Week 5: Change of Basis

We spent some time this week talking about changes of basis. Consider two bases $\{e_1, e_2\}$ and $\{v_1, v_2\}$ of $\mathbb{R}^2$. Typically these vectors would be written as columns of real numbers, but that would be confusing in this example, so we won't. Assume we have some linear map $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$, such that $\phi(e_1) = a_{11}e_1 + a_{12}e_2$, and $\phi(e_2) = a_{21}e_1 + a_{22}e_2$; all linear maps have this form, and ϕ is uniquely determined by the a_{ij}, because e_1, e_2 is a basis. We can then write ϕ as if it were multiplication by a matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

To find $\phi(w)$ for some general vector w, write $w = \lambda e_1 + \mu e_2$ and calculate:

$$\begin{pmatrix} \lambda' \\ \mu' \end{pmatrix} = A \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$$

Then $\phi(w) = \lambda' e_1 + \mu' e_2$. If we want to use the basis v_1, v_2 instead, we need to write $\phi(v_1) = b_{11}v_1 + b_{12}v_2$ and $\phi(v_2) = b_{21}v_1 + b_{22}v_2$, and then replace A by the matrix:

$$A' = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

Now for a general w, if $w = \alpha v_1 + \beta v_2$ (note that α and β will in general be different from λ and μ), then to compute $\phi(w)$ we calculate:

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = A' \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

and then $\phi(w) = \alpha' v_1 + \beta' v_2$. Note that the column vector representing w changes depending on which basis we use, although the vector w does not.
If we want to convert columns representing w in the basis $\{v_1, v_2\}$ into columns representing w in the basis $\{e_1, e_2\}$, we have a matrix to do that! We want to write down a matrix P representing the identity map (we don't want to actually change any vectors) so that it takes as inputs columns of v_i–coordinates, and outputs columns of e_i–coordinates. In particular, the column $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ represents v_1 in the v_i–coordinates. The output column is supposed to represent v_1 in the e_i–coordinates, so write $v_1 = p_{11}e_1 + p_{12}e_2$, and then we require:

$$P \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p_{11} \\ p_{12} \end{pmatrix}$$

so the first column of P is $\left(\begin{smallmatrix} p_{11} \\ p_{12} \end{smallmatrix}\right)$. Similarly, the second column is $\left(\begin{smallmatrix} p_{21} \\ p_{22} \end{smallmatrix}\right)$, where $v_2 = p_{21}e_1 + p_{22}e_2$. Then for any $w = av_1 + bv_2$, we have:

$$P \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a' \\ b' \end{pmatrix}$$

where $w = a'e_1 + b'e_2$. Now we can note that $A' = P^{-1}AP$, by comparing their actions on column vectors. For any w, left multiplication by A' replaces the column representing w

in the v_i–coordinates by the column representing $\phi(w)$ in the v_i–coordinates. On the other hand, if we take the column representing w in the v_i–coordinates and left multiply by P, we get the column representing w in the e_i–coordinates. Then multiplying by A produces the column representing $\phi(w)$ in the e_i–coordinates, and multiplying by P^{-1} gives the column representing $\phi(w)$ in the v_i–coordinates. So both A' and $P^{-1}AP$ act in the same way on $\mathbb{R}^2$, and therefore must be equal.

Note that despite using the notation e_1, e_2, at no point did we assume that these vectors were the standard basis vectors of $\mathbb{R}^2$ – this procedure works for any two bases. We can also extend this reasoning to vector spaces of arbitrary finite dimension without much alteration.