Script 19

Differentiation in $\mathbb{R}^n$

8/4: **Definition 19.1.** A **linear transformation** $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is a function such that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$,

(a) $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$;

(b) $\varphi(\lambda \mathbf{x}) = \lambda \varphi(\mathbf{x})$.

That is, φ is a linear transformation if it respects the two operations in Definition 18.2.

Lemma 19.2. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\varphi(\mathbf{0}) = \mathbf{0}$.*

Proof. Suppose for the sake of contradiction that $\varphi(\mathbf{0}) \neq \mathbf{0}$. Then

$$\begin{aligned} \mathbf{0} &= \varphi(\mathbf{x}) - \varphi(\mathbf{x}) \\ &= \varphi(\mathbf{x} - \mathbf{x}) && \text{Definition 19.1} \\ &= \varphi(\mathbf{0}) \\ &\neq \mathbf{0} \end{aligned}$$

a contradiction. □

Exercise 19.3. We denote $\mathbf{x} \in \mathbb{R}^2$ by $\mathbf{x} = (x, y)$. Determine whether the following functions are linear transformations:

(a) $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $\varphi(x, y) = x + y$.

Answer. φ is a linear transformation. □

Proof. To prove that φ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and for any $\lambda \in \mathbb{R}$, $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ and $\varphi(\lambda \mathbf{x}) = \lambda \varphi(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^2$, and let λ be an arbitrary element of $\mathbb{R}$. Then

$$\begin{aligned} \varphi(\mathbf{x} + \mathbf{y}) &= (x_1 + y_1) + (x_2 + y_2) \\ &= (x_1 + x_2) + (y_1 + y_2) \\ &= \varphi(\mathbf{x}) + \varphi(\mathbf{y}) \end{aligned}$$

and

$$\begin{aligned} \varphi(\lambda \mathbf{x}) &= \lambda x_1 + \lambda x_2 \\ &= \lambda(x_1 + x_2) \\ &= \lambda \varphi(\mathbf{x}) \end{aligned}$$

as desired. □

(b) $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, $\varphi(x, y) = (x, y + 1)$.

Answer. φ is not a linear transformation. $\square$

Proof. By the definition of φ, $\varphi(\mathbf{0}) = (0, 1) \neq \mathbf{0}$. Thus, by the contrapositive of Lemma 19.2, φ is not a linear transformation, as desired. $\square$

(c) $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(x, y) = (3x - y, x + 2y, 0)$.

Answer. φ is a linear transformation. $\square$

Proof. To prove that φ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and for any $\lambda \in \mathbb{R}$, $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ and $\varphi(\lambda\mathbf{x}) = \lambda\varphi(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^2$, and let λ be an arbitrary element of $\mathbb{R}$. Then

$$\begin{aligned}
\varphi(\mathbf{x} + \mathbf{y}) &= (3[x_1 + y_1] - [x_2 + y_2], [x_1 + y_1] + 2[x_2 + y_2], 0) \\
&= ([3x_1 - x_2] + [3y_1 - y_2], [x_1 + 2x_2] + [y_1 + 2y_1], 0) \\
&= (3x_1 - x_2, x_1 + 2x_2, 0) + (3y_1 - y_2, y_1 + 2y_2, 0) \\
&= \varphi(\mathbf{x}) + \varphi(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
\varphi(\lambda\mathbf{x}) &= (3[\lambda x_1] - [\lambda x_2], [\lambda x_1] + 2[\lambda x_2], 0) \\
&= (\lambda[3x_1 - x_2], \lambda[x_1 + 2x_2], 0) \\
&= \lambda(3x_1 - x_2, x_1 + 2x_2, 0) \\
&= \lambda\varphi(\mathbf{x})
\end{aligned}$$

as desired. $\square$

(d) $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(x, y) = (x^2, x + y, x + y^3)$.

Answer. φ is not a linear transformation. $\square$

Proof. Consider $(1, 1) \in \mathbb{R}^2$ and let $2 \in \mathbb{R}$. Then

$$\begin{aligned}
\varphi(2(1, 1)) &= (4, 4, 10) \\
&\neq (2, 4, 4) \\
&= 2(1, 2, 2) \\
&= 2\varphi(1, 1)
\end{aligned}$$

as desired. $\square$

Exercise 19.4.

(a) Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a linear transformation. What does the graph of φ look like?

Answer. A line through the origin with finite slope. $\square$

(b) Let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be a linear transformation. What does the graph of φ look like?

Answer. A plane through the origin with finite slope in both directions. $\square$

Exercise 19.5.

(a) Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ and $\psi : \mathbb{R}^m \to \mathbb{R}^\ell$ be linear transformations. Prove that $\psi \circ \varphi$ is also a linear transformation.

Proof. To prove that $\psi \circ \varphi : \mathbb{R}_n \to \mathbb{R}^\ell$ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$, $(\psi \circ \varphi)(\mathbf{x}+\mathbf{y}) = (\psi \circ \varphi)(\mathbf{x}) + (\psi \circ \varphi)(\mathbf{y})$ and $(\psi \circ \varphi)(\lambda \mathbf{x}) = \lambda(\psi \circ \varphi)(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let λ be an arbitrary element of $\mathbb{R}$. Then since φ and ψ are linear transformations themselves, we have that

$$\begin{aligned}
(\psi \circ \varphi)(\mathbf{x}+\mathbf{y}) &= \psi(\varphi(\mathbf{x}+\mathbf{y})) \\
&= \psi(\varphi(\mathbf{x}) + \varphi(\mathbf{y})) && \text{Definition 19.1} \\
&= \psi(\varphi(\mathbf{x})) + \psi(\varphi(\mathbf{y})) && \text{Definition 19.1} \\
&= (\psi \circ \varphi)(\mathbf{x}) + (\psi \circ \varphi)(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
(\psi \circ \varphi)(\lambda \mathbf{x}) &= \psi(\varphi(\lambda \mathbf{x})) \\
&= \psi(\lambda \varphi(\mathbf{x})) && \text{Definition 19.1} \\
&= \lambda \psi(\varphi(\mathbf{x})) && \text{Definition 19.1} \\
&= \lambda(\psi \circ \varphi)(\mathbf{x})
\end{aligned}$$

as desired. ☐

(b) Let $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations and let $\lambda \in \mathbb{R}$. Prove that $\varphi + \psi$ and $\lambda\varphi$ are linear transformations.

Proof. To prove that $\varphi + \psi$ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$, $(\varphi + \psi)(\mathbf{x}+\mathbf{y}) = (\varphi + \psi)(\mathbf{x}) + (\varphi + \psi)(\mathbf{y})$ and $(\varphi + \psi)(\lambda \mathbf{x}) = \lambda(\varphi + \psi)(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let λ be an arbitrary element of $\mathbb{R}$. Then since φ and ψ are linear transformations themselves, we have that

$$\begin{aligned}
(\varphi + \psi)(\mathbf{x}+\mathbf{y}) &= \varphi(\mathbf{x}+\mathbf{y}) + \psi(\mathbf{x}+\mathbf{y}) \\
&= \varphi(\mathbf{x}) + \varphi(\mathbf{y}) + \psi(\mathbf{x}) + \psi(\mathbf{y}) && \text{Definition 19.1} \\
&= \varphi(\mathbf{x}) + \psi(\mathbf{x}) + \varphi(\mathbf{y}) + \psi(\mathbf{y}) \\
&= (\varphi + \psi)(\mathbf{x}) + (\varphi + \psi)(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
(\varphi + \psi)(\lambda \mathbf{x}) &= \varphi(\lambda \mathbf{x}) + \psi(\lambda \mathbf{x}) \\
&= \lambda\varphi(\mathbf{x}) + \lambda\psi(\mathbf{x}) && \text{Definition 19.1} \\
&= \lambda(\varphi(\mathbf{x}) + \psi(\mathbf{x})) \\
&= \lambda(\varphi + \psi)(\mathbf{x})
\end{aligned}$$

as desired.

To prove that $\lambda\varphi$ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\gamma \in \mathbb{R}$, $(\lambda\varphi)(\mathbf{x}+\mathbf{y}) = (\lambda\varphi)(\mathbf{x}) + (\lambda\varphi)(\mathbf{y})$ and $(\lambda\varphi)(\gamma \mathbf{x}) = \gamma(\lambda\varphi)(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let γ be an arbitrary element of $\mathbb{R}$. Then since φ is a linear transformation itself, we have that

$$\begin{aligned}
(\lambda\varphi)(\mathbf{x}+\mathbf{y}) &= \lambda\varphi(\mathbf{x}+\mathbf{y}) \\
&= \lambda(\varphi(\mathbf{x}) + \varphi(\mathbf{y})) && \text{Definition 19.1} \\
&= \lambda\varphi(\mathbf{x}) + \lambda\varphi(\mathbf{y}) \\
&= (\lambda\varphi)(\mathbf{x}) + (\lambda\varphi)(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
(\lambda\varphi)(\gamma\mathbf{x}) &= \lambda\varphi(\gamma\mathbf{x}) \\
&= \lambda\gamma\varphi(\mathbf{x}) && \text{Definition 19.1} \\
&= \gamma\lambda\varphi(\mathbf{x}) \\
&= \gamma(\lambda\varphi)(\mathbf{x})
\end{aligned}$$

as desired. □

(c) Let $\pi_I : \mathbb{R}^m \to \mathbb{R}^k$ be the projection function from Definition 18.34. Prove that π_I is a linear transformation.

Proof. To prove that π_I is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and for any $\lambda \in \mathbb{R}$, $\pi_I(\mathbf{x}+\mathbf{y}) = \pi_I(\mathbf{x}) + \pi_I(\mathbf{y})$ and $\pi_I(\lambda\mathbf{x}) = \lambda\pi_I(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$, and let λ be an arbitrary element of $\mathbb{R}$. Then we have that

$$\begin{aligned}
\pi_I(\mathbf{x}+\mathbf{y}) &= (x_{i_1} + y_{i_k}, \dots, x_{i_k} + y_{i_k}) \\
&= (x_{i_1}, \dots, x_{i_k}) + (y_{i_1}, \dots, y_{i_k}) \\
&= \pi_I(\mathbf{x}) + \pi_I(\mathbf{y})
\end{aligned}$$

and

$$\begin{aligned}
\pi_I(\lambda\mathbf{x}) &= (\lambda x_{i_1}, \dots, \lambda x_{i_k}) \\
&= \lambda(x_{i_1}, \dots, x_{i_k}) \\
&= \lambda\pi_I(\mathbf{x})
\end{aligned}$$

as desired. □

Definition 19.6. The ***j*^th^ standard basis vector** in $\mathbb{R}^n$ is the vector $\mathbf{e}_j$ defined by

$$(\mathbf{e}_j)_k = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

For example, the standard basis vectors for $\mathbb{R}^3$ are $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, and $\mathbf{e}_3 = (0,0,1)$. Notice that if $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, then $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$.

Definition 19.7. For any linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, we denote by $[\varphi]_{ij}$ the i^{th} component of the vector $\varphi(\mathbf{e}_j)$; i.e., $[\varphi]_{ij} = \varphi_i(\mathbf{e}_j)$.

Exercise 19.8.

(a) Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let $\mathbf{x} \in \mathbb{R}^n$. Find a formula for $\varphi(\mathbf{x})$ in terms of $[\varphi]_{ij}$, the components of $\mathbf{x}$, and the standard basis vectors in $\mathbb{R}^m$.

Proof. Since $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e_n}$ by Definition 19.6 and since φ is linear, we have that

$$\begin{aligned}
\varphi(\mathbf{x}) &= \varphi(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e_n}) \\
&= \varphi(x_1\mathbf{e}_1) + \cdots + \varphi(x_n\mathbf{e}_n) \\
&= x_1\varphi(\mathbf{e}_1) + \cdots + x_n\varphi(\mathbf{e}_n) \\
&= x_1(\varphi_1(\mathbf{e}_1)\mathbf{e_1} + \cdots + \varphi_m(\mathbf{e}_1)\mathbf{e_m}) + \cdots + x_n(\varphi_1(\mathbf{e}_n)\mathbf{e_1} + \cdots + \varphi_m(\mathbf{e}_n)\mathbf{e_m}) \\
&= x_1([\varphi]_{11}\mathbf{e_1} + \cdots + [\varphi]_{m1}\mathbf{e_m}) + \cdots + x_n([\varphi]_{1n}\mathbf{e_1} + \cdots + [\varphi]_{mn}\mathbf{e_m}) \\
&= x_1\sum_{i=1}^{m}[\varphi]_{i1}\mathbf{e}_i + \cdots + x_n\sum_{i=1}^{m}[\varphi]_{in}\mathbf{e}_i
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} [\varphi]_{ij} \mathbf{e}_i \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j [\varphi]_{ij} \mathbf{e}_i
\end{aligned}
$$

□

(b) For $1 \leq i \leq m$ and $1 \leq j \leq n$, let $A_{ij} \in \mathbb{R}$. Prove that there is a unique linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ such that $[\varphi]_{ij} = A_{ij}$ for all i, j.

Proof. Let φ be defined by

$$
\varphi(\mathbf{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i
$$

for all $\mathbf{x} \in \mathbb{R}^n$. Thus, by Definition 19.7, $[\varphi]_{ij} = A_{ij}$ for all i, j.

To prove that φ is a linear transformation, Definition 19.1 tells us that it will suffice to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ and $\varphi(\lambda \mathbf{x}) = \lambda \varphi(\mathbf{x})$. Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $\mathbb{R}^n$. Then

$$
\begin{aligned}
\varphi(\mathbf{x} + \mathbf{y}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} (x_j + y_j) A_{ij} \mathbf{e}_i \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} (x_j A_{ij} \mathbf{e}_i + y_j A_{ij} \mathbf{e}_i) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i + \sum_{i=1}^{m} \sum_{j=1}^{n} y_j A_{ij} \mathbf{e}_i \\
&= \varphi(\mathbf{x}) + \varphi(\mathbf{y})
\end{aligned}
$$

and

$$
\begin{aligned}
\varphi(\lambda \mathbf{x}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} (\lambda x_j) A_{ij} \mathbf{e}_i \\
&= \lambda \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i \\
&= \lambda \varphi(\mathbf{x})
\end{aligned}
$$

as desired.

Let $\psi : \mathbb{R}^n \to \mathbb{R}^m$ be any linear transformation satisfying $[\psi]_{ij} = A_{ij}$ for all i, j. To prove that $\varphi = \psi$, it will suffice to show that $\varphi(\mathbf{x}) = \psi(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Let $\mathbf{x}$ be an arbitrary element of $\mathbb{R}^n$. Then

$$
\begin{aligned}
\varphi(\mathbf{x}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j A_{ij} \mathbf{e}_i \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} x_j [\psi]_{ij} \mathbf{e}_i && \text{Exercise 19.8a} \\
&= \psi(\mathbf{x})
\end{aligned}
$$

as desired. □

Definition 19.9. We define an $m \times n$ matrix M to be an array of scalars

$$M = \{a_{ij}\} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

So a_{ij} denotes the scalar in row i, column j of the matrix. For every linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, there is a corresponding $m \times n$ matrix $\{[\varphi]_{ij}\}$. We denote $\{[\varphi]_{ij}\}$ by $[\varphi]$. Also, by Exercise 19.8, given a matrix of scalars, there is a unique linear transformation $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ that corresponds to it.

Exercise 19.10.

(a) Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $\varphi(x, y, z) = (3x + 2y - z, 4x - 5y + 2z)$. Write down the matrix $[\varphi]$.

Answer. The matrix is

$$M = \begin{bmatrix} 3 & 2 & -1 \\ 4 & -5 & 2 \end{bmatrix}$$

□

(b) What is the linear transformation that corresponds to the following matrix?

$$\begin{bmatrix} -2 & 3 \\ 4 & 6 \\ 1 & 0 \end{bmatrix}$$

Answer. The linear transformation is $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\varphi(x, y) = (-2x + 3y, 4x + 6y, x)$$

□

Theorem 19.11. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there is a constant $M_\varphi \in \mathbb{R}$ such that for all $\mathbf{x} \in \mathbb{R}^n$, we have $\|\varphi(\mathbf{x})\| \leq M_\varphi \|\mathbf{x}\|$.*

Lemma. *Let $a_1, \ldots, a_n \in \mathbb{R}$. Then*

$$\left(\sum_{i=1}^n a_i\right)^2 \leq n \sum_{i=1}^n a_i^2$$

Proof. We have that

$$\begin{aligned} \left(\sum_{i=1}^n a_i\right)^2 &= (1a_1 + \cdots + 1a_n)^2 \\ &\leq \left(\sqrt{1^2 + \cdots + 1^2} \cdot \sqrt{a_1^2 + \cdots + a_n^2}\right)^2 && \text{Lemma 18.9b} \\ &= \sqrt{n}^2 \sqrt{\sum_{i=1}^n a_i^2}^2 \\ &= n \sum_{i=1}^n a_i^2 \end{aligned}$$

as desired. □

Proof of Theorem 19.11. Let

$$M = \max_{i,j} |[\varphi]_{ij}| \qquad\qquad M_\varphi = M\sqrt{nm}$$

Then

$$\begin{aligned}
\|\varphi(\mathbf{x})\| &= \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} x_j[\varphi]_{ij}\right)^2} \\
&\leq \sqrt{\sum_{i=1}^{m} n \sum_{j=1}^{n} (x_j[\varphi]_{ij})^2} && \text{Lemma} \\
&= \sqrt{n}\cdot\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} x_j^2[\varphi]_{ij}^2} \\
&= \sqrt{n}\cdot\sqrt{\sum_{j=1}^{n}\left(x_j^2\sum_{i=1}^{m}[\varphi]_{ij}^2\right)} \\
&\leq \sqrt{n}\cdot\sqrt{\sum_{j=1}^{n}\left(x_j^2\sum_{i=1}^{m}M^2\right)} \\
&= \sqrt{n}\cdot\sqrt{\sum_{j=1}^{n} mM^2x_j^2} \\
&= M\sqrt{nm}\cdot\sqrt{\sum_{j=1}^{n} x_j^2} \\
&= M_\varphi\|\mathbf{x}\| && \text{Definition 18.6}
\end{aligned}$$

as desired. $\square$