Notes on Matrices

Let V, W be finite-dimensional vector spaces over a field $\mathbb{F}$. Given a linear map $T : V \longrightarrow W$, we may decide to use a matrix to describe the map. To do this, we follow a certain convention. The purpose of these notes is to explain the convention.

Coordinates of a vector as a $n \times 1$ matrix

Choose a basis $\beta = (v_1, \ldots, v_n)$ for V Then any vector $x \in V$ can be written uniquely as $x = \sum_{i=1}^{n} x_i v_i$. This allows us to describe x by giving its coordinates $(x_1, \ldots, x_n) \in \mathbb{F}^n$. We define the matrix of x to be

$$\beta x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Matrix of a linear map

The linear map $T : V \longrightarrow W$ is completely determined by its values on the basis β. If we choose a basis $\gamma = (w_1, \ldots, w_k)$ for W, then the value of T on the jth basis element v_j of β is

$$T(v_j) = \sum_{i=1}^{k} a_{ij} w_i. \tag{1}$$

This defines a $k \times n$ array of numbers $a_{ij} \in \mathbb{F}$ (i indicates the row, j indicates the column), which is defined to be the matrix of T with respect to β, γ:

$$\gamma T_\beta = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kn} \end{bmatrix}$$

Note that with this definition, the coordinates of $T v_j$ appear as the jth column of the matrix. Instead of writing γT_β you could write $M(T, \beta, \gamma)$; ours is just a shorthand notation which helps to remember that β is the basis for the domain of T whereas γ is the basis for the codomain. Note also that a vector $x \in V$ defines a map $\mathbb{F} \to V$ defined by $1 \mapsto x$. If we call this map x, then we can interpret βx as the matrix of the linear map x written in the basis β for the codomain and in the standard basis for the domain $\mathbb{F}$.

Applying a matrix to a vector

A $k \times n$ matrix can be "multiplied" by or "applied" to a $n \times 1$ matrix to yield a $k \times 1$ matrix. By definition, we set
$$({}_\gamma T_\beta)({}_\beta x) = {}_\gamma(Tx),$$
in other words, the matrix of T applied to the matrix of x gives the matrix of Tx. To compute this, we use the fact
$$Tx = \sum_{j=1}^{n} x_j T(v_j) = \sum_{i=1}^{k} (\sum_{j=1}^{n} a_{ij} x_j) w_i.$$
Therefore, the ith entry in the $k \times 1$ matrix $({}_\gamma T_\beta)({}_\beta x)$ is $\sum_{j=1}^{n} a_{ij} x_j$.

Composition as matrix multiplication

Let U be another vector space, with basis $\alpha = (u_1, \ldots, u_m)$. If $S : U \longrightarrow V$ and $T : V \longrightarrow W$ are linear maps, then they can be composed to give $TS : U \longrightarrow W$. This implies that we should be able to "multiply" the $k \times n$ matrix of T by the $n \times m$ matrix of S to give the $k \times m$ matrix of TS. We define the matrix product
$$({}_\gamma T_\beta)({}_\beta S_\alpha) = {}_\gamma TS_\alpha.$$
Suppose that the matrix ${}_\beta S_\alpha$ has entries b_{ij} defined by $Su_j = \sum_{i=1}^{n} b_{ij} v_i$. Then we can compute the matrix of the composition above, using:
$$TS(u_j) = T(\sum_{p} b_{pj} v_p) = \sum_{i=1}^{k} \left(\sum_{p=1}^{n} a_{ip} b_{pj} \right) w_i.$$

Therefore, the ij^{th} entry in the $k \times m$ matrix $({}_\gamma T_\beta)({}_\beta S_\alpha)$ is $\sum_{p=1}^{n} a_{ip} b_{pj}$, which is the usual formula given for the ij^{th} entry of a matrix product (the dot product of the ith row of T with the jth column of S). It is obvious that matrix multiplication is associative, since it is defined using composition of linear maps, and composition of maps is always associative.

Change of basis

Suppose you know the matrix of $T : V \longrightarrow W$ with respect to bases β, γ for V, W, and someone hands you new bases β', γ'. How can you find the matrix of T in the new basis? We can use the simple fact that the identity maps I_V and I_W can be composed with T without changing anything, i.e.
$$T = I_W T I_V.$$
But then we know from the definition of matrix multiplication that
$${}_{\gamma'} T_{\beta'} = {}_{\gamma'}(ITI)_{\beta'} = ({}_{\gamma'} I_\gamma)({}_\gamma T_\beta)({}_\beta I_{\beta'}).$$
The matrix $P = {}_\beta I_{\beta'}$ is usually called the "change of basis matrix" and its columns are the coordinates of the β' basis when expressed in the old basis β. Setting $Q = {}_\gamma I_{\gamma'}$, we see that
$${}_{\gamma'} T_{\beta'} = Q^{-1}({}_\gamma T_\beta)P.$$