VII.C. Inner products

Let $V/\mathbb{R}$ be an n-dimensional *real* vector space. Recall from the end of §VII.B that a symmetric bilinear form

$$B\colon V \times V \to \mathbb{R}$$

is called *positive-definite* if it is nondegenerate of signature $(p,q) = (n,0)$ — or equivalently, if

$$\vec{v} \neq \vec{0} \quad \Longrightarrow \quad B(\vec{v},\vec{v}) > 0.$$

VII.C.1. DEFINITION. An **inner product** is a positive-definite symmetric bilinear form.

NOTATION: Instead of $B(\vec{u},\vec{w})$ we shall write $\langle \vec{u},\vec{w}\rangle$. Since $\langle \cdot,\cdot \rangle$ is definite,

$$0 = \|\vec{v}\|^2 := \langle \vec{v},\vec{v}\rangle \quad \Longrightarrow \quad \vec{v} = \vec{0}.$$

A given inner product determines notions of length and orthogonality on V:

$$\vec{u} \perp \vec{w} \;\Leftrightarrow\; \langle \vec{u},\vec{w}\rangle = 0\,, \quad \|\vec{u}\| = \sqrt{\langle \vec{u},\vec{u}\rangle}.$$

If $V = \mathbb{R}^n$ these should *not* be mixed with the $\|\cdot\|$ and $\perp$ coming from the dot product.[11]

For *any* symmetric bilinear form B, Sylvester's theorem guarantees the existence of a basis $\mathcal{B}$ such that

$$B(\vec{u},\vec{w}) \;=\; {}^t[\vec{u}]_{\mathcal{B}} \begin{pmatrix} \mathbb{I}_p & 0 & 0 \\ 0 & -\mathbb{I}_q & 0 \\ 0 & 0 & 0 \end{pmatrix} [\vec{w}]_{\mathcal{B}}.$$

Clearly in order for $B(\vec{u},\vec{u})$ to be always positive ($\vec{u} \neq 0$) we must have $q = 0$ and $p = n$ (all 1's along the diagonal). Therefore any given *inner product* has a basis with respect to which

$$\langle \vec{u},\vec{w}\rangle = {}^t[\vec{u}]_{\mathcal{B}}[\vec{w}]_{\mathcal{B}} = \left\{ \begin{array}{c} \text{"dot product" in that} \\ \text{coordinate system} \end{array} \right. .$$

[11] often called the "Euclidean" or "standard" inner product on $\mathbb{R}^n$.

WARNING: If $V = \mathbb{R}^n$ this is *not* necessarily the usual dot product, unless $\mathcal{B} = \hat{e}$!

If instead V is a vector space over $\mathbb{C}$, then by "inner product" we shall mean a positive-definite Hermitian form $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$. In particular, the 1^{st} entry is conjugate-linear $[\langle \alpha\vec{u}, \vec{w} \rangle = \bar{\alpha} \langle \vec{u}, \vec{w} \rangle]$, the 2^{nd} linear $[\langle \vec{u}, \alpha\vec{w} \rangle = \alpha \langle \vec{u}, \vec{w} \rangle]$, and $\langle \vec{u}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{u} \rangle}$. By Sylvester there is a basis with respect to which

$$\langle \vec{u}, \vec{w} \rangle = [\vec{u}]^*_{\mathcal{B}} [\vec{w}]_{\mathcal{B}},$$

which is "like" the complex dot product (and the same warning applies when $V = \mathbb{C}^n$). By "plugging in" the basis elements for $\vec{u}$, $\vec{w}$ one has:

VII.C.2. PROPOSITION. *For any inner product on a complex (real) vector space V, there is a basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ for V such that $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$, i.e. $\mathcal{B}$ is a unitary (orthonormal) basis under the inner product.*

WARNING: For $\mathbb{C}^n$ (resp. $\mathbb{R}^n$), $\hat{e}$ is unitary (orthonormal) *only* under the standard inner (= dot) product.

VII.C.3. REMARK. V together with some inner product $\langle \cdot, \cdot \rangle$ (and the accompanying notions of $\perp$ and length) is called an "inner product space". The Proposition simply says: ***any inner product space has a unitary (orthonormal) basis***, though (as the next Example attests) this basis is *far* from unique. The practical way to obtain one, starting from any given basis for V, is by the *Gram-Schmidt procedure* (exactly as we did it for the dot product); see for example Exercise III.B.2(a).

VII.C.4. EXAMPLE. Consider the symmetric bilinear form

$$B(\vec{x}, \vec{y}) = \langle \vec{x}, \vec{y} \rangle = {}^t\vec{x} M \vec{y} := \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \frac{1}{8} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{x_1 y_1}{8} + \frac{x_2 y_2}{2}$$

on $V = \mathbb{R}^2$ (with vectors written with respect to the standard basis). Clearly this is positive definite, hence an inner product, and $\mathcal{B} =$

$\{\vec{v}_1, \vec{v}_2\} := \left\{ \left(\begin{smallmatrix} \sqrt{8} \\ 0 \end{smallmatrix} \right), \left(\begin{smallmatrix} 0 \\ \sqrt{2} \end{smallmatrix} \right) \right\}$ is an orthonormal basis of this inner product space, i.e. $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$.

For any matrix S satisfying[12] ${}^t SMS = M$, we have

$$\langle S\vec{v}_i, S\vec{v}_j \rangle = {}^t\vec{v}_i {}^t SMS\vec{v}_j = {}^t\vec{v}_i M\vec{v}_j = \langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$$

hence $S\mathcal{B} = \{S\vec{v}_1, S\vec{v}_2\}$ is another orthonormal basis. These matrices are precisely the ones of the form

$$S = \begin{pmatrix} \cos\theta & -2\sin\theta \\ \frac{1}{2}\sin\theta & \cos\theta \end{pmatrix}.$$

So taking $\theta = -\frac{\pi}{4}$, we get $S = \left(\begin{smallmatrix} \sqrt{2}/2 & \sqrt{2} \\ -\sqrt{2}/4 & \sqrt{2}/2 \end{smallmatrix} \right)$ and $S\mathcal{B} = \left\{ \left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right), \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix} \right) \right\}$, which as you can check, is indeed orthonormal in $\langle \cdot, \cdot \rangle$.

Generalizing one more notion from the dot product setting to an arbitrary complex (resp. real) inner product space, we have the

VII.C.5. DEFINITION. Call a transformation $T\colon V \to V$ **unitary** (resp. **orthogonal**) in the given inner product if $\|T\vec{v}\| = \|\vec{v}\|$ for all $\vec{v} \in V$. (Equivalently, $\langle T\vec{v}, T\vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle \; \forall \vec{v}, \vec{w} \in V$; see Exercise (5).)

Multiplication by S in Example VII.C.4 is orthogonal in this sense.

Gram coefficients, Fourier expansion. Next let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{C}$ (resp. $\mathbb{R}$), $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ a unitary (orthonormal) basis, and $T : V \to V$ a linear transformation. Write a_{ij} for the entries of $A = [T]_{\mathcal{B}}$ so that

$$T\vec{v}_j = \textstyle\sum_i a_{ij}\vec{v}_i,$$

and notice that

$$\left\langle \vec{v}_i, T\vec{v}_j \right\rangle = \left\langle \vec{v}_i \, , \textstyle\sum_k a_{kj}\vec{v}_k \right\rangle = \textstyle\sum_k a_{kj} \left\langle \vec{v}_i, \vec{v}_k \right\rangle = \textstyle\sum_k a_{kj}\delta_{ik} = a_{ij}$$

$$\implies \quad T\vec{v}_j = \textstyle\sum_i \left\langle \vec{v}_i, T\vec{v}_j \right\rangle \vec{v}_i \quad \implies \quad T\vec{x} = \textstyle\sum_i \left\langle \vec{v}_i, T\vec{x} \right\rangle \vec{v}_i,$$

where the last implication is by linear extension. Notice that (for $T =$ Id) we've recovered the Fourier expansion formula $\vec{x} = \sum_i \langle \vec{v}_i, \vec{x} \rangle \vec{v}_i$

[12] That is, in the notation of §VII.B, $S \in \mathrm{O}_2(\mathbb{R}, B)$; we have $\langle S\vec{x}, S\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$, and S preserves the inner product/bilinear form.

from §VII.A, but for a general inner product. The coefficients $\langle \vec{v}_i, T\vec{v}_j \rangle$ are called *Gram* coefficients. WARNING: They only compute the entries of $[T]_{\mathcal{B}}$ if $\mathcal{B}$ is unitary (orthonormal).

Adjoint of a linear transformation. Now we define a new transformation,[13] the **adjoint**

$$T^\dagger : V \to V$$

of T, by demanding that for all $\vec{u}$, $\vec{w}$

$$\left\langle T^\dagger \vec{u}\,,\, \vec{w} \right\rangle = \langle \vec{u}\,,\, T\vec{w} \rangle\,. \tag{VII.C.6}$$

How do we know such a transformation exists? If it did then we could apply the expansion formula to obtain[14]

$$T^\dagger \vec{x} = \textstyle\sum_i \left\langle \vec{v}_i, T^\dagger \vec{x} \right\rangle \vec{v}_i \;=\; \sum_i \overline{\langle T^\dagger \vec{x}, \vec{v}_i \rangle} \vec{v}_i \;=\; \sum_i \overline{\langle \vec{x}, T\vec{v}_i \rangle} \vec{v}_i,$$

and the latter expression is as good as a definition (plug it in to (VII.C.6) and check). Since

$$T^\dagger \vec{v}_j = \sum \overline{\langle \vec{v}_j, T\vec{v}_i \rangle} \vec{v}_i \;=\; \sum \overline{a_{ji}} \vec{v}_i$$

the matrix $[T^\dagger]_{\mathcal{B}}$ is the conjugate transpose of $[T]_{\mathcal{B}}$ (or just the transpose if $V/\mathbb{R}$). WARNING: While this matrix point of view on the adjoint is very important, you can't just write the matrix of T with respect to an arbitrary basis, conjugate transpose it and claim you found the matrix of $T^\dagger$. (If that were possible then $T^\dagger$ would have nothing to do with the inner product.) As usual, $[T^\dagger]_{\mathcal{B}} = {}^t\overline{[T]_{\mathcal{B}}}$ only holds for $\mathcal{B}$ unitary/orthogonal (under the given inner product).

Here's a slightly more abstract point of view on what the adjoint "is": for fixed $\vec{w} \in V$, $\langle \vec{w}, \cdot \rangle$ is a linear functional on V (defined by $\langle \vec{w}, \cdot \rangle(\vec{v}) = \langle \vec{w}, \vec{v} \rangle$), i.e. an element of $V^\vee$.[15] In fact the map sending $\vec{w} \mapsto \langle \vec{w}, \cdot \rangle$ gives an isomorphism of V with its dual space.

---

[13] not to be confused with the *classical adjoint* (or *adjugate*) $\mathrm{Adj}(M)$ of a matrix M.

[14] This computation and those that follow are valid for V over either $\mathbb{R}$ or $\mathbb{C}$. You may ignore the bars (denoting complex conjugation) in the real case, since the numbers are all real.

[15] Recall from §III.D that the symbol "$\vee$" *dualizes* vector spaces and linear transformations.

QUESTION: *What transformation on V corresponds* under this isomorphism *to the dual transformation $T^\vee \colon V^\vee \to V^\vee$? That is, what dotted map makes the following diagram "commute"?*

$$\begin{tikzcd} V \arrow[r, "\cong", "\vec{w} \mapsto \langle \vec{w}, \cdot \rangle"'] \arrow[d, dashed] & V^\vee \arrow[d, "T^\vee"] \\ V \arrow[r, "\cong", "\vec{w} \mapsto \langle \vec{w}, \cdot \rangle"'] & V^\vee \end{tikzcd}$$

The answer: $T^\dagger$. In other words, the top map followed by (composed with) $T^\vee$ is the same as $T^\dagger$ followed by the bottom map:

$$\left\langle T^\dagger \vec{w}, \cdot \right\rangle = T^\vee \left\langle \vec{w}, \cdot \right\rangle$$

as elements of $V^\vee$. One sees this by applying both sides to $\vec{v} \in V$: using the definition of $T^\vee$,

$$\left(T^\vee \left\langle \vec{w}, \cdot \right\rangle\right)(\vec{v}) = \left\langle \vec{w}, \cdot \right\rangle (T\vec{v}) = \langle \vec{w}, T\vec{v} \rangle = \left\langle T^\dagger \vec{w}, \vec{v} \right\rangle = \left\langle T^\dagger \vec{w}, \cdot \right\rangle (\vec{v}).$$

CONCLUSION: *Let V be an inner product space and $T \colon V \to V$ a linear transformation. Then if we use the inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ to view V as its own dual space, $T^\dagger \colon V \to V$ plays the role of $T^\vee$.*

Since $(T^\vee)^\vee = T$ you might think $(T^\dagger)^\dagger = T$, and this is correct: write for any $\vec{w}, \vec{u} \in V$

$$\langle \vec{w}, T\vec{u} \rangle = \left\langle T^\dagger \vec{w}, \vec{u} \right\rangle = \overline{\langle \vec{u}, T^\dagger \vec{w} \rangle} = \overline{\langle (T^\dagger)^\dagger \vec{u}, \vec{w} \rangle} = \left\langle \vec{w}, (T^\dagger)^\dagger \vec{u} \right\rangle;$$

thus

$$\left\langle \vec{w},\, T\vec{u} - (T^\dagger)^\dagger \vec{u} \right\rangle = 0 \quad \forall\, \vec{u}, \vec{w}.$$

In particular, by taking $\vec{w} = T\vec{u} - (T^\dagger)^\dagger \vec{u}$ we have $\|T\vec{u} - (T^\dagger)^\dagger \vec{u}\| = 0$, and thus $T\vec{u} - (T^\dagger)^\dagger = 0$ or $T\vec{u} = (T^\dagger)^\dagger \vec{u}$ for all $\vec{u} \in V$.

VII.C.7. REMARK. The adjoint gives another way to characterize *unitary transformations* (Defn. VII.C.5). From $\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$ $(\forall \vec{x}, \vec{y} \in V)$ we get $\langle T^\dagger T \vec{x}, \vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$ hence $\langle (T^\dagger T - \mathbb{I})\vec{x}, \vec{y} \rangle = 0$. Taking $\vec{y} = (T^\dagger T - \mathbb{I})\vec{x}$ gives $\|(T^\dagger T - \mathbb{I})\vec{x}\| = 0$ hence $(T^\dagger T - \mathbb{I})\vec{x} = \vec{0}$ for all $\vec{x}$. So $T^\dagger T - \mathbb{I} = 0$ whence

$$T^\dagger T = \mathbb{I} = T T^\dagger. \tag{VII.C.8}$$

VII.C.9. EXAMPLE. Let V be the real vector space of all continuously differentiable (real-valued) functions on $\mathbb{R}$ which are periodic of period 1: i.e. $f(t) = f(t+1)$ for all t. (This is infinite-dimensional, but no matter.) Then[16]

$$\langle f, g \rangle = \int_0^1 f(t)\, g(t)\, dt$$

defines an inner product: it is clearly bilinear, and for any nonzero function f we have $\langle f, f \rangle = \int_0^1 f(t)^2 dt > 0$. Now define a transformation $T\colon V \to V$ by $T(f) = 3f - f'$ (i.e., $T = 3 - \frac{d}{dt}$). We use *integration by parts* to find the adjoint:

$$\langle f, Tg \rangle = \int_0^1 f\,(3g - g')dt \;=\; 3\int_0^1 f\,g\,dt \;-\; \int_0^1 f\,g'\,dt$$

$$= 3\int_0^1 f\,g\,dt \;+\; \int_0^1 f'\,g\,dt \;=\; \int_0^1 (3f + f')g\,dt \;=\; \left\langle T^{\dagger} f,\, g \right\rangle;$$

that is, $T^{\dagger} = 3 + \frac{d}{dt}$. (There was no $f(1)g(1) - f(0)g(0)$ term in the $\int$-by-parts because this is zero by the periodicity hypothesis!)

Exercises

(1) Let $V = \mathbb{R}^2$, and specify an inner product $\langle \cdot, \cdot \rangle$ by

$$\|\vec{x}\|^2 := (x_1 - x_2)^2 + 3x_2^2.$$

(Here $\vec{x} = x_1\hat{e}_1 + x_2\hat{e}_2$.) Find the matrix, with respect to $\hat{e}$, of the orthogonal projection (orthogonal with respect to $\langle \cdot, \cdot \rangle$, not the dot product) onto the line generated by $3\hat{e}_1 + 4\hat{e}_2$. Show that this projection is its own adjoint.

(2) Check that $T^{\dagger}\vec{x} := \sum_i \overline{\langle \vec{x}, T\vec{v}_i \rangle}\vec{v}_i$ satisfies equation (VII.C.6) as claimed in the notes. (Here $\{\vec{v}_i\}_{i=1}^n$ is a unitary/orthonormal basis.)

(3) In Exercise VII.B.2, you found an orthonormal basis $\mathcal{B}$ of $\mathcal{P}_2(\mathbb{R})$ under $\langle f, g \rangle := \int_0^1 f(t)g(t)dt$, and computed the matrix $[T]_{\mathcal{B}}$, where $(Tf)(t) := f(t-1)$.

[16] If we considered complex-valued functions, then $\langle f, g \rangle$ would be $\int_0^1 \overline{f(t)}g(t)\,dt$.

(a) Find $[T^\dagger]_{\mathcal{B}}$, and compute $T^\dagger(1)$. Check your answer by "plugging it in" to the definition for adjoint. It should be a quadratic polynomial with rather nasty coefficients.
(b) What is the adjoint of $S := \frac{d}{dt}$?

(4) (a) For $V = M_n(\mathbb{C})$, check that $\langle A, B \rangle := \mathrm{tr}(AB^*)$ defines an inner product.
(b) Now let P be a fixed invertible matrix in V, and define $T_P : V \to V$ by $T_P(A) := P^{-1}AP$. What is $T_P^\dagger$?

(5) By definition, unitary/orthogonal transformations preserve $\| \cdot \|$; show that this implies they preserve $\langle \cdot, \cdot \rangle$. [Hint: in the orthogonal (real) case you can write $\langle \vec{x}, \vec{y} \rangle = \frac{1}{2}\langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle - \frac{1}{2}\langle \vec{x}, \vec{x} \rangle - \frac{1}{2}\langle \vec{y}, \vec{y} \rangle$. What do you have to do differently in the unitary (complex) case?]