Multivariable Calculus:
Linear Algebra Basics
Supplementary Notes

Fall 2019

David Lyons
Mathematical Sciences
Lebanon Valley College

last updated: 18 October 2019

Multivariable Calculus: Linear Algebra Basics
Supplementary Notes

Fall 2019

David Lyons
Mathematical Sciences
Lebanon Valley College

Contents

Introduction

Vector spaces and linear transformations are the basic objects of the subject of *linear algebra*. The purpose of these notes is to introduce the language and techniques of linear algebra that are needed for introductory physics and multivariable calculus courses.

1 The vector space $\mathbf{R}^n$

1.1 Vectors and vector operations

An ordered list of n real numbers $(x_1, x_2, \ldots, x_n)$ is called an ***n-tuple***. The set of all n-tuples is called ***n-dimensional (real) space***.

$$\mathbf{R}^n = \{(x_1, x_2, x_3, \ldots, x_n) \colon x_i \in \mathbf{R}, 1 \leq i \leq n\}$$

The operation of ***addition*** of two n-tuples is defined by

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

and ***scalar multiplication*** of an n-tuple $(x_1, x_2, \ldots, x_n)$ by a real number (or ***scalar***) k is defined by

$$k(x_1, x_2, \ldots, x_n) = (kx_1, kx_2, \ldots, kx_n).$$

With the operations of addition and scalar multiplication, the set $\mathbf{R}^n$ is called a ***vector space***[1] and n-tuples are called ***vectors***.

The numbers $\{x_i\}$ are called the ***coordinates*** of the vector $(x_1, \ldots, x_n)$. The number n is called the ***dimension*** of the vector space $\mathbf{R}^n$. The space $\mathbf{R}^1 = \mathbf{R}$ is the real number line, $\mathbf{R}^2$ is the plane of high school geometry and algebra, and $\mathbf{R}^3$ is the mathematical abstraction for the familiar 3-space in which we live. The space $\mathbf{R}^0$ is defined to be the one point set $\mathbf{R}^0 = \{0\}$.

1.2 Notation for vectors

There are many notation styles used to denote vectors in multivariate calculus and physics courses. Symbols for vectors are sometimes bold, Roman font letters, and sometimes letters decorated with arrows[2].

$$\mathbf{x} = \vec{x} = (x_1, x_2, \ldots, x_n)$$

Sometimes, instead of parentheses, angle brackets are used to delimit n-tuples.

$$(x_1, x_2, \ldots, x_n) = \langle x_1, x_2, \ldots, x_n \rangle$$

The vector $(0, 0, \ldots, 0)$ with all entries equal to 0 is called the ***zero vector*** and is denoted $\mathbf{0}$ or $\vec{0}$.

[1] What all vector spaces have in common is a set of vectors, a set of scalars, and the operations of vector addition and scalar multiplication. We omit the complete technical definition of a vector space.

[2] In advanced mathematics, vector names usually receive no special decoration, and instead would be written simply "$x = (x_1, x_2, \ldots, x_n)$".

1.3 Standard basis vectors

The vector $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ with a 1 in the ith coordinate and zeroes in all other coordinates is called the ***i*th standard basis vector** in $\mathbf{R}^n$. In $\mathbf{R}^2$, the standard basis vectors $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ are also called $\mathbf{i}$ and $\mathbf{j}$, respectively. In $\mathbf{R}^3$, the standard basis vectors $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$ and $\mathbf{e}_3 = (0, 0, 1)$ are also called $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, respectively. Given a vector $\mathbf{x} = (x_1, \ldots, x_n)$, we have the following representation of $\mathbf{x}$ as a sum of scalar multiples of the standard basis vectors (note that the summation sign indicates ***vector*** addition).

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \tag{1.3.1}$$

1.4 Inner product

The ***inner product*** or ***dot product*** of two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ is defined to be the real number

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

In terms of inner product, the ith coordinate x_i of the vector $\mathbf{x} = (x_1, \ldots, x_n)$ is given by

$$x_i = \mathbf{e}_i \cdot \mathbf{x} \tag{1.4.1}$$

and (1.3.1) becomes

$$\mathbf{x} = \sum_{i=1}^{n} (\mathbf{e}_i \cdot \mathbf{x})\, \mathbf{e}_i. \tag{1.4.2}$$

Exercises for Section 1

1. Write each of the following vectors $\mathbf{x}$ in the form $(x_1, x_2, \ldots, x_n)$ and $\sum x_i \mathbf{e}_i$. For $n = 2, 3$, also write $\mathbf{x}$ using $\mathbf{i}, \mathbf{j}, \mathbf{k}$ notation.
   Example: Given $\mathbf{x} = 3(2, 4, -1)$, write $x = (6, 12, -3) = 6\mathbf{e}_1 + 12\mathbf{e}_2 - 3\mathbf{e}_3 = 6\mathbf{i} + 12\mathbf{j} - 3\mathbf{k}$.

   (a) $\mathbf{x} = (3, 2) - (5, -2)$

   (b) $\mathbf{x} = 2(-1, 2, 1) + 3(2, -2, 0)$

   (c) $\mathbf{x} = 2\mathbf{e}_1 - 3\mathbf{e}_2 + 4\mathbf{e}_4 - (\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3)$

2. Verify equation (1.3.1).

3. Show that for $\mathbf{x} = (x_1, \ldots, x_n)$, we have $x_i = \mathbf{e}_i \cdot \mathbf{x}$.

4. Show that
$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$$
   where δ_{ij}, called the ***Kronecker delta***, is given by
$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

5. The ***length*** or ***norm*** of vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, denoted $|\mathbf{x}|$, is defined by
$$|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2.$$
   A vector $\mathbf{x}$ is said be be ***normalized*** or have ***norm 1*** if $|\mathbf{x}| = 1$.

   (a) Show that, for any vector $\mathbf{x}$ and any $k > 0$, we have
$$|k\mathbf{x}| = k|\mathbf{x}|.$$

   (b) Show that, for any nonzero vector $\mathbf{x}$, the vector $\frac{\mathbf{x}}{|\mathbf{x}|}$ has norm 1.

2 Linear Transformations

2.1 Definition of linear transformation

A function $L\colon \mathbf{R}^n \to \mathbf{R}^m$ is called a ***linear transformation*** (or "linear map") if

(i) $L(k\mathbf{x}) = kL(\mathbf{x})$, and

(ii) $L(\mathbf{x}+\mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$

for all vectors $\mathbf{x}, \mathbf{y}$ in $\mathbf{R}^n$ and real numbers k. Properties (i) and (ii) are called ***linearity*** properties. We say that L ***preserves*** or ***respects*** vector operations of scaling and addition. Instead of $L(\mathbf{x})$, it is common practice to drop the parentheses and write $L\mathbf{x}$ when L is a linear transformation.

2.2 Formulas for linear transformations

Given a vector $\mathbf{x} = (x_1, \ldots, x_n)$ and a linear map $L\colon \mathbf{R}^n \to \mathbf{R}^m$, we have

(2.2.1)
$$\begin{aligned} L\mathbf{x} &= L(x_1, x_2, \ldots, x_n) \\ &= L\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) \\ &= \sum_{j=1}^{n} L(x_j \mathbf{e}_j) \\ &= \sum_{j=1}^{n} x_j L\mathbf{e}_j \end{aligned}$$

A consequence of this equation is that a linear map is determined by its values on the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. We can write an explicit formula for the coordinates $(y_1, y_2, \ldots, y_m)$ of the value $\mathbf{y} = L\mathbf{x}$. Let $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ denote the standard basis vectors for $\mathbf{R}^m$. Then we have

(2.2.2)
$$\begin{aligned} y_i &= \mathbf{f}_i \cdot L\mathbf{x} \\ &= \mathbf{f}_i \cdot \left(\sum_{j=1}^{n} x_j L\mathbf{e}_j\right) \\ &= \sum_{j=1}^{n} (\mathbf{f}_i \cdot L\mathbf{e}_j)\, x_j. \end{aligned}$$

This last expression shows that the values of a linear function are completely determined by the numbers

(2.2.3)
$$a_{ij} = \mathbf{f}_i \cdot L\mathbf{e}_j$$

where i is in the range $1 \le i \le m$ and j is in the range $1 \le j \le n$. Written out fully, the equations for the value $\mathbf{y} = L\mathbf{x}$ are the following.

(2.2.4)
$$\begin{array}{ccccccccccc}
y_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & a_{1j}x_j & + & \cdots & + & a_{1n}x_n \\
y_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & a_{2j}x_j & + & \cdots & + & a_{2n}x_n \\
 & \vdots & & & & & & & & & & \\
y_i & = & a_{i1}x_1 & + & a_{i2}x_2 & + & \cdots & a_{ij}x_j & + & \cdots & + & a_{in}x_n \\
 & \vdots & & & & & & & & & & \\
y_m & = & a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & a_{mj}x_j & + & \cdots & + & a_{mn}x_n
\end{array}$$

The expressions on the right side of each equation are inner products.

$$y_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \cdot \mathbf{x} \tag{2.2.5}$$

2.3 Three operations on linear transformations

Let $L, L' \colon \mathbf{R}^n \to \mathbf{R}^m$, let $M \colon \mathbf{R}^p \to \mathbf{R}^n$, and let k be a real number. Linear transformations $kL \colon \mathbf{R}^n \to \mathbf{R}^m$, $L + L' \colon \mathbf{R}^n \to \mathbf{R}^m$, and $LM \colon \mathbf{R}^p \to \mathbf{R}^m$ are defined as follows[3].

$$(kL)\mathbf{x} = k(L\mathbf{x}) \qquad \text{(scalar multiplication)} \tag{2.3.1}$$

$$(L + L')\mathbf{x} = L\mathbf{x} + L'\mathbf{x} \qquad \text{(addition)} \tag{2.3.2}$$

$$(LM)\mathbf{x} = L(M\mathbf{x}) \qquad \text{(composition)} \tag{2.3.3}$$

Note that LM is the same thing as $L \circ M$, the ordinary composition of functions. It is conventional to omit the composition symbol in the context of linear transformations.

[3] With the operations of scalar multiplication and addition, the space of linear transformations $\mathbf{R}^n \to \mathbf{R}^m$ is another example of a vector space.

Exercises for Section 2

1. Let $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ be a linear map such that $L\mathbf{e}_1 = (2,3)$ and $L\mathbf{e}_2 = (-1,-2)$. Find $L(1,2)$.

2. Let $L\colon \mathbf{R}^3 \to \mathbf{R}$ be a linear map. Find $L\mathbf{k}$ if $L\mathbf{i} = 2$, $L\mathbf{j} = -1$, and $L(1,2,3) = 0$.

3. Show that the two linearity properties in the definition of linear transformation given in §2.1 are equivalent to the single property

   $$L(a\mathbf{x} + b\mathbf{y}) = aL\mathbf{x} + bL\mathbf{y} \tag{2.3.4}$$

   for all vectors $\mathbf{x}, \mathbf{y}$ and scalars a, b.

4. Justify each equality in (2.2.1).

5. In formulas (2.2.2) and (2.2.3), what is the difference between $\mathbf{e}_i$ and $\mathbf{f}_i$? Aren't both of these vectors with 1 in the ith coordinate and 0 elsewhere?

6. The dot product has the following properties that look like the properties in the definition of linear map.

   $$\begin{aligned} \mathbf{u}\cdot(\alpha\mathbf{v}) &= \alpha\mathbf{u}\cdot\mathbf{v} \\ \mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) &= \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w} \end{aligned}$$

   for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbf{R}^n$ and scalars α. Show that these properties hold.

7. Justify the steps of the derivation (2.2.2).

8. Prove that the composition of two linear maps is a linear map.

3 Matrices

3.1 Matrix notation

A ***matrix*** is a rectangular array of numbers. Entries of a matrix are indexed by a pair of integers i, j that indicate the row and column, respectively, of the entry. Rows are counted from top to bottom and columns are counted from left to right. A matrix with m rows and n columns is called an m ***by*** n ***matrix***. We write $[a_{ij}]$ to denote[4] the matrix with entries a_{ij}.

$$[a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

A matrix with only 1 row is called a ***row matrix*** and a matrix with only 1 column is a ***column matrix***. Entries of row and column matrices are usually given with just one index. Here are examples of $1 \times m$ and $n \times 1$ matrices, respectively.

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \qquad \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

It is nearly universal practice to identify n-tuples in $\mathbf{R}^n$ with $n \times 1$ matrices. We think of a column matrix as just another way to write a vector. Here is example notation.

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

3.2 Matrices and linear transformations

Using (2.2.3), linear transformations $\mathbf{R}^n \to \mathbf{R}^m$ are in one-to-one correspondence with $m \times n$ matrices. We write $[L]$ for the matrix associated to the linear transformation L.

(3.2.1) $$L \longleftrightarrow [L] = [\mathbf{f}_i \cdot L\mathbf{e}_j]$$

Using the one-to-one correspondence $L \leftrightarrow [L]$, we can impose on matrices the three operations of scalar multiplication (2.3.1), addition (2.3.2), and composition (2.3.3). Given a scalar k, two $m \times n$ matrices A, A', and an $n \times p$ matrix B, let L, L' be the associated linear maps for A, A' (that is, $A = [L]$ and $A' = [L']$)

[4] Sometimes, instead of square brackets, parentheses are used to delimit matrices.

and let M be the associated linear map for B (so $B = [M]$). We define new matrices kA, $A + A'$, and AB in terms of the associated linear maps as follows[5].

$$kA = [kL] \quad \text{(scalar multiplication)} \tag{3.2.2}$$

$$A + A' = [L + L'] \quad \text{(addition)} \tag{3.2.3}$$

$$AB = [LM] \quad \text{(matrix multiplication)} \tag{3.2.4}$$

Here are formulas for these three basic matrix operations in terms of matrix entries of A, A' and B.

$$i, j \text{ entry of } kA = k(i, j \text{ entry of } A) \tag{3.2.5}$$

$$i, j \text{ entry of } A + A' = (i, j \text{ entry of } A) + (i, j \text{ entry of } A') \tag{3.2.6}$$

$$i, j \text{ entry of } AB = (i\text{th row of } A) \cdot (j\text{th column of } B) \tag{3.2.7}$$

In most texts, the above formulas are given as *definitions*. In these notes, these formulas are *consequences* of the definitions (3.2.2), (3.2.3), and (3.2.4). We make this choice to emphasize that matrix algebra operations are natural because they come from the corresponding natural operations on linear maps. There is a fourth basic operation, called ***transposition***, whose corresponding operation on linear maps is less easy to describe. We define the ***transpose*** of matrix A, denoted A^T, by

$$i, j \text{ entry of } A^T = j, i \text{ entry of } A. \tag{3.2.8}$$

3.3 Matrix versions of vector operations

Let $\mathbf{x}, \mathbf{y}$ be vectors in $\mathbf{R}^n$ and let $L\colon \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. matrix. We have the following.

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} \tag{3.3.1}$$

$$L\mathbf{x} = [L]\mathbf{x} \tag{3.3.2}$$

Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ be the standard basis vectors for $\mathbf{R}^n$, and let $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m$ be the standard basis vectors for $\mathbf{R}^m$. Given an $m \times n$ matrix A, we have the following.

$$A\mathbf{e}_j = j\text{th column of } A \tag{3.3.3}$$

$$\mathbf{f}_i^T A = i\text{th row of } A \tag{3.3.4}$$

$$\mathbf{f}_i^T A \mathbf{e}_j = i, j \text{ entry of } A \tag{3.3.5}$$

[5] With the operations of scalar multiplication and addition, the space of $m \times n$ matrices is yet another example of a vector space.

Exercises for Section 3

1. Perform the matrix multiplications below.

   (a)
   $$\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}$$

   (b)
   $$\begin{bmatrix} 1 & -2 & 0 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 2 & 0 \end{bmatrix}$$

2. Let $L\colon \mathbf{R}^3 \to \mathbf{R}^2$ be a linear map with $L\mathbf{e}_1 = (1,2)$, $L\mathbf{e}_2 = (-1,1)$, and $L\mathbf{e}_3 = (0,1)$.

   (a) Write the matrix for L.

   (b) Evaluate $L(2,1,3)$.

   (c) Evaluate $L(0,1,1)$.

3. Let $L\colon \mathbf{R}^2 \to \mathbf{R}^3$ be a linear map with the following matrix.
   $$\begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix}$$

   (a) Evaluate $L(1,2)$.

   (b) Evaluate $L(-2,1)$.

4. Equation (3.2.7) is a fundamental formula for matrix calculations. This exercise outlines a proof. Let L, M be linear transformations so that the composition LM is defined (the domain of L is the codomain of M). Justify each equality in the derivation below.

   (a) Use (2.2.4) to get
   $$M(j\text{th standard basis vector}) = j\text{th column of } [M]$$

   (b) Equation (2.2.5) can be interpreted as saying
   $$\text{the } i\text{th coordinate of } L\mathbf{x} = (i\text{th row of } [L]) \cdot \mathbf{x}$$

   (c) String those results together to interpret as
   $$\begin{aligned} ij \text{ entry of } [LM] &= (i\text{th standard basis vector}) \cdot LM(j\text{th standard basis vector}) \\ &= i\text{th coordinate of } LM(j\text{th standard basis vector}) \\ &= i\text{th coordinate of } L(M(j\text{th standard basis vector})) \\ &= i\text{th coordinate of } L(j\text{th column of } [M]) \\ &= (i\text{th row of } [L]) \cdot (j\text{th column of } [M]) \end{aligned}$$

5. (a) Prove (3.3.2).

   (b) Prove (3.3.3).

   (c) Prove (3.3.4).

(d) Use parts (a) and (b) to deduce (3.3.5)

6. Let $L\colon \mathbf{R}^2 \to \mathbf{R}$ and $K\colon \mathbf{R} \to \mathbf{R}^3$ be linear maps such that $L\mathbf{i} = 1$, $L\mathbf{j} = -2$ and $K(1) = (2, 1, 3)$.

   (a) Write the matrices for L and K.

   (b) Find the matrix for KL.

   (c) Find $KL(2, -1)$.

7. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ denote the standard basis vectors in $\mathbf{R}^n$.

   (a) Show that $\mathbf{e}_k\mathbf{e}_k^T$ is an $n \times n$ matrix with a 1 in the k, k entry and zeroes elsewhere.

   (b) Show that $\displaystyle\sum_{k=1}^{n} \mathbf{e}_k\mathbf{e}_k^T$ is the $n \times n$ identity matrix.

   (c) Given an $m \times n$ matrix A and an $n \times p$ matrix B, show that

   $$AB = \sum_{k=1}^{n} (\text{column } k \text{ of } A)(\text{row } k \text{ of } B).$$

   Hint: Use (3.3.3) and (3.3.4).

4 More linear and matrix algebra

4.1 More matrix algebra

The ***identity map*** $I\colon \mathbf{R}^n \to \mathbf{R}^n$, given by $I\mathbf{x} = \mathbf{x}$, is a linear transformation. Its matrix is called the $n \times n$ ***identity matrix***, and is also denoted I. Other common notations for the identity map or the identity matrix are Id and $\mathbb{1}$.

A linear transformation $L\colon \mathbf{R}^n \to \mathbf{R}^n$ is called ***invertible*** if there is another linear map $M\colon \mathbf{R}^n \to \mathbf{R}^n$ such that $L \circ M = M \circ L = I$. If M exists, M is called the ***inverse*** of L, and we write $M = L^{-1}$.

An $n \times n$ matrix is called invertible if its corresponding linear transformation is invertible. We write A^{-1} for the matrix $[L^{-1}]$ for $A = [L]$.

4.2 Two by two matrices

A 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if its ***determinant*** $\det(A) = ad - bc$ is nonzero. In this case, we have

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

4.3 Dependence and Independence

A ***linear combination*** of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in $\mathbf{R}^n$ is a vector of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r$$

where $c_1, \ldots, c_r$ are scalars. The set of all possible linear combinations of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ is called their ***span***, denoted $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r)$.

$$\text{(4.3.1)} \qquad \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r \colon c_i \in \mathbf{R}, 1 \le i \le r\}.$$

A set of vectors in $\mathbf{R}^n$ is said to be ***dependent*** if one or more of them can be written as a linear combination of the others. Otherwise, the set of vectors is said to be ***independent***.

A set $\mathcal{B}$ of vectors in $\mathbf{R}^n$ is called a ***basis*** (plural *bases*) if the vectors in $\mathcal{B}$ are independent and span all of $\mathbf{R}^n$. Here is a key fact about bases.

(4.3.2) **Basis fact.** Let $\mathcal{B} = \{\mathbf{b}_1, \cdots \mathbf{b}_n\}$ be a basis for $\mathbf{R}^n$, and let $\mathbf{x}$ be any vector in $\mathbf{R}^n$. Then there exist a unique set of constants $c_1, c_2, \ldots, c_n$ such that $\mathbf{x} = \sum_{i=1}^{n} c_i\mathbf{b}_i$.

Exercises for Section 4

1. Find the matrix for the identity map $I\colon \mathbf{R}^n \to \mathbf{R}^n$.

2. Verify the 2 by 2 matrix facts in §4.2.

3. Here is an alternative definition for independence of vectors.

   > Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are *dependent* if there exists real numbers $c_1, c_2, \ldots, c_r$, not all of which are zero, such that
   >
   > $$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r = \mathbf{0}.$$
   >
   > The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are *independent* if they are not dependent.

   Show that this definition is equivalent to the definition given in §4.3.

4. Verify that the standard basis on $\mathbf{R}^n$ is indeed a basis according to the definition in §4.3.

5. Verify (4.3.2).

5 Linear algebra in Calculus 3

5.1 Differentiability

A function $f\colon \mathbf{R}^n \to \mathbf{R}^m$ is ***differentiable*** at the point $\mathbf{x}_0$ if there exists a linear transformation $L\colon \mathbf{R}^n \to \mathbf{R}^m$ such that

$$\lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) - L\mathbf{h}}{|\mathbf{h}|} = 0. \tag{5.1.1}$$

If L exists, it is called the ***derivative of*** f ***at*** $\mathbf{x}_0$, denoted $Df(\mathbf{x_0})$.

To understand the definition of the derivative, start with the case $n = m = 1$. The derivative of f at x_0 is a *number* $f'(x_0)$ such that

$$f(x_0 + h) - f(x_0) \approx f'(x_0)h$$

for h near 0. The meaning of "approximately equals. . . for h near 0" is made precise by using a limit. To generalize to higher dimensions, interpret $f'(x_0)h$ as the value of a linear transformation that sends h to $f'(x_0)h$. The derivative $Df(\mathbf{x}_0)$ of f at $\mathbf{x_0}$ is a linear transformation such that

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) \approx Df(\mathbf{x}_0)\mathbf{h}$$

for $\mathbf{h}$ near $\mathbf{0}$. Putting $\mathbf{h} = t\mathbf{e}_j$, this reads

$$f(\mathbf{x}_0 + t\mathbf{e}_j) - f(\mathbf{x}_0) \approx Df(\mathbf{x}_0)t\mathbf{e}_j$$

for t near 0. Dividing both sides by t and taking a limit, we get an expression for $Df(\mathbf{x}_0)\mathbf{e}_j$.

$$\begin{aligned} Df(\mathbf{x}_0)\mathbf{e}_j &= \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_j) - f(\mathbf{x}_0)}{t} \\ &= \left(\frac{\partial y_1}{\partial x_j}, \frac{\partial y_2}{\partial x_j}, \dots, \frac{\partial y_m}{\partial x_j}\right) \end{aligned} \tag{5.1.2}$$

where $\mathbf{y} = (y_1, y_2, \dots, y_m) = f(x_1, x_2, \dots, x_n) = f(\mathbf{x})$. From this it follows that $Df(\mathbf{x}_0)$, if it exists, is represented by the matrix $\left[\dfrac{\partial y_i}{\partial x_j}\right]$.

$$[Df(\mathbf{x}_0)] = \left[\frac{\partial y_i}{\partial x_j}\right] \tag{5.1.3}$$

5.2 The Chain Rule

Let

$$\mathbf{R}^p \xrightarrow{g} \mathbf{R}^n \xrightarrow{f} \mathbf{R}^m$$

and suppose g is differentiable at $\mathbf{t}_0$ and f is differentiable at $\mathbf{x}_0 = g(\mathbf{t}_0)$. The ***chain rule*** says that $f \circ g$ is differentiable at $\mathbf{t}_0$, and that the derivative of the composition is the composition of the derivatives.

$$D(f\circ g)(\mathbf{t}_0) = Df(\mathbf{x}_0)Dg(\mathbf{t}_0) \tag{5.2.1}$$

This explains the "tree diagram rule" given in most multivariate calculus texts. The partial derivative $\frac{\partial y_i}{\partial t_j}$ is just the i, j entry of the product of the derivative matrices for f and g.

(5.2.2)
$$\frac{\partial y_i}{\partial t_j} = \sum_{k=1}^{n} \frac{\partial y_i}{\partial x_k} \frac{\partial x_k}{\partial t_j}$$

Exercises for Section 5

1. Verify the definition of differentiable function (5.1.1) given in is equivalent to the usual definition for $n = m = 1$ from Calculus 1.

2. Explain equation (5.1.2). *Why* does the limit on the left equal the vector on the right?

$$\lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_j) - f(\mathbf{x}_0)}{t} = \left(\frac{\partial y_1}{\partial x_j}, \frac{\partial y_2}{\partial x_j}, \ldots, \frac{\partial y_m}{\partial x_j} \right)$$

3. Explain equation (5.1.3). *How* does this equation follow from the previous?

4. Explain equation (5.2.2). How is it the same as the chain rule (5.2.1)?

Solutions to Exercises

Note: Most of the "solutions" posted here are not solutions at all, but are merely final answer keys, although some are complete. These are posted so that you can check your work; reading the answer keys is not a substitute for working the problems yourself. For homework, quizzes and exams, you need to show the steps of whatever procedure you are using—not just the final result. Sometimes you will be asked to explain your thinking in complete sentences.

Exercises for Section 1 Solutions

1. (a) $\mathbf{x} = (-2, 4) = -2\mathbf{e}_1 + 4\mathbf{e}_2 = -2\mathbf{i} + 4\mathbf{j}$
   (b) $\mathbf{x} = (4, -2, 2) = 4\mathbf{e}_1 - 2\mathbf{e}_2 + 2\mathbf{e}_3 = 4\mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$
   (c) $\mathbf{x} = (1, -2, -1, 4) = \mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3 + 4\mathbf{e}_4$
2. We have

$$\begin{aligned}
&\sum_{i=1}^{n} x_i\mathbf{e}_i \\
&= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \\
&= x_1(1, 0, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \cdots x_n(0, 0, \ldots, 0, 1) \quad \text{(defn. of standard basis vects.)} \\
&= (x_1, 0, 0, \ldots, 0) + (0, x_2, 0, \ldots, 0) + \cdots (0, 0, \ldots, 0, x_n) \quad \text{(scalar mult.)} \\
&= (x_1, x_2, \ldots, x_n) \quad \text{(vector addition)} \\
&= \mathbf{x}.
\end{aligned}$$

3. We have

$$\begin{aligned}
\mathbf{e}_i \cdot \mathbf{x} &= (0, 0, \ldots, 0, 1, 0, \ldots, 0) \cdot (x_1, x_2, \ldots, x_n) \quad \text{(defn. of standard basis vect.)} \\
&= 0x_1 + 0x_2 + \cdots 0x_{i-1} + 1x_i + 0x_{i+1} + \cdots + 0x_n \quad \text{(defn. of dot product)} \\
&= x_i.
\end{aligned}$$

4. If $i \neq j$, each summand in the dot product has a factor of 0. If $i = j$, all the terms in the dot projuct are 0 except for the i-th, which is $1 \cdot 1 = 1$.
5. (a) $|k\mathbf{x}| = \sqrt{k^2x_1^2 + k^2x_2^2 + \cdots + k^2x_n^2} = k\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = k|\mathbf{x}|$
   (b) Using the previous problem for the first equality that follows, we have $\left|\dfrac{\mathbf{x}}{|\mathbf{x}|}\right| = \dfrac{1}{|\mathbf{x}|}|\mathbf{x}| = 1$. QED

Exercises for Section 2 Solutions

1. $l(1,2) = (0,-1)$

2. $L\mathbf{k} = 0$

3. Let $L\colon \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation and let $\mathbf{x}, \mathbf{y}$ be vectors in $\mathbf{R}^n$. Suppose that the two linearity properties given in the notes hold, and let a, b be scalars. We have

$$\begin{aligned} L(a\mathbf{x} + b\mathbf{y}) &= L(a\mathbf{x}) + L(b\mathbf{y}) \quad \text{(by linearity property (ii))} \\ &= aL\mathbf{x} + bL\mathbf{y} \quad \text{(by linearity property (i))}. \end{aligned}$$

Conversely, suppose that (2.3.4) holds. Choosing $b = 0$, we have (i)

$$L(a\mathbf{x}) = L(a\mathbf{x} + b\mathbf{y}) = aL\mathbf{x} + bL\mathbf{y} = aL\mathbf{x}.$$

Choosing $a = b = 1$, we have (ii)

$$L(\mathbf{x} + \mathbf{y}) = L(a\mathbf{x} + b\mathbf{y}) = aL\mathbf{x} + bL\mathbf{y} = L\mathbf{x} + L\mathbf{y}.$$

4. 1st equality: substitution
2nd equality: equation (1.3.1)
3rd equality: linearity property (ii)
4th equality: linearity property (i)

5. The vectors $\mathbf{e}_i$ and $\mathbf{f}_i$ are members of different vector spaces unless $n = m$, in which case $\mathbf{e}_i = \mathbf{f}_i$.

6. Let $\mathbf{u} = (u_1, \dots, u_n)$, $\mathbf{v} = (v_1, \dots, v_n)$, and $\mathbf{w} = (w_1, \dots, w_n)$. We have

$$\begin{aligned} \mathbf{u} \cdot (\alpha\mathbf{v}) &= (u_1, \dots, u_n) \cdot (\alpha v_1, \dots, \alpha v_n) \quad \text{(definition of scalar mult.)} \\ &= \sum_{i=1}^{n} u_i(\alpha v_i) \quad \text{(definition of dot prod.)} \\ &= \sum_{i=1}^{n} (u_i \alpha) v_i \quad \text{(mult. is associative)} \\ &= \sum_{i=1}^{n} (\alpha u_i) v_i \quad \text{(mult. is commutative)} \\ &= (\alpha u_1, \dots, \alpha u_n) \cdot (v_1, \dots, v_n) \quad \text{(definition of dot prod.)} \\ &= (\alpha\mathbf{u}) \cdot \mathbf{v} \quad \text{(definition of scalar mult.)}. \end{aligned}$$

The proof of the other property involves similar manipulations.

7. 1st equality: by (1.4.1)
2nd equality: by (2.2.1)
3rd equality: by linearity of the dot product (2nd property in the previous problem)

8. Let $L\colon \mathbf{R}^n \to \mathbf{R}^m$ and let $M\colon R^p \to \mathbf{R}^n$ be linear transformations. First we check linearity property (i) for the composition LM. Let $\mathbf{t}, \mathbf{s}$ be ele-

ments of $\mathbf{R}^p$ and let k be a real number. We have

$$\begin{aligned} LM(k\mathbf{t}) &= L(Mk\mathbf{t}) \quad \text{(definition)} \\ &= L(kM\mathbf{t}) \quad \text{(linearity (i) for } M) \\ &= kL(M\mathbf{t}) \quad \text{(linearity (i) for } L) \\ &= kLM\mathbf{t} \quad \text{(definition).} \end{aligned}$$

For property (ii), we have

$$\begin{aligned} LM(\mathbf{t}+\mathbf{s}) &= L(M(\mathbf{t}+\mathbf{s})) \quad \text{(definition)} \\ &= L(M\mathbf{t}+M\mathbf{s}) \quad \text{(linearity (ii) for } M) \\ &= L(M\mathbf{t})+L(M\mathbf{s}) \quad \text{(linearity (ii) for } L) \\ &= LM\mathbf{t}+LM\mathbf{s} \quad \text{(definition).} \end{aligned}$$

Exercises for Section 3 Solutions

1. (a) $\begin{bmatrix} -2 & 4 \\ 3 & -5 \end{bmatrix}$

   (b) $\begin{bmatrix} -2 & 4 \\ 5 & -5 \end{bmatrix}$

2. (a) $\begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & 1 \end{bmatrix}$

   (b) $(1, 8)$

   (c) $(-1, 2)$

3. (a) $(5, 1, 2)$

   (b) $(0, -7, 1)$

4. (a) From (2.2.4), we see that $L\mathbf{e}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})$. By definition, this is the jth column of $[L]$.

   (b) This is what (2.2.5) says.

   (c) 1st equality: definition (3.2.1) of the i, j entry of the matrix for a linear transformation
   2nd equality: by (1.4.1)
   3rd equality: definition of composition of functions
   4th equality: by part (a)
   5th equality: by part (b), with $\mathbf{x} = (j\text{th column of } [M])$

5. (a)

$$\begin{aligned} \text{the } i\text{th coordinate of } L\mathbf{x} &= (i\text{th row of } [L]) \cdot \mathbf{x} \quad (\text{by } \ (2.2.5)) \\ &= [L]\mathbf{x} \quad (\text{by } \ (3.2.7)) \end{aligned}$$

   (b) This follows immediately from (3.2.7).

   (c) This follows immediately from (3.2.7).

   (d) This follows immediately from parts (b) and (c).

6. (a) $L = \begin{bmatrix} 1 & -2 \end{bmatrix}$
   $K = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$

   (b) $\begin{bmatrix} 2 & -4 \\ 1 & -2 \\ 3 & -6 \end{bmatrix}$

   (c) $(8, 4, 12)$

7. (a) Just do the matrix multiplication.

   (b) This follows immediately from part (a).

(c)

$$
\begin{aligned}
AB &= AIB \\
&= A\left(\sum_{k=1}^{n} \mathbf{e}_k \mathbf{e}_k^T\right) B \qquad \text{(by part (b))} \\
&= \sum_{k=1}^{n} (A\mathbf{e}_k)(\mathbf{e}_k^T B) \qquad \text{(by linearity)} \\
&= \sum_{k=1}^{n} (\text{column } k \text{ of } A)(\text{row } k \text{ of } B) \qquad \text{(by (3.3.3) and (3.3.4))}
\end{aligned}
$$

Exercises for Section 4 Solutions

1. By definition (3.2.1), the i, j entry of the matrix for the identity map is

$$\begin{aligned} \mathbf{e}_i \cdot I\mathbf{e}_j &= \mathbf{e}_i \cdot \mathbf{e}_j \\ &= \delta_{ij} \quad \text{(by Exercise 4 in Section 1).} \end{aligned}$$

   Thus the identity matrix is the $n \times n$ matrix with 1's on the diagonal entries (that is, entries indexed i, i) and 0's in off-diagonal entries (that is, entries indexed i, j with $i \neq j$).

   Another way to find the matrix for the identity map is to see that its rows are the vectors

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$$

   given by

$$\begin{aligned} \mathbf{x} = (x_1, x_2, \ldots, x_n) &= I(x_1, x_2, \ldots, x_n) \\ &= (\mathbf{a}_1 \cdot \mathbf{x}, \mathbf{a}_1 \cdot \mathbf{x}, \ldots, \mathbf{a}_1 \cdot \mathbf{x}). \end{aligned}$$

   From this we can see that $\mathbf{a}_i = \mathbf{e}_i$.

   A further alternative is to use the fact that the j-th column of a matrix for a linear transformation L is $L\mathbf{e}_j$.

2. There are two statements to prove. The first says that if $\det(A) \neq 0$, then A^{-1} exists and is given by the formula in the notes. To establish that this claim is true, simply check that $AA^{-1} = A^{-1}A = I$.

   The second statement says that if $\det(A) = 0$, then A is not invertible. Here is an outline of how to argue that this claim is true.

   - One way to show that a function is not invertible is to show that it is not one-to-one, that is, that there exist two different inputs that are sent by the function to the same output.
   - A linear map always sends the zero vector in the input space to the zero vector in the output space. So to establish that a linear map is not one-to-one (and therefore not invertible), it suffices to find some nonzero input vector that is sent by the linear map to the zero vector.
   - Assuming that $0 = \det(A) = ad - bc$, one can show that A sends the vector $(-b, a)$ to the zero vector. If $(-b, a)$ is not the zero vector, we have shown that A is not invertible.
   - If $(-b, a)$ is the zero vector, then one can show that A sends $(-d, c)$ to zero. If $(-d, c)$ is not zero, then A is non invertible.
   - If $(-b, a)$ and $(-d, c)$ are both the zero vector, then A is the zero matrix, so A sends *every* input vector to zero, so A is not invertible. This takes care of all possible cases, so we conclude that A is not invertible.

3. We will prove the equivalence of the two definitions of *dependence*, from which the equivalence of the two definitions of independence follows.

   Suppose that vectors $\{\mathbf{v}_i\}$ are dependent in the sense defined in this exercise. That is, suppose that there is a linear combination

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r = 0$$

for which at least one of the coefficients, say c_i, is not zero. Then we can solve for $\mathbf{v}_i$ as a linear combination of the other $\mathbf{v}_j$'s.

$$\mathbf{v}_i = -\frac{c_1}{c_i}\mathbf{v}_1 + -\frac{c_2}{c_i}\mathbf{v}_2 + \cdots + \widehat{\mathbf{v}_i} + \cdots + -\frac{c_r}{c_i}\mathbf{v}_r$$

(The symbol $\widehat{\mathbf{v}_i}$ means that the term with $\mathbf{v}_i$ is omitted from the sum.) This shows that the vectors $\{\mathbf{v}_i\}$ are dependent by the definition in §4.3.

Conversely, suppose that the vectors $\{\mathbf{v}_i\}$ are dependent by the definition in §4.3. Then one of the vectors, say $\mathbf{v}_i$, can be written as a linear combination of the others.

$$\mathbf{v}_i = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + \widehat{\mathbf{v}_i} + \cdots + c_r\mathbf{v}_r$$

Thus we have a linear combination

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + (-1)\mathbf{v}_i + \cdots + c_r\mathbf{v}_r = 0$$

so the $\{\mathbf{v}_i\}$ are dependent in the sense defined in this exercise.

4. First we verify that the standard basis vectors are independent. We will use the definition from exercise 3 above. Suppose there is a linear combination

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n = 0$$

But the vector on the left is $(c_1, c_2, \ldots, c_n)$, so it must be the case that $c_1 = c_2 = \cdots = c_n = 0$.

Next, we verify spanning. This is clear, because given any vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we can write $\mathbf{x}$ as a linear combination $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i$.

5. Because $\mathcal{B}$ is a basis, the vectors in $\mathcal{B}$ span all of $\mathbf{R}^n$, so there exist constants $c_1, \ldots, c_n$ such that $\mathbf{x} = \sum_{i=1}^n c_i\mathbf{b}_i$. Now suppose that $\mathbf{x} = \sum_{i=1}^n d_i\mathbf{b}_i$ for some constants $d_1, \ldots, d_n$. Subtracting, we have $\sum_{i=1}^n (c_i - d_i)\mathbf{b}_i = 0$. Using the definition of independence from the exercise 3 above, we conclude that $0 = (c_1 - d_1) = (c_2 - d_2) = \cdots = (c_n - d_n)$. This proves the desired uniqueness.

Exercises for Section 5 Solutions

1. A linear map $\mathbf{R}^1 \to \mathbf{R}^1$ has the form $x \mapsto mx$ for some constant m, so the assumption that $f\colon \mathbf{R}^1 \to \mathbf{R}^1$ is differentiable at x_0 in the sense of (5.1.1) means that there is some constant m such that
$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - mh}{|h|} = 0.$$
It follows that
$$\lim_{h\to 0^+} \frac{f(x_0+h) - f(x_0) - mh}{h} = 0.$$
Because the limit $\lim_{h\to 0^+} \frac{mh}{h} = m$ exists, the addition property of limits implies that
$$\lim_{h\to 0^+} \frac{f(x_0+h) - f(x_0)}{h} = m.$$
A similar argument establishes that
$$\lim_{h\to 0^-} \frac{f(x_0+h) - f(x_0)}{h} = m.$$
The agreement of both the left and right hand limits implies that the two-sided limit exists.
$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} = m$$
This is the calculus 1 definition of the differentiability of f at x_0, and further, we have $f'(x_0) = m$.

   Conversely, if we assume that $f\colon \mathbf{R}^1 \to \mathbf{R}^1$ is differentiable in the usual calculus 1 sense, we can reverse the steps of the argument above to conclude that
$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - mh}{|h|} = 0,$$
i.e., that f is differentiable at x_0 by definition (5.1.1).

2. The function $f\colon \mathbf{R}^n \to \mathbf{R}^m$ has m output components $y_1, y_2, \ldots, y_m$. We can read the limit component wise
$$\lim_{t\to 0} \frac{(y_1(\mathbf{x}_0 + t\mathbf{e}_j) - y_1(\mathbf{x}_0), y_2(\mathbf{x}_0 + t\mathbf{e}_j) - y_2(\mathbf{x}_0), \ldots, y_n(\mathbf{x}_0 + t\mathbf{e}_j) - y_n(\mathbf{x}_0))}{t}$$
and then we recognize that the ith component
$$\lim_{t\to 0} \frac{(y_i(\mathbf{x}_0 + t\mathbf{e}_j) - y_i(\mathbf{x}_0))}{t}$$
is the partial derivative $\partial y_i/\partial x_j$.

3. The jth column of the matrix for a linear transformation L is $L\mathbf{e}_j$.

4. The given formula is just the summation notation version of (3.2.7) applied to (5.2.2).

   i, j entry of $[D(f\circ g)(\mathbf{t}_0)] = (i\text{th row of } [Df(\mathbf{x}_0)])\cdot(j\text{th column of } [Dg(\mathbf{t}_0)])$.