CUCS-106-84

Linear Problems (with Extended Range) Have Linear Optimal Algorithms

Edward W. Packel

Department of Computer Science
Columbia University
New York, NY 10027

February 1984

Linear Problems (with Extended Range) Have Linear Optimal Algorithms

ABSTRACT:

Let F_1 and F_2 be normed linear spaces and $S:F_0 \dashrightarrow F_2$ a linear operator on a balanced subset F_0 of F_2. If N denotes a finite dimensional linear information operator on F_0, it is known that there need not be a *linear* optimal algorithm $\phi:N(F_0) \dashrightarrow F_2$ which is optimal in the sense that $\|\phi(N(f))-S(f)\|$ is minimized. We show that the linear problem defined by S and N can be regarded as having a linear optimal algorithm if we allow the range of ϕ to be extended in a natural way. The result depends upon imbedding F_2 isometrically in the space of continuous functions on a compact Hausdorff space X. This is done by making use of a consequence of the classical Banach-Alaoglu theorem.

AMS Subject Classifications:

Primary: 41A45, 41A65, 65J10

Secondary: 46B99, 68C05

1. Introduction

There has been considerable recent progress in applying and generalizing the information centered theory of optimal algorithms. In particular, when the problem and its information are linear, there are numerous useful and satisfying results. Thus, it has been shown that nonadaptive information is no less powerful (in terms of the error of optimal algorithms) than adaptive information of the same cardinality (see Traub and Woźniakowski [4, p. 49]). Also, for a wide variety of classical linear problems and in several general linear settings, it is known that linear optimal error algorithms exist. While it might seem reasonable to expect, in the light of the above results, that linear problems should always have linear optimal algorithms, there exist specially constructed examples to the contrary.

In this paper we resurrect the above intuition that linear problems *ought* to have optimal linear algorithms. We do this by showing that, under the minimal requirement that the range of the solution operator is a normed linear space, there must be an optimal linear algorithm if we allow its range to be extended in a natural way. Thus, linear problems do have linear optimal error algorithms as long as the solution operator is given an appropriate codomain (perhaps considerably larger than its range).

To develop this result, we will need some machinery from functional analysis, including the classical Banach-Alaoglu theorem. The presentation will be organized as follows. The next section reviews the information centered approach to algorithms in the linear framework, including some of the existing positive and negative results. The third section introduces additional notation needed to state and discuss the main result. The final section summarizes the technical material needed from functional analysis and proves the main result.

2. The Information Centered Approach to Linear Problems

A thorough development of the framework for the information centered approach may be found in Traub and Woźniakowski [4]. Here we summarize briefly the standard setting for linear problems.

Let F_1 and F_2 be normed linear spaces over the scalar field K, where K is either the real or complex numbers. Let F_0 be a balanced convex subset of F_1. In what follows, a function with domain F_0 will be said to be linear if it is the restriction of a linear function defined on F_1. Given $S:F_0\to F_2$ a linear operator and $N:F_0\to K^n$ a finite dimensional linear information operator, the inherent error, $r(S,N)$, in approximating the solution S working with the (incomplete) information N is called the ***radius of information***. This important concept is defined as follows:

Given $f\in F_0$, let $y=N(f)$ and set $V(y)=\{g\in F_0:N(g)=y\}$.
Now define $r(S,N,f)$ to be the radius of $S(V(y))$ as a subset of F_2.
Finally, define $r(S,N)=\sup\{r(S,N,f):f\in F_0\}$.

We now investigate algorithms to aprroximate $S(f)$. Since we only have limited information $y=N(f)$ on f, such algorithms can only be defined on $N(F_0)$. Of obvious importance are ***optimal algorithms*** $\phi:N(F_0)\to F_2$, where optimality means

$$\|\phi(N(f)) - S(f)\| \leq r(S,N) \quad \text{for all } f\in F_0.$$

As indicated earlier, we are interested in the existence of *linear* algorithms which are optimal. There are several reasons why linearity is desirable, which we now summarize. Linear algorithms would appear to be natural for problems in a linear setting. Indeed, many of the standard algorithms for classical numerical problems (integration and interpolation, for example) are linear. Linear algorithms tend to be simpler and easier to implement. Most importantly, linear algorithms have small combinatorial complexity and optimal linear algorithms can be formally shown to have ***nearly optimal combinatorial complexity*** (see [4, Chapter 5]). In addition to this valuable efficiency in time, linear algorithms also have small space complexity (if we ignore precomputation).

Since the result we will develop is immediate when $r(S,N)$ is infinite, we can assume for the remainder of the paper that $r(S,N)<\infty$. We now state two general positive results concerning the existence of optimal linear algorithms for linear problems. The first theorem covers the case where the solution operator S is scalar-valued.

Theorem 1: If $F_2=K$ then there exists a linear optimal error algorithm $\phi^*:N(F_0)\to K$. Thus $|\phi^*(N(f))-S(f)|\leq r(S,N)$ for all $f\in F_0$.

Proof: The case when K=R is due to Smolyak [3] and can be found in [4, p. 54]; the complex case is due to Osipenko [1].

The second result requires a slight reformulation of the general linear problem. We can, without loss of generality, assume that the balanced, convex set F_0 on which S and N are defined is generated by a linear ***restriction operator***

$T: F_1 \to F_4 = T(F_1)$ in the sense that $F_0 = \{f \in F_1 : \|T(f)\| \le 1\}$

Theorem 2: If $F_4 = T(F_1)$ is a Hilbert space and $T(\ker(N))$ is closed in F_4 then there exists a linear optimal error algorithm.

Proof: See [4, Chapter 4]. The result emerges in the context of the theory of spline algorithms, and the desired algorithm turns out to be a spline algorithm which is central and hence strongly optimal.

Theorems 1 and 2 indicate that, in the presence of appropriate structure, optimal linear algorithms can be constructed. A completely general result along these lines is ruled out by a counterexample constructed by Micchelli (see [4], p 60). We sketch below a somewhat simpler example of a linear problem which has no optimal linear algorithm.

Example:

Fix $\lambda \in (0, \sqrt{2}/2)$ and let $F_0 = \{(x_0, x_1, x_2) : \lambda |x_1| + |x_2| \le 1, |x_0| \le$?

Let $S: F_0 \to \mathbb{R}^2$ be defined by $S(x_0, x_1, x_2) = (x_1, x_2)$, where $\mathbb{R}^2$ is given its Hilbert norm.

Let $N: \mathbb{R}^3 \to \mathbb{R}$ be defined by $N(x_0, x_1, x_2) = x_1 + \lambda x_0$.

We rely heavily upon Figure 1, which pictures $S(F_0)$ and $S(V(y))$ for some critical values of $y = N(x)$, where $x = (x_0, x_1, x_2)$.

Part (a) of the figure is simply for orientation. It can be checked that the broken line has length < 1 from which it follows that $S(V(0))$ has radius 1.

It can be checked that the radius of the set $S(V(\lambda))$ in part (b) is the length of each of the broken lines. Furthermore, $S(V(\lambda))$ determines $r(S,N)$, which equals the square root of $1+\lambda^4$.

By the vertical (x_2) symmetry of the problem, an linear optimal algorithm must have the form $\phi(y) = (cy, 0)$ for some $c \in \mathbb{R}$. It then follows from the above that if ϕ is to be a linear optimal algorithm it must have the form $\phi(y) = (\lambda^2 \cdot y, 0)$.

Now using part (c) of the figure, it can be shown that with $y = 1/\lambda - \lambda$, $\|\phi(y) - (1/\lambda, 0)\| > r(S,N)$ (this is where $\lambda \in (0, \sqrt{2}/2)$ is needed). We can thus conclude that the stated problem has no optimal linear algorithm.

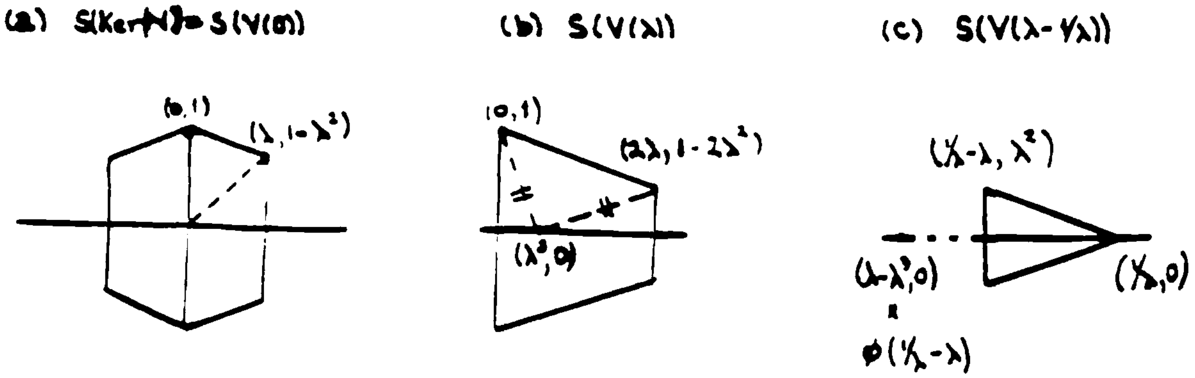

Figure 2-1: The Image under S

While it is clear that the above example is somewhat contrived, it does show that no general result about optimal linear algorithms for linear problems in the standard setting is possible. In the next section we show that a small but significant reformulation of the standard linear setting allows for optimal linear algorithms in a very general context.

3. Optimal Linear Algorithms

To state our main result, we recall some standard notation from functional analysis. A more complete exposition of these ideas, including proofs of standard results used in this and the next section can be found in Packel [2] (or any other introductory functional analysis text). Given a compact Hausdorff space X, define

$$C(X)=\{g:X\to K \mid g \text{ continuous}\}$$
$$B(X)=\{g:X\to K \mid g \text{ bounded}\}$$

where each function space is endowed with the ***sup*** norm

If E and F are normed linear spaces, they are defined to be ***isometrically isomorphic*** if there exists a linear bijection $b:E\to F$ which is norm-preserving. In this case we regard E and F as identical as far as their normed and linear structures are concerned.

We are now prepared to state the main result.

Theorem 2: Given a general linear problem defined by S and N, with $S:F_0 \subseteq F_1 \to F_2$, there exists:

i) A compact Hausdorff space X such that F_2 is isometrically isomorphic to $\hat{F}_2 \subseteq B(X)$.

ii) A linear optimal error algorithm $\phi^*: N(F_0) \to B(X)$ satisfying $\| \phi^*(N(f)) - \widehat{S(f)} \| \leq r(S,N)$ for all $f \in F_0$.

Before developing the proof, we discuss interpretations of this Theorem. Theorem 1 of the previous section showed that if the range of S is sufficiently simple (namely the scalar field K), then an optimal linear algorithm was assured. The Example sketched in that section then showed that merely expanding the range of S to R^2 destroys the guarantee of an optimal linear algorithm ***if algorithms are restricted to the range of S***. Theorem 2 now suggests that by giving S a codomain (namely B(X)) which extends beyond its range, an optimal linear algorithm (with range in this extended codomain) must exist.

As the forthcoming proof will show, the extended codomain is generally vastly larger and more complicated than the original range. In addition, its members (other than the isometric images from F_2) may have no meaningful connection with the members of F_2. Nevertheless, one interpretation of the result is that a linear problem does have an optimal linear algorithm if the solution operator is given an "appropriate" codomain. While the theorem uses the rather extreme case of B(X) for this codomain, it may be the case that linear optimality holds for less drastic extensions of the range of S. In particular, it seems reasonable to conjecture, perhaps with added hypotheses, that the Theorem might be strengthened by replacing B(X) with C(X). We leave this for now as an open problem.

4. Technical Background and Proofs

Let F be a normed linear space over the scalar field K. The *conjugate space* F^* of F is defined by

$$F^* = \{f^*\, F \to K \mid f^* \text{ continuous and linear on } F\}$$

A natural "operator" norm on F^* is defined by

$$\|f^*\| = sup\{|f^*(f)| : \|f\| = 1\}$$

A weaker topology on F^* can be defined as follows. Each $f \in F$ induces a linear functional $f^\wedge$ on F^* defined by

$$f^\wedge(f^*) = f^*(f).$$

The *weak* topology* on F^* is defined as the weakest topology such that $f^\wedge$ is continuous for every f in F.

Under the weak* topology, it thus follows that $f^\wedge \in F^{**}$ (since $f^\wedge$ is clearly linear and must be continuous on F^*). Using the natural norms on F^* and F^{**}, we also note that $\|f^\wedge\| = \|f\|$. Indeed, $\|f^\wedge\| \leq \|f\|$ since

$$\|f^\wedge\| = sup\{|f^\wedge(f^*)| : \|f^*\| = 1\} = sup\{|f^*(f)| : \|f^*\| = 1\} \leq sup\{\|f^*\|\,|f| : \|f^*\| = 1\} = |f|$$

The fact that $\|f^\wedge\| = \|f\|$ follows by a routine application of the Hahn-Banach theorem on the space F^*. The above result says that F is isometrically isomorphic to a subspace $\widehat{F}$ of F^{**} (by means of the linear isometry $^\wedge : F \to F^{**}$). We now apply these ideas to state and prove the following "folk" result about normed linear spaces. Though this result can be found in a variety of texts, we give the proof here since it is short and sets the stage for the main theorem.

Lemma 4: Let F be a normed linear space. Then there exists a compact Hausdorff space X such that F is isometrically isomorphic to a subspace $\widehat{F}$ of C(X). If F is a Banach space, then $\widehat{F}$ is a closed subspace of C(X).

Proof: Give the conjugate space F^* its weak* topology. Let X be the unit ball of F^*—$X=\{f^*\in F^*: \|f^*\|\leq 1\}$. The classical Banach-Alaoglu theorem says that X is compact in the weak* topology on X. Define $\hat{\ }: F \to C(X)$ by $f \mapsto f\hat{\ }$, where $f\hat{\ }(f^*)=f^*(f)$ for all $f^* \in X$. Then, as developed above, $\|f\hat{\ }\| = \|f\|$, so the subspace $\widehat{F}$ of C(X) defined by $\widehat{F}=\{f\hat{\ }: f\in F\}$ is isometrically isomorphic to F. If F is a Banach space, then F and $\widehat{F}$ are complete, making $\widehat{F}$ closed as a complete subspace of C(X).

Remark. The above result is not as powerful as it may seem at first glance, since very little is known about the subspaces of C(X). Our application to the proof of Theorem 2 is, to our knowledge, the first really meaningful use of this curious result. Before proceeding, we note that we can also treat F as being isometrically imbedded in B(X) since $C(X) \subseteq B(X)$.

We now prove the main theorem which we restate for the convenience of the reader.

Theorem Given a general linear problem defined by S and N, with $S: F_0 \subseteq F_1 \to F_2$, there exists:

i) A compact Hausdorff space X such that F_2 is isometrically isomorphic to $\widehat{F}_2 \subseteq B(X)$.

ii) A linear optimal error algorithm $\phi^*: N(F_0) \to B(X)$ satisfying $\|\phi^*(N(f)) - \widehat{S(f)}\| \leq r(S,N)$ for all $f\in F_0$.

Proof: i) The existence of X and $\widehat{F}_2$ follow directly from the Lemma proved previously.

ii) For each fixed $x\in X$, consider the linear problem

$S_x: F_0 \to K$ where $S_x(f)=\widehat{S(f)}(x)$.

By Theorem 1 we know there exists a linear optimal algorithm $\phi^*_x: N(F_0) \to K$ such that

$$\|\phi^*_x(N(f)) - S_x(f)\| \leq r(S_x,N) \leq r(S,N) \text{ for all } f\in F_0 \qquad (1)$$

Letting x vary over X, we now must show that the linear operator ϕ^* thus defined on $N(F_0)$ has its range in B(X). First observe that

$$|S_x(f)| = |\widehat{S(f)}(x)| \leq \|\widehat{S(f)}\| = \|S(f)\| \tag{2}$$

where the inequality follows from $\|\widehat{S(f)}\| = \sup\{|\widehat{S(f)}(x)| \cdot x \in X\}$. Using (1) and (2), we have for all $f \in F_0$,

$$\begin{aligned} |\phi^*(N(f))(x)| &\leq |\phi^*(N(f))(x) - S_x(f)| + |S_x(f)| \\ &\leq |\phi_x^*(N(f)) - S_x(f)| + |S_x(f)| \\ &\leq r(S,N) + \|S(f)\|. \end{aligned}$$

Since the final expression is independent of x, $\phi^*(N(f)) \in B(X)$ and the proof is complete.

Acknowledgement: The author wishes to thank Henryk Woźniakowski for his helpful suggestions on organizing these results.

References

[1] K. Yu. Osipenko, Best approximation of analytic functions from information about their values at a finite number of points, (English translation) Math. Notes, **19** (1976), 17-23.

[2] E. W. Packel, Functional Analysis: a Short Course, Krieger, 1980

[3] S. A. Smolyak, On optimal restoration of functions and functionals of them (in Russian), Candidate Dissertation, Moscow State Univ. (1965).

[4] J. F. Traub and H. Woźniakowski, A General Theory of Optimal Algorithms, Academic Press, 1980.