Matrix Representations (Optional)

Prerequisites:

- Linear Transformations
- Change of Basis

Suppose $\boxed{\mathrm{A}}$ is an $N \times M$ matrix. Consider the linear function $f : \mathbb{R}^M \longrightarrow \mathbb{R}^N$ determined by multiplication-by-$\boxed{\mathrm{A}}$:

$$f(\vec{x}) = \boxed{\mathrm{A}}.\vec{x}$$

Suppose that $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_M\}$ is the "standard basis" for $\mathbb{R}^M$, and $\mathcal{E}' = \{\mathbf{e}'_1, \ldots, \mathbf{e}'_N\}$ is the "standard basis" for $\mathbb{R}^N$. That is,

$$\text{for all } m \in [1..M], \quad \mathbf{e}_m = \Big(\overbrace{\underbrace{0, \ldots, 0}_{(m-1)}, 1, 0, \ldots, 0}^{M} \Big)$$

$$\text{and, for all } n \in [1..N], \quad \mathbf{e}'_n = \Big(\overbrace{\underbrace{0, \ldots, 0}_{(n-1)}, 1, 0, \ldots, 0}^{N} \Big)$$

Thus, if $\boxed{\mathrm{A}} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1M} \\ a_{21} & a_{22} & \ldots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \ldots & a_{NM} \end{bmatrix}$, then, for all $m \in [1..M]$,

$$f(\mathbf{e}_m) = \sum_{n=1}^{N} a_{n,m}\mathbf{e}'_n$$

In other words, the matrix $\boxed{\mathrm{A}}$ describes the f-images of the *standard* basis elements of $\mathbb{R}^M$, as linear combinations of the *standard* basis elements of $\mathbb{R}^N$. We could say that $\boxed{\mathrm{A}}$ is the **matrix representation** of f "in terms of" the bases $\mathcal{E}$ and $\mathcal{E}'$.

This is convenient, because it allows easy computation of f. Suppose $\vec{x} = (x_1, \ldots, x_M)$. Thus,

$$\vec{x} = \sum_{m=1}^{M} x_m \mathbf{e}_m$$

$$\text{therefore,} \quad f(\vec{x}) = f\left(\sum_{m=1}^{M} x_m \mathbf{e}_m \right)$$

$$= \sum_{m=1}^{M} x_m f(\mathbf{e}_m)$$
$$= \sum_{m=1}^{M} x_m \sum_{n=1}^{N} a_{n,m}\mathbf{e}'_n$$
$$= \sum_{n=1}^{N} \left(\sum_{m=1}^{M} x_m a_{n,m} \right) \mathbf{e}'_n$$

In other words, $f(\vec{x}) = (y_1, \ldots, y_N)$, where

$$y_n = \sum_{m=1}^{M} x_m a_{n,m}, \quad \text{for each } n \in [1...N].$$

Of course, this is just another way of saying:

$$f(\vec{x}) = \boxed{\mathrm{A}} \cdot \vec{x}$$

which is something we already knew.

Now suppose that $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_M\}$ was some *other* basis of $\mathbb{R}^M$, and $\mathcal{C}' = \{\mathbf{c}'_1, \ldots, \mathbf{c}'_N\}$ was some other basis of $\mathbb{R}^N$. We would like to find a matrix representation for f in terms of $\mathcal{C}$ and $\mathcal{C}'$. In other words, we would like to find a matrix $\boxed{\mathrm{B}} = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1M} \\ b_{21} & b_{22} & \ldots & b_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \ldots & b_{NM} \end{bmatrix}$, so that, for all $m \in [1..M]$,

$$f(\mathbf{c}_m) = \sum_{n=1}^{N} b_{n,m}\mathbf{c}'_n$$

Example 1: Let $\mathbb{L} \subset \mathbb{R}^2$ be the line through the origin at an angle of -45 degrees (ie. $\frac{-\pi}{4}$ radians). $\mathbb{L}$ is a linear subspace, with equation:

$$\mathbb{L} = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 = -x_2\}$$

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the **orthogonal projection** onto $\mathbb{L}$. Now,

$$f(1,0) = \left(\frac{1}{2}, \frac{-1}{2}\right), \qquad \text{and } f(0,1) = \left(\frac{-1}{2}, \frac{1}{2}\right)$$

so, with respect to the *standard* basis, the matrix representation of f is

$$\boxed{\mathrm{A}} = \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}$$

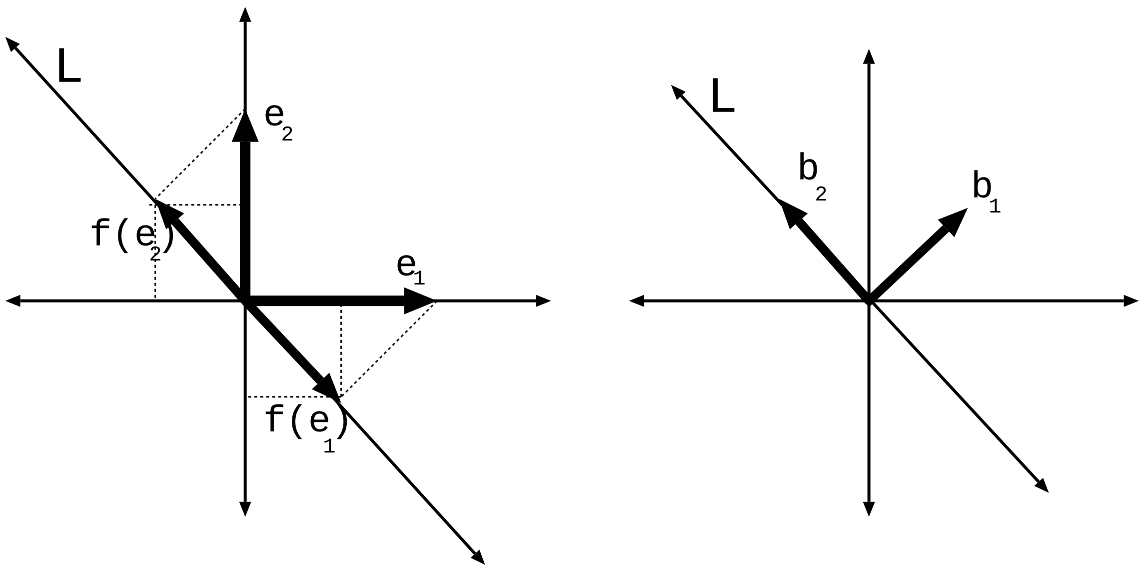

Figure 1: Projecting onto the line through 0 at angle $\frac{-\pi}{4}$.

However, suppose we use the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, where

$$\begin{aligned} \mathbf{b}_1 &= \left(\frac{1}{2}, \frac{1}{2}\right) \\ \mathbf{b}_2 &= \left(\frac{-1}{2}, \frac{1}{2}\right) \end{aligned}$$

In that case, $\begin{aligned} f(\mathbf{b}_1) &= 0, \\ f(\mathbf{b}_2) &= \mathbf{b}_2, \end{aligned}$

Hence, the matrix representation of f with respect to $\mathcal{B}$ is

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Notice how, with respect to *this* basis, the action of f as a “projection” is much more clearly visible.

Matrix Representations in Abstract Vector Spaces

We can generalize these considerations to any finite dimensional vector space. If $f : \mathbb{W} \longrightarrow \mathbb{V}$ is a linear transformation, and $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_N\}$ is a **basis** for $\mathcal{V}$, then, for any $\mathbf{w} \in \mathbb{W}$, we can write the **coordinates** of $f(\mathbf{w})$ with respect to $\{\mathbf{v}_1, \ldots, \mathbf{v}_N\}$:

$$f(\mathbf{w}) \;=\; \sum_{n=1}^{N} a_n \mathbf{v}_n$$

for some set of coefficients $\{a_1, \ldots, a_N\}$. If $\mathcal{W} = \{\mathbf{w}_1, \ldots, \mathbf{w}_M\}$ is a basis for $\mathbb{W}$, and we do this for each element of this basis, then we can make a matrix.....

Definition 2: *Matrix Representation*

Let $\mathbb{W}$ and $\mathbb{V}$ be finite-dimensional vector spaces, and $f : \mathbb{W} \longrightarrow \mathbb{V}$ be a linear transformation.

Suppose that $\mathcal{W} = \{\mathbf{w}_1, \ldots, \mathbf{w}_M\}$ is a basis for $\mathbb{W}$, and $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_N\}$ is a basis for $\mathbb{V}$. The **matrix representation** of f with respect to the bases $\mathcal{W}$ and $\mathcal{V}$ is the $N \times M$ matrix $\boxed{\mathsf{A}}$ so that, for all $m \in [1..M]$, the *mth column* of $\boxed{\mathsf{A}}$ is the *coordinates* of $f(\mathbf{w}_m)$ with respect to the basis $\mathcal{V}$. In other words, if

$$\boxed{\mathsf{A}} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1M} \\ a_{21} & a_{22} & \ldots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \ldots & a_{NM} \end{bmatrix}$$

then, for all $m \in [1..M]$,

$$f(\mathbf{w}_m) \;=\; \sum_{n=1}^{N} a_{n,m} \mathbf{v}_n$$

Remark 3: Suppose that $\mathbf{x} \in \mathbb{W}$ is any vector, and suppose that $\mathbf{x}$ has coordinates $\vec{x} = (x_1, \ldots, x_M)$ with respect to the basis $\mathcal{W}$. In other words,

$$\mathbf{x} \;=\; \sum_{m=1}^{M} x_m \mathbf{w}_m$$

Then, using the matrix $\boxed{\mathrm{A}}$, it is easy to compute the coordinates of $f(\vec{x})$ with respect to the basis $\mathcal{V}$:

$$\begin{aligned} f(\vec{x}) &= f\left(\sum_{m=1}^{M} x_m \mathbf{w}_m\right) \\ &= \sum_{m=1}^{M} x_m f(\mathbf{w}_m) \end{aligned}$$

$$= \sum_{m=1}^{M} x_m \left(\sum_{n=1}^{N} a_{n,m} \mathbf{v}_n \right)$$
$$= \sum_{n=1}^{N} \left(\sum_{m=1}^{M} x_m a_{n,m} \right) \mathbf{v}_n$$
$$= \sum_{n=1}^{N} y_n \mathbf{v}_n$$

where $y_n = \sum_{m=1}^{M} x_m a_{n,m}$. Thus, if $f(\mathbf{x}) = \mathbf{y}$, then $\mathbf{y}$ has coordinates $\vec{y} = (y_1, \ldots, y_N)$ with respect to the basis $\mathcal{V}$. But it is easy to see that the previous computation is equivalent to writing:

$$\vec{y} = \boxed{\mathrm{A}} \cdot \vec{x}$$

Thus, given a basis $\mathcal{W}$ for $\mathbb{W}$, and a basis $\mathcal{V}$ for $\mathbb{V}$, *any linear transformation f from $\mathbb{W}$ to $\mathbb{V}$ can be represented as matrix multiplication*, simply by computing the matrix representation of f with respect to $\mathcal{W}$ and $\mathcal{V}$.

This is the power of matrix representations, and one of the chief computational advantages you obtain by constructing explicit bases for $\mathbb{W}$ and $\mathbb{V}$.

Example 4: Suppose $f : \mathbb{R}^M \longrightarrow \mathbb{R}^N$ a linear transformation, and suppose that f is equivalent to multiplication by the matrix $\boxed{\mathrm{F}} \in \mathcal{M}_{N \times M}$.

Then, if $\mathcal{E}$ is the **standard basis** for $\mathbb{R}^M$ and $\mathcal{E}'$ is the standard basis for $\mathbb{R}^N$, then $\boxed{\mathrm{F}}$ is the **matrix representation** of f with respect to $\mathcal{E}$ and $\mathcal{V}$,

Change-of-Basis for Matrix Representations

In general, given a "nonstandard" basis $\mathcal{B}$ for $\mathbb{R}^M$ and a basis $\mathcal{A}$ for $\mathbb{R}^N$, how can we find the matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{A}$?

Theorem 5: *Change of Matrix Representation in Euclidean Space*

Suppose $f : \mathbb{R}^M \longrightarrow \mathbb{R}^N$ a linear transformation, equivalent to multiplication by the matrix $\boxed{\mathsf{F}} \in \mathcal{M}_{N \times M}$. (In other words, if $\mathcal{E}$ is the **standard basis** for $\mathbb{R}^M$ and $\mathcal{E}'$ is the standard basis for $\mathbb{R}^N$, then $\boxed{\mathsf{F}}$ is the matrix representation of f with respect to $\mathcal{E}$ and $\mathcal{E}'$),

Suppose that $\mathcal{B}$ is *another* basis for $\mathbb{R}^M$, and $\mathcal{A}$ another basis for $\mathbb{R}^N$. Let $\boxed{\mathsf{B}}$ be the **change-of-basis** matrix from $\mathcal{E}$ to $\mathcal{B}$, and $\boxed{\mathsf{A}}$ be the **change-of-basis** matrix from $\mathcal{E}'$ to $\mathcal{A}$. Define

$$\widetilde{\boxed{\mathsf{F}}} = \boxed{\mathsf{A}} \cdot \boxed{\mathsf{F}} \cdot \boxed{\mathsf{B}}^{-1}$$

Then $\widetilde{\boxed{\mathsf{F}}}$ is the matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{A}$.

Proof: Recall that, if $\boxed{\mathrm{B}}$ is the change-of-basis matrix from $\mathcal{E}$ to $\mathcal{B}$, then $\boxed{\mathrm{B}}^{-1}$ is the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{E}$. Thus, if $\mathbf{x} \in \mathbb{R}^M$, and the coordinates of $\mathbf{x}$ with respect to $\mathcal{B}$ are $\vec{x}^{\mathcal{B}} = (x_1^{\mathcal{B}}, \ldots, x_M^{\mathcal{B}})$, then the coordinates of $\mathbf{x}$ with respect to $\mathcal{E}$ are $\boxed{\mathrm{B}}^{-1} \cdot \vec{x}^{\mathcal{B}}$ —in other words, as M-tuples of real numbers, $\mathbf{x} = \boxed{\mathrm{B}}^{-1} \cdot \vec{x}^{\mathcal{B}}$.

Now, $\boxed{\mathrm{F}}$ is the matrix representation of f with respect to $\mathcal{E}$ and $\mathcal{E}'$, so $\mathbf{y} = \boxed{\mathrm{F}} \cdot \mathbf{x}$ is the coordinate N-tuple of $f(\mathbf{x})$, with respect to $\mathcal{E}'$.

But $\boxed{\mathrm{A}}$ is the change-of-basis matrix from $\mathcal{E}'$ to $\mathcal{A}$, so $\vec{y}^{\mathcal{A}} = \boxed{\mathrm{A}} \cdot \mathbf{y}$ is the coordinate N-tuple of $f(\mathbf{x})$ with respect to $\mathcal{A}$.

Putting it all together, we conclude that *if $\vec{x}^{\mathcal{B}}$ is the coordinate M-tuple of $\mathbf{x}$ with respect to $\mathcal{B}$, then the coordinate N-tuple of $f(\mathbf{x})$ with respect to $\mathcal{A}$ is $\vec{y}^{\mathcal{A}}$*, where

$$\vec{y}^{\mathcal{A}} = \boxed{\mathrm{A}} \cdot \mathbf{y} = \boxed{\mathrm{A}} \cdot \boxed{\mathrm{F}} \cdot \mathbf{x} = \boxed{\mathrm{A}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}^{-1} \cdot \vec{x}^{\mathcal{B}} = \widetilde{\boxed{\mathrm{F}}} \cdot \vec{x}^{\mathcal{B}}$$

in other words, $\widetilde{\boxed{\mathrm{F}}}$ is the matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{A}$.

□ [Theorem 5]

Corollary 6: Suppose $f : \mathbb{R}^M \longrightarrow \mathbb{R}^M$ a linear transformation, equivalent to multiplication by the matrix $\boxed{\mathrm{F}} \in \mathcal{M}_{M \times M}$.

Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_M\}$ is a basis for $\mathbb{R}^M$. Let $\boxed{\mathrm{B}}$ be the matrix whose columns are the elements of $\mathcal{B}$:

$$\boxed{\mathrm{B}} = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_M \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$$

Then $\widetilde{\boxed{\mathrm{F}}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}$ is the matrix representation of f with respect to $\mathcal{B}$.

Proof: $\boxed{\mathrm{F}}$ is the matrix representation of f relative to the standard basis $\mathcal{E}$. The change-of-basis matrix from $\mathcal{B}$ to $\mathcal{E}$ *is* the matrix $\boxed{\mathrm{B}}$ (check this). Now apply the previous theorem.

□ [Corollary 6]

Of course, this theorem can be generalized to abstract vector spaces...

Theorem 7: Let $\mathbb{V}$ and $\mathbb{W}$ be finite-dimensional vector spaces, and $f : \mathbb{V} \longrightarrow \mathbb{W}$ a linear transformation.

Suppose that $\mathcal{V}$ is a basis for $\mathbb{V}$, and $\mathcal{W}$ a basis for $\mathbb{W}$, and suppose that, with respect to $\mathcal{V}$ and $\mathcal{W}$, the matrix representation for f is $\boxed{\mathsf{F}}$.

Suppose that $\widetilde{\mathcal{V}}$ is *another* basis for $\mathbb{V}$, and $\widetilde{\mathcal{W}}$ another basis for $\mathbb{W}$. Let $\boxed{\mathsf{B}}$ be the **change-of-basis** matrix from $\mathcal{V}$ to $\widetilde{\mathcal{V}}$, and $\boxed{\mathsf{A}}$ be the **change-of-basis** matrix from $\mathcal{W}$ to $\widetilde{\mathcal{W}}$. Define

$$\widetilde{\boxed{\mathsf{F}}} = \boxed{\mathsf{A}} \cdot \boxed{\mathsf{F}} \cdot \boxed{\mathsf{B}}^{-1}$$

Then $\widetilde{\boxed{\mathsf{F}}}$ is the matrix representation of f with respect to $\widetilde{\mathcal{V}}$ and $\widetilde{\mathcal{W}}$.

Proof: Exercise. Try to generalize the previous proof.

□ [Theorem 7]

Here's the special case when we are only considering transformations from a vector space to itself:

Corollary 8: Let $\mathbb{V}$ be a finite-dimensional vector space, and $f : \mathbb{V} \longrightarrow \mathbb{V}$ a linear transformation. Suppose that $\mathcal{A}$ and $\widetilde{\mathcal{A}}$ are two bases for $\mathbb{V}$, and suppose that the matrix representation of f, with respect to $\mathcal{A}$, is $\boxed{\mathsf{F}}$.

If $\boxed{\mathsf{B}}$ is the **change-of-basis** matrix from $\mathcal{A}$ to $\widetilde{\mathcal{A}}$, then

$$\widetilde{\boxed{\mathsf{F}}} = \boxed{\mathsf{B}} \cdot \boxed{\mathsf{F}} \cdot \boxed{\mathsf{B}}^{-1}$$

is the matrix representation of f with respect to $\widetilde{\mathcal{A}}$

□