Coordinates

Throughout this note, V denotes an n-dimensional vector space with a fixed ordered basis $B = \{v_1, \ldots, v_n\}$. We begin with the following two exercises from the book.

Exercise 1. For each $v \in V$, there are unique scalars $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$v = \sum_{j=1}^{n} c_j v_j.$$

Exercise 2. Suppose W is another vector space, and $w_1, \ldots, w_n \in W$. There is a unique linear map $\Phi : V \to W$ such that $\Phi(v_j) = w_j$ for each j.

With these two exercises in hand, we can now define the coordinate map. We'll do so in 2 ways.

Definition 3 (Coordinates, Definition 1)**.** For $v \in V$, by Exercise 2, there are unique scalars $c_1, \ldots, c_n$ such that

$$v = \sum_{j=1}^{n} c_j v_j.$$

We define

$$[v]_B := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n.$$

The map $[\,\cdot\,]_B : V \to \mathbb{R}^n$ is called the *coordinate map*.

Exercise 4. Using only Definition 3 for $[\,\cdot\,]_B$, prove that $[\,\cdot\,]_B$ is linear and $[v_j] = e_j$ for all j.

We'll now give a second definition which is manifestly linear, but still requires one to check something.

Definition 5 (Coordinates, Definition 2)**.** By Exercise 2, there is a unique linear map $[\,\cdot\,]_B : V \to \mathbb{R}^n$ sending v_j to e_j for all j, i.e., $[v_j]_B = e_j$ for all j. We call $[\,\cdot\,]_B : V \to \mathbb{R}^n$ the *coordinate map*.

Exercise 6. Using only Definition 3 for $[\,\cdot\,]_B$, prove that

$$[v]_B = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \qquad \Longleftrightarrow \qquad v = \sum_{j=1}^{n} c_j v_j.$$

By Exercises 4 and 6, the Definitions 3 and 5 are equivalent, i.e., they define the same map.

We'll now compute explicit formulas for $[\,\cdot\,]_B$ when $V = \mathbb{R}^n$. Recall that for every linear map $\Phi : \mathbb{R}^n \to \mathbb{R}^m$, there is a unique $m \times n$ matrix A such that $\Phi = L_A$, where $L_A x = Ax$ for all $x \in \mathbb{R}^n$.

Proposition 7. *Suppose $V = \mathbb{R}^n$. Then $[\,\cdot\,]_B = L_{S^{-1}}$ where*

$$S = \left(\begin{array}{c|c|c} v_1 & \cdots & v_n \end{array} \right). \tag{1}$$

Proof. The inverse linear transformation of $[\,\cdot\,]_B$ maps e_j to v_j for all j. We know that this matrix is exactly L_S where S is as in Equation (1) above. Now $[\,\cdot\,]_B = L_S^{-1} = L_{S^{-1}}$. $\square$

We'll now compute an explicit formula for the unique linear map $L : V \to W$ given by $Lv_j = w_j$ when $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$.

Proposition 8. *Suppose $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, and $w_1, \ldots, w_n \in W$. By Exercise 2, there is a unique linear map $\Phi : V \to W$ such that $\Phi(v_j) = w_j$ for all j. Then the matrix A such that $\Phi = L_A$ is given by*

$$A = \left(\; w_1 \; \Big| \; \cdots \; \Big| \; w_n \; \right) \underbrace{\left(\; v_1 \; \Big| \; \cdots \; \Big| \; v_n \; \right)^{-1}}_{S^{-1}}.$$

Proof. We can write Φ as a composite map:

$$\begin{tikzcd} \mathbb{R}^n \arrow[rr, "\Phi"] \arrow[dr, "{[\,\cdot\,]_B}"'] & & \mathbb{R}^m \\ & \mathbb{R}^n \arrow[ur, "e_j \mapsto w_j"'] & \end{tikzcd}$$

In more detail, by Exercise 2, there is a unique linear map $\mathbb{R}^n \to W = \mathbb{R}^m$ which maps e_j to w_j for all j. We know that this linear map is given by left multiplication by the matrix

$$R := \left(\; w_1 \; \Big| \; \cdots \; \Big| \; w_n \; \right).$$

Again by Exercise 2, we have $\Phi = L_R \circ [\,\cdot\,]_B$, since both Φ and the composite map $L_R \circ [\,\cdot\,]_B$ map v_j to w_j for all j. Now by Proposition 7, we have

$$\Phi = L_R \circ [\,\cdot\,]_B = L_R \circ L_{S^{-1}} = L_{RS^{-1}}$$

where S is as in Equation (1). Hence $A = RS$ as claimed. ∎

Proposition 8 above is a special case of the following more general proposition.

Proposition 9. *Suppose V is a vector space with ordered basis $B = \{v_1, \ldots, v_n\}$, W is a vector space with ordered basis $C = \{w_1, \ldots, w_m\}$, and $\Phi : V \to W$ is a linear map. The matrix*

$$[\Phi]_B^C := \left(\; [\Phi v_1]_C \; \Big| \; \cdots \; \Big| \; [\Phi v_n]_C \; \right)$$

is the unique $m \times n$ matrix such that

$$(2) \qquad [\Phi]_B^C [v]_B = [\Phi v]_C \qquad \forall v \in V, \; \textit{i.e.,} \qquad \begin{tikzcd} V \arrow[r, "\Phi"] \arrow[d, "{[\,\cdot\,]_B}"'] & W \arrow[d, "{[\,\cdot\,]_C}"] \\ \mathbb{R}^n \arrow[r, "L_{[\Phi]_B^C}"] & \mathbb{R}^m \end{tikzcd} \quad \textit{commutes.}$$

Proof. Recall that the j-th column of a matrix A is given by Ae_j. Since $e_j = [v_j]_B$ for all j, such a matrix $[\Phi]_B^C$ exists satisfying Equation (2) if and only if its j-th column $[\Phi]_B^C e_j = [\Phi]_B^C [v]_B$ is equal to $[\Phi v_j]_C$. So defining $[\Phi]_B^C$ in this way gives the unique $m \times n$ matrix satisfying Equation (2). ∎

Exercise 10. Let V be the vector space of polynomials of degree at most 4, and let $B = \{1, x, x^2, x^3, x^4\}$. Let W be the vector space of polynomials of degree at most 3, and let $C = \{1, x, x^2, x^3\}$.

(1) Show the derivative $\frac{d}{dx} : V \to W$ given by $p(x) \mapsto p'(x)$ is a linear map, and compute $[\frac{d}{dx}]_B^C$.
(2) Show the second derivative $\frac{d^2}{dx^2} : V \to W$ given by $p(x) \mapsto p''(x)$ is a linear map, and compute $[\frac{d^2}{dx^2}]_B^C$.
(3) Show that the map $\frac{d}{dx} + \frac{d^2}{dx^2} : V \to W$ given by $p(x) \mapsto p'(x) + p''(x)$ is a linear map, and compute $[\frac{d}{dx} + \frac{d^2}{dx^2}]_B^C$. Compare your answer to $[\frac{d}{dx}]_B^C$ and $[\frac{d^2}{dx^2}]_B^C$.
Hint: Show that the sum of any two linear maps is again a llnear map.

Recall that two $n \times n$ matrices A_1, A_2 are called *similar* if there is an invertible $n \times n$ matrix S such that $S^{-1} A_1 S = A_2$.

Corollary 11. *When $V = W = \mathbb{R}^n$, $B = C = \{v_1, \dots, v_n\}$, and $\Phi = L_A$ for an $n \times n$ matrix A, then*

$$[\Phi]_B^C = S^{-1}AS$$

where the matrix S is given as in Equation (1) by $S = \left(\; v_1 \; \middle| \; \cdots \; \middle| \; v_n \; \right)$.

Proof. By Proposition 7 and Equation (2) above, we have that

$$L_{[\Phi]_B^C} = [\,\cdot\,]_B \circ L_A \circ [\,\cdot\,]_B^{-1} = L_{S^{-1}} \circ L_A \circ L_S = L_{S^{-1}AS},$$

and thus $[\Phi]_B^C = S^{-1}AS$ as claimed. □

We thus see that the equivalence relation of matrix similarlity is exactly expressing one linear transformation L_A, which is expressed in terms of the standard basis, in terms of a second basis, which is exactly the columns of S.

We now analyze the equivalence relation related to changing two bases for V and W separately.

Corollary 12. *When $V = \mathbb{R}^n$ and $B = \{v_1, \dots, v_n\}$, $W = \mathbb{R}^m$ and $C = \{w_1, \dots, w_m\}$, and $L = L_A$ for an $m \times n$ matrix A, then*

$$[L_A]_B^C = T^{-1}AS$$

where the matrices T and S are given as in Equation (2) by

$$S = \left(\; v_1 \; \middle| \; \cdots \; \middle| \; v_n \; \right) \qquad\qquad T = \left(\; w_1 \; \middle| \; \cdots \; \middle| \; w_m \; \right).$$

Proof. By Proposition 7 and Equation (2) above, we have that

$$L_{[\Phi]_B^C} = [\,\cdot\,]_C \circ L_A \circ [\,\cdot\,]_B^{-1} = L_{T^{-1}} \circ L_A \circ L_S = L_{T^{-1}AS},$$

and thus $[\Phi]_B^C = T^{-1}AS$ as claimed. □

Definition 13. We say two $m \times n$ matrices A_1, A_2 are *bi-similar*[1] if there is an invertible $m \times m$ matrix T and an invertible $n \times n$ matrix S such that $T^{-1}A_1S = A_2$.

Exercise 14. Show that two $m \times n$ matrices A_1, A_2 are row equivalent if and only if there is an $m \times m$ invertible matrix T such that $TA_1 = A_2$. Then show A_1 and A_2 are column equivalent if and only if there is an invertible $n \times n$ matrix S such that $A_1S = A_2$.

Exercise 15. Show that any elementary row and column operations commute.
Hint: They are performed by multiplying by an elementary matrix on the left and on the right respectively.

Exercise 16. Suppose that A is an $m \times n$ matrix. Show that one can obtain the matrix

$$\begin{pmatrix} I_p & 0_{p\times(n-p)} \\ 0_{(m-p)\times p} & 0_{(m-p)\times(n-p)} \end{pmatrix}$$

from A by performing a sequence of elementary row and column operations if and only if $p = \operatorname{rank}(A)$.

Theorem 17. *The following are equivalent for two $m \times n$ matrices A_1, A_2:*
(1) A_1 and A_2 are bi-similar.
(2) A_2 can be obtained from A_1 by performing a sequence of elementary row and column operations.
(3) $\operatorname{rank}(A_1) = \operatorname{rank}(A_2)$.

Proof.

[1] This is not standard notation.

(1) $\Rightarrow$ (2): Suppose A_1 and A_2 are bi-similar, so that $T^{-1}A_1S = A_2$ for some invertible $m \times m$ matrix T and some invertible $n \times n$ matrix S. Then by Exercise 14, A_2 can be obtained from A_1 by a sequence of elementary row and column operations, and (2) holds.

(2) $\Rightarrow$ (1): Suppose A_2 can be obtained from A_1 by a sequence of elementary row and column operations. Since elementary row and column operations commute by Exercise 15, we may first perform all the row operations, and then perform all the column operations to get from A_1 to A_2.

Let A_3 be the matrix obtained from A_1 by only performing the elementary row operations. By Exercise 14, there is an invertible matrix T such that $A_3 = T^{-1}A_1$. Now one can obtain A_2 from A_3 by performing only elementary column operations, so there is an invertible $n \times n$ matrix S such that $A_2 = A_3S = T^{-1}A_1S$. Hence A_1 and A_2 are bi-similar, and (1) holds.

(2) $\Leftrightarrow$ (3): Set $r_1 := \mathrm{rank}(A_1)$ and $r_2 := \mathrm{rank}(A_2)$, and define

$$E_1 := \begin{pmatrix} I_{r_1} & 0_{p\times(n-r_1)} \\ 0_{(m-r_1)\times r_1} & 0_{(m-r_1)\times(n-r_1)} \end{pmatrix} \qquad E_2 := \begin{pmatrix} I_{r_2} & 0_{p\times(n-r_2)} \\ 0_{(m-r_2)\times r_2} & 0_{(m-r_2)\times(n-r_2)} \end{pmatrix}$$

Then by Exercise 16, for $j = 1, 2$, E_j can be obtained from A_j by a sequence of elementary row and column operations.

If (2) holds, then A_2 can be obtained from A_1 by a sequence of elementary row and column operations. Since elementary row and column operations are invertible, we see that E_1 can be obtained from E_2 by elementary row and column operations. By the uniqueness statement from Exercise 16, we must have $r_1 = r_2$, and (2) holds.

If (3) holds, then $E_1 = E_2$, so again as elementary row and column operations are invertible, A_2 can be obtained from A_1 by a sequence of elementary row and column operations. Hence (2) holds. $\square$