Proofs Homework Set 12

MATH 217 — WINTER 2011

Due April 6

Given a vector space V, an **inner product** on V is a function that associates with each pair of vectors $\mathbf{v}, \mathbf{w} \in V$ a real number, denoted $\langle \mathbf{v}, \mathbf{w} \rangle$, satisfying the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and for all scalars $c \in \mathbb{R}$:

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

(ii) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

(iii) $\langle c\mathbf{u}, \mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$

(iv) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = 0$.

Note that the dot product is an inner product on $\mathbb{R}^n$ by Theorem 6.1 on page 376.

PROBLEM 12.1. Let $C[0,1]$ be the space of all continuous functions $f : [0,1] \to \mathbb{R}$. Define

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx$$

for all pairs of functions f, g in $C[0,1]$. Show that this is in fact an inner product, that is, that it satisfies the four properties listed above.

Proof. The first three properties are straightforward to verify using elementary facts about integration.

(i)

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx = \int_0^1 g(x)f(x)dx = \langle g, f \rangle$$

(ii)

$$\begin{aligned}\langle f + g, h \rangle &= \int_0^1 (f(x) + g(x))h(x)dx = \int_1^0 (f(x)h(x) + g(x)h(x))dx \\ &= \int_0^1 f(x)h(x)dx + \int_0^1 g(x)h(x)dx = \langle f, h \rangle + \langle g, h \rangle\end{aligned}$$

(iii)

$$\langle cf, g \rangle = \int_0^1 cf(x)g(x)dx = c \int_0^1 f(x)g(x)dx = c \langle f, g \rangle$$

The last property requires some more careful analysis. First, note that

$$\langle f, f\rangle = \int_0^1 f(x)^2 dx \geq 0$$

since $f(x)^2 \geq 0$ for all $x \in [0,1]$.

To see that $\langle f, f\rangle > 0$ whenever $f(x)$ is not the constant zero function, we need to think about continuous functions. First, if $f(x)$ is not the constant zero function, then $f(x_0)^2 > 0$ for some $x_0 \in [0,1]$. Let $\delta = f(x_0)^2/2$ (this is a somewhat arbitrary choice which makes everything work out later in the proof). By the definition of limit, we can find an $\varepsilon > 0$ such that $|f(x)^2 - f(x_0)^2| < \delta$ whenever $|x - x_0| < \varepsilon$. It follows that $f(x)^2 - f(x_0)^2 > -\delta = -f(x_0)^2/2$ whenever $|x - x_0| < \varepsilon$; rearranging, we find that $f(x)^2 > f(x_0)^2/2$ whenever $|x - x_0| < \varepsilon$. It then follows that

$$\begin{aligned}\langle f, f\rangle &= \int_0^1 f(x)^2 dx \geq \int_{x_0-\varepsilon}^{x_0+\varepsilon} f(x)^2 dx \\ &\geq \int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{f(x_0)^2}{2} dx = (2\varepsilon)\frac{f(x_0)^2}{2} = \varepsilon f(x_0)^2 > 0.\end{aligned}$$

Therefore, if $f(x)$ is not the constant zero function, then $\langle f, f\rangle > 0$. □

PROBLEM 12.2. Whenever V is a finite-dimensional vector space with basis $\mathcal{B}$, we can use the $\mathcal{B}$-coordinate system to define an inner product on V:

$$\langle \mathbf{u}, \mathbf{v}\rangle_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}}.$$

(a) Verify that this does indeed define an inner product on V, i.e. that the four properties listed above are true for $\langle \bullet, \bullet\rangle_{\mathcal{B}}$.

Proof. This follows from Theorem 6.1 on page 376 and the fact that the $\mathcal{B}$-coordinate transformation is a one-to-one linear transformation (see Theorem 4.8 on page 250).

(i)
$$\langle \mathbf{u}, \mathbf{v}\rangle_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{u}]_{\mathcal{B}} = \langle \mathbf{v}, \mathbf{u}\rangle_{\mathcal{B}}$$

by Theorem 6.1(a).

(ii)
$$\begin{aligned}\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle_{\mathcal{B}} &= [\mathbf{u}+\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}} = ([\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}) \cdot [\mathbf{w}]_{\mathcal{B}} \\ &= [\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}} = \langle \mathbf{u}, \mathbf{w}\rangle_{\mathcal{B}} + \langle \mathbf{v}, \mathbf{w}\rangle_{\mathcal{B}}\end{aligned}$$

by Theorem 6.1(b) and the linearity of the $\mathcal{B}$-coordinate transformation.

(iii)
$$\langle c\mathbf{u}, \mathbf{v}\rangle = [c\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}} = (c[\mathbf{u}]_{\mathcal{B}}) \cdot [\mathbf{v}]_{\mathcal{B}} = c([\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}}) = c\,\langle \mathbf{u}, \mathbf{v}\rangle$$

by Theorem 6.1(c) and the linearity of the $\mathcal{B}$-coordinate transformation.

(iv) By Theorem 6.1(d), we have that $\langle \mathbf{v}, \mathbf{v}\rangle_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}} \geq 0$ and $\langle \mathbf{v}, \mathbf{v}\rangle_{\mathcal{B}} = 0$ if and only if $[\mathbf{v}]_{\mathcal{B}} = \mathbf{0}$. Since the $\mathcal{B}$-coordinate transformation is one-to-one, we see that $[\mathbf{v}]_{\mathcal{B}} = \mathbf{0}$ if and only if $\mathbf{v} = \mathbf{0}$. Therefore, $\langle \mathbf{v}, \mathbf{v}\rangle_{\mathcal{B}} \geq 0$, with equality if and only if $\mathbf{v} = \mathbf{0}$.

$\square$

Now consider the space $M_{2\times 2}$ of 2×2 matrices with the standard basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

(b) Let Symm be the subspace of $M_{2\times 2}$ consisting of symmetric matrices (i.e. 2×2 matrices A that satisfy $A = A^T$). Find the orthogonal complement $\mathsf{Symm}^\perp$ with respect to the inner product $\langle \bullet, \bullet \rangle_{\mathcal{B}}$.

Proof. First, note that a basis for Symm consists of the three matrices

$$S_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, S_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, S_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

By the first boxed fact below in the middle of page 380, a 2×2 matrix M belongs to $\mathsf{Symm}^\perp$ if and only if $\langle M, S_1 \rangle_{\mathcal{B}} = 0$, $\langle M, S_2 \rangle_{\mathcal{B}} = 0$, and $\langle M, S_3 \rangle_{\mathcal{B}} = 0$. By definition of $\langle \bullet, \bullet \rangle_{\mathcal{B}}$, this will happen if and only if

$$[M]_{\mathcal{B}} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0, \quad [M]_{\mathcal{B}} \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = 0, \quad [M]_{\mathcal{B}} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = 0.$$

In other words, the vector $[M]_{\mathcal{B}}$ must solve the homogeneous system

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} [M]_{\mathcal{B}} = \mathbf{0}.$$

This will only happen if

$$[M]_{\mathcal{B}} \in \operatorname{Span} \left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

or, in other words,

$$M \in \operatorname{Span} \left\{ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right\}.$$

Therefore, $\mathsf{Symm}^\perp = \operatorname{Span} \left\{ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right\}$. $\square$

(c) The **trace** of a matrix, denoted $\operatorname{tr}(A)$, is the sum of the entries on the main diagonal of A. Show that $\langle A, A \rangle_{\mathcal{B}} = \operatorname{tr}(A^T A)$.

Proof. Write $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, so that $[A]_{\mathcal{B}} = \left[\begin{smallmatrix} a \\ b \\ c \\ d \end{smallmatrix}\right]$.

On the one hand, we have

$$\langle A, A \rangle_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = a^2 + b^2 + c^2 + d^2.$$

On the other hand,

$$A^T A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ba + dc \\ ba + dc & b^2 + d^2 \end{bmatrix}$$

and hence $\operatorname{tr}(A^T A) = (a^2 + c^2) + (b^2 + d^2)$.

Therefore, $\langle A, A \rangle_{\mathcal{B}} = \operatorname{tr}(A^T A)$. $\square$