Hermitian inner products.

Suppose V is vector space over $\mathbf{C}$ and
$$(\cdot,\cdot)$$
is a **Hermitian inner product on** V. This means, by definition, that
$$(\cdot,\cdot) : V \times V \to \mathbf{C}$$
and that the following four conditions hold:

(i) $(v_1 + v_2, w) = (v_1, w) + (v_2, w)$ whenever $v_1, v_2, w \in V$;

(ii) $(cv, w) = c(v, w)$ whenever $c \in \mathbf{C}$ and $v, w \in V$;

(iii) $(w, v) = \overline{(v, w)}$ whenever $v, w \in V$;

(iv) (v, v) is a positive real number for any $v \in V \sim \{0\}$.

These conditions imply that

(v) $(v, w_1 + w_2) = (v, w_1) + (v, w_2)$ whenever $v, w_1, w_2 \in V$;

(vi) $(v, cw) = \overline{c}(v, w)$ whenever $c \in \mathbf{C}$ and $v, w \in V$;

(vii) $(0, v) = 0 = (v, 0)$ for any $v \in V$.

In view of (iv) and (vii) we may set
$$||v|| = \sqrt{(v, v)} \quad \text{for } v \in V$$
and note that

(viii) $||v|| = 0 \;\Leftrightarrow\; v = 0$.

We call $||v||$ the **norm of** v. Note that

(ix) $||cv|| = |c|||v||$ whenever $c \in \mathbf{C}$ and $v \in V$.

Suppose
$$A : V \times V \to \mathbf{R} \quad \text{and} \quad B : V \times V \to \mathbf{R}$$
are such that

(1) $$(v, w) = A(v, w) + iB(v, w) \quad \text{whenever } v, w \in V.$$

One easily verifies that

(i) A and B are bilinear over $\mathbf{R}$;

(ii) A is symmetric and positive definite;

(iii) B is antisymmetric;

(iv) $A(iv, iw) = A(v, w)$ whenever $v, w \in V$;

(v) $B(v, w) = -A(iv, w)$ whenever $v, w \in V$.

Conversely, given $A : V \times V \to \mathbf{R}$ which is bilinear over $\mathbf{R}$ and which is positive definite symmetric, letting B be as in (v) and let $(\cdot,\cdot)$ be as in (1) we find that $(\cdot,\cdot)$ is a Hermitian inner product on V. The interested reader might write down conditions on B which allow one to construct A and $(\cdot,\cdot)$ as well.

Example One. Let

$$(z, w) = \sum_{j=1}^{n} z_j \overline{w_j} \quad \text{for } z, w \in \mathbf{C}^n.$$

The $(\cdot,\cdot)$ is easily seen to be a Hermitian inner product, called the **standard (Hermitian) inner product**, on $\mathbf{C}^n$.

Example Two. Suppose $-\infty < a < b < \infty$ and $\mathcal{H}$ is the vector space of complex valued square integrable functions on $[a, b]$. You may object that I haven't told you what "square integrable" means. Now I will. Sort of. To say $f : [a, b] \to \mathbf{R}$ is **square integrable** means that f is Lebesgue measurable and that

$$\int_a^b |f(x)|^2 \, dx < \infty;$$

of course I haven't told you what "Lebesgue measurable" means and I haven't told you what $\int_a^b$ means, but I will in the very near future. For the time being just think of whatever notion of integration you're familiar with.

Note that

$$\int_a^b f(x)\,dx = \int_a^b \Re f(x)\,dx + i \int_a^b \Im f(x)\,dx$$

whenever $f \in \mathcal{H}$.

Let

$$(f, g) = \int_a^b f(x)\overline{g(x)}\,dx \quad \text{whenever } f, g \in \mathcal{H}.$$

You should object at this point that the integral may not exist. We will show shortly that it does. One easily verifies that (i)-(iii) of the properties of an inner product hold and that (iv) *almost* holds in the sense that for any $f \in \mathcal{F}$ we have

$$(f, f) = \int_a^b |f(x)|^2\,dx \geq 0$$

with equality only if $\{x \in [a, b] : f(x) = 0\}$ has zero Lebesgue measure (whatever that means). In particular, if f is continuous and $(f, f) = 0$ then $f(x) = 0$ for all $x \in [a, b]$.

This Example is like Example One in that one can think of $f \in \mathcal{H}$ as a an infinite-tuple with the continuous index $x \in [a, b]$.

Henceforth V is a Hermitian inner product space.

The following simple Proposition is indispensable.

Proposition. Suppose $v, w \in V$. Then

$$||v + w||^2 = ||v||^2 + 2\Re(v, w) + ||w||^2.$$

Proof. We have

$$\begin{aligned} ||v + w||^2 &= (v + w, v + w) \\ &= (v, v) + (v, w) + (w, v) + (w, w) \\ &= (v, v) + (v, w) + \overline{(v, w)} + (w, w) \\ &= ||v||^2 + 2\Re(v, w) + ||w||^2. \end{aligned}$$

□

Corollary. The Parallelogram Law. We have

$$||v+w||^2 + ||v-w||^2 = 2\left(||v||^2 + ||w||^2\right).$$

Proof. Look at it. $\square$

Here is an absolutely fundamental consequence of the Parallelogram Law.

Theorem. Suppose V is complete with respect to $||\cdot||$ and C is a nonempty closed convex subset of V. Then there is a unique point $c \in C$ such that

$$||c|| \le ||v|| \quad \text{whenever } v \in C.$$

Remark. Draw a picture.

Proof. Let

$$d = \inf\{||v|| : v \in C\}$$

and let

$$\mathcal{C} = \{C \cap \mathbf{B}_0(r) : d < r < \infty\}.$$

Note that $\mathcal{C}$ is a nonempty nested family of nonempty closed subsets of V.

Suppose $C \in \mathcal{C}$, $d < r < \infty$ and $v, w \in C$. Because C is convex we have $\frac{1}{2}(v+w) \in C \cap \mathbf{B}_0(R)$ so

$$\frac{1}{4}||v+w||^2 = ||\frac{1}{2}(v+w)||^2 \ge d^2.$$

Thus, by the Parallelogram Law,

$$\frac{1}{4}||v-w||^2 = \frac{1}{2}\left(||v||^2 + ||w||^2\right) - \frac{1}{4}||v+w||^2 \le r^2 - d^2.$$

It follows that

$$\inf\{\mathbf{diam}\, C \cap \mathbf{B}_0(r) : d < r < \infty\} = 0.$$

By completeness there is a point $c \in V$ such that

$$\{c\} = \cap \mathcal{C}.$$

$\square$

Corollary. Suppose U is a closed linear subspace of V and $v \in V$. Then there is a unique $u \in U$ such that

$$||v-u|| \le ||v-u'|| \quad \text{whenever } u' \in U.$$

Remark. Draw a picture.

Remark. We will show very shortly that any finite dimensional subspace of V is closed.

Proof. Let $C = v - U$ and note that C is a nonempty closed convex subset of V. (Of course $-U = U$ since U is a linear subspace of U, but this representation of C is more convenient for our purposes.) By virtue of the preceding Theorem there is a unique $u \in U$ such that

$$||v-u|| \le ||v-u'|| \quad \text{whenever } u' \in U.$$

$\square$

The Cauchy-Schwartz Inequality. Suppose $v, w \in V$. Then

$$|(v,w)| \le ||v||\,||w||$$

with equality only if $\{v, w\}$ is dependent.

Proof. If $w = 0$ the assertion holds trivially so let us suppose $w \neq 0$. For any $c \in \mathbf{C}$ we have

$$0 \leq ||v + cw||^2 = ||v||^2 + 2\Re(v, cw) + ||cw||^2 = ||v||^2 + 2\Re(\bar{c}(v, w)) + |c|^2||w||^2.$$

Letting

$$c = -\frac{(v, w)}{||w||^2}$$

we find that

$$0 \leq ||v||^2 - \frac{|(v, w)|^2}{||w||^2}$$

with equality only if $||v + cw|| = 0$ in which case $v + cw = 0$ so $v = -cw$. □

Corollary. Suppose a and b are sequences of complex numbers. Then

$$\sum_{n=0}^{\infty} |a_n b_n| \leq \left(\sum_{n=0}^{\infty} |a_n|^2 \right)^{1/2} \left(\sum_{n=0}^{\infty} |b_n|^2 \right)^{1/2}.$$

Proof. For any nonnegative integer N apply the Cauchy-Schwartz inequality with $(\cdot, \cdot)$ equal the standard inner product on $\mathbf{C}^N$,

$$v = (a_0, \ldots, a_N) \quad \text{and} \quad w = (b_0, \ldots, b_N)$$

and then let $N \to \infty$. □

The Triangle Inequality. Suppose $v, w \in V$. Then

$$||v + w|| \leq ||v|| + ||w||$$

with equality only if either v is a nonnegative multiple of w or w is a nonnegative multiple of v.

Proof. Using the Cauchy-Schwartz Inequality we find that

$$||v + w||^2 = ||v||^2 + 2\Re(v, w) + ||w||^2 \leq ||v||^2 + 2||v||||w|| + ||w||^2 = (||v|| + ||w||)^2.$$

Suppose equality holds. In case $v = 0$ then $v = 0w$ so suppose $v \neq 0$. Since $|(v, w)| \geq \Re(v, w) = ||v||||w||$ we infer from the Cauchy-Schwartz Inequality that $w = cv$ for some $c \in \mathbf{C}$. Thus

$$|1 + c|||v|| = ||(1 + c)v)|| = ||v + cw|| = ||v|| + ||cw|| = (1 + |c|)||v||$$

from which we infer that

$$1 + 2\Re c + |c|^2 = |1 + c|^2 = (1 + |c|)^2 = 1 + 2|c| + |c|^2$$

which implies that c is a nonnegative real number. □

Definition. Suppose U is a linear subspace of V. We let

$$U^\perp = \{v \in V : (u, v) = 0 \text{ for all } u \in U\}$$

and note that $U^\perp$ is a linear subspace of V. It follows directly from (iv) that

$$U \cap U^\perp = \{0\}.$$

Proposition. Suppose U is a linear subspace of V. Then

$$U \subset U^{\perp\perp}$$

and $U^\perp$ is closed.
Proof. The first assertion is an immediate consequence of the definition of $U^\perp$. The second follows because $U^\perp$ is the intersection of the closed sets

$$\{v \in V : (u, v) = 0\}$$

corresponding to $u \in U$; These sets are closed because $V \ni v \mapsto (u, v)$ is continuous by virtue of the Cauchy-Schwartz Inequality. □

Orthogonal projections.

Henceforth U is closed linear subspace of V.

Definition. Keeping in mind the foregoing, we define

$$P : V \to U$$

by requiring that

$$||v - Pv|| \leq ||v - u'|| \quad \text{whenever } u' \in U.$$

That is, Pv is the closest point in U to v. We call P **orthogonal projection of V onto U.** Note that $Pu = u$ whenever $u \in U$. Thus

$$\mathbf{rng}\, P = U \quad \text{and} \quad P \circ P = P.$$

Keeping in mind that $U^\perp$ is a closed linear subspace of V we let

$$P^\perp$$

be orthogonal projection of V onto $U^\perp$.

Theorem. Suppose W is a linear subspace of V and

$$Q : V \to W$$

is such that

$$||w - Qv|| \leq ||v - w|| \quad \text{whenever } v \in V \text{ and } w \in W.$$

Then W is closed and Q is orthogonal projection of V onto W.
Proof. Suppose $\tilde{w} \in \mathbf{cl}\, W$ and $\epsilon > 0$. Choose $w \in W$ such that $||\tilde{w} - w|| \leq \epsilon$. Then

$$||\tilde{w} - Q\tilde{w}|| \leq ||\tilde{w} - w|| \leq \epsilon.$$

Owing to the arbitrariness of ϵ we infer that $||Q\tilde{w} - w|| = 0$ so $w = Q\tilde{w} \in W$ and $\mathbf{cl}\, W \subset W$. □

Theorem. We have

$$u = Pv \;\Leftrightarrow\; v - u \in U^\perp \quad \text{whenever } u \in U \text{ and } v \in V.$$

Proof. Suppose $u \in U$ and $v \in V$. For each $(t, u') \in \mathbf{R} \times U$ let

$$f(t, u') = ||(v - u) + tu'||^2$$

and note that

$$f(t, u') = ||v - u||^2 + 2t\Re(v - u, u') + t^2||u'||^2.$$

Suppose $u = Pv$. Then $f(0, u') \leq f(t, u')$ whenever $(t, u') \in \mathbf{R} \times U$. Thus $v - u \in U^\perp$.

Suppose $v - u \in U^\perp$. Then

$$||v-u||^2 = f(0, u'-u) \le f(1, u'-u) = ||v-u'||^2$$

so $u = Pv$. $\square$

Corollary. P is linear.
Proof. Suppose $v \in V$ and $c \in \mathbf{C}$. Then $cPv \in U$ and $cv - cPv = c(v - Pv) \in U^\perp$ so $P(cv) = cPv$. Suppose $v_1, v_2 \in V$. then $Pv_1 + Pv_2 \in U$ and $(v_1+v_2) - (Pv_1 + Pv_2) = (v_1 - Pv_1) + (v_2 - Pv_2) \in U^\perp$ so $P(v_1 + v_2) = Pv_1 + Pv_2$. $\square$

Corollary. Suppose $v \in V$. Then

(i) $v = Pv + P^\perp v$ and

(ii) $||v||^2 = ||Pv||^2 + ||P^\perp v||^2$.

Proof. We have $v - Pv \in U^\perp$ by the preceding Theorem and

$$v - (v - Pv) = Pv \in U \subset U^{\perp\perp}$$

so, again by the preceding Theorem only with U replaced by $U^\perp$ we find that $P^\perp v = v - Pv$. It follows that

$$||v||^2 = ||Pv + P^\perp v||^2 = ||Pv||^2 + 2\Re(Pv, P^\perp v) + ||P^\perp v||^2 = ||Pv||^2 + ||P^\perp v||^2.$$

$\square$

Corollary. We have

$$U^{\perp\perp} = U$$

and

$$(Pv, w) = (v, Pw) \quad \text{whenever } v, w \in V.$$

Proof. Let P and $P^\perp$ be orthogonal projection of V onto U and $U^\perp$, respectively. By the preceding Theorem with U replaced by $U^\perp$ we find that orthogonal projection of V onto $U^{\perp\perp}$ carries $v \in V$ to $v - P^\perp v = Pv$. Thus $U = U^{\perp\perp}$.

Suppose $v, w \in V$. Then

$$(Pv, w) = (Pv, Pw + P^\perp w) = (Pv, Pw) = (Pv + P^\perp v, Pw) = (v, Pw).$$

$\square$

Definition. We say a subset A of V is **orthonormal** if whenever $v, w \in A$ we have

$$(v, w) = \begin{cases} 1 & \text{if } v = w; \\ 0 & \text{if } v \ne w. \end{cases}$$

Exercise. Show that any orthonormal set is independent.

The Gram-Schmidt Process. Suppose $\tilde{u} \in V \sim U$, $\tilde{U} = \{u + c\tilde{u} : c \in \mathbf{C}\}$ and

$$\tilde{P}v = Pv + \frac{(v, P^\perp \tilde{u})}{||P^\perp \tilde{u}||^2} P^\perp \tilde{u} \quad \text{whenever } v \in V.$$

Then $\tilde{U}$ is closed and $\tilde{P}$ is orthogonal projection on $\tilde{U}$.
Proof. Easy exercise for the reader. $\square$

Remark. If $U = \{0\}$ then $P = 0$ so

$$\tilde{P}(v) = \frac{(v, \tilde{u})}{||\tilde{u}||^2}\tilde{u}$$

and $\tilde{P}$ is orthogonal projection on the line $\{c\tilde{u} : c \in \mathbf{C}\}$.

Corollary. Any finite dimensional subspace of V is closed and has an orthonormal basis.
Proof. Induct on the dimension of the subspace and use the Gram-Schmidt Process to carry out the inductive step. □

Proposition. Suppose U is finite dimensional and B is an orthnormal basis for U. Then

$$Pv = \sum_{u \in B} (v, u)u \quad \text{and} \quad ||Pv||^2 = \sum_{u \in B} |(v, u)|^2 \quad \text{whenever } v \in V.$$

Proof. Let

$$Lv = \sum_{u \in B} (v, u)u \quad \text{for } v \in V.$$

Suppose $v \in V$ and $\tilde{u} \in B$. The

$$\begin{aligned}(v - Lv, \tilde{u}) &= (v - \sum_{u \in B} (v, u)u, \tilde{u}) \\ &= (v, \tilde{u}) - \sum_{u \in B} (v, u)(u, \tilde{u}) \\ &= (v, \tilde{u}) - (v, \tilde{u}) \\ &= 0\end{aligned}$$

which, as B is a basis for U, implies that $v - Lv \in U^{\perp}$; thus $P = L$.

Finally, if $v \in V$ we have

$$\begin{aligned}||Lv||^2 &= (\sum_{u \in B} (v, u)u, \sum_{\tilde{u} \in B} (v, \tilde{u})\tilde{u}) \\ &= \sum_{u \in B,\ \tilde{u} \in B} (v, u)\overline{(v, \tilde{u})}(u, \tilde{u}) \\ &= \sum_{u \in B} |(u, v)|^2.\end{aligned}$$

□