A Hilbert space analogue of Heron's reflection principle

FINBARR HOLLAND AND ANCA MUSTAŢĂ[1]

Abstract. This note is concerned with the simultaneous approximation of two vectors in a Hilbert space by an element in one of its closed subspaces. The corresponding problem in elementary plane geometry admits a short and elegant solution based on reflection, probably due to Heron. Our discussion of the Hilbert space analogue follows a similar line, displaying a one-parameter family of non-linear isometries which fix the chosen subspace, and enjoy other properties possessed by linear reflections. A natural choice of parameter then yields the required minimum.

1. Introduction

Every secondary-school student learns the technique of dropping a perpendicular from a point to a straight line, thereby establishing, via the theorem of Pythagoras, the existence of a unique point on the line that is closer to the given point than any other point on the line. This is arguably the most influential theorem to emerge from elementary Euclidean Geometry, giving, as it does, prominence to the concept of perpendicularity, which is fundamental throughout Mathematics. It is also very likely the first instance of an approximation problem that dealt with existence, uniqueness and construction of a solution all at once.

This classical result opened up the Theory of Approximation in Banach spaces, and, in particular, it has a direct analogue in Hilbert space, one version of which we recall here for convenience [2]: Suppose M is a closed subspace of a Hilbert space H, with inner-product

[1] Support from Science Foundation Ireland through the Research Frontiers Programme is gratefully acknowledged.

$< \cdot, \cdot >$, and $x \in H$. Then there is a unique point $Px \in M$ such that

$$||x - Px|| = \inf\{||x - t|| : t \in M\},$$
$$< x - Px, t >= 0, \ \forall t \in M,$$

and

$$||x||^2 = ||x - Px||^2 + ||Px||^2.$$

It turns out that P is a *projection operator* whose range is M, i.e., P is a bounded self-adjoint linear mapping on H to M with $P^2 = P$. This is a key result in Hilbert space. As is well-known, not only do many profound facts about the space flow from it, such as, for instance, a description of its dual space, and a description of the space as a direct sum of one-dimensional subspaces, it also has wide applicability.

In this note we address the possibility of simultaneously approximating two or more vectors in a Hilbert space by an element in one of its closed subspaces. Given a finite subset F of vectors in a Hilbert space H, and a closed subspace M of H, can we determine an element $m \in M$ for which the elements in $\{||m - x|| : x \in F\}$ are simultaneously small? Any meaningful answer of this will of necessity involve a measure of "smallness", and we have several such measures to choose from. One natural such measure leads to the following simple result, whose proof we leave for the reader.

Theorem 1.1. *Let M be a closed subspace of a Hilbert space H. Let $x_1, x_2, \ldots, x_n$ be distinct vectors in H. Then there is a unique vector $m \in M$ such that*

$$\sum_{k=1}^{n} ||m - x_k||^2 = \inf\Big\{\sum_{k=1}^{n} ||x - x_k||^2 : x \in M\Big\}.$$

But what's the answer if we select the ℓ_1-norm rather than the ℓ_2-norm as our measure? Is the infimum of the set

$$\Big\{\sum_{k=1}^{n} ||x - x_k|| : x \in M\Big\}$$

attained? If so, what is its value?

These questions appear to be much more complex if $n > 2$. But, fortunately for us, the case $n = 2$ has a paradigm in elementary plane geometry which led to Fermat's principle of least time in optics. A preliminary version of this principle appears to have been first mooted by Heron (or Hero of Alexandria) who is thought to have lived in the first century, between 10–70 AD. Heron is probably best known to students of mathematics for his formula for the area of a triangle in terms of its side-lengths, but he is also renowned for his ingenious inventions of, for instance, precursors of the steam engine and vending machines, and his work on surveying and optics. (A concise account of Heron and his works is given in [1]; more detail about him can be elicited from the World Wide Web.) Seemingly, he discovered his area formula when he attempted to show that the "angle of incidence" in optics is equal to the "angle of reflection". So, for these reasons, it seems fair to ascribe the following result to him.

Theorem 1.2 (Heron)**.** *Suppose a, b are two complex numbers and L is a line in the complex plane. Then*

$$\max\big(|a-b|, |a-R_L b|\big) = \inf\big\{|z-a|+|z-b| : z \in L\big\},$$

where $R_L b$ denotes the reflection of b in L. Moreover, unless $a, b \in L$, the infimum is attained at a unique point in L.

Of course, the interesting case of this theorem is when a, b are on the same side of L.

In the next section we attempt to present a direct analogue of this result in a Hilbert space setting.

2. Is there a direct analogue of Heron's theorem in Hilbert space?

Given a closed subspace M of a Hilbert space H and distinct points $a, b \in H$, is the infimum of the set

$$\{||x-a|| + ||x-b|| : x \in M\}$$

attained by some point in M? If so, what is the value of the infimum? In what circumstances, if any, is the infimum attained by a unique point in M?

Can we imitate Heron's method to settle these questions? The latter question immediately raises another: What's meant by the reflection of a point in M?

If, as above, P denotes the orthogonal projection on M, and $x \in H$, $2Px - x$ presents itself as an obvious candidate for what might be termed *the reflection Rx of x in M*. It's easy to see that R is a linear isometric involution on H that fixes every element of M. In other words, it has all the characteristic properties of what is meant in plane geometry by a reflection in a line that passes through the origin. In particular, it follows that if $a, b \in H$ and $x \in M$, then

$$||a-Rb|| \le ||a-x||+||x-Rb|| = ||a-x||+||R(x-b)|| = ||a-x||+||b-x||,$$

and, of course,

$$||a-b|| \le ||a-y|| + ||b-y||, \ \forall y \in H.$$

Hence,

$$\max\big(||a-b||, ||a-Rb||\big) \le \inf\big\{||a-x|| + ||b-x|| : x \in M\big\}.$$

However, this inequality is strict, in general, as the following simple example shows.

Example 2.1. Suppose M is the subspace spanned by the unit vector $(0,0,1) \in \mathbb{R}^3$, so that the (suggested) reflection of $x = (x_1, x_1, x_3)$ in M is given by $Rx = (-x_1, -x_2, x_3)$. Let $a = (3,1,1), b = (1,2,1)$. Then

$$\inf\big\{||m-a||+||m-b|| : m \in M\big\} = \sqrt{10}+\sqrt{5} > \max\big(||a-Rb||, ||a-b||\big).$$

Proof. Clearly,

$$\begin{aligned} & \inf\big\{||m-a|| + ||m-b|| : m \in M\big\} \\ = \ & \inf\big\{\sqrt{3^2+1^2+(t-1)^2} + \sqrt{1^2+2^2+(t-1)^2} : -\infty < t < \infty\big\} \\ = \ & \sqrt{10}+\sqrt{5}, \end{aligned}$$

whereas

$$||a-b|| = \sqrt{2^2+1^2} = \sqrt{5}, \ ||a-Rb|| = \sqrt{4^2+3^2} = 5,$$

and $\max(\sqrt{5}, 5) = 5 < \sqrt{10} + \sqrt{5}$. □

Thus, the approach adopted so far is inadequate to answer the opening question of this section. In order to obtain a complete solution we find it convenient to introduce a family of nonlinear norm-preserving operators in the next section, which may be of independent interest.

3. A one-parameter family of non-linear isometries on H

From now on, M will denote a closed subspace in a Hilbert space H, and P will stand for the orthogonal projection from H to M. Let $M^{\perp}$ stand for the orthogonal complement of M and let $Q =$

$I - P$, the orthogonal projection associated with $M^{\perp}$. With each unit vector $u \in M^{\perp}$, define R_u on H by

$$R_u x = Px - ||Qx||u, \; x \in H.$$

Note the following properties of this non-linear operator.

- $||R_u x - Px|| = ||Qx|| = ||x - Px||, \; \forall x \in H$;
- $R_u x = x, \; \forall x \in M$;
- $||R_u x||^2 = ||Px||^2 + ||Qx||^2 = ||x||^2, \; \forall x \in H$;
- If $z \in M$ and $x \in H$, then $||z - R_u x|| = ||z - x||$;
- $||QR_u x|| = ||Qx||, \; \forall x \in H$;
- $R_u(R_u x) = R_u x, \; \forall x \in H$;
- If v is a unit vector in $M^{\perp}$, then

$$||R_v u - v|| = ||R_u v - u||.$$

In particular, R_u is an isometry that fixes the elements of M, and enjoys other properties possessed by a linear reflection.

4. A Hilbert space analogue of Heron's theorem

Given $y \notin M$, set $\hat{y} = Qy/||Qy||$. Then $\hat{y}$ is a unit vector in $M^{\perp}$ and generates the non-linear isometry $R_{\hat{y}}$ by

$$R_{\hat{y}} x = Px - \frac{||Qx||Qy}{||Qy||}, \; x \in H.$$

Lemma 4.1. *Let $y \notin M$. Then*

(1)
$$||z - x|| = ||z - R_{\hat{y}} x||, \; \forall z \in M;$$

(2)
$$||x - y|| \leq ||R_{\hat{y}} x - y||, \; \forall x \in H,$$

with equality if and only if $R_{\hat{y}} x = x$.

(3) *If also $x \notin M$, then*
$$||y - R_{\hat{y}} x|| = ||x - R_{\hat{x}} y||.$$

Proof. Part 1 was noted above. Part 2 is equivalent to the inequality

$$\begin{aligned} -2\Re < x, y > &\leq -2\Re < R_{\hat{y}} x, y > \\ &= -2\Re \left(< Px, y > - \frac{||Qx||}{||Qy||} < Qy, y > \right), \end{aligned}$$

i.e.,

$$\begin{aligned} 0 &\leq \Re\left(< x, y > - < Px, Py > + ||Qx||\,||Qy||\right) \\ &= \Re\left(< Qx, Qy > + ||Qx||\,||Qy||\right), \end{aligned}$$

which holds by the Cauchy-Schwarz inequality. Moreover, the equality holds if and only if $Qx = -||Qx||\hat{y}$, i.e.,

$$R_{\hat{y}}x = Px - ||Qx||\hat{y} = Px + Qx = x,$$

as claimed.

Part 3 follows from the fact that

$$\begin{aligned} < y, R_{\hat{y}}x > &= < y, Px > - \frac{||Qx||}{||Qy||} < Qy, y > \\ &= < Py, Px > - ||Qx||\,||Qy|| \\ &= \overline{< x, R_{\hat{x}}y >}. \end{aligned}$$

□

Theorem 4.2. *Suppose $x, y \in H$. Then*

$$\inf\left\{||y-z||+||z-x|| : z \in M\right\} = \begin{cases} ||x-y||, & \text{if } x \in M \text{ or } y \in M, \\ ||y - R_{\hat{y}}x||, & \text{if } x, y \notin M. \end{cases}$$

Moreover, unless $\{x, y\} \subset M$, the infimum is attained by a unique element in M.

Proof. By the triangle inequality,

$$||x-y|| \leq ||z-x|| + ||z-y||, \ \forall z \in H,$$

with equality if $z = x$ or $z = y$. This covers the first possibility.

Suppose $y \notin M$. Then, if $z \in M$, by Lemma 1,

$$\begin{aligned} ||y - R_{\hat{y}}x|| &= ||(y-z) + (z - R_{\hat{y}}x)|| \\ &\leq ||y-z|| + ||z - R_{\hat{y}}x|| \\ &= ||y-z|| + ||z-x||. \end{aligned}$$

Thus

$$||y - R_{\hat{y}}x|| \leq \inf\left\{||y-z|| + ||z-x|| : z \in M\right\}.$$

To show that the equality sign holds here, select $z_t = (1-t)y + tR_{\hat{y}}x$, where

$$t = \frac{||Qy||}{||Qx|| + ||Qy||}.$$

Claim: $z_t \in M$. Equivalently, $Pz_t = z_t$, i.e., $0 = (1-t)Qy + tQR_{\hat{y}}x$. But

$$\begin{aligned}
(1-t)Qy + tQR_{\hat{y}}x &= (1-t)Qy + t\left(QPx - \frac{Q^2y||Qx||}{||Qy||}\right) \\
&= (1-t)Qy - \frac{tQy||Qx||}{||Qy||} \\
&= \left(1 - t - \frac{t||Qx||}{||Qy||}\right) Qy \\
&= \left(\frac{||Qy|| - t(||Qy|| + ||Qx||)}{||Qy||}\right) Qy \\
&= 0,
\end{aligned}$$

as stated. Finally, since $z_t \in M$,

$$||z_t - y|| = t||y - R_{\hat{y}}x||, \ ||z_t - x|| = ||z_t - R_{\hat{y}}x|| = (1-t)||y - R_{\hat{y}}x||,$$

so that

$$||z_t - y|| + ||z_t - x|| = ||y - R_{\hat{y}}x||.$$

Hence, if $y \notin M$,

$$||y - R_{\hat{y}}x|| = \min\big\{||y - z|| + ||z - x|| : z \in M\big\}.$$

Of course, if $x \in M$, then $R_{\hat{y}}x = x$, and we capture the first case; and if $x \notin M$, then, by the lemma, $||y - R_{\hat{y}}x|| = ||x - R_{\hat{x}}y||$. So, this disposes of the second possibility.

We proceed to examine the cases of equality.

Case A: Both $x, y \in M$. In this case it's easy to see that

$$||z-x|| + ||z-y|| = ||x-y|| = \min\{||z-x|| + ||z-y|| : z \in M\}, \quad (1)$$

for every z in the line segment $[x, y]$. Conversely, if $||x - y|| = ||w - x|| + ||w - y||$, for some $w \in M \setminus [x, y]$, then, with $p = x - w, q = w - y$, we have $p \neq 0, q \neq 0$ and $||p + q|| = ||p|| + ||p||$. Equivalently, $\Re < p, q >= ||p||\,||q||$, so that, by the case of equality in the Cauchy-Schwarz inequality, $||q||p = ||p||q$. This now means that

$$w = \frac{||q||x + ||p||y}{||q|| + ||p||},$$

which conflicts with our hypothesis. In other words, there is equality in (1) if and only if $z \in [x, y]$. In particular, there is equality for infinitely many points in M unless $x = y$.

Case B: $x \in M, y \notin M$ and there is some $w \in M$ with $w \neq x$ such that

$$||x - y|| = ||w - x|| + ||w - y||.$$

An argument similar to the one just given implies that $w \in [x, y]$, which means that $y \in M$, which is impossible. Hence, the minimum is uniquely attained in this case.
Case C: $x, y \notin M$. Suppose

$$||y - R_{\hat{y}}x|| = ||w - x|| + ||w - y|| = ||w - R_{\hat{y}}x|| + ||w - y||,$$

for some $w \in M$. Again, $w \notin \{x, y\}$. This time, put $p = R_{\hat{y}}x - w$, $q = w - y$, so that, as before, $||p + q|| = ||p|| + ||q||$, whence $||q||p = ||p||q$, i.e.,

$$w = \frac{||q||R_{\hat{y}}x + ||p||y}{||q|| + ||p||} \equiv (1 - a)R_{\hat{y}}x + ay,$$

say. Since $Qw = 0$, $(1 - a)QR_{\hat{y}}x + aQy = 0$, i.e,

$$0 = -(1 - a)\frac{Qy||Qx||}{||Qy||} + aQy = (-(1 - a)\frac{||Qx||}{||Qy||} + a)Qy.$$

But $Qy \neq 0$, by hypothesis. Hence, $(1 - a)||Qx|| = a||Qy||$, so that

$$a = \frac{||Qx||}{||Qx| + ||Qy||} = 1 - t,$$

whence $w = z_t$. Thus, in Case C, the minimum is attained at a unique point in M.

In summary: unless both of x and y belong to M, the minimum is attained at a unique point in M. □

What this means is that, using the ℓ_1-norm, we can approximate simultaneously to two points by an element in M; and the approximating member of M is unique unless the given points both belong to the subspace.

REFERENCES

[1] CARL B. BOYER, *A History of Mathematics*, Princeton University Press, Princeton, NJ, 1968.
[2] WALTER RUDIN, *Real and Complex Analysis*, McGraw-Hill Series in Higher Mathematics, McGraw-Hill, Inc., 1966.

Finbarr Holland received his primary degree at University College, Cork, in 1961, and his Ph. D. at University College, Cardiff, in 1964. After spending the academic year 1964–65 at Caltech as a Research Fellow,

he took up a lectureship at his *alma-mater*. He was promoted there to Associate Professor in 1980. He co-founded The Irish Mathematical Society, and was its first President. He has been involved with mathematical contests at various levels since 1978. His main research interests lie in analysis.

Anca Mustaţă received her primary degree at the University of Bucharest in 1998, and her Ph. D. at the University of Utah in 2003. Her primary research area is Algebraic Geometry. In August 2007, she was appointed to a lectureship at University College Cork. Since 2009 she has coordinated the UCC Mathematics Enrichment Programme.

Finbarr Holland,
School of Mathematical Sciences,
University College Cork
f.holland@ucc.ie

Anca Mustaţă,
School of Mathematical Sciences,
University College Cork
a.mustata@ucc.ie

Received on 10-3-2011, and in revised form 7-4-11.