3.3.4 Corollary
If $M = \{x_1, x_2, \ldots, x_n\}$ is an orthogonal set in an inner product space $(X, \langle\cdot,\cdot\rangle)$ then

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2.$$

Proof. Exercise. ■

3.4 Best Approximation in Hilbert Spaces

3.4.1 Definition
Let K be a closed subset of an inner product space $(X, \langle\cdot,\cdot\rangle)$. For a given $x \in X \setminus K$, a **best approximation** *or* **nearest point** *to x from K is any element $y_0 \in K$ such that*

$$\|x - y_0\| \leq \|x - y\| \quad \text{for all} \quad y \in K.$$

Equivalently, $y_0 \in K$ is a best approximation to x from K if

$$\|x - y_0\| = \inf_{y \in K} \|x - y\| = d(x, K).$$

The (possibly empty) set of all best approximations to x from K is denoted by $P_K(x)$. That is,

$$P_K(x) = \{y \in K : \|x - y\| = d(x, K)\}.$$

The (generally set-valued) map P_K which associates each x in X with its best approximations in K is called the **metric projection** *or the* **nearest point map**. *The set K is called*

[1] **proximinal** *if each $x \in X$ has a best approximation in K; i.e., $P_K(x) \neq \emptyset$ for each $x \in X$;*

[2] **Chebyshev** *if each $x \in X$ has a unique best approximation in K; i.e., the set $P_K(x)$ consists of a single point.*

The following important result asserts that if K is a complete convex subset of an inner product space $(X, \langle\cdot,\cdot\rangle)$, then each $x \in X$ has one and only one element of best approximation in K.

3.4.1 Theorem
Every nonempty complete convex subset K of an inner product space $(X, \langle\cdot,\cdot\rangle)$ is a Chebyshev set.

Proof. Existence: Without loss of generality, $x \in X \setminus K$. Let

$$\delta = \inf_{y \in K} \|x - y\|.$$

By definition of the infimum, there exists a sequence $(y_n)_1^\infty$ in K such that

$$\|x - y_n\| \to \delta \quad \text{as} \quad n \to \infty.$$

We show that $(y_n)_1^\infty$ is a Cauchy sequence. By the Parallelogram Identity (Theorem 3.1.3),

$$\begin{aligned}
\|y_m - y_n\|^2 &= \|(x - y_n) - (x - y_m)\|^2 \\
&= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - (y_n + y_m)\|^2 \\
&= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \left(\frac{y_n + y_m}{2}\right)\right\|^2 \\
&\leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2,
\end{aligned}$$

since $\frac{y_n + y_m}{2} \in K$ by convexity of K. Thus,

$$\|y_m - y_n\|^2 \leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2 \to 0 \quad \text{as} \quad n, m \to \infty.$$

That is, $(y_n)_1^\infty$ is a Cauchy sequence in K. Since K is complete, there exists $y \in K$ such that $y_n \to y$ as $n \to \infty$. Since the norm is continuous,

$$\|x - y\| = \|x - \lim_{n\to\infty} y_n\| = \| \lim_{n\to\infty} (x - y_n)\| = \lim_{n\to\infty} \|x - y_n\| = \delta.$$

Thus,

$$\|x - y\| = \delta = d(x, K).$$

Uniqueness: Assume that $y,\ y_0 \in K$ are two best approximations to x from K. That is,

$$\|x - y_0\| = \|x - y\| = \delta = d(x, K).$$

By the Parallelogram Identity,

$$\begin{aligned} 0 \leq \|y - y_0\|^2 &= \|(y - x) + (x - y_0)\|^2 \\ &= 2\|x - y\|^2 + 2\|x - y_0\|^2 - \|2x - (y + y_0)\|^2 \\ &= 2\delta^2 + 2\delta^2 - 4\left\|x - \left(\frac{y + y_0}{2}\right)\right\|^2 \\ &\leq 4\delta^2 - 4\delta^2 = 0. \end{aligned}$$

Thus, $y_0 = y$. ■

3.4.2 Corollary
Every nonempty closed convex subset of a Hilbert space is Chebyshev.

The following theorem characterizes best approximations from a closed convex subset of a Hilbert space.

3.4.2 Theorem
Let K be a nonempty closed convex subset of a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$, $x \in \mathcal{H} \setminus K$ and $y_0 \in K$. Then y_0 is the best approximation to x from K if and only if

$$\Re\mathfrak{e}\langle x - y_0, y - y_0\rangle \leq 0 \ \text{ for all } \ y \in K.$$

Proof. The existence and uniqueness of the best approximation to x in K are guaranteed by Theorem 3.4.1. Let y_0 be the best approximation to x in K. Then, for any $y \in K$ and any $0 < \lambda < 1$, $\lambda y + (1 - \lambda)y_0 \in K$ since K is convex. Thus,

$$\begin{aligned} \|x - y_0\|^2 &\leq \|x - [\lambda y + (1 - \lambda)y_0]\|^2 = \|(x - y_0) - \lambda(y - y_0)\|^2 \\ &= \langle x - y_0) - \lambda(y - y_0), x - y_0) - \lambda(y - y_0)\rangle \\ &= \langle x - y_0, x - y_0\rangle - \lambda[\langle x - y_0, y - y_0\rangle + \langle y - y_0, x - y_0\rangle] \\ &\quad +\lambda^2\langle y - y_0, y - y_0\rangle \\ &= \|x - y_0\|^2 - 2\lambda\Re\mathfrak{e}(\langle x - y_0, y - y_0\rangle) + \lambda^2\|y - y_0\|^2 \\ \Rightarrow \quad 2\lambda\Re\mathfrak{e}(\langle x - y_0, y - y_0\rangle) &\leq \lambda^2\|y - y_0\|^2 \\ \Rightarrow \quad \Re\mathfrak{e}(\langle x - y_0, y - y_0\rangle) &\leq \frac{\lambda}{2}\|y - y_0\|^2. \end{aligned}$$