Isometries

If the important functions between vector spaces $V$ and $W$ are those that preserve linearity (i.e., linear transformations), then the important functions between inner product spaces should be those that preserve the inner product. To this end, we make the following natural definition: a linear transformation $T: V \to W$ between inner product spaces is isometric if it satisfies the condition

$$\langle T(X), T(Y) \rangle = \langle X, Y \rangle$$

for every $X$ and $Y$ in $V$. (Note that the inner product on the left is the one in $W$ while the one on the right is the one in $V$. Also, the term “isometric” refers to having “the same distance”: an isometric map preserves lengths of vectors as it passes from domain to codomain space. Of course, it also preserves angles.)

**Proposition** An isometric linear transformation has trivial kernel.

**Proof** Suppose $T(X) = 0$. Then $|X|^2 = \langle X, X \rangle = \langle T(X), T(X) \rangle = |T(X)|^2 = |0|^2 = 0$, so $X = 0$. //
This last proposition shows that all isometric maps are one-to-one. (Smith calls such maps isometric embeddings.) So if the linear transformation $T: \mathcal{V} \to \mathcal{W}$ is isometric and $\mathcal{V}$ and $\mathcal{W}$ have the same dimension, the map must be an isomorphism as well; we call an isometric isomorphism simply an isometry, and we say that the spaces $\mathcal{V}$ and $\mathcal{W}$ are isometric.

**Proposition** Any two inner product spaces of the same dimension are isometric.

**Proof** Suppose inner product spaces $\mathcal{V}$ and $\mathcal{W}$ are both of dimension $n$. Then we can find orthonormal bases for each, say $\{V_1, V_2, \ldots, V_n\}$ for $\mathcal{V}$ and $\{W_1, W_2, \ldots, W_n\}$ for $\mathcal{W}$. Let $T: \mathcal{V} \to \mathcal{W}$ be the linear transformation that is the extension of the assignments $T(V_i) = W_i$. Then $T$ is isometric, for any $X$ and $Y$ in $\mathcal{V}$ can be written the form

$$X = a_1 V_1 + a_2 V_2 + \cdots + a_n V_n$$

$$Y = b_1 V_1 + b_2 V_2 + \cdots + b_n V_n$$

so that
\[
\langle T(X), T(Y) \rangle = \left\langle T\left( \sum_{1 \leq i \leq n} a_i V_i \right), T\left( \sum_{1 \leq j \leq n} b_j V_j \right) \right\rangle \\
= \left\langle \sum_{1 \leq i \leq n} a_i T(V_i), \sum_{1 \leq j \leq n} b_j T(V_j) \right\rangle \\
= \sum_{1 \leq i, j \leq n} a_i b_j \langle T(V_i), T(V_j) \rangle \\
= \sum_{1 \leq i, j \leq n} a_i b_j \langle V_i, V_j \rangle \\
= \left\langle \sum_{1 \leq i \leq n} a_i V_i, \sum_{1 \leq j \leq n} b_j V_j \right\rangle \\
= \langle X, Y \rangle . \quad //
\]

**Corollary** Every \( n \)-dimensional inner product space is isomorphic to \( \mathbb{R}^n \) (endowed with the standard dot product).  //
Functionals

Recall from Exercise 9.9(3) that any finite-dimensional vector space \( \mathcal{V} \) is isomorphic to its dual space \( \mathcal{V}^* = \mathcal{L}(\mathcal{V}, \mathbb{R}) \). We will investigate next how having a scalar product helps to make an important connection between the inner product space \( \mathcal{V} \) and its dual \( \mathcal{V}^* \).

The vectors in \( \mathcal{V}^* \) are real-valued linear functions on \( \mathcal{V} \). They are often called (linear) functionals on \( \mathcal{V} \); we will denote them with letters like \( f, g \), etc., to emphasize their nature as functions on \( \mathcal{V} \).

One example of a functional makes use of the inner product on \( \mathcal{V} \). Let \( \mathbf{A} \) be a fixed vector in \( \mathcal{V} \) and define for arbitrary vectors \( \mathbf{X} \) in \( \mathcal{V} \), \( f(\mathbf{X}) = \langle \mathbf{X}, \mathbf{A} \rangle \); then \( f \) is a functional (why is it linear?). We will see that every functional on \( \mathcal{V} \) can be represented this way. For instance, consider the functional \( f: \mathbb{R}^3 \rightarrow \mathbb{R} \) given by \( f(x, y, z) = 3x - 2y + z \); we represent it in the above form by writing \( f(x, y, z) = (x, y, z) \cdot (3, -2, 1) \), i.e., \( f \) corresponds to the dot product with the fixed vector \( \mathbf{A} = (3, -2, 1) \).
Observe that we have a geometric interpretation for this as well: the kernel of the functional $f$ corresponds to the plane $3x - 2y + z = 0$ in $\mathbb{R}^3$, and $A$ is a normal vector to this plane (every vector in the kernel of $f$ is orthogonal to $A$). We now propose to work out the details of how this situation generalizes in abstract inner product spaces.

If $S$ is any set of vectors in an inner product space $\mathcal{V}$, we define the **orthogonal complement of** $S$ to be the set $S^\perp$ of vectors in $\mathcal{V}$ orthogonal to all vectors in $S$: $S^\perp = \{ X \in \mathcal{V} \mid \langle X, Y \rangle = 0 \text{ for all } Y \in S \}$.

**Proposition** If $S$ is any set of vectors in an inner product space $\mathcal{V}$, then $S^\perp$ is a subspace of $\mathcal{V}$. Further, only $0$ can lie in both $S$ and $S^\perp$.

**Proof** If $X$ and $X'$ lie in $S^\perp$, then $\langle X + X', Y \rangle = \langle X, Y \rangle + \langle X', Y \rangle = 0$ and $\langle rX, Y \rangle = r\langle X, Y \rangle = 0$ for all $Y \in S^\perp$. So $S^\perp$ is a subspace of $\mathcal{V}$. Also, if $X$ lies in both $S$ and $S^\perp$, then $\langle X, X \rangle = 0$, so $|X| = 0$, whence $X = 0$. //
**Corollary** Let $\mathcal{W}$ be a subspace of the inner product space $\mathcal{V}$. Then $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$.

**Proof** By the proposition, $\mathcal{W}^\perp$ is also subspace of $\mathcal{V}$ and their intersection is trivial. To prove the result, we need only show that $\mathcal{V} = \mathcal{W} + \mathcal{W}^\perp$. So let $\{\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_k\}$ be an orthonormal basis for $\mathcal{W}$ and suppose that $\mathbf{V}$ is any vector in $\mathcal{V}$. Let $\mathbf{X} = \langle \mathbf{V}, \mathbf{W}_1 \rangle \mathbf{W}_1 + \cdots + \langle \mathbf{V}, \mathbf{W}_k \rangle \mathbf{W}_k$ and set $\mathbf{Y} = \mathbf{V} - \mathbf{X}$. Then necessarily $\mathbf{V} = \mathbf{X} + \mathbf{Y}$ with $\mathbf{X}$ in $\mathcal{W}$. But since any $\mathbf{W}$ vector in $\mathcal{W}$ has the form

$$\mathbf{W} = a_1 \mathbf{W}_1 + \cdots + a_k \mathbf{W}_k$$

for suitable scalars $a_1, \ldots, a_k$, we can compute

$$\langle \mathbf{Y}, \mathbf{W} \rangle = \langle \mathbf{Y}, a_1 \mathbf{W}_1 + \cdots + a_k \mathbf{W}_k \rangle$$

$$= a_1 \langle \mathbf{Y}, \mathbf{W}_1 \rangle + \cdots + a_k \langle \mathbf{Y}, \mathbf{W}_k \rangle$$

$$= a_1 [\langle \mathbf{V}, \mathbf{W}_1 \rangle - \langle \mathbf{X}, \mathbf{W}_1 \rangle] + \cdots + a_k [\langle \mathbf{V}, \mathbf{W}_k \rangle - \langle \mathbf{X}, \mathbf{W}_k \rangle]$$

and since
\[ \langle X, W_i \rangle = \langle \langle V, W_1 \rangle W_1 + \cdots + \langle V, W_k \rangle W_k, W_i \rangle \]
\[ = \langle V, W_1 \rangle \langle W_1, W_i \rangle + \cdots + \langle V, W_k \rangle \langle W_k, W_i \rangle \]
\[ = \langle V, W_i \rangle \langle W_i, W_i \rangle \]
\[ = \langle V, W_i \rangle \]

it follows that \( \langle Y, W \rangle = 0 \) for every \( W \) in \( \mathcal{W} \). That is, \( Y \) lies in \( \mathcal{W}^\perp \). So \( \mathcal{V} = \mathcal{W} + \mathcal{W}^\perp \), completing the proof.  

**Corollary** Let \( \mathcal{W} \) be a \( k \)-dimensional subspace of the \( n \)-dimensional inner product space \( \mathcal{V} \). Then \( \mathcal{W}^\perp \) has dimension \( n - k \).  

The results we have just worked out provide the technical details that allow us to prove the central theorem of this discussion:
Theorem [The Riesz Representation Theorem]
Let $\mathcal{V}$ be an inner product space. Then for every functional $f \in \mathcal{V}^*$, there exists a corresponding unique vector $A \in \mathcal{V}$ with the property that

$$f(X) = \langle X, A \rangle.$$ 

That is, every functional can be represented as an inner product against some vector in $\mathcal{V}$.

Proof If $f$ is the zero map, then we can take $A = 0$ and we’re done. So from here on, we may assume that $f$ is not always zero on $\mathcal{V}$. Since $\text{Im}(f)$ is not trivial and is also a subspace of $\mathbb{R}$, it follows that $\text{Im}(f) = \mathbb{R}$. Thus, $\dim(\text{Im}(f)) = 1$, so by the Dimension Theorem, $\dim(\ker(f)) = n - 1$ where $\dim \mathcal{V} = n$.

Let $\{V_1, V_2, \ldots, V_{n-1}\}$ be an orthonormal basis for $\ker(f)$. Since the subspace $\ker(f)^\perp$ must satisfy $\mathcal{V} = \ker(f) \oplus \ker(f)^\perp$, it follows that $\ker(f)^\perp$ is one-dimensional. If $V_0$ is a unit vector in $\ker(f)^\perp$, then $\{V_0, V_1, V_2, \ldots, V_{n-1}\}$ must be an orthonormal basis for all of $\mathcal{V}$. 
So if \( X \) is any vector in \( V \), we can write

\[
X = \langle X, V_0 \rangle V_0 + \langle X, V_1 \rangle V_1 + \cdots + \langle X, V_{n-1} \rangle V_{n-1}
\]

so that

\[
f(X) = \langle X, V_0 \rangle f(V_0) + \langle X, V_1 \rangle f(V_1) + \cdots + \langle X, V_{n-1} \rangle f(V_{n-1})
\]

\[
= \langle X, V_0 \rangle f(V_0)
\]

\[
= \langle X, f(V_0) V_0 \rangle
\]

It follows that the vector \( A = f(V_0) V_0 \) satisfies the condition desired: \( f(X) = \langle X, A \rangle \).

The choice of \( A \) is unique, for if \( B \) is another vector with the property that \( f(X) = \langle X, B \rangle \), then \( \langle X, A \rangle = \langle X, B \rangle \) for all \( X \) in \( V \), so in particular,

\[
\langle A - B, A \rangle = \langle A - B, B \rangle \Rightarrow \langle A - B, A - B \rangle = 0
\]

\[
\Rightarrow |A - B|^2 = 0
\]

\[
\Rightarrow A - B = 0
\]

\[
\Rightarrow A = B
\]
The Riesz Representation Theorem, applied to the inner product space $\mathbb{R}^3$ (under the standard dot product), guarantees that for every linear map (functional) $f: \mathbb{R}^3 \to \mathbb{R}$ there is an associated vector $A = (a, b, c)$ so that $f$ has the form

$$f(x, y, z) = (x, y, z) \cdot (a, b, c) = ax + by + cz.$$ 

In fact, as the proof of the theorem indicates, $A$ determines a 1-dimensional subspace of $\mathbb{R}^3$ orthogonal to the 2-dimensional kernel of $f$, namely the plane with equation

$$(x, y, z) \cdot (a, b, c) = ax + by + cz = 0;$$

that is, $A$ is a vector normal to the plane.

While none of this is new information, we can produce some new results by looking at how the theorem applies in other inner product spaces.

Consider the inner product space $P_k(\mathbb{R})$ (with the integral metric $\langle p(x), q(x) \rangle = \int_{-1}^{1} p(x)q(x) \, dx$). In this setting, the Riesz Representation Theorem guarantees that to any functional, like the evaluation-at-$a$ map $f_a(p(x)) = p(a)$, there corresponds a fixed polynomial $q_a(x)$ so that
\((*)\) \[ p(a) = f_a(p(x)) = \langle p(x), q_a(x) \rangle = \int_{-1}^{1} p(x) q_a(x) \, dx. \]

That is, the polynomial \( q_a(x) \) has the remarkable property that computation of the integral \( \int_{-1}^{1} p(x) q_a(x) \, dx \) is identical to evaluating the polynomial \( p(x) \) at \( x = a \), and this is true simultaneously for all polynomials \( p(x) \) of degree up to \( k \).

For instance, with \( k = 2 \) and \( a = 0 \), we want a polynomial \( q_0(x) = a_0 + a_1 x + a_2 x^2 \) satisfying the condition \((*)\) above. Following the procedure outlined in the proof of the theorem, we first find a polynomial \( r(x) \) which is orthogonal to the kernel of the functional. The \( p(x) \) in this kernel satisfy \( f_0(p(x)) = p(0) = 0 \), so \( p(x) \) must have the form \( b_1 x + b_2 x^2 \). So \( r(x) \) satisfies

\[
0 = \left\langle b_1 x + b_2 x^2, r(x) \right\rangle = b_1 \int_{-1}^{1} x \cdot r(x) \, dx + b_2 \int_{-1}^{1} x^2 \cdot r(x) \, dx
\]

for every choice of \( b_1, b_2 \). Thus, both integrals above must vanish: if we put \( r(x) = c_0 + c_1 x + c_2 x^2 \), then

\[
0 = \int_{-1}^{1} x \cdot r(x) \, dx = c_0 \int_{-1}^{1} x \, dx + c_1 \int_{-1}^{1} x^2 \, dx + c_2 \int_{-1}^{1} x^3 \, dx
\]

\[
0 = \int_{-1}^{1} x^2 \cdot r(x) \, dx = c_0 \int_{-1}^{1} x^2 \, dx + c_1 \int_{-1}^{1} x^3 \, dx + c_2 \int_{-1}^{1} x^4 \, dx
\]
becomes

\begin{align*}
0 &= c_0 \cdot 0 + c_1 \cdot \frac{2}{3} + c_2 \cdot 0 \\
0 &= c_0 \cdot \frac{2}{3} + c_1 \cdot 0 + c_2 \cdot \frac{2}{5}
\end{align*}

whence \( c_1 = 0 \) and \( c_2 = -\frac{5}{3} c_0 \). So \( r(x) = c(1 - \frac{5}{3} x^2) \) for any nonzero choice of \( c \). The procedure requires that we normalize \( r(x) \), so we compute its length:

\[
\langle r(x), r(x) \rangle = c^2 \int_{-1}^{1} (1 - \frac{5}{3} x^2)^2 dx \\
= c^2 \int_{-1}^{1} (1 - \frac{10}{3} x^2 + \frac{25}{9} x^4) dx \\
= c^2 (2 - \frac{20}{9} + \frac{10}{9}) \\
= \frac{8}{9} c^2
\]

To make this length equal 1, we force \( c^2 = \frac{9}{8} \).

Finally, according to the procedure, we put

\[
q_0(x) = f_0(r(x)) \cdot r(x) = r(0) \cdot r(x) = c \cdot c(1 - \frac{5}{3} x^2) = \frac{9}{8} - \frac{15}{8} x^2
\]

It follows that every polynomial \( p(x) \) in \( P_2(\mathbb{R}) \) satisfies the formula

\[
p(0) = \int_{-1}^{1} p(x) (\frac{9}{8} - \frac{15}{8} x^2) dx.
\]