MISCELLANEOUS EXERCISES TO CHAPTER 1.

1. i) If d is the discrete metric on any set X, describe the balls $B_r(x)$ and $B_r[x]$. [Hint: consider the possible cases; $0 < r < 1$, $r = 1$ and $r > 1$.]

   ii) Let X be the family of all subsets of $A = \{1, 2, 3, 4, 5\}$. If d is the metric defined on X as in Example 6 of page 5, describe the ball $B_2(\{2,3\})$.

2. Give a counter example to the natural conjecture that in any metric space $\overline{B_r(x)} = B_r[x]$. [Show that this is, however, true in a normed linear space.]

3. Show that any finite subset of a metric space is bounded.

4. Show that the intersection of two convex sets is convex. Is the same true of the union of two convex sets?

5. Show that the terms of a Cauchy sequence in a metric space form a bounded set.

6. Let (X,d) be a metric space for $x \in X$ and $A \subseteq X$ define the *distance from* x *to* A to be

$$d(x,A) = \inf_{a \in A} d(x,A).$$

   i) Show $A \neq \phi$ implies $0 \leq d(x,A) < \infty$

   ii) Prove $x \in \text{Int}(X \backslash A)$ iff $d(x,A) > 0$.

   iii) Prove $x \in \bar{A}$ iff $d(x,A) = 0$, (hence characterise those $x \in \text{bdry } A$ in terms of distances from sets.)

7. (SEPARATION PROPERTIES)

   (a) If x,y are distinct points of the metric space (X,d), show that there exists a pair of disjoint open balls each of which is centred on one of the points. Because of this metric spaces are said to have the Hausdorff separation property.

   (b) In the metric space (X,d) let $x \notin A = \bar{A}$. Show there exists disjoint open sets G_1, G_2 with $x \in G_1$ and $A \subseteq G_2$.

   *(c) Let A_1, A_2 be any pair of disjoint closed subsets in the metric space (X,d). Show that there exists disjoint open sets G_1, G_2 with $A_i \subseteq G_i$ $(i = 1,2)$. Because of this property we say metric spaces are normal spaces.

8.* The <u>ruler function</u> $r: [0, 1] \to Q \cap [0, 1]$ defined by

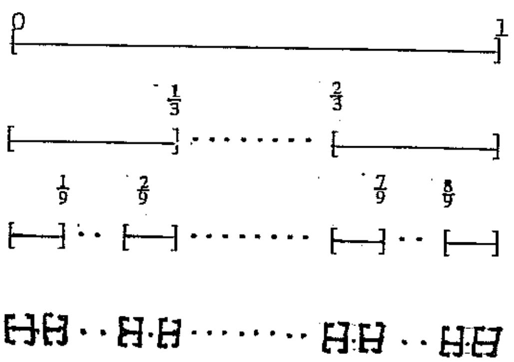
$$r(x) = \begin{cases} 0 \text{ if } x \text{ is irrational} \\ \frac{1}{q} \text{ if } x = \frac{p}{q} \end{cases}$$

where p and q are mutually prime integers (<u>i.e.</u> the greatest common divisor of p and q is 1), has the property of being continuous at each irrational point, but discontinuous at every rational point. Prove this for, at least, the two special cases of $x = \frac{1}{2}$ and $x = \sqrt{2}/2$. (This function is considered in many of the standard books including Spivak's "Calculus".)

9. In any metric space (X,d) let $x_n \xrightarrow{d} x$. Show that the set $\{x_1, x_2, \ldots, x_n, \ldots, x\}$ is compact.

10.* If K is a compact subset of the metric space (X,d) show that for any $\varepsilon > 0$ there exists a finite number of points $x_1, x_2, \ldots, x_n$ such that $K \subseteq \bigcup_{j=1}^{n} B_\varepsilon(x_j)$.

11. One of the most fascinating sets in analysis is <u>Cantor's Ternary Set</u>. It may be constructed inductively by deleting the open interval $(\frac{1}{3}, \frac{2}{3})$ from $[0,1]$, then deleting $(\frac{1}{9}, \frac{2}{9})$ and $(\frac{7}{9}, \frac{8}{9})$, followed by the deletion of $(\frac{1}{27}, \frac{2}{27})$, $(\frac{7}{27}, \frac{8}{27})$, $(\frac{19}{27}, \frac{20}{27})$, $(\frac{25}{27}, \frac{26}{27})$ followed by ... etc.



(see diagram)

Clearly, the points $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \ldots$ are never deleted by this process and so $K \neq \phi$, in fact K contains infinitely many points, despite the fact that the total "length" of non-overlapping intervals deleted is readily seen to equal

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{8}{81} + \ldots \quad = \quad \frac{1}{3}\left(1 + \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \ldots + \frac{2^n}{3^n} + \ldots\right)$$

a geometric progression whose sum is 1!

(i) Deduce that K is a closed subset of $([0,1], d_1)$.

(ii)* (For latter parts you may assume the results of this part, if you feel unable to prove them and feel they would help.) Show the following are equivalent definitions of K

(a) $$K = [0,1]\setminus \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\frac{3^n-1}{2}} \left(\frac{2m-1}{3^n}, \frac{2m}{3^n}\right)$$

(b) K consists of precisely those points in [0,1] having <u>a</u> ternary representation (<u>i.e.</u> representation to base 3) of the form $\sum_{n=1}^{\infty} \frac{a_n}{3^n}$ with $a_n = 0$ or 2 for all $n \in \mathbb{N}$.

[NOTE: $\frac{1}{9} = 0.01 = 0.00\dot{2}$ (base 3); thus $\frac{1}{9} \in K$.]

(iii) Show Int $K = \phi$

(iv) Show bdry $K = K$

(v) Show $[0,1]\setminus K$ is a dense subset of $[0,1]$

(vi)* Show K is an uncountable set with cardinality that of the continuum; <u>i.e.</u> there exists a 1-1 and onto mapping from K to [0,1]. (Thus although K has zero "length" it contains as "many" points as the original interval [0,1].)

CHAPTER 2: APPLICATIONS

This Chapter is divided into four sections. Each section extends the basic theory of Chapter 1 in a specific direction and cullminates with at least one application of the theory developed.

SECTION 1. APPROXIMATION THEORY

Preamble

Much of Mathematics depends on the "approximation" of given objects by more tractable ones.

For example:

The approximation of π by rational numbers, $\pi \simeq \frac{22}{7}, \frac{223}{71}, \ldots;$

The approximation of a given function by a polynomial, $\exp(x) \doteq 1 + x + x^2/2 + x^3/6$, $\sin x \doteq x$;

The approximation of a given function by a trigonometric polynomial or its equivalent,

$$\text{For } f(x) = \begin{cases} -1 & \text{if } -\pi \leq x \leq 0 \\ 1 & \text{if } 0 < x \leq \pi \end{cases}$$

$$f(x) \simeq \frac{4}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x\right) \text{ (c.f. Fourier Series);}$$

Finding a polynomial approximation to the solution of a given differential equation [Method of Frobenius];

For a given $n \times m$ matric A $(n > m)$ and n-vector $\underset{\sim}{b}$, finding an m-vector $\underset{\sim}{x}$ for which $A\underset{\sim}{x} - \underset{\sim}{b}$ is approximately zero, i.e., finding an approximate solution to an inconsistent ("over-specified") system of equations $A\underset{\sim}{x} = \underset{\sim}{b}$.

All of these problems share a basic similarity: *From among the elements of a given set* A, *choose one which will, in some specified sense, approximate the object, t.*

A number of natural questions arise.

(i) How "good" an approximation is possible?

(ii) For a given A and t is there a "best" approximation?

(iii) If there is a "best approximating element" is it unique or can other equally acceptable approximations be found?

(iv) Assuming it exists can we develop a procedure (algorithm) for determining the best approximation to t from among the element of A?

These are typical of the type of questions with which approximation theory deals.

From the pioneering work of Pafnuti Tchebycheff and Karl Weierstrass in the second half of the 19th century through the penetrating work of Haar, Bernstein and many others to the present day, approximation theory has become an increasingly important branch of mathematics with applications in other areas of mathematics, computing, engineering, economics and the social and life sciences.

Speaking of the approximation of t by elements of A presupposes that we have some criterion for "measuring" the proximity of $a \in A$ to t, the natural super-structure for such a criterion is that of a metric space.

Consequently, we will assume that A is a subset and t an element of some metric space (X,d). We will mainly deal with questions of type (i), (ii) or (iii) which now become.

(i) For any given $\varepsilon > 0$ is there a point a of A with $d(a,t) \leq \varepsilon$?

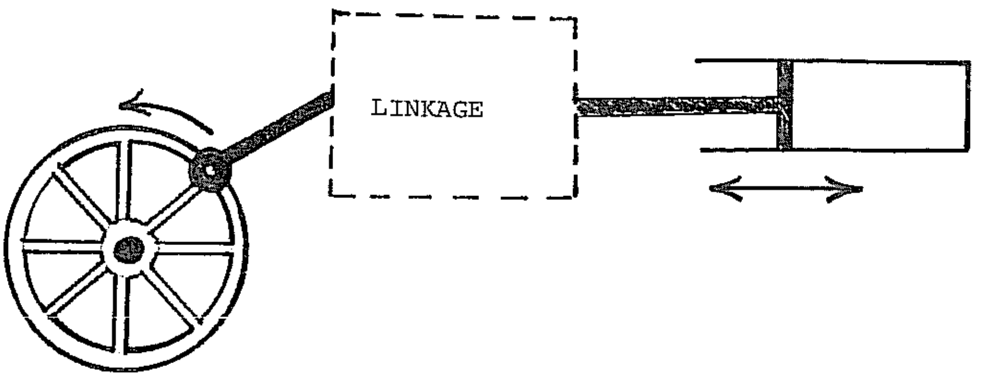
(ii) Is there a point a_0 of A with $d(a_0,t) \leq d(a,t)$ for all $a \in A$? When such an a_0 exists we refer to it as a *best approximation* to t from A or a closest point of A to t.

(iii) If it exists is the point a_0 of (ii) unique?

Except in a few simple cases the more specific and usually more involved problems associated with question (iv) have been avoided. (See Cheney, Chapter 3 onwards for examples of work of this type.)

The answers to (i), (ii) and (iii) above can vary with the choice of metric for X. Usually this choice is dictated by the context of the problem. For example, if it were required to approximate an electrical signal by another one which was required to deliver about the same average power, it would be reasonable to use the metric induced by the inner-product norm $\|\cdot\|_2$ (corresponding to the "root mean square average" of the difference in signals). If on the other hand the signal were to operate a control device sensitive to voltage changes the Tchebyacheff norm $\|\cdot\|_\infty$ would be more appropriate (measuring the maximum deviation between the two signals).

Tchebyscheff himself first introduced the "sup" norm in connection with a problem relating to the design of linkages between the wheel and piston in a steam locomotive.



The problem was to develop a linkage which would exactly convert the rectilinear motion of the piston into the circular motion of the wheel. This problem had been unsuccessfully considered by several engineers and mathematicians prior to Tchebyscheff. Tchebyscheff convinced himself that such an exact linkage was not possible. He then sought to determine the "best" linkage which could be constructed with n pivots. Clearly it is no good having such a linkage nearly exact for most of the motion and way out for a small part of it (at that point something would break). Consequently, he measured the "goodness" of the approximation by the maximum departure it produced from the desired motion, and so implicitly introduced the "sup" norm $\|\cdot\|_\infty$. *[Shortly afterwards, 1864, the French Naval Officer, Peaucellier, discovered an exact linkage involving 6 pivots. However, by that time the development of high tensile bearings, and efficient lubricants which could tolerate "slop" and withstand greater stresses, rendered the solution unnecessary.]*

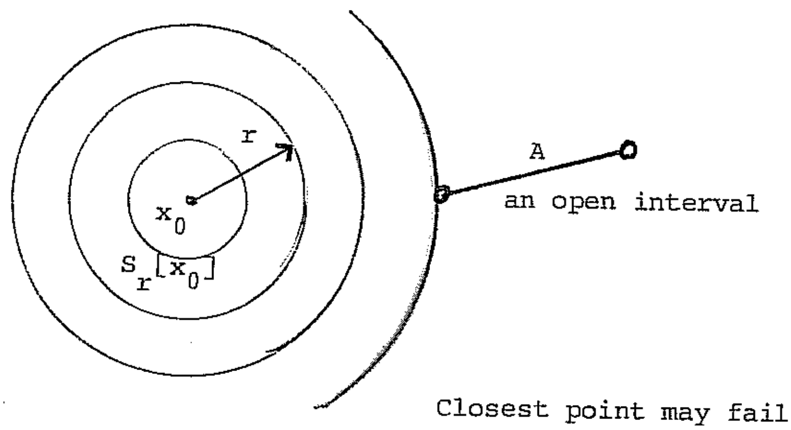
To gain some idea of the type of situations which our theory must deal with we begin by giving a GEOMETRIC INTERPRETATION OF BEST APPROXIMATION.

To find the best approximations from a subset A of X to the point $x_0 \in X$ (assuming they exist) we may intuitively proceed as follows.

Starting with a very small value of r, gradually increase r until the sphere centred on x_0, $S_r[x_0]$, first makes contact with A. (Imagine blowing up a balloon in the shāpe of the sphere.) The point(s) at which contact is first made are the closest points of A to x_0.

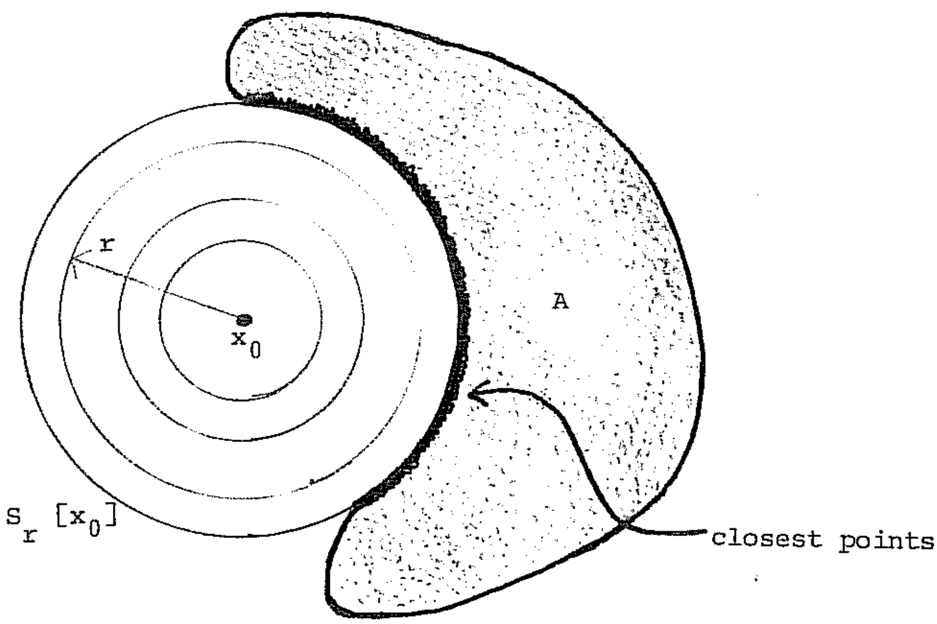
With this interpretation, the following sketches should serve to illustrate how closest points can fail to exist or fail to be unique.

I)



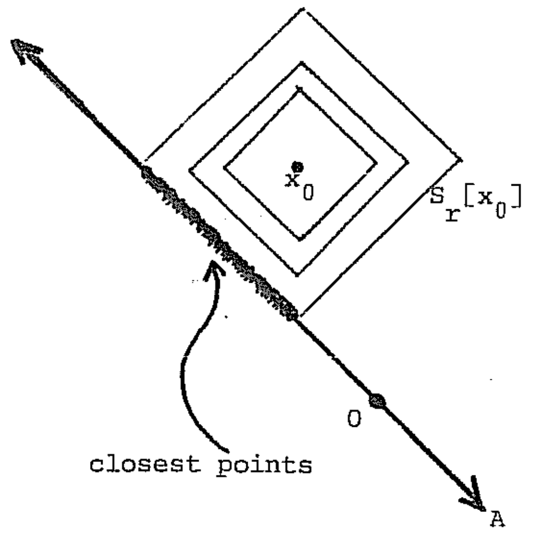
Closest point may fail to exist

II)



Even with a "nice" norm (here the euclidean norm on $\mathbb{R}^2$) without some restriction on the "shape of the set A (e.g. convexity) the closest point may fail to be unique.

III)



$\mathbb{R}^2$ with $\|\cdot\|_1$

Without a restriction on the norm (shape of the unit ball) the closest point of even a "nice" set (here a subspace in $\mathbb{R}^2$) may fail to be unique.

REMARK Our illustrations have been in $\mathbb{R}^2$ (because we can easily draw them). This does not detract from their significance, indeed it highlights the improtance of our abstract methods. We cannot easily picture the situation when for example, asking for the best approximation(s) to e^x from the set A of polynomials with degree less than 11 which involve coefficients of absolute values at most 5 in the space C[0.1] with the uniform norm. By recognising the analogy with similar questions in $\mathbb{R}^2$ (both are realization of the same abstract situation) we can, however, envisage difficulties which might arise and concepts needed for their resolution. Our abstract approach enables us to transfer some of our intuition about $\mathbb{R}^2$ to other less familiar situations.

General Theory.

We begin by observing that if x_0 is a limit point of the subset A in the metric space (X,d), then there are points of A arbitrarily close to x_0, so there can be no closest point of A to x_0 unless x_0 itself is in A, in which case x_0 is its own best approximation from A. For example, in $\mathbb{R}$ with $|\cdot|$ as norm, 1 is a limit point of the open interval (0,1) and we observe that there is no positive number strictly less than 1 which is closest to it.

A necessary condition for a subset A *to contain a closest point to each point of the space is therefore that* A *be closed.* A subset with this property; that it contains a best approximation to each point of the space; is sometimes termed "a proximal subset". Our first results will show that at least in finite dimensional normed linear spaces, the property of being closed is also a sufficient condition for a set to be proximal.

<u>Theorem 1</u>. *Let* K *be a compact subset of the metric space* (X,d), *then to each point* $x \in X$ *there corresponds a point* $k_x \in K$ *of minimum distance from* x.

That is, K contains a best approximation to each point of the space.

<u>Proof</u>. Let δ = infimum $d(k,x) : k \in K$, then by the definition of infimum (greatest lower bound) there exists a sequence of points of K, (k_n) such that $d(k_n,x) \to \delta$ as $n \to \infty$, and further $d(k,x) \geq \delta$ for all $k \in K$. By the compactness of K we may extract a subsequence of (k_n); $k_{n_1}, k_{n_2}, k_{n_3}, \ldots, k_{n_m}, \ldots$ convergent to some point $k_x \in K$. Now $\delta \leq d(k_x,x) \leq d(k_x,k_{n_m}) + d(k_{n_m},x)$ for any $m \in N$ and so taking the limit as $m \to \infty$ and noting that $d(k_x,k_{n_m}) \to 0$ while $d(k_{n_m},x) \to \delta$ we have that

$$\delta \leq d(k_x,x) \leq \delta$$

or

$$d(k_x,x) = \delta = \inf \{d(k,x) : k \in K\} \leq d(k,x) \text{ for all } k \in K,$$

as required. ∎

EXERCISE: Give an alternative proof of Theorem 1, based on Theorem 7 of §1.6 (page 69), by first showing that the mapping $K \to \mathbb{R}$: $k \mapsto d(k,x)$ is continuous.

As a consequence of Theorem 1 and the work in §1.6 we have

<u>Theorem 2</u>. *Let* M *be a finite dimensional subspace of the normed linear space* $(X, \|\cdot\|)$ *and let* $A \subseteq M$ *be a closed subset of* X. *Then* A *contains a best approximation to each point of* X.

<u>Proof</u>. Given $x \in X$, choose any point $a_1 \in A$, then if there is a closest point a_0 of A to x we have

$$\|x - a_0\| \leq \|x - a_1\|$$

and so a_0 would belong to

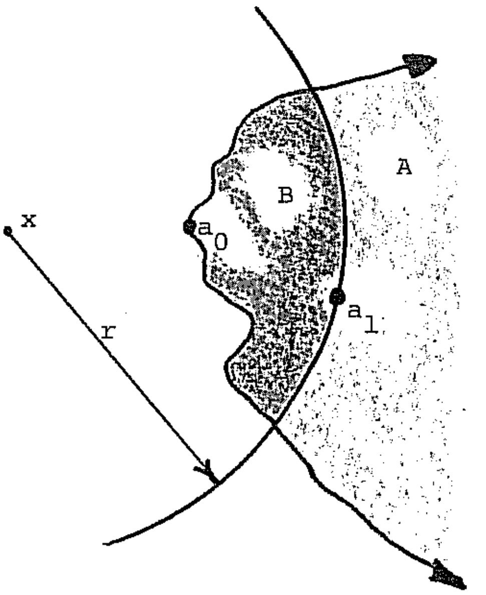
$$B = \{y \in A : \|x - y\| \leq \|x = a_1\|\}.$$

~~Thus it suffices to show that there is a closest point of B to x.~~

Now

$$B = A \cap \{y \in X : \|x - y\| \le \|x - a_1\|\}$$

$$= A \cap B_r[x], \text{ where } r = \|x - a_1\|$$



and so B is a closed (intersection of two closed sets) bounded ($B \subseteq B_r[x]$) subset of the finite dimensional subspace M. Hence, by Theorem 8 of §1.6 (page 75), B is compact and the desired conclusion follows from Theorem 1.

■

To illustrate this result consider the following examples (in each case the above theorem guarantees the existence of a solution).

(I) Since an irrational number can be approximated arbitrarily well by rationals, there is no rational q for which $d(q,\pi) = |q - \pi|$ is a minimum; that is, there is no best rational approximation to π. However, we can ask; from among the set of positive rational numbers with denominator less than or equal to 10, say, which best approximates π in the sense that $|q - \pi|$ is a minimum. More generally for $N \in \mathbb{N}$ let $A_N = \left\{\frac{p}{q} : p,q \in \mathbb{N} \text{ and } 0 < q \le N\right\}$, then $A_N \cap [3,4]$ contains only finitely many elements, and so there exists $q_N \in A_N$, for which $|q_N - \pi| \le |q - \pi|$ for all $q \in A_N$; that is, q_N is a best approximation to π from A_N. For example, it is readily checked that $q_{10} = 22/7$.

Exercises

1) Show that in this case the best approximation to π from A_N is unique. (HINT: observe that if there were two distinct best approximations they would be of the form $r_1 = \pi - \varepsilon$ and $r_2 = \pi + \varepsilon$. But then, $\pi = \frac{1}{2}(r_1 + r_2)$ contradicting its irrationality.)

2) (Optional) Determine q_{20}. If you have a programmable calculator available, you might like to try finding q_N for some larger values of N. (NOTE: $\pi \doteq 3.1415926536$.)

The question of how, in general, to determine q_N is an old one. It was studied by Huyghens (1682) when designing toothed wheels for a planitarium and was first solved by Wallis in 1685 using the theory of continued fractions. (See, for example, Chriystal "Textbook of Algebra" Vol. II Ch XXXII, first published 1889; seventh edition, Chelsea 1964.)

For interest, a few results are tabulated below.

$$q_{100} = \frac{311}{99}$$

$$q_{1000} = \frac{355}{113}$$

$$q_{25000} = \frac{78433}{24966}$$

(II) Let C be the one-dimensional (basis element, the function $f_1(x) \equiv 1$) subspace of $(C[1,2], \|\cdot\|_\infty)$ consisting of all constant functions. A best approximation to $f(x) = \frac{1}{x}$ from C is the function $f_0(x) \equiv C_0$, where C_0 is such that

$$\Delta(t) = \|\frac{1}{x} - t\|_\infty = \max_{1 \le x \le 2} |\frac{1}{x} - t|$$

has a minimum at $t = C_0$.

Exercise: Determine C_0 and show it is unique.

(III) <u>Exercise:</u> Let L_0 be the one-dimensional (basis element the function x) subspace of $C[0,1]$ consisting of all polynomials p of degree less than or equal to one for which $p(0) = 0$; that is, all "straight lines through the origin".

(a) Find the best approximation to $f(x) = e^x$ from L_0 in $(C[0,1], \|\cdot\|_2)$.

That is, find m such that

$$\left[\int_0^1 (e^x - mx)^2 \, dx \right]^{\frac{1}{2}}$$

is smallest.

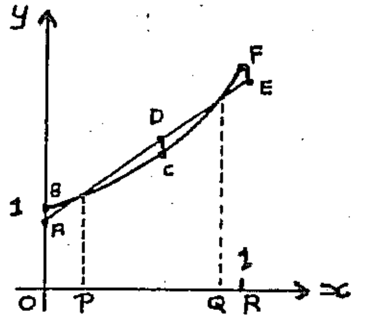
*(b) Find the best approximation to $f(x) = e^x$ from L_0 in $(C[0,1], \|\cdot\|_\infty)$.

That is, find m such that

$$\max_{0 \le x \le 1} |e^x - mx|$$

is smallest.

(IV) Let P_1 be the two dimensional subspace of $C[0,1]$ consisting of all polynomials of degree less than one equal to one.
It can be shown that the element (straight line) $y = mx + b$ is a best approximation from P_1 to $f(x) = e^x$ in $(C[0,1], \|\cdot\|_\infty)$ if and only if the line intersects $y = e^x$ in two points P, Q (see diagram) in such a way that the maximum absolute difference in value between the line and $y = e^x$ in each of the three regions OP, PQ and QR are equal, i.e., $|AB| = |CD| = |EF|$.



This is a special case of the very elegant <u>*Alternation Theorem*</u> *(Proved by Borel in 1905 and generalised by Young in 1907 to functions satisfying a Haar condition):*

$$P_n(x) = \sum_{m=0}^{n} a_m x^m$$

is the best n'th degree polynomial approximation to f *on* [a,b] *with respect to the Tchebyscheff norm if and only if there exists* $x_0 < x_1 < x_2 < \ldots < x_n$ *in* [a,b] *such that* $r(x_i) = -r(x_{i+1}) = \pm\|r\|_\infty$ *where* $r = f - P_n$.

<u>Exercise</u>

*(a) By use of diagrams, etc. give a heuristic argument which demonstrates the truth of the above statement about approximating e^x.

(b) Using the above facts show that the straight line of best approximation to $y = e^x$ in $(C[0,1], \|\cdot\|_\infty)$ is

$$y = (e-1)x + \frac{e - (e-1)\ln(e-1)}{2}$$

(V) Consider the problem of finding "best" solutions to the inconsistent systems of linear equations, $A\underset{\sim}{x} = \underset{\sim}{b}$ where A is an $n \times m$ matrix, $\underset{\sim}{x} \in R^m$, $\underset{\sim}{b} \in R^n$ (and often $n > m$, that is the system is overspecified). By a best solution to such a system we mean a vector $\underset{\sim}{x}$ for which $\|A\underset{\sim}{x}-\underset{\sim}{b}\|$ is a minimum, where $\|.\|$ is some suitable norm on R^n. This may be reformulated as follows. From basic linear algebra, we know that A corresponds to a linear mapping from R^m into R^n, whose range

$A(\mathcal{R}^m)$, is a subspace of $\mathcal{R}^n$ (spanned by the columns of A). Thus, setting $M = A(\mathcal{R}^m)$, the above problem amounts to finding $\underset{\sim}{m} \in M$ for which $\|\underset{\sim}{m}-\underset{\sim}{b}\|$ is a minimum, i.e. $\underset{\sim}{m}$ is a closest point of M to $\underset{\sim}{b}$. A "best" solution is then any $\underset{\sim}{x}$ for which $A\underset{\sim}{x} = \underset{\sim}{m}$ and at least one such $\underset{\sim}{x}$ exists as $\underset{\sim}{m}$ belongs to the range of A.

In the case of the euclidean norm $\|.\|_2$ the problem is essentially the statistical problem of linear regression analysis <u>or</u> that of finding a least-square "line" of best fit, through a set of data points in n-dimensional space (see later). We now obtain a solution for a simple instance of this general problem in the case of the Tchebyscheff norm, $\|.\|_\infty$.

<u>Tchebycheff solution for an "overspecified" system of equations in one unknown</u>

Given the system of equations, in the unknown x;

$$\begin{aligned} a_1x &= b_1 \\ a_2x &= b_2 \\ a_3x &= b_3 \\ &\cdots\cdots \\ a_nx &= b_n \end{aligned}$$

we seek x such that

$$\max\{|a_1x-b|,|a_2x-b|,\ldots,|a_nx-b_n|\}$$

is a minimum, <u>i.e.</u> we seek to minimise the maximum of the residuals

$$|a_ix - b_i| \quad (i=1,2,\ldots,n).$$

For this reason such problems are often referred to as *minimax* problems.

[<u>Note</u>: This problem is equivalent to seeking an element of the 1-dimensional subspace spanned by $\underset{\sim}{a} = (a_1,a_2,\ldots,a_n)$ which is closest to the vector $\underset{\sim}{b} = (b_1,b_2,\ldots,b_n) \in \mathcal{R}^n$ with respect to the norm $\|.\|_\infty$.]

To see how such a problem might arise in practice, imagine we wish to determine the "elastic constant (x)" of a given spring, by taking readings of the extension (b) caused by various loads (a), and assuming Hooke's law $b = xa$ applies.

For example, suppose we have taken two such readings and obtained the following values.

load (a gms wt)	extension caused (b cms)
2	2
½	1

Each reading determines a value of x, however because of "errors", Hooke's law produces the pair of inconsistent equations

$$2x = 2$$
$$\tfrac{1}{2}x = 1.$$

In such a situation it is not unreasonable to select the value of x which minimises the maximum of the residuals $|2x - 2|$ and $|\tfrac{1}{2}x - 1|$. That is x such that $\delta(x) = \max\{|2x - 2|, |\tfrac{1}{2}x - 1|\}$ is a minimum. Observing that, for any real number a

$$|a| = \max\{a, -a\}$$

we see that

$$\delta(x) = \max\{|a_1x - b_1|, |a_2x - b_2|, \ldots, |a_nx - b_n|\}$$

may be replaced by the more convenient expression

$$\delta(x) = \max\{a_1x-b_1, b_1-a_1x, a_2x - b_2, b_2-a_2x, \ldots, a_nx-b_n, b_n-a_nx\}.$$

The problem is then to find x for which $\delta(x)$ is a minimum.

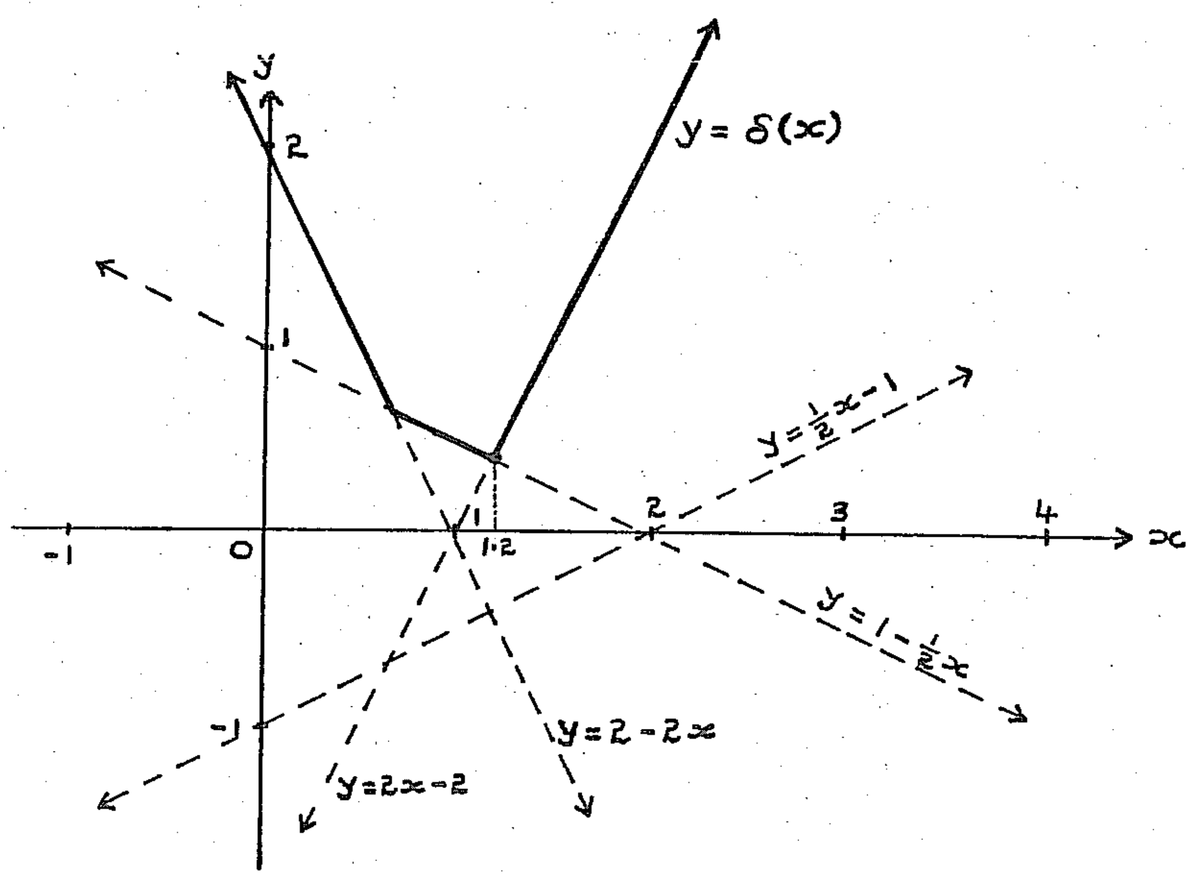
For the "spring" problem above, we must minimize

$$\delta(x) = \max\{2x-2, 2-2x, \tfrac{1}{2}x - 1, 1-\tfrac{1}{2}x\}.$$

By first graphing each of the straight lines

$$y = 2x - 2, \quad y = 2 - 2x, \quad y = \tfrac{1}{2}x - 1, \quad y = 1 - \tfrac{1}{2}x$$

it is not difficult to draw the graph of $y = \delta(x)$



from which, not insignificantly, we see that $\delta(x)$ has its minimum at the intersection of the two lines $y = 1 - \frac{1}{2}x$ and $y = 2x - 2$, whence our estimate for the spring's elastic constant is $x = 1\frac{1}{5}$ (cms/gm wt).

While this graphical method may be used to solve any similar problem, it is not readily *automated* in a way which allows it to be easily programmed for use on a computer, say, nor will it easily generalize to include more than one unknown, in which case the lines become planes (hyperplanes) in three (or higher) dimensional spaces.

Many more suitable algorithms have been developed, one of which is described below.

Descent from Vertex to Vertex.

We have seen that the problem, minimize

$$\delta(x) = \max\{|a_1x - b_1|, |a_2x - b_2|, \ldots, |a_nx - b_n|\}$$

can be rephrased as, minimize

$$\delta(x) = \{r_1(x), r_2(x), \ldots, r_{2n}(x)\}$$

where

$$r_1(x) = c_1x - d_1 \equiv a_1x - b_1$$
$$r_2(x) = c_2x - d_2 \equiv -a_1x + b_1$$
$$r_3(x) = c_3x - d_3 \equiv a_2x - b_2$$
$$r_4(x) = c_4x - d_4 \equiv -a_2x + b_2$$

etc.

(I.e., for $i = 1, 2, \ldots, n$, $r_{2i-1}(x) = c_{2i-1}x - d_{2i-1} \equiv a_ix - b_i$

and $r_{2i}(x) = c_{2i}(x) - d_{2i} \equiv -a_ix + b_i$.)

Thus for the above "spring Problem"

$$r_1(x) = 2x - 2$$
$$r_2(x) = 2 - 2x$$
$$r_3(x) = \tfrac{1}{2}x - 1$$
$$r_4(x) = 1 - \tfrac{1}{2}x.$$

As we develop the algorithm each step will be illustrated by reference to this particular problem.

Selecting any value x_0 as starting point -

(1) Determine the set M of indices i for which $r_i(x_0) = \delta(x_0)$

(i.e., identify those lines $y = r_i(x)$ which intersect with the graph of $\delta(x)$ when $x = x_0$.)

Note, $M \neq \emptyset$ as $\delta(x_0) = \max\{r_1(x_0), r_2(x_0) \ldots\}$.

E.g., for $x_0 = 1$ we have

$$\delta(1) = \max\{0, 0, -\tfrac{1}{2}, \tfrac{1}{2}\} = \tfrac{1}{2}$$

while

$$r_1(1) = 0 \text{ so } 1 \notin M$$

$$r_2(1) = 0 \text{ so } 2 \notin M$$

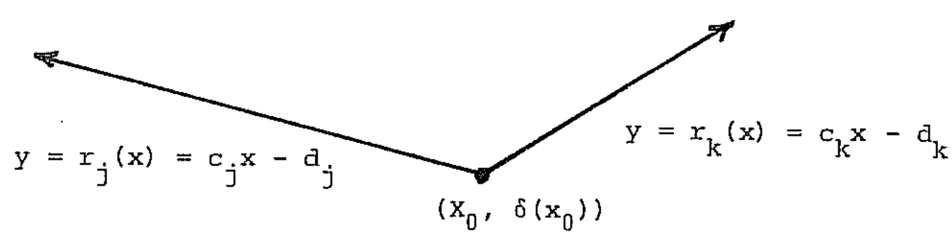
$$r_3(1) = -\tfrac{1}{2} \text{ so } 3 \notin M$$

$$r_4(1) = \tfrac{1}{2} \text{ so } 4 \in M$$

whence $M = \{4\}$.

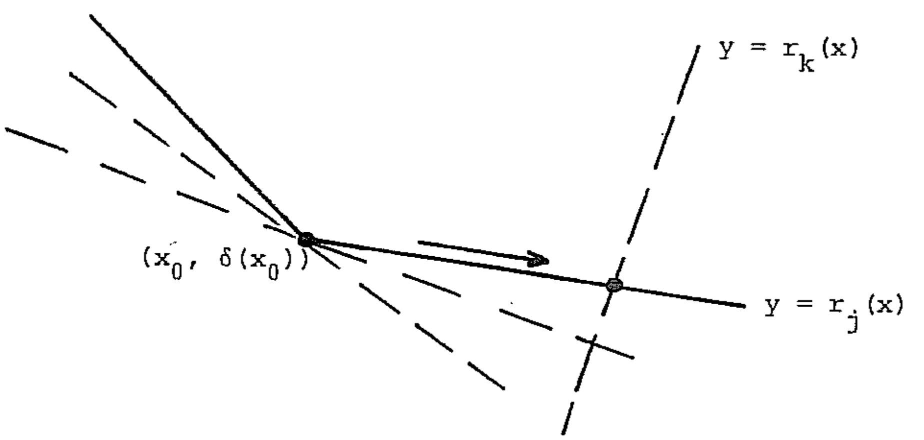
(2) If M contains indices j and k for which $c_j c_k < 0$, then x_0 is the desired value of x which minimized $\delta(x)$.

(To see this note that c_j and c_k have opposite sign and so at x_0 the two lines $y = r_j(x)$ and $y = r_k(x)$ must look like



Since the graph of $\delta(x)$ passes through $(x_0, \delta(x_0))$, the intersection of these two lines, and since $y = \delta(x)$ must lie "above" each of the lines $(\delta(x) = \max\{r_1(x), r_2(x), \ldots\})$, $\delta(x)$ must have a local minimum at x_0, but then by its convexity this must be a global minimum (see subsequent exercise) as required.

If M does not contain such a pair of indices (as, for example, is the case in our numeric example, where M = {4}), then determine $j \in M$ for which $|c_j|$ is smallest (in our example we have only one choice, viz $j = 4$) and depending on whether c_j is positive or negative determine the nearest value of x to the left or right of x_0 for which $r_k(x) = r_j(x)$ for some $k \neq j$. (I.e., slide along the line, $y = r_j(x)$, through $(x_0, \delta(x_0))$ with shallowest slope in the "down-hill" direction,



which must coincide with the graph of $\delta(x)$, until it intersects another line at which point $\delta(x)$ may change "direction".)
E.g., for our example, $c_4 = -\frac{1}{2}$ so we seek the value of x to the right of 1 and nearest to it for whcih

$$1 - \tfrac{1}{2}x = r_k(x) \qquad (k = 1 \text{ or } 2 \text{ or } 3).$$

Now

$$1 - \tfrac{1}{2}x = r_1(x) = 2x - 2 \Rightarrow x = \frac{6}{5} \text{ (to right of 1 as sought)}$$

$$1 - \tfrac{1}{2}x = r_2(x) = 2 - 2x \Rightarrow x = \frac{2}{3} \text{ (to left of 1)}$$

$$1 - \tfrac{1}{2}x = r_3(x) = \tfrac{1}{2}x - 1 \Rightarrow x = 2 \text{ (to right of 1 as sought)}$$

so the sought after value of x is $x = \frac{6}{5}$.

Using the value of x so obtained as our next estimate (x_1) return to step (1).

E.g. In step 1 with x_0 replaced by $x_1 = \frac{6}{5}$ we have

$$\delta\left(\frac{6}{5}\right) = \max\left\{\frac{2}{5}, -\frac{2}{5}, -\frac{2}{5}, \frac{2}{5}\right\}$$

so

$$M = \{1, 4\}$$

and proceeding to step (2) we see

$$c_1 c_4 = 2 \times -\tfrac{1}{2} = -1 < 0.$$

So $x = \frac{6}{5}$ is the required value of x (see before).

This and several similar algorithms are available for solving the problem in one unknown and extensions of the theory by Pólya, Haar and others provide algorithms for handling problems in more than one unknown. All of these procedures are intimately related to the general theory of convex programming which is of considerable importance in economic theory and operations research type problems.

Exercises:

1) Determine a Tchebycheff solution to the "over-specified" system of equations

$$2.0x = 1.2$$
$$4.0x = 2.1$$
$$5.0x = 2.6.$$

2) Find x such that $\|\tilde{y}\|_\infty$ is a minimum where

$$\tilde{y} = Ax - \tilde{b} \quad \text{and} \quad A = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \qquad \tilde{b} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

---

3) If you have the facilities you might try developing a programme which uses the above algorithm for solving an overspecified system of equations in one unknown.

4) Give an example to show that the vector x for which $\|Ax - b\|_\infty$ is minimum need not be unique.

5) DEFINITION: A function f: $\mathbb{R} \to \mathbb{R}$ is termed <u>convex</u> if its supergraph $\{(x,y) \in \mathbb{R}^2: \ y \geq f(x)\}$ is a convex subset of $\mathbb{R}^2$.

a) Show that a function of the form $y = |ax + b|$ is convex, where a, b are constants.

b) Hence, show that the function $y = \delta(x)$ is convex. (HINT: use Exercise 4 on page 79.)

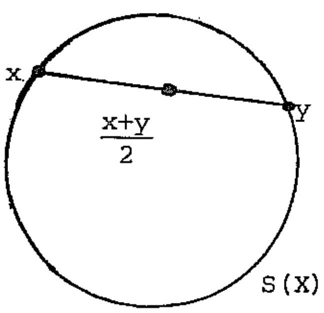
6)* Show that a local minimum of a convex function is necessarily a global minimum. (HINT: Draw a diagram.)

We have seen that compactness is a sufficient condition for a set A to be proximal. As illustrated by II on page 86, to ensure that the best approximation from A to any point of X is unique we must further restrict the nature of A. For example, by requiring that A be convex. A subset with the above property; that it contains a unique best approximation to each point of the space is termed a <u>Tchebyscheff set</u>.

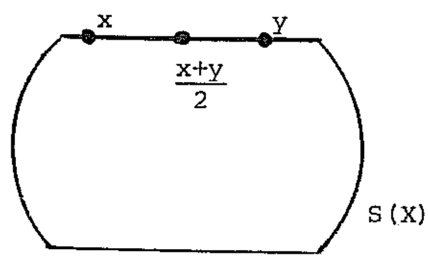
The question of characterizing Tchebyscheff subsets remains, in general, unanswered (though it has been the subject of considerable research). For example, for many spaces (including Hilbert spaces) it is not known whether convexity is a necessary condition for a closed set to be Tchebyscheff, a characterization of Tchebyscheff sets in $(C[a,b], \|\cdot\|_\infty)$ *has only recently been given (1975) by A.L. Brown and others.*

As demonstrated by illustration III on page 86 the convexity of A alone is not a sufficient condition for A to be a Tchebyscheff set. We must also ~~place restrictions on the nature of the norm (or possibly impose stronger~~ restrictions on A).

DEFINITION: A normed linear space $(X, \|\cdot\|)$ is *strictly convex* (or rotund) if whenever two elements x and y are such that $\|x\| = \|y\| = \|\frac{x+y}{2}\|$ we must have $x = y$. Excluding the trivial case of $x = y = 0$ we may, by dividing throughout by the common value of the norm, assume without loss of generality that $\|x\| = \|y\| = 1$ and then X is strictly convex if and only if $\|\frac{x+y}{2}\| \lneq 1$ whenever x and y are two distinct elements of the unit sphere $S(x) = \{x \in X : \|x\| = 1\}$. Geometrically this states that the unit shpere does not contain any non-trivial line segments $[x,y]$ - can you prove this?



The unit sphere of a strictly convex space



The unit sphere of a space which is not strictly convex

EXAMPLES:

1) From the parallelogram rule

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

we see that if $\|x\| = \|y\| = \|\frac{x+y}{2}\|$, then $\|x - y\| = 0$ or $x = y$. Thus, *every inner-product space is strictly convex.*

2) *The space ℓ_∞^2 is not strictly convex.* To see this note that $\underset{\sim}{x} = (-1,1)$ and $\underset{\sim}{y} = (1,1)$ are points in $S(\ell_\infty^2)$ and that $\|\frac{x+y}{2}\|_\infty = \|(0,1)\|_\infty = 1$.

EXERCISE: Show that none of the spaces ℓ_1^2, $(C[a,b], \|\cdot\|_\infty)$, or ℓ_∞ are strictly convex.

REMARK: *It may be proved that all of the spaces ℓ_p with $1 < p < \infty$ are strictly convex spaces.*

The importance of strict convexity for our purpose lies in the following result.

THEOREM 3: *Let $(X, \|\cdot\|)$ be a strictly convex normed linear space and ~~let A be a convex subset of X. Then, given any point $x_0 \in X$, if there~~ is a best approximation to x_0 from A it is unique.*

Proof: Assume a_1 and a_2 are two best approximations to x_0 from A; that is

$$\|x_0 - a_1\| = \|x_0 - a_2\| \leq \|x_0 - a\| \qquad \text{for all } a \in A.$$

Let $r = \|x_0 - a_1\|$ and let $x = x_0 - a_1$, $y = x_0 - a_2$, then $\|x\| = \|y\| = r$. Since A is convex, $\frac{a_1 + a_2}{2} \in A$ and so

$$r = \|x_0 - a_1\| \leq \|x_0 - \frac{a_1 + a_2}{2}\| .$$

But $x_0 - \frac{a_1 + a_2}{2} = \frac{x + y}{2}$ and so we have

$$\begin{aligned} r &\leq \|\frac{x + y}{2}\| \\ &= \|x_0 - \frac{a_1 + a_2}{2}\| \\ &= \|\frac{(x_0 - a_1) + (x_0 - a_2)}{2}\| \\ &\leq \frac{\|x_0 - a_1\| + \|x_0 - a_2\|}{2} \\ &= r. \end{aligned}$$

Thus, $\|x\| = \|y\| = r = \|\frac{x + y}{2}\|$ and so by the strict convexity of X we must have $x = y$ or $x_0 - a_1 = x_0 - a_2$ and so $a_1 = a_2$ as required. ∎

As an immediate corollary we have

THEOREM 4: *Every compact convex subset of a strictly convex normed linear space is a Tchebyscheff set.*

EXERCISES:

1) Show that every closed convex subset (in particular, every subspace) of a finite dimensional inner-product space is a Tchebyscheff subset.
2) If A is a convex subset of a normed linear space $(X, \|\cdot\|)$ and if a_1, a_2 are two distinct best approximations from A to the point $x_0 \in X$, show that every point of the line segment $[a_1, a_2]$ is also a best approximation from A to x_0. *Intuitively, this shows that the sphere, centre x_0, which "just touches A" contains the line segment $[a_1, a_2]$ and suggest why there can not be two distinct best approximations in a strictly convex space.*

By further restricting the natrue of the norm function it is possible to relax the conditions on A while still ensuring that it is proximal (indeed Tchebyscheff). Thus effecting a kind of "trade-off":- more properties on the norm for less on the set, and possibly vice-versa.

DEFINITION: A normed linear space $(X,\|.\|)$ is *uniformly convex* if whenever (x_n) and (y_n) are two sequences of elements from X such that $\|x_n\| \to 1$, $\|y_n\| \to 1$ and $\left\|\frac{x_n+y_n}{2}\right\| \to 1$ as $n \to \infty$, we must have $\|x_n-y_n\| \to 0$.

EXAMPLE: Rearranging the parallelogram rule we have

$$\|x_n-y_n\|^2 = 2\|x_n\|^2 + 2\|y_n\|^2 - \|x_n+y_n\|^2.$$

Thus if $\|x_n\|$, $\|y_n\|$ and $\left\|\frac{x_n+y_n}{2}\right\| \to 1$ we have

$$\|x_n - y_n\|^2 = 2\|x_n\|^2 + 2\|y_n\|^2 - 4\left\|\frac{x_n+y_n}{2}\right\|^2$$

$$\to 2 + 2 - 4 = 0$$

and so we conclude that *every inner-product space is uniformly convex.*

REMARK: *The notion of uniform convexity was introduced by J.A. Clarkson in 1936 as part of an endeavour to find a larger class of spaces for which certain Hilbert space results would remain valid.*

PROPOSITION 5: *Every uniformly convex space is strictly convex.*

Proof. Let x and y be such that $\|x\| = \|y\| = \|\frac{x+y}{2}\|$ (=1, without loss of generality). We must show x=y.

Consider the constant sequences $x_n \equiv x$ and $y_n \equiv y$ $(n=1,2,\ldots)$ then $\|x_n\| \to 1$ (indeed $\|x_n\| = 1$ for all n) and similarly $\|y_n\| \to 1$, $\|\frac{x_n+y_n}{2}\| \to 1$, so by uniform convexity $\|x_n-y_n\| \to 0$, but $\|x_n-y_n\|=\|x-y\|$ for all n and so we conclude that $\|x-y\| = 0$ or x=y as required.

■

The converse of this proposition is not true in general. However, for finite dimensional spaces the notions of uniform and strict convexity coincide.

**EXERCISES: 1) Show that a strictly convex finite dimensional normed linear space is uniformly convex.

[Hint: If $\|x_n\|, \|y_n\|$ and $\left\|\frac{x_n+y_n}{2}\right\| \to 1$ but $\|x_n - y_n\| \not\to 0$ there exist subsequences (x_{n_k}) and (y_{n_k}) with $\|x_{n_k} - y_{n_k}\| \geq \epsilon > 0$ for some $\epsilon > 0$ and all k. Use the compactness of closed bounded sets to extract further subsequences $(x_{n_{k_j}})$ and $(y_{n_{k_j}})$ with $x_{n_{k_j}} \to x$ and $y_{n_{k_j}} \to y$. Conclude that $1 = \|x\| = \|y\| = \|\frac{x+y}{2}\|$, but that $\|x-y\| \geq \epsilon$, contradicting the strict convexity of X.]

*2) (a) Let X be the space of all bounded sequences and let T be the 1-1 function which maps the sequence $x_0, x_1, x_2, \ldots, x_n, \ldots$ to the sequence $x_0, x_1/2, x_2/4, \ldots, x_n/2^n, \ldots$

For each $\underset{\sim}{x} \in X$ show that $T(\underset{\sim}{x})$ is a square summable sequence and conclude that

$$\begin{aligned}\|\underset{\sim}{x}\| &= \|\underset{\sim}{x}\|_\infty + \|T\underset{\sim}{x}\|_2 \\ &= \sup_n |x_n| + \left(\sum_{n=1}^{\infty} 2^{-2n} x_n^2\right)^{\frac{1}{2}}\end{aligned}$$

defines a norm on X (indeed an equivalent norm to $\|.\|_\infty$).

(b) Show that with this norm X is strictly convex.

(c) Show that X with the above norm is not uniformly convex.

[Hint: consider the sequences $(\underset{\sim}{x}_n)$ and $(\underset{\sim}{y}_n)$ where

$$\underset{\sim}{x}_n = \underbrace{0,0,\ldots,0}_{\text{first } n \text{ terms zero}},\underbrace{1,1,\ldots,1}_{\text{a block of } n \text{ one's}},0,0,\ldots$$

and $\underset{\sim}{y}_n = \underset{\sim}{x}_{n+1}$.]

Before proceeding to our main result on uniformly convex spaces we need the following lemma and its corollary.

LEMMA 6: *If* $(X,\|.\|)$ *is a uniformly convex normed linear space and* (t_n) *is such that* $\|t_n\| \to 1$ *and* $\left\|\frac{t_n+t_m}{2}\right\| \to 1$ *as* $n,m \to \infty$, *then* (t_n) *is a Cauchy sequence.*

Proof. Assume (t_n) is not a Cauchy sequence; that is, for some $\epsilon_0 > 0$ and each $N \in \mathbb{N}$ there exists a pair of indicies n_N and m_N both greater than N with $\|t_{n_N} - t_{m_N}\| \geq \epsilon_0$. (Otherwise, for each $\epsilon_0 > 0$ there would exist an $N \in \mathbb{N}$ such that $\|t_n - t_m\| < \epsilon_0$ for all $n,m > N$ and so (t_n) would be a Cauchy sequence.) Since $n_N, m_N > N$ we have n_N and $m_N \to \infty$ as $N \to \infty$ and so

$$\|t_{n_N}\|,\ \|t_{m_N}\| \text{ and } \left\|\frac{t_{n_N}+t_{m_N}}{2}\right\| \to 1 \text{ as } N \to \infty, \text{ by the assumptions on } (t_n).$$

Now, let $x_N = t_{n_N}$ and $y_N = t_{m_N}$ $(N=1,2,3,\ldots)$, then

$$\|x_N\|,\ \|y_N\| \text{ and } \left\|\frac{x_N+y_N}{2}\right\| \to 1 \text{ as } N \to \infty.$$

So by the uniform convexity of X we must have

$\|x_N - y_N\| \to 0$, but $\|x_N - y_N\| = \|t_{n_N} - t_{m_N}\| \geq \epsilon_0$ for all $N \in \mathbb{N}$, a contradiction, showing the assumption; (t_n) is not a Cauchy sequence, must be false. ∎

Corollary 7: *If* $(X,\|.\|)$ *is a uniformly convex normed linear space and* (z_n) *is such that* $\|z_n\| \to k$ *and* $\left\|\frac{z_n+z_m}{2}\right\| \to k$ *as* $n,m \to \infty$, *then* (z_n) *is a Cauchy sequence.*

Proof. If k=0 then $z_n \to 0$ and the result follows. Thus suppose $k \neq 0$ and let $t_n = z_n/k$, then (t_n) is such that $\|t_n\| \to 1$, $\left\|\frac{t_n + t_m}{2}\right\| \to 1$ as $n,m \to \infty$ and so, by lemma 6, (t_n) is such that $\|t_n - t_m\| \to 0$ as $n,m \to \infty$. But,

$$\|z_n - z_m\| = k\|t_n - t_m\|$$

so $\|z_n - z_m\| \to 0$ as $n,m \to \infty$ and we conclude that (z_n) is a Cauchy sequence as required. ∎

THEOREM 8: *A closed convex subset of a uniformly convex Banach space contains a unique best approximation to every point of the space.* That is, every closed convex subset of a uniformly convex Banach space is a Tchebyscheff subset.

Proof. Let x_0 be a given point in the uniformly convex Banach space $(X, \|.\|)$ and $A \subseteq X$ a closed convex subset. If $d = \inf_{a \in A} \|a - x_0\|$, then there exists a sequence of points (a_n) of A such that $\|a_n - x_0\| \to d$, further, by the convexity of A, $\frac{1}{2}(a_n + a_m) \in A$ so

$$d \leqslant \|\tfrac{1}{2}(a_n + a_m) - x_0\|$$
$$\leqslant \|\tfrac{1}{2}(a_n - x_0)\| + \|\tfrac{1}{2}(a_m - x_0)\| \quad \text{(by the triangle inequality)}$$
$$\to \tfrac{1}{2}d + \tfrac{1}{2}d$$

i.e. $\|\tfrac{1}{2}[(a_n - x_0) + (a_m - x_0)]\| \to d$.

Thus, by Corollary 7 above, $(a_n - x_0)$ is a Cauchy sequence, i.e. $\|(a_n - x_0) - (a_m - x_0)\| = \|a_n - a_m\| \to 0$ as $n, m \to \infty$ and so (a_n) is itself a Cauchy sequence. But X is complete, hence, there exists $a_0 \in X$ such that $a_n \to a_0$, further, since A is closed $a_0 \in A$. Whence, $d \leqslant \|a_0 - x_0\| \leqslant \|a_0 - a_n\| + \|a_n - x_0\| \to 0 + d$ as $n \to \infty$, so $\|a_0 - x_0\| = d$ and a_0 is a best approximation to x_0 from A. That this best approximation is unique follows immediately from Proposition 5 and Theorem 3. ∎

Best approximations from subspaces, particularly of a Hilbert space.

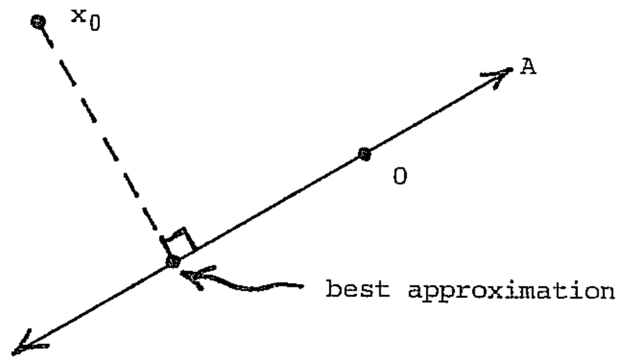
We have already encountered this situation in example (V) on page 91, where a solution was obtained (descent from vertex to vertex algorithm) for the case of a one-dimensional subspace of ℓ_∞^n.

We begin with an observation (due to Pták in 1958) that a_0 *is a best approximation to* x *from the subspace* A *of the normed linear space* $(X, \|.\|)$ *if and only if* $\|x_0 - a_0\| \leqslant \|(x_0 - a_0) + \lambda a\|$ *for all* $a \in A$ and $\lambda \in \mathbb{R}$. That is, if and only if $x_0 - a_0$ is orthogonal to each element of A in the generalized sense of Exercise 1)(ii) on page 22. (To prove this observation it suffices to note that $-a_0 + \lambda a$ is an element of A (as A is a subspace) and that for each $\lambda \in R$ every element a' of A may be written in this form by taking $a = \frac{a' + a_0}{\lambda}$.

So, $\|x_0 - a_0\| \le \|(x_0 - a_0) + \lambda a\| \Leftrightarrow \|x_0 - a_0\| \le \|x_0 - a'\|$ for all $a' \in A$.

$\Leftrightarrow a_0$ is a best approximation to x_0 from A.)

Intuitively, we may interpret this observation as: A best approximation to x_0 from A is, in the appropriate sense, the foot of a perpendicular from x_0 to the subspace A.



Since a subspace A is necessarily a convex subset, in the case of an inner-product space, if a best approximation from A to x_0 exists it must necessarily be unique [Example 1 on page 100 and Theorem 3] and the above observation may be restated as follows.

PROPOSITION 10: *For an inner-product space* X, a_0 *is the best approximation from the subspace* A *to* x *if and only if* $\langle x_0 - a_0, a\rangle = 0$ *for all* $a \in A$.

This follows from exercise 1(ii) on p.22, where you were asked to show that in an inner-product space the generalized notion of orthogonality use above corresponds to the usual one: x is orthogonal to a if $\langle x,a\rangle = 0$. For completeness we now prove this.

$\langle x,a\rangle = 0$ *if and only if* $\|x\| \le \|x + \lambda a\|$ *for all* $\lambda \in R$