Exercise 9.2 Prove that BA is linear if A and B are linear transformations.

Proof. Assume the codomain of A is the domain of B. Then for any $\mathbf{x}_1, \mathbf{x}_2$ in the domain of X and scalar c, the linearity of A and B give us

$$BA(c\mathbf{x}_1 + \mathbf{x}_2) = B(cA\mathbf{x}_1 + A\mathbf{x}_2) = cBA\mathbf{x}_1 + BA\mathbf{x}_2.$$

Hence, BA is linear. □

Prove also that A^{-1} is linear and invertible.

Proof. Assume A is an invertible (bijective) linear transformation with inverse A^{-1}. Then for any $\mathbf{y}_1, \mathbf{y}_2$ is the domain of A^{-1} (same as the codomain of A) and scalar c, we find

$$\begin{aligned} A^{-1}(c\mathbf{y}_1 + \mathbf{y}_2) &= A^{-1}(cAA^{-1}\mathbf{y}_1 + AA^{-1}\mathbf{y}_2) \\ &= A^{-1}A(cA^{-1}\mathbf{y}_1 + A^{-1}\mathbf{y}_2) \\ &= cA^{-1}\mathbf{y}_1 + A^{-1}\mathbf{y}_2. \end{aligned}$$

Hence, A^{-1} is linear. For any $\mathbf{x}$ in the codomain of A^{-1} (same as the domain of A), we know that $A^{-1}A\mathbf{x} = \mathbf{x}$, so A^{-1} is surjective. If $A^{-1}\mathbf{y}_1 = A^{-1}\mathbf{y}_2$, then applying A, we find

$$y_1 = AA^{-1}\mathbf{y}_1 = AA^{-1}\mathbf{y}_2 = \mathbf{y}_2.$$

Hence, A^{-1} is also injective, therefore invertible.

□

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 $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that $A\mathbf{x}_1 = A\mathbf{x}_2$. Since A is linear,

$$\mathbf{0} = A\mathbf{x}_1 - A\mathbf{x}_2 = A(\mathbf{x}_1 - \mathbf{x}_2).$$

Therefore, $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, so in fact $\mathbf{x}_1 = \mathbf{x}_2$. Hence, A is injective.

□

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 $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$.

Proof. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis for $\mathbb{R}^n$. Given $A \in L(\mathbb{R}^n, \mathbb{R}^1)$, define the vector

$$\mathbf{y} = A\mathbf{e}_1 + \cdots + A\mathbf{e}_n = \begin{bmatrix} A\mathbf{e}_1 & \cdots & A\mathbf{e}_n \end{bmatrix}^T.$$

Each $\mathbf{x} \in \mathbb{R}^n$ has a unique representation as

$$\mathbf{x} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T,$$

for some $a_1, \dots, a_n \in \mathbb{R}$. Then we find

$$\begin{aligned} A\mathbf{x} &= A(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n) \\ &= a_1 A\mathbf{e}_1 + \cdots + a_n A\mathbf{e}_n \\ &= \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T \cdot \begin{bmatrix} A\mathbf{e}_1 & \cdots & A\mathbf{e}_n \end{bmatrix}^T \\ &= \mathbf{x} \cdot \mathbf{y}. \end{aligned}$$

This proves existence, we now show uniqueness. Suppose there are two vectors $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^n$ such that $\mathbf{x} \cdot \mathbf{y}_1 = A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}_2$ for all $\mathbf{x} \in \mathbb{R}$. In particular, take

$$\mathbf{y}_1 = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T \quad \text{and} \quad \mathbf{y}_2 = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}^T.$$

Then for $j = 1, \dots, n$, we find

$$a_j = \mathbf{e}_j \cdot \mathbf{y}_1 = \mathbf{e}_j \cdot \mathbf{y}_2 = b_j.$$

Hence, $y_1 = y_2$.

□

Prove also that $\|A\| = |\mathbf{y}|$.

Proof. For any $\mathbf{x} \in \mathbb{R}^n$ with $|x| = 1$, the Cauchy-Schwarz inequality gives us

$$|A\mathbf{x}| = |\mathbf{x} \cdot \mathbf{y}| \le |\mathbf{x}||\mathbf{y}| = |\mathbf{y}|.$$

Since this inequality holds for all unit vectors, then

$$\|A\| = \sup_{\substack{x \in \mathbb{R}^n \\ |x|=1}} |Ax| \le |\mathbf{y}|.$$

In particular, consider the the unit vector $\mathbf{y}/|\mathbf{y}|$. We find

$$\|A\| \ge |A\mathbf{y}| = |\mathbf{y} \cdot (\mathbf{y}/|\mathbf{y}|)| = \frac{|\mathbf{y} \cdot \mathbf{y}|}{|\mathbf{y}|} = \frac{|\mathbf{y}|^2}{|\mathbf{y}|} = |\mathbf{y}|$$

Thus, $\|A\| = |\mathbf{y}|$.

□

Exercise 9.6 If $f(0,0) = 0$ and

$$f(x,y) = \frac{xy}{x^2+y^2} \quad \text{if } (x,y) \neq (0,0),$$

prove that $(D_1 f)(x,y)$ and $(D_2 f)(x,y)$ exist at every point of $\mathbb{R}^2$, although f is not continuous at $(0,0)$.

Proof. Note that $f(x,y)$ is symmetric in the arguments x and y, so it is sufficient to show that $(D_1 f)(x,y)$ exists at every point of $\mathbb{R}^2$. For all $x \in \mathbb{R}$, we have $f(x,0) = 0$, so $(D_1 f)(x,0) = 0$. If we fix $y \neq 0$, then $f(x,y)$ is simply a differentiable function on $\mathbb{R}$ in the variable x, with derivative $(D_1 f)(x,y)$. Hence, all the partial derivatives of f exist.

As previously noted, $f(x,0) = 0$ for all $x \in \mathbb{R}$, which means that $f(x,0) \to 0$ as $x \to 0$. On the other hand, for $x \neq 0$, we have

$$f(x,x) = \frac{x^2}{2x^2} = \frac{1}{2.}$$

So $f(x,x) \to 1/2$ as $x \to 0$. Thus, the limit of $f(x,y)$ as $(x,y) \to (0,0)$ does not exist, implying that f is discontinuous at the origin.

$\square$

Exercise 9.8 Suppose that f is a differentiable real function in an open set $E \subseteq \mathbb{R}^n$, and that f has a local maximum at a point $\mathbf{x} \in E$. Prove that $f'(\mathbf{x}) = 0$.

Proof. Let $\mathbf{x} = (x_1, \ldots, x_n)$. For each $j = 1, \ldots, n$, define the function

$$f_j(x) = f(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n),$$

Define $E_j \subseteq$ to be the projection of E onto the jth coordinate. Then $f_j : E_j \to \mathbb{R}$ is a differentiable function with a local maximum at $x_j \in E_j$. Then as an instance of the 1-dimensional case, we know that

$$(D_j f)(\mathbf{x}) = f_j'(x_j) = 0.$$

Therefore,

$$[f'(\mathbf{x})] = \begin{bmatrix}(D_1 f)(\mathbf{x}) & \cdots & (D_n f)(\mathbf{x})\end{bmatrix}^T = \begin{bmatrix}0 & \cdots & 0\end{bmatrix}^T,$$

so in fact $f'(\mathbf{x}) = 0$.

$\square$

Exercise 9.9 If $\mathbf{f}$ is a differentiable mapping of a connected open set $E \subseteq \mathbb{R}^n$ into $\mathbb{R}^m$, and if $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ for every $\mathbf{x} \in E$, prove that $\mathbf{f}$ is constant in E.

Proof. Choose some point $\mathbf{x}_0 \in E$, and define the set

$$U = \{\mathbf{x} \in E : \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)\}.$$

We claim that U is open. For any $\mathbf{x} \in U \subseteq E$, there is some radius $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{x}) \subseteq E$. Since $\mathbf{f}'$ is zero on the convex set $B_\varepsilon(\mathbf{x})$ and $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)$, then $\mathbf{f}$ is constantly $\mathbf{f}(\mathbf{x}_0)$ on the open ball. Hence, $B_\varepsilon(\mathbf{x}) \subseteq U$, and we conclude that U is open.

Now define the set

$$V = E \setminus U = \{\mathbf{x} \in E : \mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{x}_0)\}.$$

Again, we claim that V is open, and give a similar argument. Any point $\mathbf{x} \in V$ must have some open ball centered at $\mathbf{x}$ and contained in E, on which $\mathbf{f}$ is constantly $\mathbf{f}(\mathbf{x})$. In particular, $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{x}_0)$, implying that the open ball is contained in V, which tells us that V is open.

Since we now have disjoint open sets U and V with U nonempty such that $E = U \cup V$ (and U is nonempty), then in fact V must be empty. Hence, $\mathbf{f}$ is constant on E, in particular equal to $\mathbf{f}(\mathbf{x}_0)$.

□

Exercise 9.10 If f is a real function defined in a convex open set $E \subseteq \mathbb{R}^n$, such that $(D_1 f)(\mathbf{x}) = 0$ for every $\mathbf{x} \in E$, prove that $f(\mathbf{x})$ depends only on $x_2, \ldots, x_n$.

Proof. Fix some $x_2, \ldots, x_n \in \mathbb{R}$ and define the set

$$D = \{x \in \mathbb{R} : (x, x_2, \ldots, x_n) \in E\}.$$

Then there is an injective mapping $D \to E$ given by $x \mapsto (x, x_2, \ldots, x_n)$. We see that D is convex (i.e., an interval in $\mathbb{R}$) since any pair of points $x, y \in D$ have a corresponding pair of points $(x, x_2, \ldots, x_n), (y, x_2, \ldots, x_n) \in E$. The segment between the pair in E is contained in E, and the points on that segment vary only in the first coordinate. Hence, there is a corresponding segment between x and y in D, implying D is convex.

Then function the function $D \to \mathbb{R}$ defined by $x \mapsto f(x, x_2, \ldots, x_n)$ is differentiable with derivative $(D_1 f)(x, x_2, \ldots, x_n) = 0$ for all $x \in D$. Then since the derivative of this function is zero on the real interval D, it must be constant on D. That is, f is constant on all the points in E of the form $(x, x_2, \ldots, x_n)$. Hence the value of $f(\mathbf{x})$ is independent of the first coordinate of $\mathbf{x}$.

□

Show that the convexity of E can be replaced by a weaker condition, but that some condition is required.

All that is required for the above proof is that E contain every interval between pairs of points which differ only in the first coordinate. This is precisely the condition which ensures D to be an interval in $\mathbb{R}$.