law.

The **distance** between \mathbf{x} and \mathbf{y} is $\text{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{y} - \mathbf{x}\|$.

Two vectors \mathbf{x} and \mathbf{y} are called **orthogonal** if $\mathbf{x} \cdot \mathbf{y} = 0$.

If \mathbf{x} is perpendicular to \mathbf{y} then it is perpendicular to $-\mathbf{y}$ and $\text{dist}(\mathbf{x}, \mathbf{y}) = \text{dist}(\mathbf{x}, -\mathbf{y})$:

$$\text{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2 = (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - \mathbf{x} \cdot \mathbf{y} - \mathbf{y} \cdot \mathbf{x} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x} \cdot \mathbf{y},$$

$$\text{dist}(\mathbf{x}, -\mathbf{y}) = \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\mathbf{x} \cdot \mathbf{y}.$$

It follows that $\mathbf{x} \cdot \mathbf{y} = 0$ if \mathbf{x} and \mathbf{y} are perpendicular so it is the same as orthogonal.

The **Pythagorean law**: $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

Two subspaces W and V are said to be orthogonal if every vector in W is orthogonal to every vector in V . A vector \mathbf{z} is said to be orthogonal to a subspace W if it is orthogonal to every vector in W . The set of all vectors that are orthogonal to a subspace $W \subset \mathbf{R}^n$ is called the **orthogonal complement** of W and is denoted by

$$W^\perp = \{\mathbf{z} \in \mathbf{R}^n; \mathbf{z} \cdot \mathbf{y} = 0, \text{ for every } \mathbf{y} \in W\}$$

Ex If W is plane through the origin in \mathbf{R}^3 and L is the line through the origin perpendicular to W , then $W^\perp = L$. In fact, clearly $L \subset W^\perp$ since L is perpendicular to W and any vector not in L is not perpendicular to W . Similarly $L^\perp = W$. If $L = \{\mathbf{x} \in \mathbf{R}^3; \mathbf{x} = \alpha(1, 1, 1), \alpha \in \mathbf{R}\}$ then $L^\perp = \{\mathbf{y} \in \mathbf{R}^3; \alpha(1, 1, 1) \cdot \mathbf{y} = \alpha(y_1 + y_2 + y_3) = 0\}$.

(1) If $W = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ then $\mathbf{z} \in W^\perp$ if and only if $\mathbf{z} \cdot \mathbf{v}_1 = \cdots = \mathbf{z} \cdot \mathbf{v}_k = 0$.

(2) W^\perp is a subspace. (3) $(W^\perp)^\perp = W$.

Th Let A be an $m \times n$ matrix. Then $(\text{Row } A)^\perp = \text{Nul } A$ and $(\text{Col } A)^\perp = \text{Nul } A^T$.

Pf Since $C(A) = \text{Row}(A^T)$ the second statement follows from the first. If $\mathbf{x} \in \text{N}(A)$

$$\begin{array}{l} \mathbf{r}_1 \rightarrow \\ \mathbf{r}_2 \rightarrow \\ \vdots \\ \mathbf{r}_m \rightarrow \end{array} \begin{bmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

so \mathbf{x} is orthogonal to $\text{Row } A$ since its orthogonal to $\mathbf{r}_1, \dots, \mathbf{r}_m$.

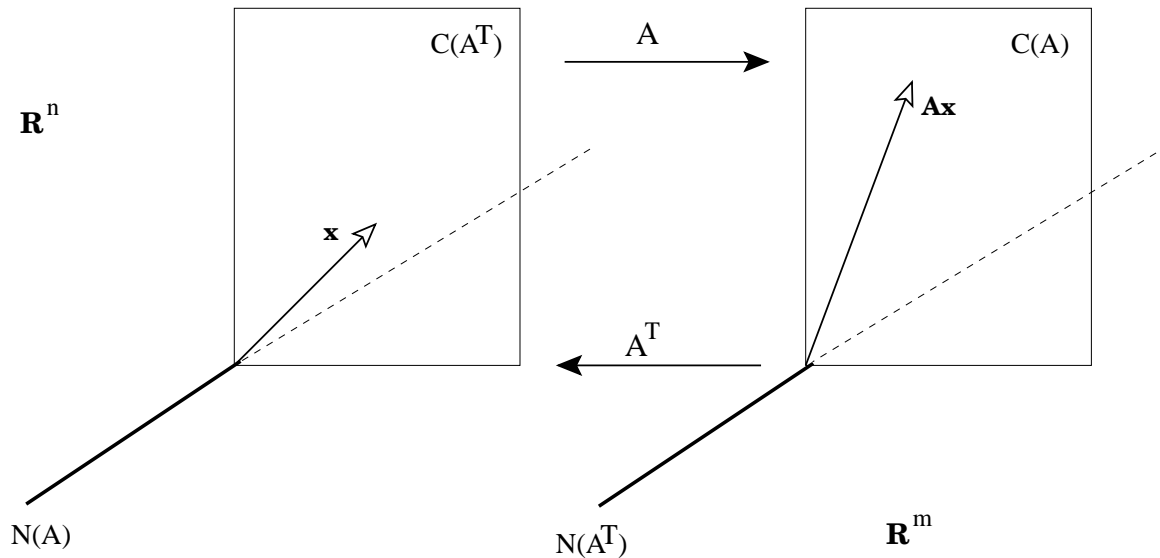
Ex Let $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \end{bmatrix}$. Basis for $\text{Nul } A = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$, so $\text{Nul } A$ is a plane in

\mathbf{R}^3 . Basis for $\text{Row } A = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$ so $\text{Row } A$ is a line in \mathbf{R}^3 . Basis for $\text{Col } A = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$, so

$\text{Col } A$ is a line in \mathbf{R}^2 . Basis for $\text{Nul } A^T = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$, so $\text{Nul } A^T$ is a line in \mathbf{R}^2 .

Th A matrix transformation A is invertible from the row space to the column space.

Pf First we note that the dimension of the row space is equal to the dimension of the column space which is equal to the rank of the matrix.



In order to prove the statement in the theorem we must show that for each $\mathbf{b} \in C(A)$ there is one and only one $\mathbf{x} \in C(A^T)$ such that $A\mathbf{x} = \mathbf{b}$. Let us first prove the uniqueness of $\mathbf{x} \in C(A^T)$. If we had there were $\mathbf{x}, \mathbf{y} \in C(A^T)$ such $A\mathbf{x} = A\mathbf{y} = \mathbf{b}$, then it follows that $A(\mathbf{x} - \mathbf{y}) = \mathbf{0}$. Hence we must have that $\mathbf{x} - \mathbf{y} \in N(A)$. But since both \mathbf{x} and \mathbf{y} are in $C(A^T)$ it follows that so is the difference and hence $\mathbf{x} - \mathbf{y} \in N(A)^\perp$. But the only vector that is in both $N(A)$ and $N(A)^\perp$ is the $\mathbf{0}$ vector so $\mathbf{x} - \mathbf{y} = \mathbf{0}$. This proves the uniqueness. To prove the existence we actually have to know about the orthogonal projection which comes in the next chapter. Since $\mathbf{b} \in C(A)$ we always know that there is $\mathbf{y} \in \mathbf{R}^m$ such that $A\mathbf{y} = \mathbf{b}$. However, a priori \mathbf{y} need not be in $C(A^T)$. In the next section we will show that we can project \mathbf{y} along the nullspace $N(A)$ onto the row space $C(A^T)$, i.e. we can write $\mathbf{y} = \mathbf{x} + \mathbf{w}$, where $\mathbf{w} \in N(A)$ and $\mathbf{x} \in C(A^T)$.