FUNDAMENTAL THEOREM OF LINEAR ALGEBRA

Recall some definitions:

- If $\mathbf{A} \in \mathrm{Mat}_{k\times n}$, the *column space of* $\mathbf{A}$ is the span of the columns of $\mathbf{A}$ (in $\mathbb{R}^k$):
$$col(\mathbf{A}) = Span\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \quad \text{if} \quad \mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}.$$
- The *row space of* $\mathbf{A}$ is the column space of $\mathbf{A}^\top$:
$$row(\mathbf{A}) = col(\mathbf{A}^\top).$$
- The *nullspace of* $\mathbf{A}$ is the kernel of the map $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$:
$$null(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}.$$
- If V is a subspace of $\mathbb{R}^n$, then the *orthogonal space* or (*orthogonal complement*) of V is:
$$V^\perp = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{v}\cdot\mathbf{x} = 0 \text{ for all } \mathbf{v} \in V\} = \mathbf{v}_1^\perp \cap \cdots \cap \mathbf{v}_\ell^\perp \quad \text{if } \{\mathbf{v}_i\} \text{ is a basis for } V.$$

Fundamental Theorem of Linear Algebra. *Let* $\mathbf{A}$ *be a* $k \times n$ *matrix. Then*
$$null(\mathbf{A}) = row(\mathbf{A})^\perp \qquad \textit{and} \qquad null(\mathbf{A}^\top) = col(\mathbf{A})^\perp.$$
Moreover, if $\mathbf{A}$ *has rank* r *(i.e., the echelon form of* $\mathbf{A}$ *has* r *pivotal* 1*s), then*
$$\dim(row(\mathbf{A})) = \dim(col(\mathbf{A})) = r,$$
$$\dim(null(\mathbf{A})) = n - r, \quad \textit{and} \quad \dim(null(\mathbf{A}^\top)) = k - r.$$

Remark. The dimensions given in the statement above follow immediately from the dimension formula, covered in last week's lecture. We'll only prove the first part of the theorem.

Proof. Let $\mathbf{b}_i^\top$, $1 \le i \le k$, be the rows of $\mathbf{A}$, i.e.,
$$\mathbf{A} = \begin{bmatrix} \mathbf{b}_1^\top \\ \vdots \\ \mathbf{b}_k^\top \end{bmatrix}$$

First, suppose $\mathbf{x} \in row(\mathbf{A})^\perp$. Then $\mathbf{v}\cdot\mathbf{x} = 0$ for every $\mathbf{v} \in row(\mathbf{A})$; in particular, $\mathbf{b}_i \cdot \mathbf{x} = 0$ for every i. By the definition of matrix multiplication, $[\mathbf{A}\mathbf{x}]_i = 0$ for $1 \le i \le k$. Therefore $\mathbf{x} \in null(\mathbf{A})$, which implies $row(\mathbf{A})^\perp \subset null(\mathbf{A})$.

Now, suppose $\mathbf{x} \in null(\mathbf{A})$. Then, again by the definition of matrix multiplication, $\mathbf{b}_i \cdot \mathbf{x} = 0$ for all i. If $\mathbf{b} \in row(\mathbf{A})$, then there exist scalars $c_1, \dots, c_k$ such that $\mathbf{b} = c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k$. Thus we compute:
$$\mathbf{b}\cdot\mathbf{x} = (c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k)\cdot\mathbf{x} = c_1(\mathbf{b}_1\cdot\mathbf{x}) + \cdots + c_k(\mathbf{b}_k\cdot\mathbf{x}) = 0 + \cdots + 0 = 0.$$
Therefore $\mathbf{x} \in row(\mathbf{A})^\perp$, which implies $null(\mathbf{A}) \subset row(\mathbf{A})^\perp$.

We conclude $null(\mathbf{A}) = row(\mathbf{A})^\perp$. The equality $null(\mathbf{A}^\top) = col(\mathbf{A})^\perp$ follows immediately by applying the same argument to $\mathbf{A}^\top$, since by definition $col(\mathbf{A}) = row(\mathbf{A}^\top)$. $\square$

DIRECT SUMS

The notion of a *direct sum* of subspaces is very important in linear algebra. First, let's define the general sum: if V and W are subspaces of $\mathbb{R}^n$, then their *sum* is

$$V + W = \{\mathbf{v} + \mathbf{w} \mid \mathbf{v} \in V,\ \mathbf{w} \in W\}.$$

This definition says nothing about how V and W are related. One may be contained in the other, or they may have some large overlap. But if $V \cap W = \{\mathbf{0}\}$, then the sum has a special property, which you are to prove in the homework: every $\mathbf{x}$ in $V + W$ has a *unique* expression as a vector in V plus a vector in W. To show this special relationship, we write $V + W = V \oplus W$, and call $V \oplus W$ the *direct sum* of V and W.

Let's see an important example:

Example. If V is a subspace of $\mathbb{R}^n$, then $\mathbb{R}^n = V \oplus V^\perp$.

Proof. Let $\mathbf{v}_1, \ldots, \mathbf{v}_j$ be a basis for V. Then $V^\perp$ is the space of vectors $\mathbf{x}$ that solve the system

$$\begin{cases} \mathbf{v}_1 \cdot \mathbf{x} = 0 \\ \qquad \vdots \\ \mathbf{v}_j \cdot \mathbf{x} = 0 \end{cases},$$

by the Fundamental Theorem of Linear Algebra. By the rank-nullity theorem, the dimension of $V^\perp$ is $n - j$. Every vector in $V^\perp$ is linearly independent from V, so a basis for $V^\perp$ together with a basis for V must span $\mathbb{R}^n$. But clearly $V \cap V^\perp = \{\mathbf{0}\}$, so $\mathbb{R}^n = V \oplus V^\perp$. $\square$

Remark. It now follows immediately from the Fundamental Theorem that, if $\mathbf{A} \in \mathrm{Mat}_{k\times n}$,

$$\mathbb{R}^n = null(\mathbf{A}) \oplus row(\mathbf{A}), \qquad \text{and} \qquad \mathbb{R}^k = null(\mathbf{A}^\top) \oplus col(\mathbf{A}).$$

GRAPHING

We draw the graph of a function $f : \mathbb{R} \to \mathbb{R}$ in $\mathbb{R}^2$. But we're using $\mathbb{R}^2$ in a particular way: we definitely want to use the standard basis, and we want the x-direction to represent one copy of $\mathbb{R}$ and the y-direction to represent another copy of $\mathbb{R}$. To make this distinction, we'll write $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$, meaning that $\mathbb{R}^2$ is the direct sum of the x- and y-directions. The *graph of f* is the subset

$$G_f = \left\{ \begin{pmatrix} x \\ f(x) \end{pmatrix} \in \mathbb{R}^2 \mid x \in \mathbb{R} \right\}.$$

We thus generalize the notion of graph as follows: suppose $f : \mathbb{R}^n \to \mathbb{R}^k$. Then we need $n + k$ dimensions to draw the graph of f. Write $\mathbb{R}^{n+k} = \mathbb{R}^n \oplus \mathbb{R}^k$, so that $\mathbb{R}^n$ is identified with the first n variables in $\mathbb{R}^{n+k}$, and $\mathbb{R}^k$ with the last k variables. Then the graph of f is

$$G_f = \left\{ \begin{pmatrix} \mathbf{x} \\ f(\mathbf{x}) \end{pmatrix} \in \mathbb{R}^n \oplus \mathbb{R}^k = \mathbb{R}^{n+k} \mid \mathbf{x} \in \mathbb{R}^n \right\}.$$

Why is it so important to think about what we're doing when graphing? Because we're going to need to be able to switch perspectives. For example, any line through the origin in $\mathbb{R}^2$ can be thought of as the graph of a linear function $\mathbb{R} \to \mathbb{R}$. If it's neither horizontal nor vertical, then we can think of the function as defining y in terms of x, or x in terms of y. If it is either horizontal or vertical, however, there's only one option for the independent variable. With more variables, there are (sometimes) more choices. This is mostly a heads-up; the real work of knowing how to change perspectives this way will come in a couple of weeks. The following theorem should begin to give you some insight.

Linear Implicit Function Theorem. *Let $T : \mathbb{R}^n \to \mathbb{R}^k$ be a linear transformation, and assume T is onto. Set $m = n - k$. Then there is a linear function $S : \mathbb{R}^m \to \mathbb{R}^k$ defined implicitly by the kernel of T, i.e., $ker(T)$ is the "graph" of S.*

To get S, relabel the coordinates in $\mathbb{R}^n$ (if necessary) so that the first k columns of $[T]$ are pivotal. Then row reduce $[T]$ to get 1s as coefficients for the first k variables in the corresponding system of equations. Write the equations, and transfer all non-pivotal variables to the other side of each equation.

There may be more than one choice for the pivotal variables. For example, in

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

any two of the columns are linearly independent, so any two can be made pivotal by relabeling. This is a 3×2 matrix, so there will be 2 pivotal variables and 1 non-pivotal variable; i.e., we'll get a map $\mathbb{R} \to \mathbb{R}^2$, whose graph is a line. Leaving them as they are, we get x_1 and x_2 in terms of x_3. The three possible functions $\mathbb{R} \to \mathbb{R}^2$ are:

$$\begin{cases} x_1 = -x_3 \\ x_2 = x_3 \end{cases}, \qquad \begin{cases} x_1 = -x_2 \\ x_3 = x_2 \end{cases}, \qquad \text{and} \qquad \begin{cases} x_2 = -x_1 \\ x_3 = -x_1 \end{cases}.$$