Linear operators, matrices, change of coordinates: a brief HOWTO

A function $\mathcal{A}\colon V \to W$ from one vector space V (source space) to another vector space W (target space) with the same set of scalars is called *a linear operator (a linear transformation)*, if for all vectors $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$ and for any $c \in F$ we have

1. $\mathcal{A}(\mathbf{v}_1 + \mathbf{v}_2) = \mathcal{A}(\mathbf{v}_1) + \mathcal{A}(\mathbf{v}_2)$;
2. $\mathcal{A}(c\mathbf{v}) = c\mathcal{A}(\mathbf{v})$.

A linear operator is completely determined by images of basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ of V; indeed, if

$$\mathbf{v} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n,$$

then the above properties immediately imply that

$$\mathcal{A}(\mathbf{v}) = x_1\mathcal{A}(\mathbf{e}_1) + \ldots + x_n\mathcal{A}(\mathbf{e}_n).$$

If the vector spaces are finite-dimensional, for any bases $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ of V and $\mathbf{f}_1$, $\mathbf{f}_2$, ..., $\mathbf{f}_m$ of W we can define the *matrix of $\mathcal{A}$ with respect to these bases (relative to that bases)*. Namely, take the images $\mathcal{A}(\mathbf{e}_1)$, ..., $\mathcal{A}(\mathbf{e}_n)$ of the basis vectors in V, and expand them using the basis in W:

$$\begin{aligned}
\mathcal{A}(\mathbf{e}_1) &= a_{11}\mathbf{f}_1 + a_{21}\mathbf{f}_2 + \ldots + a_{m1}\mathbf{f}_m,\\
\mathcal{A}(\mathbf{e}_2) &= a_{12}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + \ldots + a_{m2}\mathbf{f}_m,\\
&\ldots\\
\mathcal{A}(\mathbf{e}_n) &= a_{1n}\mathbf{f}_1 + a_{2n}\mathbf{f}_2 + \ldots + a_{mn}\mathbf{f}_m.
\end{aligned}$$

The matrix of $\mathcal{A}$ relative to the given bases is

$$M_{\mathcal{A}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \ldots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix};$$

in other words, we write coordinates of vectors $\mathcal{A}(\mathbf{e}_i)$ with respect to the basis $\mathbf{f}_1$, ..., $\mathbf{f}_m$ in columns, and arrange these columns in a $m \times n$-matrix. Vice versa, any matrix determines a linear operator given by the above formulas.

How does that apply for the coordinate realisation? The answer is given by the most simple formula. If $\mathbf{x}$ is the column of coordinates of $\mathbf{v}$ relative to the basis in V, and $\mathbf{y}$ is the column of coordinates of $\mathcal{A}(\mathbf{v})$ relative to the basis in W, then

$$\mathbf{y} = M_{\mathcal{A}}\mathbf{x},$$

in other words, action of the linear operator on a vector corresponds to multiplication of the column of coordinates of this vector by the matrix of the operator on the left.

An immediate corollary of the previous formula is the following important property of matrices of linear operators.Assume that we have three vector spaces U, V, and W, with bases $\mathbf{d}_1$, ..., $\mathbf{d}_k$, $\mathbf{e}_1$, ..., $\mathbf{e}_n$, and $\mathbf{f}_1$, ..., $\mathbf{f}_m$ respectively. Let $\mathcal{A}$ be a linear operator from U to V, and $\mathcal{B}$ be a linear operator from V to W. Then we can define their composition

$\mathcal{C}$ that acts on any vector $\mathbf{u}$ as $\mathcal{C}(\mathbf{u}) = \mathcal{B}(\mathcal{A}(\mathbf{u}))$. Let $M_{\mathcal{A}}$, $M_{\mathcal{B}}$, and $M_{\mathcal{C}}$ be matrices of $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ relative to the bases of our spaces. Then

$$M_{\mathcal{C}} = M_{\mathcal{B}} M_{\mathcal{A}}$$

To a linear operator $\mathcal{A}$ from a vector space V to the same space, we can assign its matrix relative to a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V (as before, we expand images of basis vector in our basis, and arrange columns of coordinates in a matrix). Though in general we can choose coordinates (bases) for the source and the target space independently, when they coincide, it makes sense to keep the coordinate the same on the source side and the target side.

The *identity operator* Id_V on a vector space V maps any vector to itself; $\mathrm{Id}(\mathbf{v}) = \mathbf{v}$. Its matrix relative to any basis is the identity matrix I.

Take two bases $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of the same space V. Define a linear operator $\mathcal{A}_{\mathbf{ef}}$ by the formula $\mathcal{A}_{\mathbf{ef}}(\mathbf{e}_i) = \mathbf{f}_i$. This operator is called the transition operator (from the first basis to the second one). Its matrix relative to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is called the transition matrix, and is sometimes denoted by $M_{\mathbf{ef}}$.

Let us list the main properties of transition matrices.

0. $M_{\mathbf{fe}} = M_{\mathbf{ef}}^{-1}$. Moreover, every invertible matrix is a transition matrix between two bases. For instance, the transition matrix from the basis of standard unit vectors in $\mathbb{R}^n$ to some basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is the matrix whose columns are the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

1. Take some vector $\mathbf{v}$; let $\mathbf{x}$ be its column of coordinates relative to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, and let $\mathbf{y}$ be its column of coordinates relative to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$. Then

$$\mathbf{x} = M_{\mathbf{ef}} \mathbf{y}.$$

In other words, the column of "old" coordinates of a vector is obtained from a column of "new" coordinates multiplying it by the transition matrix on the left. Note that the transition matrix is used to get "new" vectors from the "old" ones, but "old" coordinates from the "new" ones.

2. Let C be the transition matrix between two bases of V, and D the transition matrix between two bases of W. Furthermore, let M_1 and M_2 be the matrices of the same operator $A \colon V \to W$, where M_1 is computed relative to the first basis of V and the first basis of W, and M_2 — relative to the second basis of V and the second basis of W. Then

$$M_2 = D^{-1} M_1 C.$$

In particular, when $V = W$,

$$M_2 = C^{-1} A C,$$

in other words, to compute the matrix of a linear operator in the "new" basis, you should multiply the matrix in the "old" basis by the transition matrix on the right, and by its inverse on the left.

Two matrices A and B are called *similar*, if there exists an invertible matrix C such that $B = C^{-1}AC$. The last property tells us, that two matrices are similar if and only if they represent the same linear operator in two different bases. Note that in the home assignments it was proved that similar matrices have equal determinants and traces, so the determinant and the trace are numerical invariants of linear operators.