Math 416 Lecture Note: Week 6

Daesung Kim

1 The matrix representation of a linear transformation

Let V be a finite dimensional vector space over $\mathbb{R}$ and $\beta = \{v_1, \cdots, v_n\}$ a basis for V. For any $v \in V$, there exists a unique linear combination

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

Thus, for fixed basis β, every vector v in V corresponds to a collection of scalars $a_1, a_1, \cdots, a_n$.

Definition 1.1. Let V be a finite dimensional vector space over $\mathbb{R}$ and $\beta = \{v_1, \cdots, v_n\}$ a basis for V. If $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$, we define the coordinate vector of v relative to β by

$$[v]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Example 1.2. Let $V = \mathcal{M}_{2\times 2}(\mathbb{R})$ and β be the standard basis for V, that is,

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Then, for any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have $[A]_\beta = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$.

Let V and W be finite dimensional vector spaces over $\mathbb{R}$ with bases $\beta = \{v_1, \cdots, v_n\}$ and $\gamma = \{w_1, \cdots, w_n\}$. Let $T : V \to W$ be a linear transformation. Then, there exists a unique expression

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

for each $j = 1, 2, \cdots, n$.

Definition 1.3. Using the notation above, we define the matrix representation of T with respect to β and γ by $[T]_\beta^\gamma = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{R})$. If $\beta = \gamma$, then we use the notation $[T]_\beta^\gamma = [T]_\beta$.

Remark 1.4. When we consider the coordinate vector and the matrix representation, the order of bases is important. In this context, we think a basis as an ordered set. It means that we give an order to each element in a set. Even though $A = B$ as sets, they could be not equal as ordered sets.

Example 1.5. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T(x, y) = (2x - y, x + y, x - y)$. Let β and γ be the standard bases for $\mathbb{R}^2$ and $\mathbb{R}^3$. Then, $T(1, 0) = 2e_1 + e_2 + e_3$ and $T(0, 1) = -e_1 + e_2 - e_3$. Thus,

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Example 1.6. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T(x, y) = (2x - y, x + y, x - y))$. Let $\beta = \{f_1 = (1, 0), f_2 = (1, 1)\}$ and γ be the standard basis for $\mathbb{R}^3$. Then, $T(1, 0) = 2e_1 + e_2 + e_3$ and $T(1, 1) = e_1 + 2e_2$. Thus,

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 0 \end{pmatrix}.$$

Definition 1.7. Let V and W be vector spaces over $\mathbb{R}$ and $T, U : V \to W$ linear transformations. For each $c \in \mathbb{R}$, we define $T + U, cT : V \to W$ by $(T + U)(v) = T(v) + U(v)$ and $(cT)(v) = cT(v)$ for all $v \in V$.

Definition 1.8. The collection of all linear transformations from V to W is denoted by $\mathcal{L}(V, W)$. If $V = W$, then we simply denote by $\mathcal{L}(V)$.

Theorem 1.9. *Let V and W be vector spaces over $\mathbb{R}$, then $\mathcal{L}(V, W)$ is a vector space over $\mathbb{R}$ with the addition and scalar multiplication defined above.*

Proof. Let $\mathcal{F}(V, W)$ the set of all functions from V to W, then $\mathcal{F}(V, W)$ is a vector space with the same addition and scalar multiplication. Since the set of all linear transformations from V to W is a subset of $\mathcal{F}(V, W)$, it suffices to show that $\mathcal{L}(V, W)$ is a subspace of $\mathcal{F}(V, W)$. It is trivial that $(cT + U)(0) = 0$. Let $T, U \in \mathcal{L}(V, W)$. Let $x, y \in V$ and $a \in \mathbb{R}$, then

$$\begin{aligned} (cT + U)(ax + y) &= cT(ax + y) + U(ax + y) \\ &= c(aT(x) + T(y)) + (aU(x) + U(y)) \\ &= a(cT(x) + U(x)) + (cT(y) + U(y)) \\ &= a(cT + U)(x) + (cT + U)(y). \end{aligned}$$

Thus, $cT + U$ is linear. □

Theorem 1.10. *Let V and W be finite dimensional vector spaces over $\mathbb{R}$ bases β and γ and $T, U : V \to W$ linear transformations. Then, the map $\Phi : \mathcal{L}(V, W) \to \mathcal{M}_{m\times n}(\mathbb{R})$ defined by $T \mapsto [T]_\beta^\gamma$ is linear, where $\dim(V) = n$ and $\dim(W) = m$. That is, we have*

(i) $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$, *and*

(ii) $[cT]_\beta^\gamma = c[T]_\beta^\gamma$ *for all* $c \in \mathbb{R}$.

Proof. Let $\beta = \{v_1, v_2, \cdots, v_n\}$ and $\gamma = \{w_1, w_2, \cdots, w_m\}$ be bases for V and W respectively. For each $j = 1, 2, \cdots, m$, we have

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i, \qquad U(v_j) = \sum_{i=1}^m b_{ij} w_i$$

for some $a_{ij}, b_{ij} \in \mathbb{R}$. Then,

$$(T + U)(v_j) = T(v_j) + U(v_j) = \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m b_{ij} w_i = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i,$$

which implies

$$[T + U]_\beta^\gamma = (a_{ij} + b_{ij}) = (a_{ij}) + (b_{ij}) = [T]_\beta^\gamma + [U]_\beta^\gamma.$$

Similarly,

$$(cT)(v_j) = cT(v_j) = c\sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m (ca_{ij}) w_i$$

and

$$[cT]_\beta^\gamma = (ca_{ij}) = c(a_{ij}) = c[T]_\beta^\gamma.$$

□

2 Composition of linear transformations and matrix multiplication, part 1

Notation 2.1. For two maps U and T, we use the notation UT to denote the composition of maps $U \circ T$, that is, $UT(x) = U \circ T(x) = U(T(x))$.

Theorem 2.2. *Let V, W, Z be vector spaces over $\mathbb{R}$, and $T : V \to W$ and $U : W \to Z$ be linear. Then, $UT : V \to Z$ is linear.*

Proof. It is easy to see that $UT(0) = U(T(0)) = U(0) = 0$. Let $x, y \in V$ and $c \in \mathbb{R}$, then

$$UT(cx + y) = U(T(cx + y)) = U(cT(x) + T(y)) = cU(T(x)) + U(T(y)) = cUT(x) + UT(y).$$

Thus, UT is linear. $\square$

Example 2.3. Let $T, U : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x, y) = (x, -y)$ and $U(x, y) = (y, x)$. Then, T is the reflection about x-axis and U is the reflection about the line $x = y$. Consider $UT : \mathbb{R}^2 \to \mathbb{R}^2$, then

$$UT(x, y) = U(T(x, y)) = U(x, -y) = (-y, x)$$

is the rotation by $\pi/2$ counterclockwise. Let β be the standard basis for $\mathbb{R}^2$, then

$$[T]_\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad [U]_\beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad [UT]_\beta = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Definition 2.4. Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $B \in \mathcal{M}_{n\times p}(\mathbb{R})$. We define the product of A and B by $AB \in \mathcal{M}_{m\times p}(\mathbb{R})$,

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

Example 2.5.

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 4 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & -1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 1 \\ 5 & -5 \end{pmatrix}$$

Definition 2.6. We define the Kronecker delta δ_{ij} by

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & \text{otherwise}. \end{cases}$$

The $n \times n$ identity matrix is $I_n = (\delta_{ij})$.

Theorem 2.7. *Let $A, C \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $B, D \in \mathcal{M}_{n\times p}(\mathbb{R})$, then*

(i) $(A + C)B = AB + CB$ *and* $A(B + D) = AB + AD$.

(ii) $c(AB) = (cA)B = A(cB)$ *for* $c \in \mathbb{R}$.

(iii) $I_m A = A = A I_n$.

(iv) *Let V be an n-dimensional vector space over $\mathbb{R}$ with a basis β and $I_V : V \to V$ be defined by $I_V(x) = x$ for all $x \in V$. Then, $[I_V]_\beta = I_n$.*

Proof. Exercise. $\square$

Theorem 2.8. *Let V, W, Z be finite dimensional vector spaces over $\mathbb{R}$ with bases α, β, γ, and $T : V \to W$ and $U : W \to Z$ be linear. Then, $[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta \in \mathcal{M}_{p\times n}(\mathbb{R})$.*

Proof. Let $\alpha = \{v_1, v_2, \cdots, v_n\}, \beta = \{w_1, w_2, \cdots, w_m\}, \gamma = \{z_1, z_2, \cdots, z_p\}$ and

$$T(v_j) = \sum_{k=1}^{m} b_{kj} w_k, \qquad U(w_k) = \sum_{i=1}^{p} a_{ik} z_i.$$

Then,

$$(UT)(v_j) = U(T(v_j)) = U(\sum_{k=1}^{m} b_{kj} w_k) = \sum_{k=1}^{m} b_{kj} U(w_k) = \sum_{k=1}^{m} b_{kj} \sum_{i=1}^{p} a_{ik} z_i = \sum_{i=1}^{p} \left(\sum_{k=1}^{m} a_{ik} b_{kj} \right) z_i$$

$\square$

Example 2.9. Let $T, U : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x, y) = (x, -y)$ and $U(x, y) = (y, x)$. Let β be the standard basis for $\mathbb{R}^2$, then

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = [U]_\beta [T]_\beta = [UT]_\beta = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Note that $[U]_\beta [T]_\beta \neq [T]_\beta [U]_\beta$. In fact, $TU(x, y) = T(y, x) = (y, -x)$ is the rotation by $\pi/2$ clockwise.

3 Composition of linear transformations and matrix multiplication, part 2

Notation 3.1. (i) For $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and $k \in \mathbb{N}$, $A^k = A \cdot A \cdots A$ (k times) and $A^0 = I_n$.

(ii) O is the zero matrix, whose entries are all zero.

(iii) For $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $j = 1, 2, \cdots, n$, $[A]_j$ the j-th column of A.

Theorem 3.2. *Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $B \in \mathcal{M}_{n\times p}(\mathbb{R})$.*

(i) $[A]_j = A \cdot e_j$ *where e_j is the standard basis for $\mathbb{R}^n$.*

(ii) $[AB]_j = A \cdot [B]_j$.

Proof. Exercise. □

Theorem 3.3. *Let V and W be finite dimensional vector spaces with bases β and γ. Let $T : V \to W$ be linear. Then for each $v \in V$, we have*

$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta.$$

Proof. Let $\beta = \{v_1, v_2, \cdots, v_n\}$ and $\gamma = \{w_1, w_2, \cdots, w_m\}$, then

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \qquad v = b_1 v_1 + \cdots + b_n v_n$$

for some $a_{ij}, b_j \in \mathbb{R}$. Then,

$$T(v) = T(b_1 v_1 + \cdots + b_n v_n) = b_1 T(v_1) + \cdots + b_n T(v_n) = \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} b_j \right) w_i$$

□

Definition 3.4. For $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, we define $L_A : \mathbb{R}^n \to \mathbb{R}^m$ by $L_A(x) = Ax$, where x is a column vector in $\mathbb{R}^n$. The map L_A is called a left-multiplication transformation.

Example 3.5. Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, then $L_A : \mathbb{R}^3 \to \mathbb{R}^2$ and

$$L_A(x, y, z) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ y + 2z \end{pmatrix}.$$

Remark 3.6. Let β, γ be the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$. Recall that the map $\mathcal{F} : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to \mathcal{M}_{m\times n}(\mathbb{R})$ defined by $\mathcal{F}(T) = [T]_\beta^\gamma$ is a linear map by Theorem 1.10.

Theorem 3.7. *Let $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$ and β, γ be the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$. Define $\mathcal{G} : \mathcal{M}_{m\times n}(\mathbb{R}) \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ by $\mathcal{G}(A) = L_A$.*

(i) *$\mathcal{G}$ is well-defined. ($L_A : \mathbb{R}^n \to \mathbb{R}^m$ is linear for all $A \in \mathcal{M}_{m\times n}(\mathbb{R})$.)*

(ii) *$\mathcal{G}$ is linear. ($L_{A+B} = L_A + L_B$ and $L_{cA} = cL_A$ for $c \in \mathbb{R}$ for all $A \in \mathcal{M}_{m\times n}(\mathbb{R})$.)*

(iii) *$\mathcal{G}$ is one-to one. (For $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$, $L_A = L_B$ if and only if $A = B$.)*

(iv) *$\mathcal{G}$ is onto. (If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then there exists a unique matrix $C \in \mathcal{M}_{m\times n}(\mathbb{R})$ such that $T = L_C$.)*

(v) *$\mathcal{F} \circ \mathcal{G} = Id$. ($[L_A]_\beta^\gamma = A$ for all $A \in \mathcal{M}_{m\times n}(\mathbb{R})$.)*

Proof. (i) Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. First, $L_A(0) = A \cdot 0 = 0$. For $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

$$L_A(cx + y) = A(cx + y) = cAx + Ay = cL_A(x) + L_A(y).$$

Thus, L_A is linear, that is, $\mathcal{G}(A) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

(ii) For $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

$$L_{A+B}(x) = (A + B)x = Ax + Bx = L_A(x) + L_B(y)$$

and

$$L_{cA}(x) = (cA)x = c(Ax) = cL_A(x).$$

This implies that $L_{A+B} = L_A + L_B$ and $L_{cA} = cL_A$ so that $\mathcal{G}$ is linear.

(iii) For $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$, suppose that $L_A = L_B$. For each $i = 1, 2, \cdots, n$,

$$L_A(e_i) = Ae_i = [A]_i = [B]_i = Be_i = L_B(e_i),$$

which implies $A = B$.

(iv) Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear (i.e. $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$), then let $C = [T]_\beta^\gamma$. We claim that $L_C = T$. For $x \in \mathbb{R}^n$,

$$L_C(x) = [T]_\beta^\gamma [x]_\beta = [T(x)]_\gamma = T(x).$$

(v) Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. It suffices to show that $[L_A]_\beta^\gamma e_i = Ae_i$ for each $i = 1, 2, \cdots, n$. Indeed,

$$[L_A]_\beta^\gamma e_i = [L_A(e_i)]_\gamma = [Ae_i]_\gamma = Ae_i.$$

□

Theorem 3.8. *If $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $B \in \mathcal{M}_{n\times p}(\mathbb{R})$, then $L_{AB} = L_A L_B$. We also have $L_{I_n} = I_{\mathbb{R}^n}$.*

Proof. For each $i = 1, 2, \cdots, n$,

$$L_{AB}(e_i) = AB(e_i) = [AB]_i = A[B]_i = A(Be_i) = L_A(L_B(e_i)) = L_A L_B(e_i).$$

□

Corollary 3.9. *Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $B \in \mathcal{M}_{n\times p}(\mathbb{R})$, and $C \in \mathcal{M}_{p\times q}(\mathbb{R})$, then $(AB)C = A(BC)$.*

Proof. It suffices to show that $L_{(AB)C} = L_{A(BC)}$. By the previous theorem, we have

$$L_{(AB)C} = L_{AB}L_C = (L_A \circ L_B) \circ L_C = L_A \circ (L_B \circ L_C) = L_A L_{BC} = L_{A(BC)}.$$

□

References

[FIS] Freidberg, Insel, and Spence, *Linear Algebra*, 4th edition, 2002.

[Bee] Beezer, *A First Course in Linear Algebra*, Version 3.5, 2015.

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu