An approximation theoretic characterisation of inner product spaces

A. L. *Brown*

ABSTRACT

Two approximation theoretic properties of Hilbert spaces used in the proof of a result concerning Chebyshev sets obtained by Frerking and Westphal [4] are discussed. Investigation of one leads to a characterisation of inner product spaces of dimension at least three; it is an improvement of one due to Berens [3]. The other property is shared by all finite dimensional spaces [3] and here a topological result of which that fact is a simple consequence is proved.

1980 Mathematics Subject Classification (1985): 46C05; 41A65, 54C60.

Considerable but so far unsuccessful efforts have been made to determine whether or not every Chebyshev subset of a Hilbert space is convex. A number of conditional results have been obtained and one of the most satisfying of them is one of those due to L.P. Vlasov [v.6]:

(I) *If* X *is a Banach space,* X^* *is strictly convex and* K *is a Chebyshev subset of* X *such that the metric projection of* X *onto* K *is continuous then* K *is convex.*

This result can be compared with a striking recent Hilbert space result of J. Frerking and U. Westphal [4] which is more conveniently stated in the contrapositive form:

(II) *If* X *is a Hilbert space and* K *is a non-convex Chebyshev subset of* X *then the set of points of discontinuity of the metric projection of* X *onto* K *contains a non-trivial Lipschitz curve.*

The result (II) is a improvement of one due to Balaganskii [2]; its proof depends upon four approximation theoretic properties of Hilbert spaces. In order to state them as (i) - (iv) below we must introduce some notation.

Let X be a real normed linear space. The open ball, closed ball and sphere centre $x \in X$ and of radius r will be denoted by $B(x,r)$, $B'(x,r)$ and $S(x,r)$ respectively. The unit sphere $S(0,1)$ will also be denoted by S_X. The identity mapping on X will be denoted by I. If K is a subset of X and $x \in X$ then

$$d(x,K) = \inf\{\|x-k\| : k \in K\}$$

is the distance of x from K. The *metric projection* of X onto K is the set valued mapping P_K defined by

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

K is said to be *proximinal* if $P_K(x) \neq \phi$ for each $x \in X$, and *Chebyshev* if $P_K(x)$ is a single point for each $x \in X$ (in which case P_K is regarded as a mapping of X onto K). A second set valued mapping Φ_K is defined by

$$\Phi_K(x) = \bigcap_{\epsilon>0} \overline{co}(K \cap B'(x,d(x,K)+\epsilon)).$$

Thus $co\, P_K(x) \subseteq \Phi_K(x)$ for each $x \in X$. If X is of finite dimension and K is closed then $co\, P_K(x) = \Phi_K(x)$ for each $x \in X$. If X is reflexive then $\Phi_K(x) \neq \phi$ for each $x \in X$.

The first of the four properties is possessed by any uniformly convex space:

(i) *If X is a uniformly convex space and K is a Chebyshev subset of X then P_K is continuous at $x \in X$ if and only if* $P_K(x) = \Phi_K(x)$.

If X is a Hilbert space and K is a subset of X then Φ_K is an accretive (monotone) set valued mapping. If Φ is any accretive set valued mapping on a normed linear space X then for each $\lambda > 0$ the set valued mapping $\lambda I + \Phi$ is injective and its inverse $(\lambda I + \Phi)^{-1}$ is Lipschitz with constant $\frac{1}{\lambda}$ on its domain. Thus

(ii) *If X is a Hilbert space and K is a subset of X then for each*

$\lambda > 0$ *the set valued mapping* $\lambda I + \Phi_K$ *is injective, and*

(iii) $(\lambda I + \Phi_K)^{-1}$ *is continuous on its domain.*

However, if X is a Hilbert space then Φ_K is maximal accretive (maximal monotone) and this has the important consequence that

(iv) *If* X *is a Hilbert space and* K *is a subset of* X *then for each* $\lambda > 0$ *the set valued mapping* $\lambda I + \Phi_K$ *is surjective.*

The truth of (ii), (iii) and (iv) for a single $\lambda > 0$ would be sufficient for the proof of (II). (We have not considered weaker conditions which might suffice). The comparison of (I) and (II) naturally prompts the question of whether (or to what extent) the properties (ii), (iii) and (iv) extend to any spaces which are not Hilbert spaces. Here we show that (ii) does not do so even for a single λ, and we prove a topological theorem which gives a variant proof of the fact that property (iv) is shared by all finite dimensional spaces.

THEOREM 1 *Let* X *be a real normed linear space of dimension at least three. Then the following four conditions are equivalent:*

(1) X *is an inner product space,*

(2) *For each* $x \in S_X$ *there exist positive constants* λ, ϵ *such that*

$$0 \notin \lambda tx + \mathrm{co}(S_X \cap (tx + S_X))$$

for all $t \in (0,\epsilon)$,

(3) *For each proximinal subset* K *of* X *there exists* $\lambda > 0$ *such that* $\lambda I + \mathrm{co}\, P_K$ *is injective.*

(4) *For each* $x \in S_X$ *there exists* $\epsilon > 0$ *such that* $S_X \cap (tx + S_X)$ *is contained in a hyperplane for all* $t \in (0,\epsilon)$.

Proof That (1) and (4) are equivalent is a characterisation of inner product spaces due to P. Gruber [v.1,(15.17)]; it is natural to include it here for comparison as it is easily seen that (4) $\Longrightarrow$ (2). The implication

(1) $\Longrightarrow$ (4) is quite elementary. The statement (ii) above includes the implication (1) $\Longrightarrow$ (3). It remains to prove that (2) $\Longrightarrow$ (1) and (3) $\Longrightarrow$ (1).

If X is not an inner product space then it contains a three dimensional subspace which is not an inner product space. If (2) and (3) are false for some three dimensional subspace of X, in place of X, then they are false for X. It may therefore be supposed that X is of dimension three.

If $x \in X$ then y is said to be orthogonal to x (in the sense of Birkhoff) and one writes $y \perp x$, if $d(y,\mathbb{R}s) = \|y\|$. If $\|x\| = \|y\| = 1$ then $y \perp x$ if and only if the line $y + \mathbb{R}x$ supports S_X at y.

Suppose that X is not an inner product space. Then by a characterisation due to R.C. James [v.1,(14.2)] there exists $x \in S_X$ such that the set $\{y \in X : y \perp x\}$, which is symmetric about 0, does not contain a hyperplane. It follows that

$$0 \in \operatorname{int}(\operatorname{co}\{y \in S_X : y \perp x\})$$

(where 'int' denotes the interior relative to X). Therefore there exist points y_1, y_2, y_3, y_4 in S_X, each orthogonal to x such that

$$0 \in \operatorname{int}(\operatorname{co}\{y_1, y_2, y_3, y_4\}).$$

Now, for each i = 1,2,3,4, if $t \in (0,1]$, so $\|tx\| \leq 1$, then

$$\|y_i - tx\| \geq 1 \quad \text{and} \quad \left\|\frac{y_i + tx}{\|y_i + tx\|} - tx\right\| \leq 1.$$

So there exists $\theta \in [0,t]$ such that

$$\frac{y_i + \theta x}{\|y_i + \theta x\|} \in S_X \cap (tx + S_X).$$

It follows that, given $\epsilon > 0$, there exists $t_0 \in (0,1]$ such that

$$d(y_i, S_X \cap (tx + S_X)) < \epsilon$$

for all $t \in (0, t_o]$ and for i = 1,2,3,4. For some $\eta > 0$

$$B(0,2\eta) \subseteq \operatorname{co}\{y_1, y_2, y_3, y_4\}$$

and therefore for some $t_0 > 0$

$$B(0,\eta) \subseteq \operatorname{co}(S_X \cap (tx + S_X))$$

for all $t \in (0,t_0]$. If $\lambda > 0$ and $0 < t < \min\{\frac{\eta}{\lambda}, t_o\}$ then

$$-\lambda tx \in B(0,\eta) \subseteq co(S_X \cap (tx+S_X)).$$

This proves that (2) is not satisfied.

Now for $t \in (0,t_0]$ let $K_t = S_X \cap (tx+S_X)$. Then $P_{K_t}(0) = K_t$, $P_{K_t}(tx) = K_t$ and if $\lambda \leq \frac{\eta}{t}$

$$0 \in B(0,\eta) \subseteq co\ K_t = (\lambda I+co\ P_{K_t})\ (0),$$

$$0 \in \lambda tx + co\ K_t = (\lambda I+co\ P_{K_t})(tx).$$

Now let $(t_n)_{n\geq 1}$ be a sequence in $(0,t_0]$ which is convergent to 0, let $(z_n)_{n\geq 1}$ be a sequence in X such that $\|z_n\| = 6n$, for $n = 1,2,\ldots$, and let

$$K = \bigcup_{n=0}^{\infty} (z_n+K_{t_n}).$$

It is now easily seen that $\lambda I + co\ P_K$ is not injective for any $\lambda > 0$. Thus (3) is not satisfied. This completes the proof that (2) $\Longrightarrow$ (1) and (3) $\Longrightarrow$ (1). The characterisation (3) of inner product spaces is an improvement of one obtained by Berens [3].

The next theorem involves Cech (or Alexander or sheaf theoretic) cohomology. Let G by any non-trivial abelian group. A topological space is said to be *acyclic* if its Cech cohomology with coefficients in G is trivial. The theorem is, modulo the Vietoris-Begle mapping theorem [5, p.344], a very simple result. It is the natural extension to set valued mappings of a result for (ordinary) mappings which follows directly from the fact that a sphere is not a retract of its ball. The set of subsets of the space $\mathbb{R}^n$ will be denoted by $\mathcal{P}(\mathbb{R}^n)$. For the purposes of the notation let $\mathbb{R}^n$ be equipped with the euclidean norm.

THEOREM 2 *Let* $\Phi : \mathbb{R}^n \longrightarrow \mathcal{P}(\mathbb{R}^n)\backslash\{\emptyset\}$ *be a bounded upper semi-continuous set valued mapping with the property that* $\Phi(x)$ *is an acyclic closed subset of* $\mathbb{R}^n$ *for each* $x \in \mathbb{R}^n$. *Then* $I + \Phi$ *is surjective.*

Proof Suppose on the contrary that there exists a $\Phi : \mathbb{R}^n \longrightarrow \mathcal{P}(\mathbb{R}^n)\backslash\{\emptyset\}$ satisfying the conditions of the theorem but such that $0 \notin x + \Phi(x)$ for each $x \in \mathbb{R}^n$. It is easily shown that, by the upper semi-continuity and boundedness of Φ, the range $(I+\Phi)(\mathbb{R}^n) = \bigcup_{x\in\mathbb{R}^n}(x+\Phi(x))$ is a closed subset of $\mathbb{R}^n$. Therefore, for some $r > 0$, $\|x+y\| \geq r$ for all $x \in \mathbb{R}^n$ and $y \in \Phi(x)$. Let $K = \sup\{\|y\| : y \in \Phi(x), x \in \mathbb{R}^n\}$. Choose $R_1 \geq K+r$ and $R_2 > R_1$. Let $f : [0,R_2] \longrightarrow [0,1]$ be a continuous function such that

$$f(\rho) = \begin{cases} 0, & \text{for } \rho \in [0,R_1], \\ 1, & \text{for } \rho = R_2, \end{cases}$$

and define $\Psi : B'(0,R_2) \longrightarrow \mathcal{P}(\mathbb{R}^n)\backslash\{\emptyset\}$ by

$$\Psi(x) = f(\|x\|)x + (1-f(\|x\|))(x + \Phi(x)).$$

Then Ψ is also upper semi-continuous and $\Psi(x)$ is an acyclic closed subset of $\mathbb{R}^n$ for each $x \in B'(0,R_2)$. Furthermore $\|x+y\| \geq r$ for all $x \in B'(0,R_2)$ and all $y \in \Psi(x)$, and $\Psi(x) = \{x\}$ for each $x \in S(0,R_2)$.

Let

$$\mathcal{G}(\Psi) = \bigcup_{x\in B'(0,R_2)} \{x\} \times \Psi(x) \subseteq B'(0,R_2) \times (\mathbb{R}^n\backslash B(0,r))$$

denote the 'graph' of Ψ. The graph is a subset of the product space. The projection $p : \mathcal{G}(\Psi) \longrightarrow B'(0,R_2)$ is closed, continuous and surjective and $p^{-1}(x) = \{x\} \times \Psi(x)$ is acyclic for each $x \in B'(0,R_2)$. Therefore by the Vietoris-Begle mapping theorem p induces an isomorphism

$$p^* : \check{H}^*(B'(0,R_2)) \longrightarrow \check{H}^*(\mathcal{G}(\Psi))$$

of Cech cohomology. Thus $\mathcal{G}(\Psi)$ has trivial Cech cohomology i.e. has the cohomology of a point. If we let

$$\text{diag } S(0,R_2) = \{(x,x) : x \in S(0,r_2)\}$$

then there is a commutative diagram

$$\begin{array}{ccc} \text{diag } S(0,R_2) & \longrightarrow & B(0,R_2) \times (\mathbb{R}^n\backslash B(0,r)) \\ \downarrow & & \downarrow \\ S(0,R_2) & \longrightarrow & \mathbb{R}^n\backslash B(0,r) \end{array}$$

of horizontal inclusions and vertical projections to the second factor. The left hand projection is a homeomorphism, the right hand projection is a homotopy equivalence, and the lower horizontal inclusion is a homotopy equivalence. Therefore all the mappings induce isomorphisms of Cech cohomology. However there are inclusions

$$\operatorname{diag} S(0,R_2) \longrightarrow \mathcal{G}(\Psi) \longrightarrow B(0,R_2) \times (\mathbb{R}^n \setminus B(0,r))$$

and the composite

$$\check{H}^*(B(0,R_2) \times (\mathbb{R}^n \setminus B(0,r))) \longrightarrow \check{H}^*(\mathcal{G}(\Psi)) \longrightarrow \check{H}^*(\operatorname{diag} S(0,R_2))$$

of the induced mappings is a non-trivial isomorphism. This contradicts the fact that $\check{H}^*(\mathcal{G}(\Psi))$ is trivial. The proof of the theorem is complete.

A convex subset of a normed linear space is contractible and so acyclic. Suppose that X is a finite dimensional normed linear space. If K is a bounded closed subset of X and $\lambda > 0$ then $\operatorname{co} P_K = \Phi_K$ is upper semi-continuous and it follows immediately from Theorem 2 that $\lambda I + \operatorname{co} P_K$ is surjective. The link to the general case is provided by a proposition.

PROPOSITION *Let $\lambda > 0$. If X is a real normed linear space and $\lambda I + \Phi_K$ is surjective for each bounded closed non-empty subset K of X then $\lambda I + \Phi_K$ is surjective for each closed non-empty subset of X.*

Proof The proof expresses the fact that $(\lambda I + \Phi_K)^{-1}$ is locally bounded. It is convenient to calculate in terms of $\frac{\lambda}{\lambda+1}I + \frac{1}{\lambda+1}\Phi_K$. If $z_0 \in X$, $x \in X$ and $z \in \frac{\lambda}{\lambda+1}x + \frac{1}{\lambda+1}\Phi_K(x)$ then

$$\begin{aligned}(1+\lambda)\|x-z\| &= \|x-((\lambda+1)z-\lambda x)\| \\ &\le d(x,K) \\ &\le \|x-z\| + \|z-z_0\| + d(z_0,K)\end{aligned}$$

and therefore

$$\|x-z_0\| \le \frac{\lambda+1}{\lambda}\|z-z_0\| + \frac{1}{\lambda}\, d(z_0,K).$$

It follows that

$$(*) \qquad \left(\frac{\lambda}{\lambda+1}I + \frac{1}{\lambda+1}\Phi_K\right)^{-1}(B(z_0,\epsilon)) \subseteq B\left(z_0, \frac{\lambda+1}{\lambda}\epsilon + \frac{1}{\lambda}d(z_0,K)\right).$$

If $R > 2\left(\frac{\lambda+1}{\lambda}\epsilon + \frac{1}{\lambda}d(z_0,K) + d(z_0,K)\right)$ and $K' = K \cap B'(z_0,R)$ then $d(z_0,K) = d(z_o,K')$, the inclusion (*) holds with K' in place of K and Φ_K, $\Phi_{K'}$ coincide on $B(z_0,\frac{\lambda+1}{\lambda}\epsilon + \frac{1}{\lambda}d(z_0,K))$. Consequently

$$B(z_0,\epsilon) \subseteq \left(\frac{\lambda}{\lambda+1}I + \frac{1}{\lambda+1}\Phi_K\right)(x)$$

if and only if

$$B(z_0,\epsilon) \subseteq \left(\frac{\lambda}{\lambda+1}I + \frac{1}{\lambda+1}\Phi_{K'}\right)(X).$$

The proposition now follows.

Theorem 2 and the Proposition immediately give the following theorem, the original proof of which invoked a result for convex valued set valued mappings.

THEOREM (Berens [3]) *If K is a closed non-empty subset of a finite dimensional normed linear space X then, for each $\lambda > 0$, the set valued mapping $\lambda I + \text{co}\, P_K$ is surjective.*

REFERENCES

[1] AMIR, D., (1986). Characterizations of inner product spaces, Birkhauser Verlag, Basel etc.

[2] BALAGANSKII, V.S., (1982). Approximative properties of sets in Hilbert space, Mat.Zametki, 31, No 5, 785-800. = Math.Notes 31, No 5, 397-404 (1982).

[3] BERENS, H., (1982). Uber die Beste Approximation in $\mathbb{R}^n$, Archiv der Math., 39, 376-382.

[4] FRERKING, J. AND WESTPHAL, U. On a property of metric projections onto closed subsets of Hilbert spaces, to appear.

[5] SPANIER, E.H., (1966). *Algebraic Topology*, McGraw-Hill, New York.

[6] VLASOV, L.P., (1973). Approximative properties of sets in normed linear spaces, Uspekhi Mat. Nauk 28, no 6, 3-66. = Russian Math.Surveys 28 (1973), no 6, 1-66.

Department of Mathematics and Statistics
The University, Newcastle Upon Tyne, NE1 7RU