Some key facts about transpose

Let A be an $m \times n$ matrix. Then A^T is the matrix which switches the rows and columns of A. For example

$$\begin{pmatrix} 1 & 5 & 3 & 4 \\ 2 & 7 & 0 & 9 \\ 1 & 3 & 2 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 & 1 \\ 5 & 7 & 3 \\ 3 & 0 & 2 \\ 4 & 9 & 6 \end{pmatrix}$$

We have the following useful identities:

$$\boxed{\begin{array}{ccc} (A^T)^T = A & (A+B)^T = A^T + B^T & (kA)^T = kA^T \\ (AB)^T = B^T A^T & (A^T)^{-1} = (A^{-1})^T & \vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} \end{array}} \quad \text{Transpose Facts 1}$$

A deeper fact is that

$$\boxed{\mathrm{Rank}(A) = \mathrm{Rank}(A^T).} \quad \text{Transpose Fact 2}$$

Remember that $\mathrm{Rank}(B)$ is $\dim(\mathrm{Im}(B))$, and we compute Rank as the number of leading ones in the row reduced form.

Recall that $\vec{u}$ and $\vec{v}$ are **perpendicular** if and only if $\vec{u} \cdot \vec{v} = 0$. The word **orthogonal** is a synonym for perpendicular.

If V is a subspace of $\mathbb{R}^n$, then $V^\perp$ is the set of those vectors in $\mathbb{R}^n$ which are perpendicular to every vector in V. $V^\perp$ is called the ***orthogonal complement*** to V; I'll often pronounce it "Vee perp" for short.

You can (and should!) check that $V^\perp$ is a subspace of $\mathbb{R}^n$. It is geometrically intuitive that

$$\boxed{\dim V^\perp = n - \dim V} \quad \text{Transpose Fact 3}$$

and that

$$\boxed{(V^\perp)^\perp = V.} \quad \text{Transpose Fact 4}$$

We will prove both of these facts later in this note.

In this note we will also show that

$$\boxed{\mathrm{Ker}(A) = \mathrm{Im}(A^T)^\perp \quad \mathrm{Im}(A) = \mathrm{Ker}(A^T)^\perp} \quad \text{Transpose Fact 5}$$

As is often the case in math, the best order to state results is not the best order to prove them. We will prove these results in reverse order: 5, 4, 3, 2. The identities in the first box are left for you to check; see Theorem 5.3.9 in your textbook if you need help.

All of these facts do appear in your textbook, but scattered about and sometimes in the problem sections: Transpose Fact 2 is Exercises 3.3.71 through 3.3.74; Transpose Fact 3 is Theorem 5.1.8; Transpose Fact 4 is Exercise 5.1.23; Tranpose Fact 5 is Theorem 5.4.1.

The kernel of A is perpindicular to the image of A^T.

Observation If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ span V, then $\vec{u}$ is perpendicular to V if and only if it is perpendicular to $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$.

To see this, notice that if $\vec{u}$ is perpendicular to $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$, and $\vec{v}$ is any other vector in V, then we can write $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$ and hence

$$\vec{u} \cdot (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m) = c_1(\vec{u} \cdot \vec{v}_1) + c_2(\vec{u} \cdot \vec{v}_2) + \cdots + c_m(\vec{u} \cdot \vec{v}_m).$$

We now show that the kernel of A is the orthogonal space to the image of A^T and the image of A is the orthogonal space to the kernel of A^T, which is Transpose Fact 5.

I'll work through an example which shows all the key points. Take

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Note that the following are equivalent:

(1) $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is in the kernel of A.
(2) We have $x + 2y + 3z = 0$, $4x + 5y + 6z = 0$ and $7x + 8y + 9z = 0$.
(3) $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is perpendicular to $\left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 4 \\ 5 \\ 6 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 7 \\ 8 \\ 9 \end{smallmatrix}\right)$.
(4) $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is perpendicular to the span of $\left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 4 \\ 5 \\ 6 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 7 \\ 8 \\ 9 \end{smallmatrix}\right)$.
(5) $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is perpendicular to the image of $\left(\begin{smallmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{smallmatrix}\right)$.
(6) $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is perpendicular to the image of A^T.

Going between (3) and (4) is Observation 1.

So we have checked that $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is in the kernel of A if and only if $\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ is perpendicular to the image of A^T. The argument is just the same for a general matrix.

The complement of the complement is the original space

Let V be any subspace of $\mathbb{R}^n$. We're going to show that $(V^\perp)^\perp = V$, just like you would expect from drawing pictures in three dimensions.

Find a map A with image V. (We actually never proved that you can always do this! Assuming that every subspace has a basis $\vec{v}_1$, $\vec{v}_2$, $\dots$, $\vec{v}_m$, take A to be the matrix whose columns are $\vec{v}_i$. I added a note to the reading column for Sept. 30 on why every subspace has a basis.)

$$(V^\perp)^\perp = \left(\mathrm{Im}(A)^\perp\right)^\perp = \left(\mathrm{Ker}(A^T)\right)^\perp = \mathrm{Im}((A^T)^T) = \mathrm{Im}(A) = V.$$

The dimension of the complement is what you expect it to be.

Let V be any subspace of $\mathbb{R}^n$. We're going to show that $\dim V^\perp = n - \dim V$.

Let P be the orthogonal projection onto V; this is a map from $\mathbb{R}^n \to \mathbb{R}^n$. Then $\mathrm{Ker}(P) = V^\perp$ and $\mathrm{Im}(P) = V$. By the rank-nullity theorem,

$$\dim(\mathrm{Im}(P)) + \dim(\mathrm{Ker}(P)) = n \quad \text{so} \quad \dim(V) + \dim(V^\perp) = n \quad \text{as promised.}$$

We'll talk in class about how to compute P. For the record, here is the most concise formula. Let $\vec{v}_1$, $\vec{v}_2$, $\dots$, $\vec{v}_m$ be a basis for V and let A be a matrix with columns $\vec{v}_1$, $\vec{v}_2$, $\dots$, $\vec{v}_m$. So A is an injective map $\mathbb{R}^m \to \mathbb{R}^n$ with image V. Then

$$P = A(A^T A)^{-1} A^T.$$

This formula raises two questions: Why is $A^T A$ invertible? Why does this formula give the orthogonal projection? We'll answer both of these, but maybe you can figure them out first.

Rank of A equals rank of A^T

Let A be an $m \times n$ matrix. From the rank-nullity theorem,

$$\dim \mathrm{Im}(A^T) = m - \dim \mathrm{Ker}(A^T).$$

From Transpose Fact 5, $\mathrm{Ker}(A^T) = \mathrm{Im}(A)^\perp$ and, from Transpose Fact 4, $\dim \mathrm{Im}(A)^\perp = m - \dim \mathrm{Im}(A)$. Putting these together,

$$\dim \mathrm{Im}(A^T) = m - (m - \dim \mathrm{Im}(A)) = \dim \mathrm{Im}(A).$$