CHANGE OF BASIS FOR LINEAR SPACES

1. Motivating Example

Consider P_2 the space of polynomials of degree at most 2. Let $T : P_2 \to P_2$ be defined by

$$T(p)(x) = p'(x) - p(x).$$

This is a linear transformation. We would like to associate to T a matrix (as this can make computations easier. To do so we need to pick a basis of P_2. Let

$$\mathcal{B} = (1, x, x^2)$$

be the standard basis, so

$$L_{\mathcal{B}}(a_0 + a_1 x + a_2 x^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}.$$

We compute

$$\begin{aligned} T(a_0 + a_1 x + a_2 x^2) &= (a_0 + a_1 x + a_2 x^2)' - (a_0 + a_1 x + a_2 x^2) \\ &= a_1 + 2a_2 x - a_0 - a_1 x - a_2 x^2 \\ &= (a_1 - a_0) + (2a_2 - a_1)x - a_2 x^2. \end{aligned}$$

This gives

$$\begin{array}{ccc} p(x) = a_0 + a_1 x + a_2 x^2 & \xrightarrow{T} & (a_1 - a_0) + (2a_2 - a_1)x - a_2 x^2 = T(p)(x) \\ \big\downarrow L_{\mathcal{B}} & & \big\downarrow L_{\mathcal{B}} \\ [p(x)]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} & \xrightarrow{T_{\mathcal{B}}} & \begin{bmatrix} a_1 - a_0 \\ 2a_2 - a_1 \\ -a_2 \end{bmatrix} = [T(p)(x)]_{\mathcal{B}} \end{array}$$

Here $T_{\mathcal{B}} : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation. It has matrix

$$[T_{\mathcal{B}}] = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{bmatrix}.$$

2. $\mathcal{B}$-matrix of a linear transformation

We generalize the preceding example. Let V be a linear space with basis $\mathcal{B} = (v_1, \ldots, v_n)$. For a linear transform $T : V \to V$ define a matrix, called the *$\mathcal{B}$-matrix of T* by

$$[T]_{\mathcal{B}} = \big[[T(v_1)]_{\mathcal{B}} \quad | \quad \cdots \quad | \quad [T(v_n)]_{\mathcal{B}} \big]$$

That is, the columns of the matrix are precisely the $\mathcal{B}$-coordinate vectors of the images under T of the elements of the basis $\mathcal{B}$. This is a direct generalization of the matrix of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with respect to a basis $\mathcal{B}$ of $\mathbb{R}^n$.

Let's write $B = [T]_{\mathcal{B}}$. This matrix ensures that the following "diagrams" hold

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ \downarrow{\scriptstyle L_{\mathcal{B}}} & & \downarrow{\scriptstyle L_{\mathcal{B}}} \\ \mathbb{R}^n & \xrightarrow{B} & \mathbb{R}^n \end{array} \qquad \begin{array}{ccc} p & \xrightarrow{T} & T(p) \\ \downarrow{\scriptstyle L_{\mathcal{B}}} & & \downarrow{\scriptstyle L_{\mathcal{B}}} \\ [p]_{\mathcal{B}} & \xrightarrow{B} & B[p]_{\mathcal{B}} = [T(p)]_{\mathcal{B}} \end{array}$$

where here the bottom arrow is multiplication by B. In other words, to find the value of $T(p)$ one computes

$$T(p) = L_{\mathcal{B}}^{-1}(B[p]_{\mathcal{B}}) = L_{\mathcal{B}}^{-1}([T]_{\mathcal{B}}[p]_{\mathcal{B}}).$$

EXAMPLE: Consider P_2 with basis $\mathcal{B} = (1, x, x^2)$ and $T(p) = p' - p$ we compute

$$T(1) = -1 \Rightarrow [T(1)]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \text{ and } T(x) = 1 - x \Rightarrow [T(x)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\text{and } T(x^2) = 2x - x^2 \Rightarrow [T(x^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$$

$$\Rightarrow [T]_{\mathcal{B}} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{bmatrix}.$$

This is $[T_{\mathcal{B}}]$ from before.

3. Image and Kernel

Fix a linear space V with basis $\mathcal{B} = (v_1, \ldots, v_n)$. The transformations $L_{\mathcal{B}}$ and $L_{\mathcal{B}}^{-1}$ give a dictionary between V and $\mathbb{R}^n$. In particular, they give a dictionary between the image and kernel of T and of $[T]_{\mathcal{B}}$.

EXAMPLE: Consider the linear transformation $T : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ defined by

$$T(A) = A \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} A.$$

Determine a basis of $\ker(T)$ and $\mathrm{Im}\,(T)$. Using the basis $\mathcal{B} = (e_{11}, e_{12}, e_{21}, e_{22})$ where

$$e_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, e_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, e_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, e_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

One computes,

$$B = [T]_{\mathcal{B}} = \left[[T(e_{11})]_{\mathcal{B}} \quad | \quad [T(e_{12})]_{\mathcal{B}} \quad | \quad [T(e_{21})]_{\mathcal{B}} \quad | \quad [T(e_{22})]_{\mathcal{B}} \right] = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 1 & -2 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{For example: } [T(e_{12})]_{\mathcal{B}} = \left(\begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \right)_{\mathcal{B}} = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}.$$

Moreover,

$$\operatorname{rref}(B) = \begin{bmatrix} 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence, the pivot columns are 1st and 3rd and free columns are 2nd and 4th. The kernel of B consists of solutions to $x_1 - 2x_2 - x_4 = 0$ and $x_3 = 0$. That is, by taking $x_2 = 1, x_4 = 0$ and $x_2 = 0, x_4 = 0$ and solving these simple systems one obtains

$$\ker(B) = \ker(\operatorname{rref}(B)) = \operatorname{span}\left(\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right).$$

This corresponds to

$$ker(T) = \operatorname{span}\left(\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right).$$

Likewise, the pivot columns form a basis of the image of B and so

$$\operatorname{Im}(B) = \operatorname{span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right).$$

Hence,

$$\operatorname{Im}(T) = \operatorname{span}(\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix} \right).$$

4. Change of Basis Matrix

Fix a linear space V with bases $\mathcal{B} = (v_1, \ldots, v_n)$ and $\mathcal{U} = (w_1, \ldots, w_n)$. Observe, that the map

$$L_{\mathcal{U}} \circ L_{\mathcal{B}}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$$

is the composition of isomorphisms and hence is an isomorphism. We call

$$S_{\mathcal{B}\to\mathcal{U}} = [L_{\mathcal{U}} \circ L_{\mathcal{B}}^{-1}] \in \mathbb{R}^{n \times n}$$

the *change of basis matrix from $\mathcal{B}$ to $\mathcal{U}$* Clearly, this is an invertible matrix. We have the diagrams

$$\begin{array}{ccc} V & \xrightarrow{L_{\mathcal{B}}} & \mathbb{R}^n \\ & {\searrow}^{L_{\mathcal{U}}} & \downarrow S_{\mathcal{B}\to\mathcal{U}} \\ & & \mathbb{R}^n \end{array} \qquad \begin{array}{ccc} p & \xrightarrow{L_{\mathcal{B}}} & L_{\mathcal{B}}(p) \\ & {\searrow}^{L_{\mathcal{U}}} & \downarrow S_{\mathcal{B}\to\mathcal{U}} \\ & & L_{\mathcal{U}}(p) = S_{\mathcal{B}\to\mathcal{U}}[p]_{\mathcal{B}}. \end{array}$$

EXAMPLE: Let $V \subset C^\infty$ be the space $V = \operatorname{span}(1, \cos(2x), \sin(2x))$. Compute $S_{\mathcal{U}\to\mathcal{B}}$ for the bases (I will leave it to you to check these are both bases)

$$\mathcal{B} = (1, \cos(2x), \sin(2x))$$

$$\mathcal{U} = (1, \cos^2(x), \sin(x)\cos(x))$$

To do so observe, that a basic trigonometric identity tells us that

$$\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 \text{ and } \sin(2x) = 2\cos(x)\sin(x).$$

In particular,

$$\begin{aligned} L_{\mathcal{U}}\left(L_{\mathcal{B}}^{-1}\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right)\right) &= L_{\mathcal{U}}(a + b\cos(2x) + c\sin(2x)) \\ &= L_{\mathcal{U}}(a - b + 2b\cos^2(x) + 2c\cos(x)\sin(x)) \\ &= \begin{bmatrix} a-b \\ 2b \\ 2c \end{bmatrix}. \end{aligned}$$

Hence,

$$S_{\mathcal{B}\to\mathcal{U}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

5. Change of basis for subspaces

EXAMPLE: Consider V to be the subspace of $\mathbb{R}^3$ given by $x_1 + x_2 + x_3 = 0$. This has bases

$$\mathcal{B} = \left(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right) \text{ and } \mathcal{U} = \left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right).$$

We compute

$$L_{\mathcal{U}}(L_{\mathcal{B}}^{-1}(\vec{e}_1)) = L_{\mathcal{U}}\left(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}\right) = L_{\mathcal{U}}\left(-2\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) = -2\vec{e}_1 + \vec{e}_2$$

and

$$L_{\mathcal{U}}(L_{\mathcal{B}}^{-1}(\vec{e}_2)) = L_{\mathcal{U}}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right) = L_{\mathcal{U}}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) = \vec{e}_1 - \vec{e}_2$$

Hence,

$$S_{\mathcal{B}\to\mathcal{U}} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}$$

Observe,

$$\begin{bmatrix} -2 & 1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}$$

Here the left hand matrix has columns the elements of $\mathcal{B}$ and the righthand side is the matrix with columns the elements of $\mathcal{U}$.

This last fact can be generalized as follows:

Theorem 5.1. *If $V \subset \mathbb{R}^n$ has basis $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_m)$ and $\mathcal{U} = (\vec{u}_1, \ldots, \vec{u}_m)$ then*

$$\left[\vec{b}_1 \mid \cdots \mid \vec{b}_m\right] = \left[\vec{u}_1 \mid \cdots \mid \vec{u}_m\right] S_{\mathcal{B}\to\mathcal{U}}.$$

6. Change of basis matrix and linear transformations

Fix a linear space V with with $\dim(V) = n$. Suppose that $T : V \to V$ is linear transformation from V to V. If $\mathcal{B} = (v_1, \ldots, v_n)$ and $\mathcal{U} = (u_1, \ldots, u_n)$ form two bases of V, then it is natural to ask what is the relationship between $[T]_\mathcal{B}$ and $[T]_\mathcal{U}$. That is, what is the relationship between the matrix of T with respect to the two bases.

Theorem 6.1. *In the above situation, if $S = S_{\mathcal{B}\to\mathcal{U}}$ is the change of basis matrix, then*

$$[T]_\mathcal{U} S = S[T]_\mathcal{B} \text{ and } [T]_\mathcal{U} = S[T]_\mathcal{B} S^{-1} \text{ and } [T]_\mathcal{B} = S^{-1}[T]_\mathcal{B} S.$$

A heuristic to remember the order of multiplication is the following: $[T]_\mathcal{U}$ "eats" a $\mathcal{U}$-coordinate vector and so is multiplied on the right by $S = S_{\mathcal{B}\to\mathcal{U}}$ (as this outputs $\mathcal{U}$-coordinate vectors). This product "eats" $\mathcal{B}$-vectors and outputs $\mathcal{B}$-coordinate vectors. Similarly, $[T]_\mathcal{B}$ outputs $\mathcal{B}$-coordinate vectors and so has to be multiplied on the left by $S = S_{\mathcal{B}\to\mathcal{U}}$ (which "eats" $\mathcal{B}$-vectors). This product "eats" $\mathcal{B}$-coordinate vectors and outputs $\mathcal{U}$-coordinate vectors.

EXAMPLE: Consider $V = span(1, \cos(2x), \sin(2x)) \subset C^\infty$ with basis

$$\mathcal{B} = (1, \cos(2x), \sin(2x)) \text{ and } \mathcal{U} = (1, \cos^2(x), \sin(x)\cos(x)).$$

One checks the map $D : V \to V$ given by $D(f) = f'$ is a well defined linear transformation. Indeed,

$$D(a + b\cos(2x)) + c\sin(2x)) = -2b\sin(2x) + 2c\cos(2x) \in V.$$

Clearly,

$$[D]_\mathcal{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & -2 & 0 \end{bmatrix}.$$

As $D(1) = 0, D(\cos^2(x)) = -2\cos(x)\sin(x)$ and

$$D(\cos(x)\sin(x)) = \cos^2(x) - \sin^2(x) = -1 + 2\cos^2(x),$$

$$[D]_\mathcal{U} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 2 \\ 0 & -2 & 0 \end{bmatrix}$$

We check

$$[D]_\mathcal{U} S_{\mathcal{B}\to\mathcal{U}} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 2 \\ 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 4 \\ 0 & -4 & 0 \end{bmatrix}$$

$$S_{\mathcal{B}\to\mathcal{U}}[D]_\mathcal{B} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 4 \\ 0 & -4 & 0 \end{bmatrix},$$

these agree as expected.