MTH6140 Linear Algebra II

Notes 7 16th December 2010

7 Inner product spaces

Ordinary Euclidean space is a 3-dimensional vector space over $\mathbb{R}$, but it is more than that: the extra geometric structure (lengths, angles, etc.) can all be derived from a special kind of bilinear form on the space known as an inner product. We examine inner product spaces and their linear maps in this chapter. Throughout the chapter, all vector spaces will be *real* vector spaces, i.e. the underlying field is always the real numbers $\mathbb{R}$.

7.1 Inner products and orthonormal bases

Definition 7.1 An *inner product* on a real vector space V is a function $b : V \times V \to \mathbb{R}$ satisfying

- b is bilinear (that is, b is linear in the first variable when the second is kept constant and *vice versa*);
- b is symmetric (that is, $b(v,w) = b(w,v)$ for all $v,w \in V$);
- b is *positive definite*, that is, $b(v,v) \geq 0$ for all $v \in V$, and $b(v,v) = 0$ if and only if $v = 0$.

We usually write $b(v,w)$ as $v \cdot w$. An inner product is sometimes called a *dot product* (because of this notation).

Definition 7.2 When a real vector space is equipped with an inner product, we refer to it as an *inner product space*.

Geometrically, in a real vector space, we define $v \cdot w = |v|.|w| \cos\theta$, where $|v|$ and $|w|$ are the lengths of v and w, and θ is the angle between v and w. Of course this definition doesn't work if either v or w is zero, but in this case $v \cdot w = 0$. But it is much easier to reverse the process. Given an inner product on V, we define

$$|v| = \sqrt{v \cdot v}$$

for any vector $v \in V$; and, if $v, w \neq 0$, then we define the angle between them to be θ, where

$$cos\theta = \frac{v \cdot w}{|v|.|w|}.$$

Definition 7.3 A basis $(v_1, \ldots, v_n)$ for an inner product space is called *orthonormal* if $v_i \cdot v_j = \delta_{ij}$ (the Kronecker delta) for $1 \leq i, j \leq n$.
In other words, $v_i \cdot v_i = 1$ for all $1 \leq i \leq n$, and $v_i \cdot v_j = 0$ for $i \neq j$.

Remark: If vectors $v_1, \ldots, v_n$ satisfy $v_i \cdot v_j = \delta_{ij}$, then they are necessarily linearly independent. For suppose that $c_1 v_1 + \cdots + c_n v_n = 0$. Taking the inner product of this equation with v_i, we find that $c_i = 0$, for all i.

7.2 The Gram-Schmidt process

Given any basis for V, a constructive method for finding an *orthonormal* basis is known as the *Gram–Schmidt process*.

Let $w_1, \ldots, w_n$ be any basis for V. The Gram–Schmidt process works as follows.

- Since $w_1 \neq 0$, we have $w_1 \cdot w_1 > 0$, that is, $|w_1| > 0$. Put $v_1 = w_1/|w_1|$; then $|v_1| = 1$, that is, $v_1 \cdot v_1 = 1$.
- For $i = 2, \ldots, n$, let $w_i' = w_i - (v_1 \cdot w_i) v_1$. Then

  $$v_1 \cdot w_i' = v_1 \cdot w_i - (v_1 \cdot w_i)(v_1 \cdot v_1) = 0$$

  for $i = 2, \ldots, n$.
- Now apply the Gram–Schmidt process recursively to $(w_2', \ldots, w_n')$.

Since we replace these vectors by linear combinations of themselves, their inner products with v_1 remain zero throughout the process. So if we end up with vectors $v_2, \ldots, v_n$, then $v_1 \cdot v_i = 0$ for $i = 2, \ldots, n$. By induction, we can assume that $v_i \cdot v_j = \delta_{ij}$ for $i, j = 2, \ldots, n$; by what we have said, this holds if i or j is 1 as well.

Definition 7.4 The inner product on $\mathbb{R}^n$ for which the standard basis is orthonormal is called the *standard inner product* on $\mathbb{R}^n$.

Example 7.1 In $\mathbb{R}^3$ (with the standard inner product), apply the Gram–Schmidt process to the vectors

$$w_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \ w_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \ w_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

We have $w_1 \cdot w_1 = 9$, so in the first step we put

$$v_1 = \tfrac{1}{3} w_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}.$$

Now $v_1 \cdot w_2 = 1$ and $v_1 \cdot w_3 = \frac{1}{3}$, so in the second step we find

$$\begin{aligned} w_2' &= w_2 - v_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}, \\ w_3' &= w_3 - \tfrac{1}{3} v_1 = \begin{bmatrix} 8/9 \\ -2/9 \\ -2/9 \end{bmatrix}. \end{aligned}$$

Now we apply Gram–Schmidt recursively to w_2' and w_3'. We have $w_2' \cdot w_2' = 1$, so

$$v_2 = w_2' = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}.$$

Then $v_2 \cdot w_3' = \frac{2}{3}$, so

$$w_3'' = w_3' - \tfrac{2}{3} v_2 = \begin{bmatrix} 4/9 \\ -4/9 \\ 2/9 \end{bmatrix}.$$

Finally, $w_3'' \cdot w_3'' = \frac{4}{9}$, so

$$v_3 = \frac{3}{2} w_3'' = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}.$$

Check that the three vectors v_1, v_2, v_3 we have found really do form an orthonormal basis.

7.3 Adjoints and orthogonal linear maps

Definition 7.5 Let V be an inner product space, and $T : V \to V$ a linear map. Then the *adjoint* of T is the linear map $T^* : V \to V$ defined by

$$v \cdot T^*(w) = T(v) \cdot w.$$

Proposition 7.1 If $T : V \to V$ is represented by the matrix A relative to an orthonormal basis of V, then T^* is represented by the transposed matrix $A^\top$.

Now we define two important classes of linear maps on V.

Definition 7.6 Let T be a linear map on an inner product space V.

(a) T is *self-adjoint* if $T^* = T$.

(b) T is *orthogonal* if it is invertible and $T^* = T^{-1}$.

We see that, if T is represented by a matrix A (relative to an orthonormal basis), then

- T is self-adjoint if and only if A is symmetric;
- T is orthogonal if and only if $A^\top A = I$, i.e. if and only if $A^{-1} = A^\top$.

We will examine self-adjoint maps (or symmetric matrices) further in the next chapter. Here we look at orthogonal maps.

Theorem 7.2 The following are equivalent for a linear map T on an inner product space V:

(a) T is orthogonal;

(b) T preserves the inner product, that is, $T(v) \cdot T(v) = v \cdot v$;

(c) T maps an orthonormal basis of V to an orthonormal basis.

Proof We have

$$T(v) \cdot T(w) = v \cdot T^*(T(w)),$$

by the definition of adjoint (see Definition 7.5); so (a) and (b) are equivalent.

Suppose that $(v_1, \ldots, v_n)$ is an orthonormal basis, that is, $v_i \cdot v_j = \delta_{ij}$. If (b) holds, then $T(v_i) \cdot T(v_j) = \delta_{ij}$, so that $(T(v_1), \ldots, T(v_n)$ is an orthonormal basis, and (c) holds. Conversely, suppose that (c) holds, and let $v = \sum x_i v_i$ and $w = \sum y_i v_i$ for some orthonormal basis $(v_1, \ldots, v_n)$, so that $v \cdot w = \sum x_i y_i$. We have

$$T(v) \cdot T(w) = \left(\sum x_i T(v_i)\right) \cdot \left(\sum y_i T(v_i)\right) = \sum x_i y_i,$$

since $T(v_i) \cdot T(v_j) = \delta_{ij}$ by assumption; so (b) holds.

Corollary 7.3 T is orthogonal if and only if the columns of the matrix representing T relative to an orthonormal basis themselves form an orthonormal basis.

Proof The columns of the matrix representing T are just the vectors $T(v_1),\dots,T(v_n)$, written in coordinates relative to $v_1,\dots,v_n$. So this follows from the equivalence of (a) and (c) in the theorem. Alternatively, the condition on columns shows that $A^\top A = I$, where A is the matrix representing T; so $T^*T = I$, and T is orthogonal.

Example 7.2 Our earlier example of the Gram–Schmidt process produces the orthogonal matrix

$$\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

whose columns are precisely the orthonormal basis we constructed in the example.