**Supplement on Representation of Functions in Different Coordinates**

Though you may believe you understand the basics of vectors, matrices, bases, and coordinates relative to a basis, there are some subtle aspects that warrant additional explanation. For example, if your concept of a vector in $\mathbf{R}^n$ is "an ordered $n$-tuple" or some similar definition, then this doesn't really hold up objectively.

If you were to change units, for example, the components of the vector might be completely different but still represent the same vector. In physics, the acceleration due to Earth's gravity is a vector pointing downward, but is it 32 ft/sec$^2$ or 9.8 m/sec$^2$? The numerical value of the downward component can be different things depending on what coordinates you choose. Nonetheless, the physics is still the physics. It doesn't change just because you decide to use different units.

The same ambiguity applies to the description of functions. If we have a function represented as $y = f(x) = x^2$, think about what happens if you change coordinates by letting $x = u - 3$ and $y = 12v$. The former is a shift in the horizontal axis, and the latter is a linear change of scale. Substituting, we get $12v = (u-3)^2$ or $v = \frac{1}{12}(u-3)^2 = \tilde{f}(u)$. In the new coordinates, the function has a different algebraic representation (formula), but it still describes the same parabola. There's an interesting way to think about this in terms of the functions that determine the coordinate changes. If we write $x = h(u) = u - 3$ and $y = k(v) = 12v$, then $\tilde{f} = k^{-1} \circ f \circ h$. This can be schematically understood via the following diagram (where the variables are appended for guidance):

$$\begin{array}{ccc} \{\mathbf{R}, x\} & \xrightarrow{f} & \{\mathbf{R}, y\} \\ h\uparrow & & \uparrow k \\ \{\mathbf{R}, u\} & \xrightarrow{\tilde{f}} & \{\mathbf{R}, v\} \end{array}$$

A given function can, in fact, be represented in arbitrarily many different ways. We require only that the appropriate coordinate changes be understood and that the relationship between different representations be determined by a diagram such as the one above. If so, we'll say that $f$ and $\tilde{f}$ are <u>equivalent</u>.

This is somewhat simplified in the case where the coordinate change is the same in both the domain and range of a given function. If this change is given by a function $h$, we get the simpler relation $\tilde{f} = h^{-1} \circ f \circ h$ and the corresponding simplified schematic:

$$\begin{array}{ccc} \mathbf{R} & \xrightarrow{f} & \mathbf{R} \\ h\uparrow & & \uparrow h \\ \mathbf{R} & \xrightarrow{\tilde{f}} & \mathbf{R} \end{array}$$

Let's focus on how this plays out in the context of vectors and matrices. In our standard view of $\mathbf{R}^n$ we can think of a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + \cdots x_n\mathbf{e}_n$ where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, …, $\mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$. We refer to the numbers $\{x_1, x_2, \ldots, x_n\}$ as the <u>standard coordinates</u> of the vector. We say that these are the coordinates of the vector relative to the <u>standard basis</u> $\mathcal{E} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$.

We have shown that if $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis for $\mathbf{R}^n$, then any vector $\mathbf{x}$ in $\mathbf{R}^n$ can be expressed <u>uniquely</u> as $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{S}[\mathbf{x}]_{\mathcal{B}}$. The matrix $\mathbf{S} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow \end{bmatrix}$ is called the <u>change of basis</u>

matrix. It is necessarily invertible. The vector $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ gives the coordinates of **x** relative to the basis $\mathcal{B}$.

Note that $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}} = \mathbf{S}[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{x}$. These tell us how to change coordinates.

---

**Example**: In $\mathbf{R}^2$, the vectors $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ form a basis $\mathcal{B}$ for $\mathbf{R}^2$. If we write $\mathbf{S} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, then for a vector such as $\mathbf{x} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$, we can calculate its coordinates relative to the basis $\mathcal{B}$ by $[\mathbf{x}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{x} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 13 \\ -11 \end{bmatrix} = \begin{bmatrix} 13/3 \\ -11/3 \end{bmatrix}$. You can verify that $\frac{13}{3}\mathbf{v}_1 - \frac{11}{3}\mathbf{v}_2 = \mathbf{x}$.

---

An $n$ by $n$ matrix **A** represents a linear function from $\mathbf{R}^n$ to $\mathbf{R}^n$, and matrix multiplication corresponds to composition of linear functions, i.e. $(\mathbf{AB})\mathbf{x} = \mathbf{A}(\mathbf{Bx})$. We can use the facts that $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}} = \mathbf{S}[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{x}$ in conjunction with our earlier observations about coordinate changes to define not only the idea of the coordinates of a vector relative to a basis, but also the idea of the matrix of a linear function relative to a basis. Specifically, if we let $\{\mathbf{R}^n, \mathcal{E}\}$ represent $\mathbf{R}^n$ with the coordinates of vectors expressed in terms of the standard basis $\mathcal{E}$, and let $\{\mathbf{R}^n, \mathcal{B}\}$ represent $\mathbf{R}^n$ with the coordinates of vectors expressed in terms of a different basis $\mathcal{B}$, and if we use the notation $[\mathbf{A}]_{\mathcal{B}}$ for the matrix representation of the linear function corresponding to matrix **A** but relative to the basis $\mathcal{B}$, then we have the following schematic:

$$\begin{array}{ccc} \{\mathbf{R}^n, \mathcal{E}\} & \xrightarrow{\mathbf{A}} & \{\mathbf{R}^n, \mathcal{E}\} \\ \mathbf{S}\uparrow & & \uparrow\mathbf{S} \\ \{\mathbf{R}^n, \mathcal{B}\} & \xrightarrow{[\mathbf{A}]_{\mathcal{B}}} & \{\mathbf{R}^n, \mathcal{B}\} \end{array}$$

From this we observe that $\boxed{[\mathbf{A}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{AS}}$. This is an extremely important result.

$[\mathbf{A}]_{\mathcal{B}}$ can also be calculated directly for a given basis $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ as $[\mathbf{A}]_{\mathcal{B}} = \begin{bmatrix} \uparrow & & \uparrow \\ [\mathbf{Av}_1]_{\mathcal{B}} & \cdots & [\mathbf{Av}_n]_{\mathcal{B}} \\ \downarrow & & \downarrow \end{bmatrix}$.

---

**Example**: Suppose we want to find the matrix **A** for the linear transformation $T$ representing orthogonal projection in $\mathbf{R}^2$ onto the line (through the origin) that is rotated 30° from the horizontal. The vector $\mathbf{v}_1 = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$ is in the direction of this line, and the vector $\mathbf{v}_2 = \begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix}$ is perpendicular to this. Relative to this basis, we see that $T(\mathbf{v}_1) = \mathbf{v}_1$ and $T(\mathbf{v}_2) = \mathbf{0}$. The matrix of $T$ relative to the basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is therefore $[\mathbf{A}]_{\mathcal{B}} = \big[[T(\mathbf{v}_1)]_{\mathcal{B}} \quad [T(\mathbf{v}_2)]_{\mathcal{B}}\big] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. The change of basis matrix is $\mathbf{S} = \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$ and $\mathbf{S}^{-1} = \frac{1}{4}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix}$. Since $[\mathbf{A}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{AS}$, it follows that:

$$\mathbf{A} = \mathbf{S}[\mathbf{A}]_{\mathcal{B}}\mathbf{S}^{-1} = \frac{1}{4}\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} \sqrt{3} & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix}.$$