Coordinatization

In $\mathbb{R}^3$, we have the *standard basis* $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. When we write a vector in coordinate form, say

$$\mathbf{v} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}, \tag{1}$$

it is understood as

$$\mathbf{v} = 3\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}.$$

The numbers 3, (-2) and 5 are the coordinates of $\mathbf{v}$ relative to the standard basis $\xi = (\mathbf{i}, \mathbf{j}, \mathbf{k})$. It has always been understood that a coordinate representation such as that in (1) is with respect to the **ordered** basis ξ. A little thought reveals that it need not be so. One could have chosen the same basis elements in a different order, as in the basis $\xi' = (\mathbf{i}, \mathbf{k}, \mathbf{j})$. We employ notation indicating the coordinates are with respect to the different basis ξ':

$$[\mathbf{v}]_{\xi'} = \begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix}, \qquad \text{to mean that} \qquad \mathbf{v} = 3\mathbf{i} + 5\mathbf{k} - 2\mathbf{j},$$

reflecting the order in which the basis elements fall in ξ'. Of course, one could employ similar notation even when the coordinates are expressed in terms of the standard basis, writing $[\mathbf{v}]_\xi$ for (1), but whenever we have coordinatization with respect to the standard basis of $\mathbb{R}^n$ in mind, we will consider the $[\cdot]_\xi$ wrapper to be optional.

Of course, there are many non-standard bases of $\mathbb{R}^n$. In fact, any linearly independent collection of n vectors in $\mathbb{R}^n$ provides a basis. Say we take

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 4 \end{bmatrix} \mathbf{u}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 2 \end{bmatrix} \mathbf{u}_4 = \begin{bmatrix} 4 \\ -1 \\ 2 \\ -1 \end{bmatrix}.$$

As per the discussion above, these vectors are being expressed relative to the standard basis of $\mathbb{R}^4$. A quick check of a related 4-by-4 determinant reveals these vectors are linearly independent, and hence form a basis of $\mathbb{R}^4$.

```
octave:80> A = [−1 0 3 0; 1 1 −1 4; −1 1 0 2; 4 −1 2 −1]'
A =

  −1   1  −1   4
   0   1   1  −1
   3  −1   0   2
   0   4   2  −1
```

```
octave:81> det(A)
ans = −17
```

Let us denote the ordered basis as $\alpha = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$. Every vector of $\mathbb{R}^4$ can be expressed uniquely as a linear combination of elements in the basis α. Gaussian elimination is a tool that reveals the weights.

Example 1:

Find the coordinatization relative to the basis α of the vector $\mathbf{v} = (8, -3, 13, -5)$.

We can reduce the relevant augmented matrix to RREF:

```
octave:85> augA = [A [8; −3; 13; −5]]
augA =

  −1    1   −1    4    8
   0    1    1   −1   −3
   3   −1    0    2   13
   0    4    2   −1   −5

octave:86> rref(augA)
ans =

   1   0   0   0    2
   0   1   0   0   −1
   0   0   1   0    1
   0   0   0   1    3
```

Thus, $\mathbf{v} = 2\mathbf{u}_1 - \mathbf{u}_2 + \mathbf{u}_3 + 3\mathbf{u}_4$, and

$$[\mathbf{v}]_\alpha = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 3 \end{bmatrix}.$$

■

Matrix of a Linear Transformation

Let $T\colon \mathcal{V} \to \mathcal{W}$ be a linear transformation from the n-dimensional vector space $\mathcal{V}$ having ordered basis $\alpha = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ into to the m-dimensional vector space $\mathcal{W}$, having ordered basis $\beta = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m)$. As we have seen, as soon as one knows the outputs from T for a basis, all of its behavior is known. Let us suppose we know the outputs of T on the basis α. More specifically, let us assume we know how each $T(\mathbf{v}_k)$, an element of $\mathcal{W}$, is expressed relative to the basis β. Specifically, we know the weights a_{ik} in the expression

$$T(\mathbf{v}_k) \;=\; a_{1k}\mathbf{w}_1 + a_{2k}\mathbf{w}_2 + \cdots + a_{mk}\mathbf{w}_m, \qquad \text{so that} \qquad [T(\mathbf{v}_k)]_\beta \;=\; \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}.$$

Then the **matrix of T in the ordered bases α and β** is

$$\big[T\big]_\alpha^\beta \;=\; \begin{bmatrix} \big[T(\mathbf{v}_1)\big]_\beta & \big[T(\mathbf{v}_2)\big]_\beta & \cdots & \big[T(\mathbf{v}_n)\big]_\beta \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \;=\; \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{2}$$

Note that, if $\mathcal{V} = \mathcal{W}$, it is still possible to have different bases α and β, but if the same basis is used twice, we can simplify the notation and write $\big[T\big]_\alpha^\alpha$ as $\big[T\big]_\alpha$. You may realize from this discussion that there are many different matrix representations for the same linear transformation T, potentially one for each pair of bases α and β.

Example 2:

In $\mathbb{R}^2$ and $\mathbb{R}^3$ consider ordered bases $\alpha = (\mathbf{v}_1, \mathbf{v}_2)$ and $\beta = (\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$, respectively, where

$$\mathbf{v}_1 \;=\; \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 \;=\; \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{w}_1 \;=\; \begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 \;=\; \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}, \quad \mathbf{w}_3 \;=\; \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}.$$

Suppose $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ is linear, and produces the following images for the elements in α:

$$T(\mathbf{v}_1) \;=\; \begin{bmatrix} 5 \\ -1 \\ 11 \end{bmatrix} \qquad \text{and} \qquad T(\mathbf{v}_2) \;=\; \begin{bmatrix} 5 \\ 2 \\ -7 \end{bmatrix}.$$

Find the matrix $\big[T\big]_\alpha^\beta$.

Since the augmented matrices

$$\begin{bmatrix} 4 & 0 & -1 & 5 \\ -2 & 3 & 2 & -1 \\ 0 & -2 & 3 & 11 \end{bmatrix} \;\sim\; \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix}, \qquad \text{we have} \qquad \big[T(\mathbf{v}_1)\big]_\beta \;=\; \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad \text{and}$$

$$\begin{bmatrix} 4 & 0 & -1 & 5 \\ -2 & 3 & 2 & 2 \\ 0 & -2 & 3 & -7 \end{bmatrix} \;\sim\; \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \qquad \text{so} \qquad \big[T(\mathbf{v}_2)\big]_\beta \;=\; \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Thus,

$$[T]_\alpha^\beta = \begin{bmatrix} [T(\mathbf{v}_1)]_\beta & [T(\mathbf{v}_2)]_\beta \\ \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \\ 3 & -1 \end{bmatrix}.$$

■

One expects that, having found the matrix $[T]_\alpha^\beta$, it can be used to find $[T(\mathbf{v})]_\beta$ for any input vector $\mathbf{v}$. Notice that one obtains

$$[T(\mathbf{v}_1)]_\beta = \begin{bmatrix} 2 & 1 \\ -1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}.$$

That is, the result $[T(\mathbf{v}_1)]_\beta$ comes from matrix multiplication after expressing $\mathbf{v}_1$ as $(1, 0)$, which is its coordinatization relative to the basis α. Speaking generally, it is the case that

$$[T(\mathbf{v})]_\beta = [T]_\alpha^\beta [\mathbf{v}]_\alpha. \tag{3}$$

Example 3:

$\mathcal{P}_3$, the vector subspace of $C(\mathbb{R})$ consisting of at-most-3$^{\text{rd}}$-degree polynomials, has ordered basis $\alpha = (1, x, x^2, x^3)$. Vectors/members of $\mathcal{P}_3$ can be coordinatized with respect to this basis. For instance, the polynomial

$$p(x) = 2x^3 - x + 7 \qquad \text{has coordinates} \qquad [p]_\alpha = \begin{bmatrix} 7 \\ -1 \\ 0 \\ 2 \end{bmatrix}.$$

The operation of differentiation, d/dx, is linear, and for each input $p(x)$ from $\mathcal{P}_3$ the image/result $d/dx(p) = p'(x)$ is in $\mathcal{P}_2$ and can be represented by coordinates relative to the ordered basis $\beta = (1, x, x^2)$ of $\mathcal{P}_2$. We have

$$[d/dx]_\alpha^\beta = \begin{bmatrix} [d/dx(1)]_\beta & [d/dx(x)]_\beta & [d/dx(x^2)]_\beta & [d/dx(x^3)]_\beta \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Using this, you can find the derivative of any polynomial in $\mathcal{P}_3$. For $p(x) = 2x^3 - x + 7$, for instance, its derivative is

$$[d/dx]_\alpha^\beta [p]_\alpha = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 7 \\ -1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix},$$

which should be understood as $p'(x) = 6x^2 - 1$.

■

An important result is the following, which affirms that matrix multiplication is defined in the right way.

Theorem 1: Suppose $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$ are finite-dimensional vector spaces with ordered bases α, β and γ, respectively. Suppose that $S\colon \mathcal{U} \to \mathcal{V}$ and $T\colon \mathcal{V} \to \mathcal{W}$ are linear transformations. Then

- The composition $T \circ S$ given by $(T \circ S)(\mathbf{x}) = T(S(\mathbf{x}))$ is a linear transformation from $\mathcal{U}$ to $\mathcal{W}$.
- The matrix of $T \circ S$ in the ordered bases α and γ is the product of the matrix of S in the ordered bases α and β with the matrix of T in the ordered bases β and γ; that is

$$\left[T \circ S\right]_{\alpha}^{\gamma} = \left[T\right]_{\beta}^{\gamma} \left[S\right]_{\alpha}^{\beta}.$$

Change of Basis

A special instance is the identity operator $\mathrm{id}\colon \mathbb{R}^n \to \mathbb{R}^n$, given by $\mathrm{id}(\mathbf{x}) = \mathbf{x}$. If ξ represents the standard basis for $\mathbb{R}^n$, and $\alpha = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is another ordered basis, then

$$\left[\mathrm{id}\right]_{\alpha}^{\xi} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}.$$

We call this a **change-of-basis**, or **transition matrix**, as it maps a vector from its coordinate representation relative to basis α to its representation in another (in this case, the standard) basis, ξ:

$$\left[\mathbf{v}\right]_{\xi} = \left[\mathrm{id}\right]_{\alpha}^{\xi} \left[\mathbf{v}\right]_{\alpha}.$$

What this means is that *any* nonsingular matrix $\mathbf{X}$ can be viewed as a change-of-basis matrix, mapping from coordinates relative to the ordered basis α consisting of the columns of $\mathbf{X}$ to standard coordinates. The matrix which maps from standard coordinates back to coordinates in the ordered basis α is naturally $\mathbf{X}^{-1}$. That is, if $\left[\mathrm{id}\right]_{\alpha}^{\xi} = \mathbf{X}$, then $\left[\mathrm{id}\right]_{\xi}^{\alpha} = \mathbf{X}^{-1}$.

As with general linear maps, one might use two nonstandard bases $\alpha = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and β for $\mathbb{R}^n$. The matrix of the identity map $\mathrm{id}\colon \mathbb{R}^n \to \mathbb{R}^n$ in the ordered bases α and β, since $\mathrm{id}(\mathbf{v}_j) = \mathbf{v}_j$, is found by tailoring the formula (2) appropriately:

$$[\mathrm{id}]_\alpha^\beta \;=\; \begin{bmatrix} [\mathbf{v}_1]_\beta & [\mathbf{v}_2]_\beta & \cdots & [\mathbf{v}_n]_\beta \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}.$$

This, too, serves to change coordinates

$$[\mathbf{v}]_\beta \;=\; [\mathrm{id}]_\alpha^\beta [\mathbf{v}]_\alpha$$

and is thus a change-of-basis matrix. Note that, whenever the same basis is used, we have $[\mathrm{id}]_\alpha = \mathbf{I}$, the usual identity matrix.

Now, suppose $\mathbf{A}$ is diagonalizable. Then there exists a basis α of $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$. Constructing a matrix $\mathbf{X}$ from that basis, we have

$$\mathbf{A} \;=\; \mathbf{X\Lambda X}^{-1} \;=\; [\mathrm{id}]_\alpha^\xi\, \mathbf{\Lambda}\, [\mathrm{id}]_\xi^\alpha.$$

That is, we can view this factorization as a change first from standard coordinates to ones in α (the action of $\mathbf{X}^{-1}$), then rescalings in those coordinates (using eigenvalues, the diagonal elements of $\mathbf{\Lambda}$, as scalars), followed by a return from coordinates in α to standard ones. More generally, every matrix $\mathbf{A}$ is is similar to a Jordan form matrix

$$\mathbf{A} \;=\; \mathbf{BJB}^{-1} \;=\; [\mathrm{id}]_\beta^\xi\, \mathbf{J}\, [\mathrm{id}]_\xi^\beta,$$

where β is the ordered basis arising from the columns of the matrix $\mathbf{B}$.

Next, let's assume only that the real matrix $\mathbf{A}$ is m-by-n, and take $\mathbf{U}$, $\mathbf{V}$ to be the usual matrices of the singular value decompositon of $\mathbf{A}$. Write α for the ordered basis of $\mathbb{R}^n$ comprised of the columns of $\mathbf{V}$, β for the ordered basis of $\mathbb{R}^m$ arising from the columns of $\mathbf{U}$, ξ_m for the standard basis of $\mathbb{R}^m$, and ξ_n for the standard basis of $\mathbb{R}^n$. Then

$$\mathbf{A} \;=\; \mathbf{U\Sigma V}^{\mathrm{T}} \;=\; [\mathrm{id}]_\beta^{\xi_m}\, \mathbf{\Sigma}\, [\mathrm{id}]_{\xi_n}^\alpha.$$

Similarity

Finally, we mentioned before that there can be many different matrix representations for one linear operator $T\colon \mathcal{V} \to \mathcal{W}$. In the case where the domain $\mathcal{V}$ and codomain $\mathcal{W}$ are the same, so that one can talk about a matrix $[T]_\alpha^\alpha$ (abbreviated as $[T]_\alpha$), there are still many matrix representations of T, one for each different basis α of $\mathcal{V}$. However, any two are related, are, in fact similar. Specifically, if α and β are both ordered bases of $\mathcal{V}$, then

$$[T]_\alpha \;=\; [\mathrm{id}]_\beta^\alpha\, [T]_\beta\, [\mathrm{id}]_\alpha^\beta \;=\; [\mathrm{id}]_\beta^\alpha\, [T]_\beta\, ([\mathrm{id}]_\beta^\alpha)^{-1} \;=\; \mathbf{C}\, [T]_\beta\, \mathbf{C}^{-1},$$

where the matrix $\mathbf{C}$ relating the two is the transition matrix from β-coordinates to α-coordinates.