Linear Functions, Bases, and Matrices

Learning goal: to see how matrices make their reappearance as soon as we have bases.

Let \( f: V \to W \) be a linear function between two vector spaces. Notice how we can say that without having any other qualities—no bases or dimensions or anything. Just linear qualities of things. The vectors themselves exist independent of there being a basis, and the function transforms them.

Let’s see what we get when we add some bases, though. Start with \( V \) being finite dimensional with basis \( v_1, v_2, \ldots, v_n \). Now any vector \( x \) has a unique expression as \( x = x_1v_1 + x_2v_2 + \ldots + x_nv_n \). For convenience sake, we will express this as a (column) vector \( (x_1, x_2, \ldots, x_n) \).

Right away this gives us a one-to-one correspondence with \( F^n \), where \( F \) is the field for our vector space. For each column vector represents a vector by multiplying the basis vectors, and each vector has an expression and corresponding column vector.

In fact, the function taking vectors \( n \) \( V \) to columns in \( F^n \) is also a linear function, for taking sums or scalar products does exactly the same thing to the coefficient vectors. So in a very strict sense:

**Theorem:** every finite dimensional vector space of dimension \( n \) is isomorphic to \( F^n \).

Note: there are LOTS of such correspondences, one for each choice of basis. If we have several bases in operation, and we need to distinguish which one a particular expression is working with, we will use a subscript: \((x_1, x_2, \ldots, x_n)_v \) or \((x_1, x_2, \ldots, x_n)_w \).

**Example:** in the basis \( 1, x, x^2, x^3 \) for \( P_3 \), the polynomial \( 2x^3 - x + 1 \) is denoted with the vector \((1, -1, 0, 2) \). In the basis \( x^3, x^2, x, 1 \), it has the vector \((2, 0, -1, 1) \). In the basis \( x^3 - 1, x^2 - 2, x - 3, 4 \), it has the vector \((2, 0, -1, 0) \).

Now the real interesting stuff happens when both \( V \) and \( W \) are finite dimensional with bases \( v_1, \ldots, v_n \) and \( w_1, \ldots, w_m \) respectively, and \( f \) is a linear function between them. We already know that \( f \) is determined by what it does to a basis, and we know that if we have a vector \( n \) \( V \) we can express it as a combination of \( v \)'s, and one in \( W \) as a combination of \( w \)'s.

So let’s start by taking \( f(v_1) \) and expressing it as a linear combination of \( w \)'s. In fact, let \( f(v_1) = a_{11}w_1 + a_{12}w_2 + \ldots + a_{1m}w_m \). Similarly, let \( f(v_2) = a_{21}w_1 + \ldots + a_{2m}w_m \), and so forth. We can use the \( m \times n \) matrix \( A = (a_{ij}) \) to keep track of all these coefficients. Now what happens to a general vector \( x \) with \( f(x) = y \)?

Well, \( f(x) = f(x_1v_1 + \ldots + x_nv_n) = x_1f(v_1) + \ldots + x_nf(v_n) \). Expanding each of these gives \( f(x) = x_1(a_{11}w_1 + \ldots + a_{1m}w_m) + \ldots + x_n(a_{n1}w_1 + \ldots + a_{nm}w_m) \). Group by \( w \) now: \( f(x) = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m)w_1 + \ldots + (a_{n1}x_1 + \ldots + a_{nm}x_m)w_m \). We can use a column vector to collect these coefficients—but then we notice something wonderful. *This column vector is exactly the product of the matrix \( A \) with the column vector \( x \).* That is, \( y_w = Ax_v \).
Example: use the basis $x^3, x^2, x, 1$ for $P_3$ as both the domain and range, and let the linear function be differentiation. Then $D(x^3) = 3x^2$, so $a_{11} = 0$, $a_{21} = 3$, $a_{31} = 0$, and $a_{41} = 0$. We can continue constructing the matrix in this manner. The matrix is

$$
\begin{pmatrix}
0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 1 & 0
\end{pmatrix}.
$$

- Now take some polynomial, say $2x^3 - x + 1$ represented by $(2, 0, -1, 1)$ (in the basis chosen). Multiply this by the differentiation matrix to get $(0, 6, 0, -1)$, which represents $6x^2 - 1$ in the given basis, which is (of course) the derivative of the original polynomial.

- Are there any polynomials with $y' = y$? In other words, does the derivative function have an eigenvector with eigenvalue 1? Well, from the above matrix, the characteristic polynomial is $\lambda^4$, so zero is the only eigenvalue. So no, there is no third degree polynomial that is its own derivative.

- The matrix is singular. This makes it non-invertible. Indeed, differentiation is not invertible on polynomials. For the derivative of any constant is zero, so what is the antiderivative of zero?

Example: (change of variables) Let’s say we have two bases $v$ and $w$ on a vector space. We can, of course, express members of the second basis in terms of the first. For example, we can have $w_1 = a_{11}v_1 + a_{12}v_2 + \ldots + a_{1n}v_n$, or $(a_{11}, a_{12}, \ldots, a_{1n})_v$. In other words, the matrix $A$ expresses (row by row) the $w$’s in terms of the $v$’s. Now take some vector $x$. What is the relationship between its expressions in terms of the $v$’s and the $w$’s? That is, what is the relationship between $x_v$ and $x_w$? Let’s pick a simple one, like $x = w_1 = (1, 0, \ldots, 0)_w$. This should be $(a_{11}, a_{12}, \ldots, a_{1n})_v$. In other words, it’s the top row of $A$. To express $w_2$ in terms of the $v$’s, we get the second row of $A$. To express a linear combination of $w$’s in terms of $v$’s, we need to take the same linear combination of rows of $A$. In other words, if $A$’s rows express the $w$’s in terms of the $v$’s, then we have the relationship $x_v = A^T x_w$. Unsurprisingly, $x_w = (A^T)^{-1} x_v$. How do we know the matrix is invertible? Because every vector has to have an expression in terms of both bases. We know it’s square, and the range must be all column vectors, so it has full rank and is invertible.

Example: given bases for $V$ and $W$, each linear function gives rise to a matrix, and by expressing vectors in terms of their coordinates in the two bases and using matrix multiplication, we find that each matrix represents a linear function. In fact, this relationship is one-to-one, and linear to boot—the sum of two linear functions gives rise to the sum of their matrices, and the same goes for a scalar multiple. So given bases for both spaces there is an isomorphism between linear function and $m \times n$ matrices.

Example: (compositions and products). Let $V$ have basis $v_1, \ldots, W$ have basis $w_1, \ldots, and Z$ have basis $z_1, \ldots$. Let $f: V \rightarrow W$ be a linear function with matrix $A$ (in terms of the $v$ and $w$ bases), and let $g: W \rightarrow Z$ be a linear function with matrix $B$ (in terms of the $w$ and $z$ bases).

What is the matrix for $g(f(x))$ (in terms of the $v$ and $z$ bases)? Well, we can follow $v_1$ through all the rigamarole and then $v_2$, etc. Or we could be smart. Let $y = f(x)$. Then we know $y_w = Ax_v$.

And to send this into $Z$? Multiply by $B$ of course! So the matrix for $g(f(x))$ is $BA$. 
Example: in the previous example, what happens if we use a different basis for $\mathbf{W}$ only? The matrices $A$ and $B$ will change, but the result $BA$ can’t! How does that work? Well, let $M$ be the matrix that expresses the basis $\mathbf{w}'$ in terms of the basis $\mathbf{w}$. We know that $y_{\mathbf{w}'} = (M^T)^{-1} y_{\mathbf{w}}$. To figure out the new matrix to replace $A$, we simply note that the function is the same, only the basis has changed. So what we will do is transform in the original basis, and then transform the expression to be in the right basis. In other words, $y_{\mathbf{w}'} = M^T y_{\mathbf{w}} = M^T A x_v$.

Now what should replace $B$? We need to start with $y_{\mathbf{w}}$ and get $g(y)$. The easiest thing to do would be to transform back to $\mathbf{w}$ coordinates and then transform by $B$. So the replacement for $B$ must be $BM^T$.

Then the matrix that represents $g(f(x))$ in the $\mathbf{v}$, $\mathbf{w}'$, and $\mathbf{z}$ bases must be $(BM^T)((M^T)^{-1}A)$, and the stuff in the middle cancels to give $BA$. 