

Linear Algebra and Vector Calculus

Notes and Exercises

Somnath Sikdar

February 2, 2020

Contents

1	Linear Algebra Basics	2
1.1	Linear Functions	2
1.2	Image and Kernel of a Linear Function	2
1.3	Quadratic Forms	3
1.3.1	Definiteness	4

Chapter 1

Linear Algebra Basics

1.1 Linear Functions

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

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

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

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

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

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

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

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

□

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

1.2 Image and Kernel of a Linear Function

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

Theorem 1.2. Let $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a linear function. Then the following hold:

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

Proof. The proof of (1) and (2) are similar. For proving (1), let $z_1, z_2 \in \text{Im}(L)$ and $\alpha, \beta \in \mathbb{R}$. Then there exist $x_1, x_2 \in \mathbb{R}^m$ such that

$$L(\alpha x_1 + \beta x_2) = \alpha z_1 + \beta z_2,$$

implying that $\alpha z_1 + \beta z_2 \in \text{Im}(L)$. Since $\text{Im}(L)$ is closed under vector addition and multiplication by scalars, it is a vector space. Since it is a subset of \mathbb{R}^n , it must be a subspace of \mathbb{R}^n .

To prove (3), let us assume that $\dim(\text{Ker}(L)) = k$ and that b_1, \dots, b_k is a basis of $\text{Ker}(L)$. Extend this basis to a basis $b_1, \dots, b_k, b_{k+1}, \dots, b_m$ for \mathbb{R}^m . Then for every vector $x \in \mathbb{R}^m$ there exist scalars β_1, \dots, β_m such that $x = \sum_{i=1}^k \beta_i b_i$. Moreover since L is linear,

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

This shows that every vector in $\text{Im}(L)$ can be expressed as a linear combination of the vectors $L(b_{k+1}), \dots, L(b_m)$. To show that they form a basis of $\text{Im}(L)$, it is sufficient to show that they are linearly independent. Suppose not. Then there exist scalars $\beta_{k+1}, \dots, \beta_m$, not all zero, such that $\sum_{i=k+1}^m \beta_i L(b_i) = 0_n$. By the linearity of L , we have $L(\sum_{i=k+1}^m \beta_i b_i) = 0_n$ and hence $0_m \neq \sum_{i=k+1}^m \beta_i b_i \in \text{Ker}(L)$. This is a contradiction since the vectors b_{k+1}, \dots, b_m are not in the space spanned by the vectors b_1, \dots, b_k . Thus the vectors $\beta_{k+1}, \dots, \beta_m$ must be independent and form a basis of $\text{Im}(L)$. This proves (3). \square

1.3 Quadratic Forms

A function $q: \mathbb{R}^m \rightarrow \mathbb{R}$ is called a *quadratic form* if there exists a real symmetric $m \times m$ matrix A such that for all $x \in \mathbb{R}^m$ such that

$$q(x) = x^T A x.$$

The right-hand side of the equation may be written as:

$$\begin{aligned} (x_1, \dots, x_m) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{m1} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} &= (x_1, \dots, x_m) [x_1 A_{*1} + \cdots + x_m A_{*m}] \\ &= \sum_{i=1}^m \sum_{j=1}^m x_i a_{ij} x_j. \end{aligned}$$

The symbol A_{*j} denotes the j th column of A . The name “quadratic” form arises from the last expression which is what a quadratic expression in m variables looks like.

Example 1.1. The function

$$q(h_1, h_2, h_3) = h_1^2 + h_2^2 + h_3^2 + 2h_1h_2 + 2h_2h_3 + 2h_1h_3 = (h_1 + h_2 + h_3)^2$$

can be expressed as $\mathbf{h}^T A \mathbf{h}$, where A is the all-ones 3×3 matrix.

1.3.1 Definiteness

A quadratic form $q: \mathbb{R}^m \rightarrow \mathbb{R}$ is

1. *positive definite* if for all non-zero $\mathbf{x} \in \mathbb{R}^m$, $q(\mathbf{x}) > 0$;
2. *negative*

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{1.1}$$

definite if for all non-zero $\mathbf{x} \in \mathbb{R}^m$, $q(\mathbf{x}) > 0$;

3. *indefinite* if $q(\mathbf{x})$ takes on both positive and negative values.

Example 1.2. Let $q(\mathbf{h}) = -h_1^2 + 2h_2^2 - h_3^2$ such that the matrix associated with q is

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The $q(0, 1, 0) = 2$ and $q(1, 0, 1) = -2$. Thus the quadratic form q is indefinite.

Example 1.3. The quadratic form $q(h_1, h_2, h_3) = h_1^2 + h_2^2 + 4h_3^2 + 2h_1h_2 - 4h_2h_3 - 4h_1h_3$ is positive definite, since we may re-write $q(h_1, h_2, h_3)$ as $(h_1 + h_2 - 2h_3)^2$ which is strictly positive for all $(h_1, h_2, h_3) \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$.

To characterize the definiteness of quadratic forms, we make use of the following properties of real symmetric matrices.

Lemma 1.1. *The eigenvectors of a real symmetric matrix A are real.*

Proof. Suppose that A is an $m \times m$ matrix and suppose that some eigenvalue λ is complex. Let \mathbf{x} be an eigenvector corresponding to it, which can be a complex vector. Then

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{1.2}$$

Taking complex conjugates on each side, we get:

$$A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}. \tag{1.3}$$

Pre-multiply (1.2) by \bar{x} and (1.3) by x^\top to obtain:

$$\bar{x}^\top Ax = \bar{x}^\top \lambda x \tag{1.4}$$

$$x^\top A \bar{x} = x^\top \bar{\lambda} \bar{x}. \tag{1.5}$$

Take the transpose of equation (1.5), we obtain: $\bar{x}^\top Ax = \bar{x}^\top \bar{\lambda} x$, where we made use of the fact that A is symmetric. Now subtracting this equation from (1.4), we obtain:

$$x^\top (\lambda - \bar{\lambda}) \bar{x} = 0.$$

Since $x^\top \bar{x}$ is the sum of products of complex conjugates, it is not zero unless each component of x is zero. Since this is not the case (x is an eigenvector), we must have $\lambda = \bar{\lambda}$ and hence λ is real. \square