Chapter 9: Functions of Several Variables

Author: Meng-Gen Tsai
Email: plover@gmail.com

Exercise 9.1. *If S is a nonempty subset of a vector space X, prove (as asserted in Section 9.1) that the span of S is a vector space.*

Denote the span of S by $\mathrm{span}(S)$.

Proof.

(1) Since $S \neq \varnothing$, there is $\mathbf{z} \in S$. So $1\mathbf{z} = \mathbf{z} \in \mathrm{span}(S) \neq \varnothing$. (In fact, $\mathrm{span}(S) \supseteq S$.)

(2) If $\mathbf{x}, \mathbf{y} \in \mathrm{span}(S)$, then there exist elements $\mathbf{x}_1, \ldots, \mathbf{x}_m$, $\mathbf{y}_1, \ldots, \mathbf{y}_n \in S$ and scalars $a_1, \ldots, a_m$, $b_1, \ldots, b_n$ such that

$$\begin{aligned}\mathbf{x} &= a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m,\\ \mathbf{y} &= b_1\mathbf{y}_1 + \cdots + b_n\mathbf{y}_n.\end{aligned}$$

Then
$$\mathbf{x} + \mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m + b_1\mathbf{y}_1 + \cdots + b_n\mathbf{y}_n$$
is a linear combination of the elements of S. For any scalar c,

$$c\mathbf{x} = (ca_1)\mathbf{x}_1 + \cdots + (ca_m)\mathbf{x}_m$$

is again linear combination of the elements of S.

(3) By (1)(2), $\mathrm{span}(S)$ is a vector space.

□

Note. Any subspace of X that contains S must also contain $\mathrm{span}(S)$.

Exercise 9.2. *Prove (as asserted in Section 9.6) that BA is linear if A and B are linear transformations. Prove also that A^{-1} is linear and invertible if A is invertible.*

Proof. Use the notation in Definitions 9.6.

(1) *Show that BA is linear if A and B are linear transformations.* Let X, Y, Z be vector spaces, $A \in L(X, Y)$ and $B \in L(Y, Z)$.

(a) Given any $\mathbf{x}_1, \mathbf{x}_2 \in X$.

$$\begin{aligned}
(BA)(\mathbf{x}_1 + \mathbf{x}_2) &= B(A(\mathbf{x}_1 + \mathbf{x}_2)) \\
&= B(A\mathbf{x}_1 + A\mathbf{x}_2) && (A \text{ is a linear transformation}) \\
&= B(A\mathbf{x}_1) + B(A\mathbf{x}_2) && (B \text{ is a linear transformation}) \\
&= (BA)\mathbf{x}_1 + (BA)\mathbf{x}_2.
\end{aligned}$$

(b) For any $\mathbf{x} \in X$ and scalar c,

$$\begin{aligned}
(BA)(c\mathbf{x}) &= B(A(c\mathbf{x})) \\
&= B(cA\mathbf{x}) && (A \text{ is a linear transformation}) \\
&= cB(A\mathbf{x}) && (B \text{ is a linear transformation}) \\
&= c(BA)\mathbf{x}.
\end{aligned}$$

By (a)(b), $BA \in L(X, Z)$.

(2) *Show that A^{-1} is linear if A is invertible.*

(a) Given any $\mathbf{y}_1, \mathbf{y}_2 \in X$. Since A is surjective, there exist $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that

$$\begin{aligned}
\mathbf{y}_1 &= A\mathbf{x}_1 \\
\mathbf{y}_2 &= A\mathbf{x}_2.
\end{aligned}$$

So

$$\begin{aligned}
A^{-1}\mathbf{y}_1 &= A^{-1}(A\mathbf{x}_1) = \mathbf{x}_1 \\
A^{-1}\mathbf{y}_2 &= A^{-1}(A\mathbf{x}_2) = \mathbf{x}_2
\end{aligned}$$

(by Definitions 9.4). Hence

$$\begin{aligned}
A^{-1}(\mathbf{y}_1 + \mathbf{y}_2) &= A^{-1}(A\mathbf{x}_1 + A\mathbf{x}_2) \\
&= A^{-1}(A(\mathbf{x}_1 + \mathbf{x}_2)) && (A \text{ is a linear transformation}) \\
&= \mathbf{x}_1 + \mathbf{x}_2 && (\text{Definitions 9.4}) \\
&= A^{-1}\mathbf{y}_1 + A^{-1}\mathbf{y}_2.
\end{aligned}$$

(b) For any $\mathbf{y} \in X$ and scalar c, there is a corresponding $\mathbf{x} \in X$ such that $\mathbf{y} = A\mathbf{x}$ since A is surjective. So $A^{-1}\mathbf{y} = \mathbf{x}$ by Definition 9.4. Hence

$$\begin{aligned}
A^{-1}(c\mathbf{y}) &= A^{-1}(cA\mathbf{x}) \\
&= A^{-1}(A(c\mathbf{x})) && (A \text{ is a linear transformation}) \\
&= c\mathbf{x} && (\text{Definitions 9.4}) \\
&= cA^{-1}\mathbf{y}.
\end{aligned}$$

By (a)(b), $A^{-1} \in L(X)$.

(3) *Show that A^{-1} is invertible if A is invertible.* It suffices to show that A^{-1} is injective and surjective.

(a) *Show that A^{-1} is injective.* Given any $\mathbf{y}_1, \mathbf{y}_2 \in X$. Since A is surjective, there exist $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that

$$\begin{aligned} \mathbf{y}_1 &= A\mathbf{x}_1 \\ \mathbf{y}_2 &= A\mathbf{x}_2. \end{aligned}$$

Suppose $A^{-1}\mathbf{y}_1 = A^{-1}\mathbf{y}_2$. So $A^{-1}(A\mathbf{x}_1) = A^{-1}(A\mathbf{x}_2)$, or $\mathbf{x}_1 = \mathbf{x}_2$, or $\mathbf{y}_1 = A\mathbf{x}_1 = A\mathbf{x}_2 = \mathbf{y}_2$.

(b) *Show that A^{-1} is surjective.* For any $\mathbf{x} \in X$, there exists $A\mathbf{x} \in X$ such that $A^{-1}(A\mathbf{x}) = \mathbf{x}$ by Definitions 9.4.

□

Exercise 9.3. *Assume $A \in L(X, Y)$ and $A\mathbf{x} = \mathbf{0}$ only when $\mathbf{x} = \mathbf{0}$. Prove that A is then* 1-1.

Proof. Suppose $A\mathbf{x} = A\mathbf{y}$. Since A is a linear transformation, $A(\mathbf{x} - \mathbf{y}) = A\mathbf{x} - A\mathbf{y} = \mathbf{0}$. By assumption, $\mathbf{x} - \mathbf{y} = \mathbf{0}$ or $\mathbf{x} = \mathbf{y}$. □

Exercise 9.4. *Prove (as asserted in Section 9.30) that null spaces and ranges of linear transformations are vector spaces.*

Proof. Use the notation in Definitions 9.30. Suppose X, Y are vector spaces, and $A \in L(X, Y)$, as in Definition 9.6.

(1) *Show that $\mathscr{N}(A)$ is a vector space in X.*

(a) Note that $\mathbf{0} \in X$. Since $A\mathbf{0} = \mathbf{0}$, $\mathbf{0} \in \mathscr{N}(A) \neq \varnothing$.

(b) Suppose $\mathbf{x}_1, \mathbf{x}_2 \in \mathscr{N}(A)$. Then

$$\begin{aligned} A(\mathbf{x}_1 + \mathbf{x}_2) &= A\mathbf{x}_1 + A\mathbf{x}_2 && (A \text{ is a linear transformation}) \\ &= \mathbf{0} + \mathbf{0} && (\mathbf{x}_1, \mathbf{x}_2 \in \mathscr{N}(A)) \\ &= \mathbf{0}. \end{aligned}$$

So $\mathbf{x}_1 + \mathbf{x}_2 \in \mathscr{N}(A)$.

(c) Suppose $\mathbf{x} \in \mathscr{N}(A)$ and c is a scalar. Then

$$\begin{aligned} A(c\mathbf{x}) &= cA\mathbf{x} && (A \text{ is a linear transformation}) \\ &= c\mathbf{0} && (\mathbf{x} \in \mathscr{N}(A)) \\ &= \mathbf{0}. \end{aligned}$$

So $c\mathbf{x} \in \mathscr{N}(A)$.

By (a)(b)(c), $\mathscr{N}(A)$ is a vector space.

(2) *Show that $\mathscr{R}(A)$ is a vector space in Y.*

(a) Note that $\mathbf{0} \in X$. So $A\mathbf{0} = \mathbf{0} \in \mathscr{R}(A) \neq \varnothing$.

(b) Suppose $\mathbf{y}_1, \mathbf{y}_2 \in \mathscr{R}(A)$. Then there exist $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that $A\mathbf{x}_1 = \mathbf{y}_1$ and $A\mathbf{x}_2 = \mathbf{y}_2$. Hence

$$\begin{aligned} \mathbf{y}_1 + \mathbf{y}_2 &= A\mathbf{x}_1 + A\mathbf{x}_2 \\ &= A(\mathbf{x}_1 + \mathbf{x}_2) \qquad (A \text{ is a linear transformation}). \end{aligned}$$

So $\mathbf{y}_1 + \mathbf{y}_2 \in \mathscr{R}(A)$.

(c) Suppose $\mathbf{y} \in \mathscr{R}(A)$ and c is a scalar. Then there exists $\mathbf{x} \in X$ such that $A\mathbf{x} = \mathbf{y}$. Hence

$$\begin{aligned} c\mathbf{y} &= cA\mathbf{x} \\ &= A(c\mathbf{x}) \qquad (A \text{ is a linear transformation}). \end{aligned}$$

So $c\mathbf{y} \in \mathscr{R}(A)$.

By (a)(b)(c), $\mathscr{R}(A)$ is a vector space.

□

Exercise 9.5. *Prove that to every $A \in L(\mathbb{R}^n, \mathbb{R}^1)$ corresponds a unique $\mathbf{y} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$. Prove also that $\|A\| = |\mathbf{y}|$. (Hint: Under certain conditions, equality holds in the Schwarz inequality.)*

Proof.

(1) Recall that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$ (Definitions 9.1). Given any $\mathbf{x} \in \mathbb{R}^n$, write $\mathbf{x} = (x_1, \ldots, x_n)$ as $\mathbf{x} = \sum x_j \mathbf{e}_j$.

(2) *Show that $\mathbf{y}$ exists.* Since A is a linear transformation,

$$\begin{aligned} A\mathbf{x} &= A\left(\sum x_j \mathbf{e}_j\right) \\ &= \sum x_j A\mathbf{e}_j \\ &= (x_1, \ldots, x_n) \cdot (A\mathbf{e}_1, \ldots, A\mathbf{e}_n) \\ &= \mathbf{x} \cdot \sum (A\mathbf{e}_j)\mathbf{e}_j. \end{aligned}$$

Define $\mathbf{y} = \sum (A\mathbf{e}_j)\mathbf{e}_j \in \mathbb{R}^n$ so that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$.

(3) *Show that $\mathbf{y}$ is unique.* Suppose there exists some $\mathbf{z} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{z}$. So

$$\begin{aligned} 0 &= A\mathbf{x} - A\mathbf{x} \\ &= \mathbf{x} \cdot \mathbf{y} - \mathbf{x} \cdot \mathbf{z} \\ &= \mathbf{x} \cdot (\mathbf{y} - \mathbf{z}) \end{aligned}$$

for any $\mathbf{x} \in \mathbb{R}^n$. In particular, take $\mathbf{x} = \mathbf{y} - \mathbf{z} \in \mathbb{R}^n$ to get

$$0 = (\mathbf{y} - \mathbf{z}) \cdot (\mathbf{y} - \mathbf{z}) = |\mathbf{y} - \mathbf{z}|^2$$

or $\mathbf{y} - \mathbf{z} = \mathbf{0}$ or $\mathbf{y} = \mathbf{z}$.

(4) *Show that* $\|A\| = |\mathbf{y}|$. By the Schwarz inequality (Theorem 1.37(d)),

$$|A\mathbf{x}| = |\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}| \leq |\mathbf{y}|$$

as $|\mathbf{x}| \leq 1$. Take the sup over all $|\mathbf{x}| \leq 1$ to get

$$\|A\| \leq |\mathbf{y}|.$$

If $\mathbf{y} = \mathbf{0}$, then $\|A\| = |\mathbf{y}| = 0$. If $\mathbf{y} \neq \mathbf{0}$, then the equality holds when $\mathbf{x} = \frac{\mathbf{y}}{|\mathbf{y}|} \in \mathbb{R}^n$. (Here $|\mathbf{x}| = 1$.)

$\square$