Linear Algebra and Vector Calculus
Notes and Exercises

Somnath Sikdar

January 28, 2020

Contents

Chapter 1

Linear Algebra Basics

1.1 Linear Functions

A function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ is a linear function if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^m$ and for all $a, b \in \mathbf{R}$

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

It follows (by induction) that for all $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r \in \mathbf{R}^m$ and all $a_1, \dots, a_r \in \mathbf{R}$

$$L(a_1\boldsymbol{x}_1 + \cdots + a_r\boldsymbol{x}_r) = a_1L(\boldsymbol{x}_1) + \cdots + a_rL(\boldsymbol{x}_r).$$

Theorem 1.1. *A linear function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ is completely determined by its effect on the standard basis vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ of $\mathbf{R}^m$. An arbitrary choice of vectors $L(\boldsymbol{e}_1), \dots, L(\boldsymbol{e}_m)$ of $\mathbf{R}^n$ determines a linear function from $\mathbf{R}^m$ to $\mathbf{R}^n$.*

Proof. Given any vector $\boldsymbol{x} \in \mathbf{R}^m$, we can express it as a unique linear combination $\sum_{i=1}^m \alpha_i \boldsymbol{e}_i$ of the basis vectors. By the linearity of L, $L(\boldsymbol{x}) = \sum_i \alpha_i L(\boldsymbol{e}_i)$ which is completely specified by $L(\boldsymbol{e}_1), \dots, L(\boldsymbol{e}_m)$.

Let $\boldsymbol{b}_1, \dots, \boldsymbol{b}_m$ be any vectors in $\mathbf{R}^n$. Define a map L from $\mathbf{R}^m$ to $\mathbf{R}^n$ as follows: for $\boldsymbol{x} = \sum_{i=1}^m \alpha_i \boldsymbol{e}_i \in \mathbf{R}^m$, $L(\boldsymbol{x}) = \sum_{i=1}^m \alpha_i \boldsymbol{b}_i$. Then $L(\boldsymbol{e}_i) = \boldsymbol{b}_i$ for all $1 \leq i \leq m$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^m$ and all $a, b \in \mathbf{R}$:

$$L(a\boldsymbol{x} + b\boldsymbol{y}) = \sum_{i=1}^m (ax_i + by_i)\boldsymbol{b}_i = a\sum_i x_i\boldsymbol{b}_i + b\sum_i y_i\boldsymbol{b}_i = aL(\boldsymbol{x}) + bL(\boldsymbol{y})$$

$\square$

Note that the domain of definition of a linear function must be a vector space. An non-linear function can be defined on a subset of a vector space.

1.2 Image and Kernel of a Linear Function

The image $\mathrm{Im}(L)$ of a linear function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ is the set of vectors in $\mathbf{R}^n$ that L maps $\mathbf{R}^m$ to. In symbols, $\mathrm{Im}(L) := \{L(\boldsymbol{x}) \in \mathbf{R}^n \colon \boldsymbol{x} \in \mathbf{R}^m\}$. The kernel $\mathrm{Ker}(L)$ of L is the set of vectors in $\mathbf{R}^m$ that L maps to the zero vector in $\mathbf{R}^n$: $\mathrm{Ker}(L) := \{\boldsymbol{x} \in \mathbf{R}^m \colon L(\boldsymbol{x}) = \mathbf{0}_n\}$.

Theorem 1.2. *Let $L\colon \mathbf{R}^m \to \mathbf{R}^n$ be a linear function. Then the following hold:*

1. $\mathrm{Im}(L)$ *is a subspace of* $\mathbf{R}^n$*;*
2. $\mathrm{Ker}(L)$ *is a subspace of* $\mathbf{R}^m$*;*
3. $\dim(\mathrm{Im}(L)) + \dim(\mathrm{Ker}(L)) = \dim(\mathbf{R}^m) = m$.

Proof. The proof of (1) and (2) are similar. For proving (1), let $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathrm{Im}(L)$ and $\alpha, \beta \in \mathbf{R}$. Then there exist $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbf{R}^m$ such that

$$L(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) = \alpha \boldsymbol{z}_1 + \beta \boldsymbol{z}_2,$$

implying that $\alpha \boldsymbol{z}_1 + \beta \boldsymbol{z}_2 \in \mathrm{Im}(L)$. Since $\mathrm{Im}(L)$ is closed under vector addition and multiplication by scalars, it is a vector space. Since it is a subset of $\mathbf{R}^n$, it must be a subspace of $\mathbf{R}^n$.

To prove (3), let us assume that $\dim(\mathrm{Ker}(L)) = k$ and that $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ is a basis of $\mathrm{Ker}(L)$. Extend this basis to a basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k, \boldsymbol{b}_{k+1}, \ldots, \boldsymbol{b}_m$ for $\mathbf{R}^m$. Then for every vector $\boldsymbol{x} \in \mathbf{R}^m$ there exist scalars $\beta_1, \ldots, \beta_m$ such that $\boldsymbol{x} = \sum_{i=1}^{k} \beta_i \boldsymbol{b}_i$. Moreover since L is linear,

$$\begin{aligned} L(\boldsymbol{x}) &= \sum_{i=1}^{m} \beta_i L(\boldsymbol{b}_i) \\ &= \sum_{i=k+1}^{m} \beta_i L(\boldsymbol{b}_i). \end{aligned}$$

This shows that every vector in $\mathrm{Im}(L)$ can be expressed as a linear combination of the vectors $L(\boldsymbol{b}_{k+1}), \ldots, L(\boldsymbol{b}_m)$. To show that they form a basis of $\mathrm{Im}(L)$, it is sufficient to show that they are linearly independent. Suppose not. Then there exist scalars $\beta_{k+1}, \ldots, \beta_m$, not all zero, such that $\sum_{i=k+1}^{m} \beta_i L(\boldsymbol{b}_i) = \mathbf{0}_n$. By the linearity of L, we have $L(\sum_{i=k+1}^{m} \beta_i \boldsymbol{b}_i) = \mathbf{0}_n$ and hence $\mathbf{0}_m \neq \sum_{i=k+1}^{m} \beta_i \boldsymbol{b}_i \in \mathrm{Ker}(L)$. This is a contradiction since the vectors $\boldsymbol{b}_{k+1}, \ldots, \boldsymbol{b}_m$ are not in the space spanned by the vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$. Thus the vectors $\beta_{k+1}, \ldots, \beta_m$ must be independent and form a basis of $\mathrm{Im}(L)$. This proves (3). $\square$