Mathematics 468
Linear algebra primer

(I will use the words "point" and "vector" interchangeably to refer to elements of $\mathbf{R}^n$.)

Definition. A *vector subspace* of $\mathbf{R}^n$ is a set V of points of $\mathbf{R}^n$ so that

- For all $\vec{v}_1$ and $\vec{v}_2$ in V, $\vec{v}_1 + \vec{v}_2$ is in V ("vector addition"), and
- For all $\vec{v} \in V$ and $c \in \mathbf{R}$, $c\vec{v}$ is in V ("scalar multiplication").

These conditions imply that V contains the origin $\vec{0}$, for example.

Examples. Any line through the origin is a vector subspace of $\mathbf{R}^n$, as is any plane through the origin, etc.

Theorem. V is a vector subspace of $\mathbf{R}^n$ if and only if there are vectors $\vec{v}_1, \dots, \vec{v}_k$ in $\mathbf{R}^n$ so that

$$V = \{c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \mid c_i \in \mathbf{R}\}.$$

The sum $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is called a *linear combination* of the vectors $\vec{v}_1, \dots, \vec{v}_k$.

Examples.

- If $\vec{v}_1$ is any nonzero vector, then the set $\{c_1\vec{v}_1 \mid c_1 \in \mathbf{R}\}$ is the set of all scalar multiples of $\vec{v}_1$: all points on the line in the $\vec{v}_1$ direction through the origin.
- If $\vec{v}_1$ and $\vec{v}_2$ are two vectors that don't point in the same direction (i.e., if neither one is a scalar multiple of the other), then the set $\{c_1\vec{v}_1 + c_2\vec{v}_2 \mid c_1, c_2 \in \mathbf{R}\}$ is the plane determined by $\vec{v}_1$ and $\vec{v}_2$.
- If $\vec{v}_1 = (1, 0, 0)$ and $\vec{v}_2 = (0, 0, 1)$, then

  $$\{c_1\vec{v}_1 + c_2\vec{v}_2 \mid c_i \in \mathbf{R}\} = \{(c_1, 0, c_2) \mid c_i \in \mathbf{R}\}$$

  is the xz-plane.

Definition. If $V = \{c_1\vec{v}_1 + \cdots + c_k\vec{v}_k\}$, then we say that V is *spanned* by the vectors $\vec{v}_1, \dots, \vec{v}_k$. If V can be spanned by k vectors but not by fewer than k, then V has *dimension* k.

Examples. The dimension of a line is one. The dimension of a plane is two. The xz-plane in $\mathbf{R}^3$ is spanned by the vectors

$$(1, 0, 0), (0, 0, 1), (1, 0, 2),$$

and also by the vectors

$$(1, 0, 0), (0, 0, 1).$$

It can't be spanned by fewer than two vectors, so its dimension is two.

Definition. Let V be a vector subspace of $\mathbf{R}^n$, of dimension k. A *basis* for V is any set of k vectors which spans V. If $\{\vec{v}_1, \dots, \vec{v}_k\}$ is a basis for V, then every vector in V can be written uniquely as a linear combination of the $\vec{v}_i$'s.

Examples. Consider the line in $\mathbf{R}^2$ through the origin with slope 2. It is one-dimensional, so any single vector that spans it will form a basis. Here are several bases:

$$\{(1, 2)\},\ \{(2, 4)\},\ \{(-1, -2)\},\ \{(\pi, 2\pi)\}.$$

For instance, every point on the line can be written in the form $(t\pi, 2t\pi)$ for some number t. Any two of the vectors $(1, 0, 0)$, $(0, 0, 1)$, $(1, 0, 2)$ form a basis for the xz-plane.

Definition. Given two vector spaces V and W, a function $f : V \to W$ is *linear* if

- $f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2)$, and
- $f(c\vec{v}_1) = cf(\vec{v}_1)$, for all vectors $\vec{v}_1, \vec{v}_2 \in V$ and all scalars $c \in \mathbf{R}$.

Example. If $V = \mathbf{R}^n$ and $W = R^k$, then a good example of a linear function is multiplication by a $k \times n$ matrix A: if I write an element $\vec{v}$ of V as an n-dimensional column vector $\vec{v} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$, then $A\vec{v}$ is a k-dimensional column vector. It is pretty easy to check that this defines a linear function.

This example says that every matrix determines a linear function; to some extent, the converse is true. Given vector spaces V and W with $\dim V = n$ and $\dim W = k$, and given a linear function $f : V \to W$, then choosing bases for V and for W lets me find a matrix corresponding to f. Choosing different bases will, in general, result in a different matrix. (This is why it's best to think of linear functions as linear functions, and not as determined by matrices.)

In detail: if $\{\vec{v}_1, \dots \vec{v}_n\}$ is a basis for V, then for each j, $f(\vec{v}_j)$ is a vector in W. If $\{\vec{w}_1, \dots, \vec{w}_k\}$ is a basis for W, then $f(\vec{v}_j)$ can be expressed uniquely as a linear combination of the vectors $\vec{w}_i$: write

$$f(\vec{v}_j) = a_{1j}\vec{w}_1 + \cdots + a_{kj}\vec{w}_k.$$

Then the function f corresponds to the matrix A with (i, j)-entry a_{ij}. f sends the vector $\vec{v}_j$ to the "jth column of A", and it sends a linear combination of the $\vec{v}_j$'s to a linear combination of the columns of A.

Example. Let V be the line of slope 2 through the origin in $\mathbf{R}^2$: all points of the form $(t, 2t)$. Let W be the real line. Define $f : V \to W$ by $f(t, 2t) = t$ (take a point on the line V and send it to the point on W with the same x-coordinate). If I want to write this as a matrix, I have to choose bases for V and W. Depending on this choice, I'll get a different matrix:

- Basis for V: $\{(1, 2)\}$. Basis for W: $\{1\}$. Then to find the matrix for f, I see where the basis for V goes, and write this in terms of the basis for W: $f(1, 2) = 1$. So the matrix is $[1]$. (The matrix in this case is 1×1, so it's just a number.)
- Basis for V: $\{(-1, -2)\}$. Basis for W: $\{1\}$. Then $f(-1, -2) = -1 = -1(1)$. So the matrix is $[-1]$.
- Basis for V: $\{(2, 4)\}$. Basis for W: $\{1\}$. Then $f(2, 4) = 2 = 2(1)$. So the matrix is $[2]$.
- Basis for V: $\{(2, 4)\}$. Basis for W: $\{3\}$. Then $f(2, 4) = 2 = \frac{2}{3}(3)$. So the matrix is $[\frac{2}{3}]$.

These matrices all correspond to the same function, just parametrized differently. There is often no "best" way to choose bases for V and W, so there is no canonical choice of a matrix to represent a linear function. So, as I said earlier, it is best to think of linear functions as linear functions, not as matrices.

Definition. Given a linear function $f : V \to W$

- the *null space* of f, also known as the *kernel* of f, is the set $\{\vec{v} \in V \mid f(\vec{v}) = \vec{0}\}$. This is a vector subspace of V.
- The *nullity* of f is the dimension of the null space.
- The *image* of f is the set $\{\vec{w} \in W \mid \vec{w} = f(\vec{v}) \text{ for some } \vec{v} \in V\}$. This is a vector subspace of W.
- The *rank* of f is the dimension of the image.

Theorem. Given a linear function $f : V \to W$, $\text{rank}(f) + \text{nullity}(f) = \dim V$.

Examples.

- If you define $f : \mathbf{R}^n \to \mathbf{R}^k$ as multiplication by some $k \times n$ matrix A, then the dimension of V is the number of column of A. You can figure out the rank of A by row-reducing A to get a matrix in row-echelon form; then the rank is the number of nonzero rows. The nullity of A is the difference between these numbers.

- Define $f : \mathbf{R}^2 \to \mathbf{R}$ by $f(x, y) = 2x - y$. Then the image of f is all of $\mathbf{R}$: for every $t \in \mathbf{R}$, I can easily find x and y so that $2x - y = t$ (e.g., let $x = t/2$ and let $t = 0$). So the rank of f is one, the dimension of its image, $\mathbf{R}$. By the theorem, the nullity must be one. Indeed, the null space of A is all points (x, y) with $2x - y = 0$, or all points (x, y) with $2x = y$. This is the line through the origin with slope 2, a one-dimensional subspace of $\mathbf{R}^2$.

- If I use the "standard" bases for $\mathbf{R}^2$ and $\mathbf{R}$—$\{(1,0),(0,1)\}$ and $\{1\}$, respectively—then the function f that I just defined is given by the matrix $\begin{bmatrix} 2 & -1 \end{bmatrix}$: f sends $(1, 0)$ to 2 (the first column of the matrix), and f sends $(0, 1)$ to -1 (the second column). This is already row-reduced, essentially, and has one nonzero row, so has rank 1.

Theorem. Suppose $f : V \to W$ is a linear function, with $\dim V = n$ and $\dim W = k$.

(a) $\text{rank}(f)$ is no larger than either n or k.

(b) f is one-to-one (also known as "injective") $\Leftrightarrow$ the kernel of f equals $\{\vec{0}\}$ $\Leftrightarrow$ the nullity of f is zero.

(c) f is onto (also known as "surjective") $\Leftrightarrow$ the image of f equals W the rank of f is k.

(d) f is bijective $\Leftrightarrow n = k = \text{rank}(f) \Leftrightarrow n = k$ and $\text{nullity}(f) = 0$.

If f is a bijection, then there is an inverse function $g : W \to V$; it turns out that if f is linear, then so is g. In this case, f is called an *isomorphism* (of vector spaces). Two vector spaces V and W are *isomorphic* if there is an isomorphism from one to the other.

Theorem. Two vector spaces V and W are isomorphic if and only if they have the same dimension. If $\dim V = \dim W = \text{n}$, then $f : V \to W$ is an isomorphism $\Leftrightarrow \text{rank}(f) = n \Leftrightarrow \text{nullity}(f) = 0$.

Theorem. If $\dim V = \dim W = n$, then a linear function f from V to W can be represented by an $n \times n$ matrix A. Then f is an isomorphism $\Leftrightarrow$ the matrix A is invertible $\Leftrightarrow \det(A) \neq 0$.

So if the dimensions of V and W are equal, you can tell whether a linear function is an isomorphism by computing a single number: the determinant of the matrix A.