Coordinates, Change of Basis, and Matrix of a Linear Transformation

Notes by Tom Zaslavsky

The book's treatments are okay but some can be simplified.

1. Coordinates

Finding the coordinates of a vector ("coordinatizing" it) means solving a certain linear system. You have a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and a vector $\mathbf{v}$, all in some arbitrary vector space V, and you want to find the coordinates of $\mathbf{v}$ in the basis $\mathcal{B}$. The complete meaning of that is the following equation:

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_n\mathbf{b}_n.$$

The coordinates of $\mathbf{v}$ are $c_1, c_2, \dots, c_n$, which we write in a vector called the coordinate vector of $\mathbf{v}$, namely,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

No matter what vector space V is, the coordinate vector is always in $\mathbb{R}^n$, where n is always the dimension of V. Your task in finding the coordinates is to solve the first equation. There are plenty of examples in the book, so I recommend you turn to Section 6.6.

Think of coordinates this way: Each basis in V creates a coordinate system for V. That means, if you know the coordinates, you know where the vector is. For instance, in $V = \mathbb{R}^n$, the standard basis creates the standard coordinate system (the one we usually use for $\mathbb{R}^n$). Some other vector spaces also have a standard basis, for instance $\mathcal{P}_d(t)$ or $\mathbb{R}_{m\times n}$, but most vector spaces don't, so we can't depend on having a standard basis. Even when there is one, many other bases are valuable, as we'll see in Chapter 8.

Coordinatization is a linear transformation from V to $\mathbb{R}^n$. (I leave the proof to you.) We call it $T_{\mathcal{B}}$. The formula is $T_{\mathcal{B}}(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$, i.e., we transform $\mathbf{v}$ in V into its coordinate vector in $\mathbb{R}^n$. (Our book does not discuss this explicitly.)

2. Matrix of a Linear Transformation

The book's treatment in and around Theorem 6.7.1 is good. (The diagram 6.7.1 is not so good because it assumes there is a standard basis.) The basic idea is that you want a matrix for a linear transformation T between any two vector spaces U and V (they may be the same one), not just $\mathbb{R}^m$ and $\mathbb{R}^n$, so you have to have a coordinate system in each vector space. A coordinate system means a basis; that's how we get coordinates (see the preceding). So, we have to have a basis $\mathcal{B}$ in U and a basis $\mathcal{C}$ in V. (Note that even if $U = V$ we may want $\mathcal{B}$ and $\mathcal{C}$ to be different bases; that is okay.) Then the matrix that gives the effect of T on the coordinates in U and V is $[T]^{\mathcal{C}}_{\mathcal{B}}$ in Theorem 6.7.1.

Note that this matrix doesn't act directly with U and V. It works with the coordinates, which belong to $\mathbb{R}^m$ (where $m = \dim U$) and $\mathbb{R}^n$ (where $n = \dim V$), respectively. So we have a diagram like this:

$$\begin{array}{ccc} \mathbf{u}\in U,\mathcal{B} & \xrightarrow{\quad T\quad} & T(\mathbf{u})\in V,\mathcal{C} \\ \Big\downarrow T_{\mathcal{B}} & & \Big\downarrow T_{\mathcal{C}} \\ [\mathbf{u}]_{\mathcal{B}}\in\mathbb{R}^m,\mathcal{E}_m & \xrightarrow{\quad [T]^{\mathcal{C}}_{\mathcal{B}}\quad} & [T(\mathbf{u})]_{\mathcal{C}}\in\mathbb{R}^n,\mathcal{E}_n \\ & & = [T]^{\mathcal{C}}_{\mathcal{B}}\cdot[\mathbf{u}]_{\mathcal{B}} \end{array}$$

$T_{\mathcal{B}}$ and $T_{\mathcal{C}}$ are the coordinatization transformations, which are linear transformations (see above).

The diagram means that if you start with vector $\mathbf{u} \in U$ and apply T (the top arrow), you get to $T(\mathbf{u}) \in V$. If you then take coordinates (the right-hand down arrow), you wind up with the coordinates of $T(\mathbf{u})$, i.e., the column vector $[T(\mathbf{u})]_{\mathcal{C}} \in \mathbb{R}^n$. (In the diagram I mention the basis for each vector space.) However, if you first take coordinates of $\mathbf{u}$, you get the column vector $[\mathbf{u}]_{\mathcal{B}} \in \mathbb{R}^m$. Multiply it by the matrix $[T]^{\mathcal{C}}_{\mathcal{B}}$, giving $[T]^{\mathcal{C}}_{\mathcal{B}} \cdot [\mathbf{u}]_{\mathcal{B}}$, and lo! and behold! you have the same column vector $[T(\mathbf{u})]_{\mathcal{C}}$. That is, the linear transformation T is turned into a matrix multiplication on coordinate vectors. This is often very convenient.

(We call this a "commutative diagram" because you get the same result whether you take the right-down path or the down-right path from U to $\mathbb{R}^m$. You don't need to memorize that name.)

3. Change of Basis

Change of basis means changing from one coordinate system to another in the same vector space. Remember, a coordinate system is the system of coordinates obtained from a basis. That's why changing coordinate system is called change of basis. The book treats this in Section 6.6.

The situation is that we have one vector space, V, and two bases, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$. Theorem 6.6.1 in the book is good for this. The first main point is that there is a single matrix, which they write $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, that changes coordinates with respect to $\mathcal{B}$ into coordinates with respect to $\mathcal{C}$ by multiplication. Specifically, $[\mathbf{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P} \cdot [\mathbf{v}]_{\mathcal{B}}$ (Theorem 6.6.1). Note that there are no "standard bases" involved in this process. We call $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ the *transition matrix*, or the *change-of-basis matrix*, from $\mathcal{B}$ to $\mathcal{C}$.

The second main point is how to find the P matrix. I showed this in class, without bothering with standard bases. The formula is simple: Column i of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ equals $[\mathbf{b}_i]_{\mathcal{C}}$, the coordinate vector in the second basis $\mathcal{C}$ of the i-th basis vector in the first basis $\mathcal{B}$. I will show this in matrix form:

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \big[[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}}\ \cdots\ [\mathbf{b}_n]_{\mathcal{C}}\big].$$

(A way to think of basis change is to put $U = V$ and $T =$ the identity transformation in the diagram above, but you don't have to think of it that way.)