Journal of the Orissa Mathematical Society ISSN:0975-2323
Volume 40, No.01-02, 2021, 124-137

APPROXIMATION AND OPTIMIZATION

C NAHAK

Department of Mathematics, IIT Kharagpur, Kharagpur, India-721302

Abstract. This paper deals with the concept and relation between the theory of approximation and optimization problems.

Keywords: Approximation; Optimization; Unperturbational theory

2020 AMS Subject Classification: 90C26, 90C30, 90C32, 90C46, 90C47

1. Approximation in norm spaces

Approximation theory is concerned with the approximation of functions of a certain kind (for instance, continuous function on some interval) by other (probably simpler) functions (for example, polynomials) such situation already aries in calculus: if a function has a Taylor series, we may regard and the partial sums of series as approximation.

How to get the quality of such approximation. To get the quality we have the following results:

Given a set X of functions to approximate and a set Y of functions by which the elements of X are to be approximated, one may consider the problems of existence, uniqueness and constructing of a "best approximation" in the sense of such a criterion. A natural setting for the problem of approximation is as follows.

E-mail addresses: cnahak@maths.iitkgp.ac.in

Let $X = (X, \|\cdot\|)$ be a normed space and suppose that any given $x \in X$ is to be approximated by a $y \in Y$, where Y is a fixed subspace of X. Suppose δ denotes the distance from x to Y. By definition $\delta = \delta(x, y) = \inf_{y \in Y} \|x - y\| = \inf_{y \in Y} d(x, Y)$. Clearly, δ depends on both x and Y, which we keep fixed, so that the simple notation δ is in order.

If there exists a $y_0 \in Y$ such that

$$\|x - y_0\| = \delta$$

then y_0 is called a best approximation to x out of Y.

We see that a best approximation y_0 in element of minimum distance from the given x. Such a $y_0 \in Y$ may or may not exits, this raises the problem of existence. The problem of uniqueness is of practical interest too, since for a given x and Y there many be more than one best approximation, as we can see.

In many applications, Y will be finite dimensional. Then we have the following.

Theorem 1.1 (Existence Theorem (Best approximation))**.** *If Y is a finite dimensional subspaces of normed space $X = (X, \|\cdot\|)$, then for each $x \in X$ there exits a best approximation to x out of Y.*

Example 1.1. Suppose $X = C[a, b]$. A finite dimensional subspaces of X is

$$Y = span\{x_0, x_1, ..., x_n\}, x_j(t) = t^j \quad (n \text{ fixed}, j = 0, 1, 2, \cdots n).$$

This is the set of all polynomials of degree at most n, together with $x = 0$ (for which no degree is defined in the usual discussion of degree).

By Theorem 1.1, for a given continuous function on $[a, b]$ there exits a polynomial P_n of degree at most n such that for every $y \in Y$,

$$\max_{t \in [a,b]} |x(t) - p_n(t)| \leq \max_{t \in [a,b]} |x(t) - y(t)|.$$

Approximation in C[a,b] is called uniform approximation.

Note that the finite dimensionality of Y in Theorem 1.1 is essential. In fact, let Y be the set of all polynomials on $[0, \frac{1}{2}]$ of any degree, considered as a subspace of $C[0, \frac{1}{2}]$. Then $\dim(Y) = \infty$. Let $x(t) = \frac{1}{1-t}$. Then for every $\epsilon > 0$ there is an N such that, setting

$$y_n(t) = 1 + t + ... + t^n,$$

we have $\|x - y_n\| < \epsilon$ for all $n > N$. Hence $\delta(x, Y) = 0$. However, since x is not a polynomial, we see that there is no y_0 satisfying $\delta = \delta(x, Y) = \|x - y_0\|.=0$

Uniqueness, strict convexity

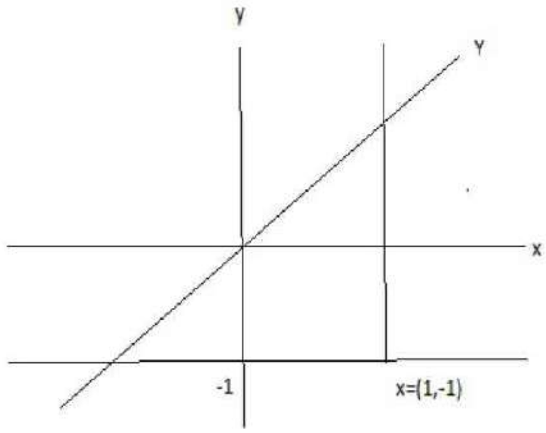
Example 1.2. If $X = R^3$ and Y is xy-plane ($z = 0$), then we know that for a given point $x_0 = (x_{10}, y_{10}, z_{10})$ is best approximation out of Y is the point $x_0 = (x_{10}, y_{10}, 0)$, the distance from x_0 to Y is $\delta = |z_{10}|$ and that best approximation y_0 is unique.

Example 1.3. Let $X = (X, \|\cdot\|_1)$ be the vector space of order pair $x = (x_1, y_1)$ of real numbers with norm defined by

$$\|x\|_1 = |x_1| + |y_1|.$$

Let us take a point $x = (1, -1)$ and the subspace Y defined as

$$Y = \{y = (\eta, \eta) : \eta \text{ is real}\}.$$



Then for all $y \in Y$, we have

$$\|x - y\|_1 = |1 - \eta| + |-1 - \eta| \geq 2.$$

The distance from x to Y is $\delta(x, Y) = 2$, and all $y = (\eta, \eta)$ with $|\eta| \leq 1$ are the best approximations to x out of Y. This illustrates that even in such a simple space, for given x and Y we can have several best approximations, even infinitely may of them.

Lemma 1.1. ***(Convexity:)*** *In a normed space $(X, \| \cdot \|)$ the set M of best approximation to a given point x, out of a subspace Y of X is convex.*

The theory of optimization presented in this note is derived from a few simple, intuitive geometric relations. The existence of these relations to infinite-dimensional spaces is the motivation for the mathematics of functional analysis which, in a sense, often enables us to extend our three dimensional spaces to complete infinite dimensional problems. This is the conceptual utility of functional analysis. We briefly describe a few of the important geometric principles of optimization that are developed in detail in optimization theory.

(i) **The projection theorem:** This theorem is one of the simplest and nicest results of optimization theory. In arbitrary three-dimensional Euclidean space, it states that the shortest line from a point to a plane is furnished by the perpendicular from the point to the plane, as given by :

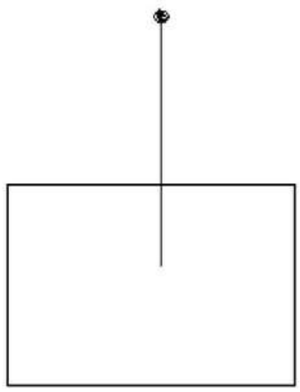


Fig. 1

This results has a direct extension in higher dimensional and infinite dimensional Hilbert space. In the generalized form, this optimization principle forms the basis of all least-squares approximation, control and estimation procedures.

(ii) **The Hanh-Banach theorem:** One of the many results and concepts in functional analysis, the one theorem dominating the theme of this subject is the Hanh-Banach theorem. This theorem takes several forms.

One version extends the projection theorem to problems having nonquadratic objectives. In this way the simple geometric interpretation is presented for those more complex problems. Another version of the Hanh-Banach theorem states (in simple form) that given a sphere and a point not in the sphere there is hyperplane separating the point and the sphere.

(iii) **Duality:** There are several duality principles in optimization theory that relates a problem expressed in terms of vectors in a space to a problem expressed in terms of hyperplanes in the space. This concept of duality is the theme in optimization theory. Many of the duality principles are based on the geometric relation illustrated in Fig 2.

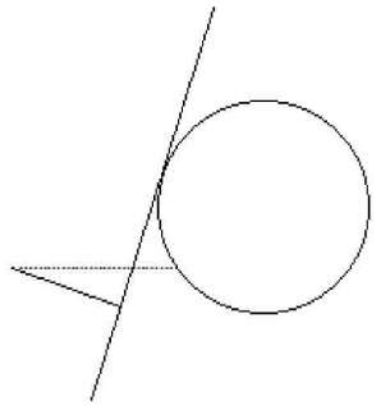


Fig. 2

The shortest distance from a point to a convex set is equal to the maximum of the distances from the point to a hyperplane separating the point from the convex set. Thus, the original minimization over vectors can be converted to maximization over hyperplanes.

(iv) Differentials perhaps the most familiar optimization technique is the method of differential calculus setting the derivative of the objective function equal to zero. The technique is discussed for a single or, perhaps, finite number of variables in most elementary courses of differential calculus. Its extension to infinite-dimensional space is straight forward and in that form, it can be applied to a variety of interesting optimization problems. Most of the classical theory of calculus of variations can be viewed as a consequences of this principle. The geometric interpretation of the technique for one-dimensional problem is obvious. At a maximum or minimum the target to the graph of a function is horizontal. In higher dimensions the geometric interpretation is familiar at a maximum or minimum the tangent hyperplane to the graph is horizontal. This again we are led to observe the fundamental role of hyperplanes in optimization.

2. Motivation and Preliminary results

For some time, approximation theory and optimization theory have developed independently, in parallel. In 1960s it was observed that optimization, i.e., the minimization or maximization of a function, contains approximation as a particular case. Indeed, approximation is the minimization or maximization of a particular function on a normed linear space X, namely the function

$$f(y) = \| x_0 - y \|, \ (y \in X)$$

. Thus in the 1970s there appeared naturally the idea of studying them together in this spirit, as reflected for example by titles like book of Laurent [6] Approximation and optimisation, Holmes [3], A course on optimization and best approximation, Glashoff and Gustafson [2], Linear optimization and approximation etc. The same point of view also appeared in parts of other monographs on optimization theory. On the other hand going in the opposite direction, Cheney and Goldstein [1] have extended a result on the existence of best approximations to a result on the existence of optimal solutions of minimization problems. Starting with [4], [5], there was suggested and systematically carried out a program of work in this direction, namely, to show that many methods and results of approximation theory are so strong that they can be generalized to yield new methods and results in optimization theory. In the present work we shall study these two theories and their interactions, going from approximation to optimization and vice versa.

It has long been known that duality is a powerful tool in the study of approximation and optimization problems. For problems of approximation in a normed linear space X, namely of minimization or maximization of the distance to a given subset of X, "duality" means simply their study with the aid of the elements of the conjugate space X^*. In a general setting "duality" in optimization means the simultaneous study of a pair of optimization problems, related in some way, namely, the initial problem called the "Primal Problem" of maximization or minimization of a function on a subset of a norm linear space X, and the "Dual Problem" of minimization or maximization of a function on a subset of a norm linear space W with the aim of obtaining more information on primal problem (or its "optimal values" or its "optimal solutions" etc.). In general (with few exceptions) W is a set of functions on X, or alternatively W is an arbitrary set, but paired with X with the aid of a function on the cartesian product $X \times W$ called a "Coupling function". If G is a subset of a norm linear space X, and $g_0 \in G$ for which $\text{dist}(x_0, G) := \inf_{g \in G} \| x_0 - g \|$ is attained i.e. $\| x_0 - g_0 \| = \inf_{g \in G} \| x_0 - g \|$ or equivalently, such that

$$\| x_0 - g_0 \| \leq \| x_0 - g \|, \ (g \in G),$$

is called an element of best approximation of (or a nearest point to) x_0 in G. We shall denote by $\mathcal{P}_G(x_0)$ the set of all nearest points to x_0 in G i.e.

$$\mathcal{P}_G(x_0) = \left\{ g_0 \in G \,\middle|\, \| x_0 - g_0 \| = \inf_{g \in G} \| x_0 - g \| \right\}.$$

We shall denote by max (res, min) any sup (res, inf) that is attained. Clearly $\mathcal{P}_G(g_0) = \{g_0\}, \forall g_0 \in G$. In finite dimensional space X, if $G \subseteq X$ is closed, $\mathcal{P}_G(x_0) \neq \emptyset, \forall x_0 \in X$, but in infinite dimensional norm linear space X, we may have $\mathcal{P}_G(x_0) = \emptyset$ i.e. elements of best approximation of x_0 need not exist, even if for closed set G with "very good " geometric properties.

One can also express $\mathcal{P}_G(x_0)$ with the aid of (closed) ball

$$B(x_0, d) = \left\{ y \in X \middle| \parallel x_0 - y \parallel \leq d \right\}$$

with centre x_0 and radius $d = dist(x_0, G)$, namely,

$$\mathcal{P}_G(x_0) = G \cap B(x_0, dist(x_0, G))$$

We shall be concerned with the following two main problems:

(1) Find convenient formula for $dist(x_0, G)$

(2) Give the characterizations of elements of best approximation i.e. necessary and sufficient condition in order that an element $g_0 \in G$, $g_0 \in \mathcal{P}_G(x_0)$. For these problems, " duality " means simply their study with the aid of the elements of the conjugate space X^*. More natural is that best approximation by a convex set is a particular case of convex optimization, namely, it is the infimization of the convex function $f : X \to \mathbb{R}$

$$f(y) :=\parallel x_0 - y \parallel, \; x_0 \in X$$

The following is the basic formula for the distance to a convex set.

Theorem 2.1. *Let X be a normed linear space, G be a convex subset of X, and $x_0 \in \overline{G}^c$. Then*

$$dist(x_0, G) = \max_{\phi \in X^*, \|\phi\|=1} \{\phi(x_0) - \sup \phi(G)\}$$

In other words, we have

$$dist(x_0, G)) \geq \phi(x_0) - \sup \phi(G)\,, (\phi \in X^*, \parallel \phi \parallel = 1)$$

and there exists $\phi_0 \in X^*$ such that $\parallel \phi_0 \parallel = 1$ and $dist(x_0, G) = \phi_0(x_0) - \sup \phi_0(G)$.

The formula $dist(x_0, G) = \max_{\phi \in X^*, \|\phi\|=1} \{\phi(x_0) - \sup \phi(G)\}$ can take simpler form for linear spaces G.

Theorem 2.2. *Let X be a norm linear space, G be a linear subspace of X, $x_0 \in X \setminus \overline{G}$ and $g_0 \in G$. We have $g_0 \in \mathcal{P}_G(x_0)$ if and only if there exists an $f \in X^*$ such that*

(1) $\| f \| = 1$;

(2) $f(g) = 0$;

(3) $f(x_0 - g_0) = \| x_0 - g_0 \|$

.

Proof. Assume that $g_0 \in \mathcal{P}_G(x_0)$. Then since $x_0 \in X \setminus \overline{G}$, we have $dist(x_0, G)$ $= \| x_0 - g_0 \| > 0$ and hence, by a corollary of Hahn-Banach theorem, there exists $f_0 \in X^*$ such that $\| f_0 \| = \frac{1}{\|x_0 - g_0\|}$, $f_0(g) = 0$ $(g \in G)$, $f_0(x_0) = 1$. Then the functional $f = \| x_0 - g_0 \| f_0 \in X^*$ satisfies all the above three properties. conversely, if there is an $f_0 \in X^*$ satisfying $(1) - (3)$ then for any $g \in G$ we have

$$\| x_0 - g_0 \| = f(x_0 - g_0) = | f(x_0 - g_0) | \leq \| f \| \| x_0 - g_0 \| = \| x_0 - g_0 \|$$

and hence $g_0 \in \mathcal{P}_G(x_0)$, which completes the proof. □

We also have the following characterization of the Theorem 1.

Theorem 2.3. *Let X be a norm linear space, G a convex subset of X and $x_0 \in \overline{G}^c$ i.e. $x_0 \in X \setminus \overline{G}$. For an element $g_0 \in G$, the following statements are equivalent:*

$1.^o$ $g_0 \in \mathcal{P}_G(x_0)$.

$2.^o$ *There exists $\phi_0 \in X^*$, with $\| \phi_0 \| = 1$ and $\phi_0(x_0) - \sup \phi_0(G) = \| x_0 - g_0 \|$*

3.^{o} *There exists* $\phi_0 \in X^*$, *with* $\| \phi_0 \| = 1$ *and* $\phi_0(x_0 - g) \geq \| x_0 - g_0 \|$ $(g \in G)$
4.^{o} *There exists* $\phi_0 \in X^*$, *with* $\| \phi_0 \| = 1$ *and* $\phi_0(g_0) = \max \phi_0(G)$, $\phi_0(x_0 - g_0)$
$= \| x_0 - g_0 \|$ $(g \in G) - (B)$

Moreover, we can take the same ϕ_0 in statements 2.^{o}, 3.^{o}, and 4.^{o}.
Any function satisfying $\| \phi_0 \| = 1$ and (B) is called a "maximal function" of the element $x_0 - g_0$. The usefulness of Theorem 3 for applications in various concrete norm linear space is due to the fact that for these spaces the general form of maximal functions of the elements of the space is well known and simple.

3. Unperturbational Theory

Assume that we are given a constrained "primal" problem, called a problem of " problem of infimization ".

$$(P): \ \alpha = \inf f(G), \tag{3.1}$$

where the "constraint set" G is a subset of a norm linear space X and $f : X \to \bar{\mathbb{R}}$ is called "the objective function" when $G = X$, the problem (P) is called "unconstrained". Any $g_0 \in G$ for which the inf in 3.1 is attained i.e. such that

$$f(g_0) == \inf f(G) \tag{3.2}$$

is called a (global) "optimal solution" of problem (P). The set of all optimal solution will be denoted by $S_G(f)$, i.e.

$$S_G(f) = \left\{ g_0 \in G \middle| f(g_0) = \inf f(G) \right\} \tag{3.3}$$

naturally, one can also write min instead of inf in (3.2) and (3.3). If G is convex set and f is a convex function, then (P) of (3.1) is called a problem of infimization.

As we have observed above, best approximation may be regarded as a particular case of infimization, by taking X to be a norm linear space, $x_0 \in X$, and $f : X \to \mathbb{R}$ the convex function

$$f(y) :=\| x_0 - y \|, \ (y \in X) \tag{3.4}$$

indeed, then

$$\inf f(G) = dist(x_0, G),$$

and the optimal solution $g_0 \in G$ of problem (P), of this case, are the elements of best approximation of x_0 by G. Therefore, it is natural that many results on infimization can be applied to best approximation. Moreover, in the converse direction, although the extension from the particular function f of (3.4) to a function $f : X \to \overline{\mathbb{R}} = [-\infty, \infty]$ on a norm linear space X is a rather big step, it turns out that many results and methods of the theory of best approximation can be extended to results on the infimization of function.

Also $x_0 \in X$, and f is the function (3.4), then for any $g_0 \in X$ we have

$$\partial f(g_0) = \left\{ \phi \in X^* \middle| \phi_0(x_0 - g_0) =\| x_0 - g_0 \|, \| \phi \| \leq 1 \right\} \tag{3.5}$$

where,

$$\partial f(z_0) = \left\{ \phi \in X^* \middle| \phi(x) - \phi(z_0) + f(z_0) \leq f(x), \ (x \in X) \right\}$$

is the sub-gradient of f at z_0. We shall be concerned with the following two main problems.

(1) Find convenient formula for $\inf f(G)$

(2) Necessary and sufficient conditions in order that an element $g_0 \in G$ satisfies (3.2) i.e., $g_0 \in S_G(f)$.

References

[1] Cheney, E.W. and Goldstein, A.A., Tchebycheff approximation and related extremal problems, *Journal of Mathematics and Mechanics*, **14**(1), (1965), 87–98.

[2] Glashoff, K. and Gustafson, S.A., *Linear optimization and approximation: an introduction to the theoretical analysis and numerical treatment of semi-infinite programs.* Vol. 45. Springer Science & Business Media, 2012.

[3] Holmes, R.B., *A course on optimization and best approximation.* Vol. 257. Springer, 2006.

[4] Ivan S., *Duality for nonconvex approximation and optimization.* Springer Science & Business Media, 2007.

[5] Ivan S., *The theory of best approximation and functional analysis.* Society for industrial and applied mathematics, 1974.

[6] Laurent, P.J., *Approximation et optimisation,* (Vol. 375832942), Paris: Hermann, 1972.