M2AN. MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS - MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

CHARLES A. MICCHELLI
FLORENCIO I. UTRERAS

Smoothing and interpolation in a convex subset of a Hilbert space : II. The semi-norm case

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 25, nº 4 (1991), p. 425-440

<http://www.numdam.org/item?id=M2AN_1991__25_4_425_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol 25, n° 4, 1991, p 425 a 440)

SMOOTHING AND INTERPOLATION IN A CONVEX SUBSET OF A HILBERT SPACE : II. THE SEMI-NORM CASE (*)

Charles A MICCHELLI (1), Florencio I UTRERAS (2)

Communicated by P J LAURENT

Abstract — *We improve upon results of our previous paper on interpolation subject to convex constraints In this paper we focus on the case of constrained best interpolation when the object function is chosen to be* $\|Tx\|^2$ *where T is a bounded linear operator defined on a Hilbert space X onto another Hilbert space Y with a finite dimensional kernel (We simply say T is* correct *from X to Y) We prove that under rather general circumstances this problem can be separated into first finding an orthogonal projection onto some constraint set and then solving a* finite dimensional min-max *problem whose saddle point determines the solution of our problem*

Resume — *On presente des resultats permettant d ameliorer des theoremes obtenus dans un article precedent Dans cet article on etudie le probleme d'interpolation optimale sous contraintes obtenue quand on minimise une semi-norme* $\|Tx\|^2$ *Ici T est un operateur lineaire borne et surjectif defini dans un espace de Hilbert X dans un autre espace de Hilbert Y ayant un noyau de dimension finie On demontre que, sous des hypotheses assez générales, ce probleme peut être decompose en une projection orthogonale sur un certain ensemble convexe suivie de la resolution d'un probleme de* min-max *en dimension finie le point de selle determinant la solution de notre probleme*

1. INTRODUCTION

Let X, Y be a Hilbert space and T be a bounded linear operator defined on X with range Y and a finite dimensional kernel

$$\dim(\mathrm{Ker}(T)) < +\infty \tag{1.1}$$

(*) Received in November 1989, revised in August 1990

(1) The work of this author was partially supported by a DGICYT Grant of the Spanish Ministry of Education and Science

Department of Mathematical Sciences, IBM T J Watson Research Center, P O Box 218, Yorktown Heights, NY 10598

(2) The work of this author was partially supported by Fondo Nacional de Cienca y Technologia under Grant 1296186

Departamento de Matematicas y Ciencas de la Computacion, Facultad de Ciencias Fisicas y Matematicas, Universidad de Chile, Casilla 17013 Correo 3, Santiago, Chile

For simplicity we say that T is *correct* from X to Y. Assume that Z is a finite dimensional Hilbert space, I a bounded linear operator from X into Z, C a closed convex cone and $z^0 \in Z \setminus \{0\}$ a given « data value ».

We are interested in the solution of the *Best Interpolation Problem :*

$$p := \inf \left\{ \frac{1}{2} \|Tx\|^2 : x \in C, Ix = z^0 \right\}, \tag{1.2}$$

where we assume that the data z^0 is consistent with the constraint cone, that is, $C \cap I^{-1}(z^0) \neq \phi$.

In the case that $C = X$ it is well-known that there exists a unique solution to (1.2) for any $z^0 \in Z$ with $C \cap I^{-1}(z^0) \neq \phi$ provided that

$$\text{Ker}(T) \cap \text{Ker}(I) = \{0\}, \tag{1.3}$$

[2]. Moreover, the unique solution to (1.2) which we denote by θ is determined by the equation

$$(T^* T)(\theta) = I^*(\lambda) \tag{1.4}$$

for some $\lambda \in Z$ chosen so that $I\theta = z^0$.

Our aim in this paper is to show that under the « interior moment cone » hypothesis, that is, $z^0 \in$ interior $I(C)$ of [4], see also [8], and Section 2 the solution σ to (1.2) satisfies the equation

$$(T^* T)(\sigma_c) = \Pi_K(I^*(\hat{\lambda}) + \hat{\mu}) - \hat{\mu}, \tag{1.5}$$

where $\hat{\lambda} \in Z$, $\hat{\mu} \in \text{Ker}(T)$ are chosen to insure that $I(\sigma_c) = z^0$ and Π_K is the orthogonal projection onto some cone K in X. Specifically, we let P be the orthogonal projection of $(X, \|\cdot\|_X)$ onto Ker (T) and choose

$$K := \{u \in X : S^* Su + Pu \in C\} \tag{1.6}$$

where

$$S = (T^*)^\dagger P^\perp . \tag{1.7}$$

Here we use the notation $A^\dagger$ for the pseudoinverse of A. The operator Π_K is now defined to be the orthogonal projection of X onto K relative to the norm

$$|x|^2 := \|Sx\|_Y^2 + \|Px\|_X^2, \quad x \in X. \tag{1.8}$$

This result improves upon a previous result of the authors [4], given for the case $T =$ identity, $X = Y$. In that case σ_c is the orthogonal projection onto C of some element in the range of I^*, that is,

$$\sigma_c = P_c(I^*(\tilde{\lambda})) \tag{1.9}$$

where P_c is the orthogonal projection of X onto C. Here $\tilde{\lambda} \in Z$ can be determined as a solution of the finite dimensional minimization problem

$$d := \min \left\{ \frac{1}{2} \| P_c(I^*(\lambda)) \|_X^2 - \langle \lambda, z^0 \rangle : \lambda \in Z \right\} . \tag{1.10}$$

As pointed out in [4], the gradient of the objective function in (1.10) is easily obtained in terms of $P_c(I^*(\lambda))$ so that efficient minimization methods to solve (1.10) are available. In many interesting applications, however, the introduction of the operator T becomes necessary. In Section 2 we proceed to study the effect of introducing T into the problem. We prove under the interior data cone condition that the parameter $\tilde{\lambda} \in Z$, $\tilde{\mu} \in \text{Ker}\,(T)$ in (1.5) can be determined as the *saddle point* of the max-min problem

$$q := \max_{\lambda \in Z} \min_{\mu \in \text{Ker}\,(T)} \left\{ \frac{1}{2} |\Pi_{c^*}(I^*(\lambda) + \mu)|^2 - \frac{1}{2} |I^*(\lambda)|^2 - \langle I(\mu) - z^0, \lambda \rangle \right\} \tag{1.11}$$

$$= \max_{\lambda \in Z} \min_{\mu \in \text{Ker}\,(T)} \left\{ \frac{1}{2} |\mu|^2 - \frac{1}{2} |\Pi_K(I^*(\lambda) + \mu)|^2 + \langle \lambda, z^0 \rangle \right\} \tag{1.12}$$

where C^* is the conjugate cone of C (relative to $\| \cdot \|_X$). Therefore, as in (1.10), we reduce (1.2) to the solution of a *finite dimensional unconstrained extremal problem.* Section 3 contains results concerning smoothing under convex constraints.

We apply our main characterization for (1.2) to only one model problem at the end of Section 2. Generally speaking, any application of the general principle we develop here, for instance, to monotone or convex interpolation or to area-matching of density functions as in [9] require some effort. For practical reasons this issue deserves more work. Nevertheless, the main content of our result is to demonstrate how to separate the « global » convexity constraint in (1.2) embodied by the requirement, $x \in C$, from the finite linear constraints $Ix = z^0$ which, in fact, can be large in number. Our computational experience with the case T = identity in [4] indicates that a significant reduction of computational cost results by this separation of constraints.

2. THE CONSTRAINED INTERPOLATION PROBLEM

Let T be a bounded linear operator with domain $D(T) = X$ and suppose that $R(T)$, the range of T, is Y. As indicated in the introduction we are interested in the case $0 \neq \dim\,(\text{Ker}\,(T)) < +\infty$. Let P be the orthogonal projection of X onto Ker (T) and denote by $P^{\perp}$ the orthogonal projection onto $R(T^*) = [\text{Ker }(T)]^{\perp}$. Now, let $T^{\dagger}$ be the *pseudoinverse* of T, i.e., $T^{\dagger} : R(T) \to D(T)$ is the bounded linear operator defined as $x = T^{\dagger} y$ where x has *least* norm among all $v \in D(T)$ such that

$$\|y - Tv\| = \min \{ \|y - T\omega\| : \omega \in D(T)\} . \tag{2.1}$$

We list some facts about pseudoinverse which can be found in [7].

First we record the equation

$$\ker (T^{\dagger}) = R(T)^{\perp} , \quad R(T^{\dagger}) = \ker (T)^{\perp} \tag{2.2}$$

and then the useful formula

$$TT^{\dagger} y = y , \quad y \in R(T) . \tag{2.3}$$

It is known that since $R(T)$ is (a closed subspace of) Y, $T^{\dagger}$ is a bounded linear operator. Also, it is important to recall the identities

$$(T^{\dagger})^{\dagger} = T \tag{2.4}$$

and

$$(T^{\dagger})^* = (T^*)^{\dagger} . \tag{2.5}$$

Next, we introduce the operator

$$S = (T^*)^{\dagger} P^{\perp} . \tag{2.6}$$

According to (2.3), applied to the adjoint of T and $y := P^{\perp} x \in R(T^*)$ we have

$$T^* Sx = P^{\perp} x , \quad x \in X . \tag{2.7}$$

Hence, Ker (S) = Ker $(P^{\perp})$ = Ker (T) and we are led to

PROPOSITION 2.1 : *Let T be correct from X to Y. Then*

$$G := S^* S + P \tag{2.8}$$

is a bounded linear operator from X into itself which is invertible and has a dense range.

Proof : Since Ker $(S^* S)$ = Ker (S) = Ker (T) which is orthogonal to Ker $(P) = [\text{Ker }(T)]^{\perp}$, we conclude that Ker $(G) = \{0\}$. Similarly, if $\omega \in R(G)^{\perp}$ then $0 = (\omega, GP^{\perp} \omega) = (P^{\perp} \omega, S^* SP^{\perp} \omega)$ and so $P^{\perp} \omega = 0$. Also, since $0 = (\omega, G\omega) = (\omega, P\omega)$ we get $P\omega = 0$ and therefore $\omega = 0$. Thus we have proved that $\overline{R(G)} = X$, as claimed.

Let us now observe that the cone K introduced in (1.7) may be expressed in the equivalent form

$$K = G^{-1}(C) . \tag{2.9}$$

Moreover, it is now apparent that the semi-norm $| \cdot |$, introduced in (1.8) is in fact a norm since

$$\begin{aligned} |x|^2 &\Leftrightarrow \|Sx\|_Y^2 + \|Px\|_X^2 = 0 \\ &\Leftrightarrow Sx = 0 \text{ and } Px = 0 \\ &\Leftrightarrow x \in \text{Ker }(S) \cap \text{Ker }(P) \\ &\Leftrightarrow x \in [\text{Ker }(T)]^\perp \cap \text{Ker }(T) = \{0\} . \end{aligned}$$

We require some further facts concerning the operators S and T.

LEMMA 2.2 : *Let T be correct from X to Y. Then for S defined by* (2.8) *we have*

$$S^* ST^* T = T^* TS^* S = P^\perp$$

and so, in particular $S^* S$ and $T^* T$ commute.

Proof : First we note that

$$S^* ST^* T = P^\perp T^\dagger (T^*)^\dagger P^\perp T^* T$$

which, on account of the fact that $P^\perp$ is the orthogonal projection $R(T^*)$, becomes

$$\begin{aligned} &= P^\perp (T^* T)^\dagger T^* T \\ &= P^\perp . \end{aligned}$$

Similarly, we have

$$T^* TS^* S = T^* TP^\perp T^\dagger (T^*)^\dagger P^\perp .$$

Since $TP^\perp = T(P^\perp + P) = T$ this becomes

$$\begin{aligned} &= T^* TT^\dagger (T^*)^\dagger P^\perp \\ &= (T^* T)(T^* T)^\dagger = P^\perp \end{aligned}$$

which proves the lemma.

Next we give a useful alternative description of the projection.

LEMMA 2.3 : *Let T be correct from X to Y. Suppose that C^* is the conjugate cone of C relative to the standard norm on X. Furthermore, denote* *by* Π_{C^*} *the orthogonal projection of* X, *relative to the norm* (1.8), *onto* C^*. *Then*

$$\Pi_K(x) + \Pi_{C^*}(x) = x\,, \quad x \in X\,. \tag{2.10}$$

In other words, K *is the conjugate cone of* C^* *in* $(X, |\cdot|)$.

Proof : According to our definition $u \in K$ if and only if $S^* Su + Pu \in c$ which is equivalent to

$$\langle S^* Su + Pu, \omega \rangle_X \leq 0\,, \quad \omega \in C^*$$

where $\|x\|_X^2 = \langle x, x \rangle_X$ is the original inner product on X. Hence, $u \in K$ if and only if

$$\langle Su, S\omega \rangle_Y + \langle Pu, P\omega \rangle_X \leq 0\,, \quad \omega \in C^*\,.$$

However, the left hand side of the above inequality is precisely the inner product of u and ω relative to the norm (1.9). This proves the claim (2.10).

We are now ready to show that characterizing the solutions of (1.2) can be reduced to an existence problem. For this reason we introduce the operator

$$H : Z \times \text{Ker}\ (T) \to X$$

defined by

$$H(\lambda, \mu) = I^*(\lambda) + \mu\,.$$

THEOREM 2.4 : *Let* T *be correct from* X *to* Y. *Assume that there exists a* $\gamma^0 = (\lambda^0, \mu^0) \in Z \times \text{Ker}\ (T)$ *such that*

$$x^0 := G\Pi_K(H(\gamma^0)) \tag{2.11}$$

satisfies both

$$I(x^0) = z^0 \tag{2.12}$$

and

$$P\Pi_K(H(\gamma^0)) = \mu^0\,. \tag{2.13}$$

Then x^0 *is the solution to the problem.*

$$p = \min \left\{ \frac{1}{2} \|Tx\|^2 : x \in C, Ix = z \right\}.$$

Proof : Let us first remark that since $\Pi_K(H(\gamma^0)) \in K$ we conclude that $x^0 \in C$.

Now, consider an arbitrary element $\omega \in C$ such that $I\omega = z^0$. We will show that

$$\| Tx^0\| \leq \| T\omega \|^2 .$$

For this purpose, we use Lemma 2.2 and observe that

$$\begin{aligned} T^* Tx^0 &= T^* TS^* S\Pi_K(H(\gamma^0)) + T^* TP\Pi_K(\gamma^0) \\ &= P^\perp \Pi_K(H(\gamma^0)) \\ &= \Pi_K(H(\gamma^0)) - P\Pi_K(H(\gamma^0)) \\ &= \Pi_K(H(\gamma^0)) - \mu^0 . \end{aligned}$$

Therefore we have established that

$$\langle Tx^0, T\omega \rangle = \langle \Pi_K(H(\gamma^0)) - \mu^0, \omega \rangle$$

and so invoking Lemma 2.3 gives us

$$\begin{aligned} &= \langle H(\gamma^0) - \Pi_{C^*}(H(\gamma^0)) - \mu^0, \omega \rangle \\ &= \langle I^*(\lambda^0), \omega \rangle - \langle \Pi_{C^*}(H(\gamma^0)), \omega \rangle \\ &= \langle \lambda^0, z^0 \rangle - \langle \Pi_{C^*}(H(\gamma^0)), \omega \rangle . \end{aligned}$$

Hence we have obtained the inequality

$$\langle Tx^0, T\omega \rangle \geq \langle \lambda^0, z^0 \rangle \tag{2.14}$$

because $\Pi_{C^*}(H(\gamma^0)) \in C^*$. Moreover, equality holds in (2.14) for $\omega = x^0$, if we can demonstrate that

$$\langle \Pi_{C^*}(H(\gamma^0)), x^0 \rangle = 0 . \tag{2.15}$$

For this purpose, we use the definition of x^0 and note that

$$\begin{aligned} &\langle \Pi_{C^*}(H(\gamma^0)), x^0 \rangle \\ &= \langle S\Pi_{C^*}(H(\gamma^0)), S\Pi_K(H(\gamma^0)) \rangle + \langle P\Pi_K(H(\gamma^0)), P\Pi_K(H(\gamma^0)) \rangle \\ &= (\Pi_{C^*}(H(\gamma^0)), \Pi_K(H(\gamma^0))) \\ &= 0 \end{aligned}$$

since Lemma 2.3 guarantees that K is the conjugate cone of C^* relative to the norm (1.9). Thus we have demonstrated that

$$\langle Tx^0, Tx^0 \rangle = \langle \lambda^0, z^0 \rangle$$

and

$$\langle Tx^0, T\omega \rangle \geq \| Tx^0 \|^2 ,$$

for any $\omega \in C$ such that $I\omega = z^0$. Consequently, the Cauchy-Schwartz inequality implies that

$$\|Tx^0\| \leq \|T\omega\|$$

which proves the desired result.

To make use of Theorem 2.4 we consider the following finite dimensional variational problem. For every $\lambda \in Z$ and $\mu \in \operatorname{Ker}(T)$ we define

$$J(\lambda, \mu) = \frac{1}{2} |\Pi_{C^*}(I^*(\lambda) + \mu)|^2 - \frac{1}{2} |I^*(\lambda)|^2 - \langle \mu, I^*(\lambda) \rangle + \langle \lambda, z^0 \rangle . \tag{2.17}$$

Using Lemma 2.3 we first rewrite J in the form

$$J(\lambda, \mu) = \frac{1}{2} |I^*(\lambda) + \mu|^2 - \frac{1}{2} |\Pi_K(I^*(\lambda) + \mu)|^2 - \frac{1}{2} |I^*(\lambda)|^2 - \langle \mu, I^*(\lambda) \rangle + \langle \lambda, z^0 \rangle . \tag{2.18}$$

Then, using the fact that $\mu \in \operatorname{Ker}(T)$ implies $S\mu = 0$, we get

$$\begin{aligned}(\mu, I^*(\lambda)) &= \langle S\mu, SI^*(\lambda) \rangle + \langle P\mu, PI^*(\lambda) \rangle \\ &= \langle P\mu, I^*(\lambda) \rangle = \langle \mu, I^*(\lambda) \rangle\end{aligned}$$

and so

$$J(\lambda, \mu) = \frac{1}{2} |\mu|^2 - \frac{1}{2} |\Pi_K(I^*(\lambda) + \mu)|^2 + \langle \lambda, z^0 \rangle . \tag{2.19}$$

Therefore we conclude from (2.19) that J is concave in λ, while from (2.17), we see it convex in μ. This leads us to consider the sup-inf problem

$$q = \sup_{\mu \in Z} \inf_{\mu \in \operatorname{Ker}(T)} J(\lambda, \mu) . \tag{2.20}$$

The following lemma clarifies the connection between J and the solution to (1.2).

LEMMA 2.5 : *Let T be correct from X to Y. Assume that there exists a saddle point of J, that is, there is a $\gamma^0 = (\lambda^0, \mu^0) \in Z \times \operatorname{Ker}(T)$ such that*

$$J(\lambda, \mu^0) \leq J(\lambda^0, \mu^0) \leq J(\lambda^0, \mu), \; (\lambda, \mu) \in Z \times \operatorname{Ker}(T) . \tag{2.21}$$

Then $\gamma^0 = (\lambda^0, \mu^0)$ satisfies the hypothesis of Theorem 2.4 and consequently

$$x^0 = G(\Pi_K(H(\gamma^0)))$$

is the unique solution to (1.2).

Proof: If (2.21) is satisfied then for any $\lambda \in Z$, $\nabla_{(\lambda,0)} J(\gamma^0) = 0$, that is, the directional derivative of J at γ^0 in the direction of $(\lambda, 0)$ is zero. Using (2.19) we conclude that

$$\begin{aligned}\nabla_{(\lambda,0)} J(\gamma^0) &= -(\Pi_K(I^*(\lambda^0)+\mu^0), I^*(\lambda)) + \langle \lambda, z^0 \rangle \\ &= -\langle (S^* S + P)\, \Pi_K(I^*(\lambda^0)+\mu^0), I^*(\lambda) \rangle + \langle \lambda, z^0 \rangle \\ &= \langle z^0 - I(G\Pi_K(H(\gamma^0))), \lambda \rangle\end{aligned}$$

which means that $z^0 = I(G\Pi_K(H(\gamma^0)))$. On the other hand, we also have for all $\mu \in \operatorname{Ker}(T)$

$$\begin{aligned}0 &= \nabla_{(0,\mu)} J(\gamma^0) \\ &= (\mu^0 - \Pi_K(H(\gamma^0)), \mu)\,.\end{aligned}$$

Since Ker (T) = Ker (S) we obtain

$$0 = \langle \mu^0 - \Pi_K(H(\gamma^0)), \mu \rangle$$

and so $\mu^0 - \Pi_K(H(\gamma^0)) \in [\operatorname{Ker} T]^{\perp}$. In other words,

$$\mu^0 = \mathrm{P}(\Pi_K(H(\gamma^0))) \tag{2.22}$$

which proves the lemma.

In order to complete our characterization of a solution to (1.2) we must provide conditions under which J has a saddle point. To do this let us recall the notion of *recession direction* of a (finite) convex function $g : \mathbb{R}^s \to \mathbb{R}$ (cf., [5], p. 60).

DEFINITION 2.6: *An everywhere finite convex function g defined on $\mathbb{R}^s$ has a recession direction $y \in \mathbb{R}^s \setminus \{0\}$ if there exists a constant $M \in \mathbb{R}$ such that for any $\alpha \geqslant 0$ one has*

$$g(\alpha y) \leqslant M\,.$$

Notice that the convexity of g implies that for any $\alpha \in \mathbb{R}^s$ there exists $M \in \mathbb{R}$ such that for any $\alpha \geqslant 0$ one has

$$g(x + \alpha y) \leqslant M\,.$$

The usefulness of this notion for us rests on the following theorem which is proved in [5], specifically see Theorems (37.3) and (37.6).

THEOREM 2.7 (*cf. Rockafeller* [5]): *Let Λ be a (finite) concave-convex function on $\mathbb{R}^m \times \mathbb{R}^n$. Either of the following conditions implies that the saddle value of Λ exists, that is,* $\inf\sup \Lambda = \sup\inf \Lambda$. *If both conditions hold, then Λ has a saddle point, that is, there is an $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ such that* $\Lambda(x, y) = \inf\sup \Lambda = \sup\inf \Lambda$.

(a) *The convex functions* $\Lambda(x, .)$, *for* $x \in \mathbb{R}^m$ *have no common direction of recession.*

(b) *The convex functions* $-\Lambda(., y)$, *for* $y \in \mathbb{R}^n$ *have no common direction of recession.*

With the help of this fact we shall now prove under rather general conditions the existence of a saddle point of J, that is, the existence of an extended real number q such that

$$\begin{aligned} q &= \inf_{\mu \in \operatorname{Ker}(T)} \sup_{\lambda \in Z} J(\lambda, \mu) \\ &= \sup_{\lambda \in Z} \inf_{\mu \in \operatorname{Ker}(T)} J(\lambda, \mu) . \end{aligned} \tag{2.23}$$

THEOREM 2.8 : *Let T be correct from X to Y. Assume that*

$$C \cap \operatorname{Ker}(T) \cap \operatorname{Ker}(I) = \{0\} . \tag{2.24}$$

then the saddle value q exists.

Proof : According to Theorem 2.7, it is sufficient to prove that the convex functions $J(\lambda, .)$ $\lambda \in Z$ have no *common* direction of recession other than zero.

To verify this fact we assume $\mu \in \operatorname{Ker} T$ is a recession direction for $J(\lambda, .)$. Thus there is an $M \in \mathbb{R}$ such that for $\alpha > 0$

$$\frac{1}{2}\alpha^2 |\Pi_{C^*}(\alpha^{-1} I^*(\lambda) + \mu)|^2 - \frac{1}{2}|I^*(\lambda)|^2 - \alpha\langle \mu, I^*(\lambda)\rangle + \langle \lambda, z^0\rangle \leq M .$$

Dividing both sides of this inequality by α^2 and letting $\alpha \to \infty$ we get $|\Pi_{C^*}(\mu)| = 0$ and so $\mu \in C$. Next we use the fact

$$J(\lambda, \alpha\mu - I^*(\lambda)) = -\frac{1}{2}|I^*(\lambda)|^2 - \alpha\langle \mu, I^*(\lambda)\rangle + \|I^*(\lambda)\|^2 + \langle \lambda, z^0\rangle$$

is also bounded for all $\alpha \geq 0$. Hence we conclude that

$$\langle \mu, I^*(\lambda)\rangle \geq 0 . \tag{2.25}$$

Since λ was arbitrary chosen in Z we conclude $I\mu = 0$. In summary, we have $\mu \in C \cap \operatorname{Ker}(I) \cap \operatorname{Ker}(T)$ and so by hypothesis $\mu = 0$ which completes the proof.

Our final and main result of this section depends on the following fact :

LEMMA 2.9 : *Suppose $z^0 \in I^0(C)$ ($:=$ interior $I(C)$). Then for any μ*

$$\lim_{\lambda \to \infty} J(\lambda, \mu) = -\infty .$$

Proof: Suppose to the contrary that there is a sequence $\{\lambda^k\} \subseteq Z$, $\lim_{k \to \infty} \lambda^k = \infty$, and a constant M such that

$$- J(\lambda^k, \mu) \leq M . \tag{2.26}$$

Set $\delta^k := \lambda^k / \| \lambda^k \|$ and let $\{\delta^{k'}\}$ be a subsequence which converges to some $\delta^\infty \in Z \setminus \{0\}$. Dividing both sides of (2.26) by $\| \lambda^k \|^2$ and sending $k' \to \infty$ we conclude from (2.19) that $I^*(\delta^\infty) \in C^*$.

Returning to (2.26), we also have

$$- \langle \lambda^k, z^0 \rangle \leq M - \frac{1}{2} | \mu |^2$$

and so $\langle \delta^\infty, z^0 \rangle \geq 0$ as well. It was shown in [8] that $z^0 \in I^0(C)$ if and only if $\{\delta : \delta \in Z, \ \langle \delta, z^0 \rangle \geq 0, \ I^* \delta \in C^*\} = 0$. Therefore it follows that $\delta^\infty = 0$, a contradiction which proves the lemma.

This lemma is the last ingredient we need for

THEOREM 2.10: *Let $T : X \to Y$ be correct from X to Y, two Hilbert spaces. Let $I : X \to Z$ be a bounded linear operator from X into Z, $\dim Z < \infty$ and C a closed convex cone in X. Suppose further that*

$$C \cap \operatorname{Ker}(T) \cap \operatorname{Ker}(I) = \{0\} \tag{2.27}$$

and

$$z^0 \in I^0(C) . \tag{2.28}$$

Then there exists a $\gamma^0 = (\lambda^0, \mu^0) \in Z \times \operatorname{Ker}(T)$ such that the unique solution of the variational problem

$$p = \min \left\{ \frac{1}{2} \| Tx \|^2 : x \in C, Ix = z^0 \right\}$$

is given by

$$x^0 = G(\Pi_K(H(\gamma^0))) \tag{2.29}$$

where

$$I(x^0) = z^0 \tag{2.30}$$

and

$$P \Pi_K(H(\gamma^0)) = \mu^0 . \tag{2.31}$$

Moreover, $\gamma^0 = (\lambda^0, \mu^0)$ is a saddle point of J with saddle value p, that is,

$$p = J(\lambda^0, \mu^0) = \min_{\mu \in \operatorname{Ker}(T)} \max_{\lambda \in Z} J(\lambda, \mu) = \max_{\lambda \in Z} \min_{\mu \in \operatorname{Ker}(T)} J(\lambda, \mu) . \tag{2.32}$$

Proof According to Lemma 2 9, our hypothesis (2 28) implies for *any* $\mu \in \operatorname{Ker}(T)$, $-J(.,\mu)$ does not have a direction of recession We have already verified in Theorem 2 8 that condition (2 26) implies the family of convex functions $J(\lambda)$, $\lambda \in Z$ does not have a common direction of recession Hence we have confirmed both a) and b) of Theorem 2 7 This proves the existence of a saddle point of J and so Lemma 2 5 and Theorem 2 4 establishes (2 29) to (2 31) It only remains to establish that $p = J(\lambda^0, \mu^0)$ which we do next

Recall that since $T^* TG = T^* TS^* S = P^{\perp}$ we get

$$\begin{aligned}
\| Tx^0 \|^2 &= \langle T^* Tx^0, x^0 \rangle \\
&= \langle P^{\perp} \Pi_K(H(\gamma^0)), G\Pi_K(H(\gamma^0)) \rangle \\
&= (P^{\perp} \Pi_K(H(\gamma^0)), \Pi_K(H(\gamma^0))) \\
&= |\Pi_K(H(\gamma^0))|^2 - |P\Pi_K(H(\gamma^0))|^2 \\
&= |\Pi_K(H(\gamma^0))|^2 - |\mu^0|^2
\end{aligned} \tag{2 33}$$

On the other hand,

$$\begin{aligned}
\langle \lambda^0, z^0 \rangle &= \langle I^*(\lambda^0), G\Pi_K(H(\gamma^0)) \rangle = (I^*(\lambda^0), \Pi_K(H(\gamma^0))) \\
&= (H(\gamma^0), \Pi_K(H(\gamma^0))) - (\mu^0, \Pi_K(H(\gamma^0))) \\
&= |\Pi_K(H(\gamma^0))|^2 - |\mu^0|^2
\end{aligned} \tag{2 34}$$

and so combining these equations give us

$$\begin{aligned}
J(\lambda^0, \mu^0) &= \frac{1}{2} |\mu^0|^2 - \frac{1}{2} |\Pi_K(H(\gamma^0))|^2 + |\Pi_K(H(\gamma^0))|^2 - |\mu^0|^2 \\
&= \frac{1}{2} |\Pi_K(H(\gamma^0))|^2 - \frac{1}{2} |\mu^0|^2 \\
&= \frac{1}{2} \| Tx^0 \|^2
\end{aligned} \tag{2 35}$$

which proves the theorem

Remark 2 11 Note that when $C = X$, the characterization Theorem 2 9 takes the familar form

$$T^* T(x^0) = I^*(\lambda^0) + \mu^0 \tag{2 36}$$

where $\mu^0 \in \operatorname{Ker}(T)$ and $I^*(\lambda^0) \in \operatorname{Ker}(T)^{\perp}$ (*cf* [2], chapter 5) These conditions, with the equation $I(x^0) = z^0$, provide a system of nonsingular linear equations for determining $x^0 = G(I^*(\lambda^0) + \mu^0)$ when $\operatorname{Ker}(T) \cap \operatorname{Ker}(I) = \{0\}$ Alternatively, (λ^0, μ^0) can be obtained as the saddle point of

$$\frac{1}{2}|\mu|^2 - \frac{1}{2}|I^*(\lambda) + \mu|^2 + \langle z^0, \lambda \rangle . \tag{2.37}$$

Generally, we can expect that the use of algorithms for determining saddle points coupled with a numerical method to find Π_K will lead to an efficient general purpose numerical method for solving (1.2). However, this program requires further work for its successful implementation.

As an example of Theorem 2.8 we consider the problem of smooth *monotone* interpolation. This leads us, for instance, to the problem

$$\min \left\{ \int_0^1 (u^{(2)}(t))^2 \, dt : u \in H^2[0,1], u(t_i) = y_i , \right. \\ \left. i = 0, 1, \ldots, n, u'(t) \geq 0, \quad 0 \leq t \leq 1, \text{a.e.} \right\} \tag{2.38}$$

where $0 \leq t_0 < \cdots < t_n \leq 1$. This problem is equivalent, with the replacement $v = u'$, to (1.2) where

$$(Tv)(t) = v'(t), \quad C = \{v : v(t) \geq 0, 0 \leq t \leq 1, \text{a.e.}\},$$

$$Iv = \left(\int_{t_0}^{t_1} v(t)\, dt, \ldots, \int_{t_{n-1}}^{t_n} v(t)\, dt \right), \quad z = (y_1 - y_0, \ldots, y_n - y_{n-1}),$$

$$X = H^1[0,1], \quad Y = L^2[0,1], \quad \text{and} \quad Z = \ell^2(\mathbb{R}^n).$$

T is correct from X to Y and all other requirements in Theorem 2.8 are fulfilled. A straightforward computation shows that

$$(Gf)(t) = \int_0^1 G(t,x) f(x)\, dx , \quad f \in L^2[0,1],$$

where $G(t,x)$ is the strictly positive kernel

$$G(t,x) = \min(t,x) + \frac{1}{3} + t + x - \frac{1}{2}(x^2 + t^2).$$

Thus the cone K in this case is

$$K = \{g : Gg \geq 0\}.$$

According to Theorem 2.8 the solution of the optimization problem (2.37) is given as

$$u'_{\text{opt}}(t) = \int_0^1 G(t,x)\, \Pi_K(f_0)(x)\, dx$$

for some piecewise linear function f_0 with breakpoints at $t_j \in (0,1)$. We do not describe Π_K in detail except to note the following: Given *any*

$f \in L^2[0, 1]$, there is a nonnegative measure μ_f with support on $[0, 1]$ (depending on f) such that

$$\Pi_{K^*}(f)\ (x) = \int_0^1 G(t, x)\, d\mu_f(t)\,. \tag{2.38}$$

The dual extremal problem described in Theorem 2.8 gives information for finding f_0 and computing u'_{opt} numerically.

As for the reason for equation (2.38), we observe the following general fact.

LEMMA 2.9 : *Let* $K(x, t)$ *be continuous function for* $x, t \in [0, 1]$ *and suppose there does not exist a nontrivial measure* $d\mu(t)$ *such that* $\int_0^1 K(x, t)\, d\mu(x) = 0$ *for all* $t \in [0, 1]$ *then given any* $f \in L^2[0, 1]$ *there exists a nonnegative measure* μ_f *with support on* $[0, 1]$ *such that*

$$(\Pi_{K^*} f)\ (t) = \int_0^1 K(x, t)\, d\mu_f(x)$$

is the orthogonal projection of $L^2[0, 1]$ *onto the closed cone*

$$K = \left\{ g : g \in L^2[0, 1],\ \int_0^1 K(x, t)\, g(t)\, dt \geqslant 0, 0 \leqslant t \leqslant 1 \right\}.$$

Proof : Let $g^0 = \Pi_K f \in K$ be the best approximation of f in K. Then $(f - g_0, g) = \int_0^1 (f(t) - g_0(t))\, g(t)\, dt \geqslant 0$ whenever $g \in K$. Define the set $W = \left\{ \int_0^1 K(x, t)\, d\mu(x) : \ \mu \text{ a nonnegative measure} \right\}$. W is a *closed* convex cone in $L^2[0, 1]$ because any sequence of measures $d\mu_n$ such that the functions $\int_0^1 K(x, t)\, d\mu_n(x)$ converge in $L^2[0, 1]$ must have bound variation independent of n. As a result, if $f - g_0 \notin W$ there would exist an $h \in L^2[0, 1]$ such that $(g - g_0, h) < 0$ while necessarily $\int_0^1 K(x, t)\, h(t)\, dt \geqslant 0$ for all x. But then h would be in K which contradicts the definition of g_0.

3. BEST CONSTRAINED SMOOTHING

In this section we present a result similar to Theorem 2.8 for smoothing under convex constraints. We let T, I, X, Y and C be as before and consider the extremal problem

$$p_\rho = \inf \left\{ \frac{1}{2} \|Tx\|_X^2 + \frac{\rho}{2} \|Ix - z\|_Y^2 \right\} \tag{3.1}$$

for $\rho > 0$. This problem arises when the data vector is « noisy » hence the demand that $x \in C$ satisfies $Ix = z^0$, as in (1.2), would generally lead to an inappropriate estimate for x. Now, for (3.1) we study the concave-convex function $J_\rho(\lambda, \mu)$

$$J_\rho(\lambda, \mu) := \frac{1}{2} |\mu|^2 - \frac{1}{2} |\Pi_K(I^* \lambda + \mu)|^2 + \\ + \langle \lambda, z^0 \rangle - \frac{1}{2\rho} \|\lambda\|^2 - \frac{1}{2} \|z^0\|^2 \quad (3.2)$$

where $|\,.\,|$ and Π_K are as in the preceding section. This leads us to

THEOREM 3.1 : *Let T be correct from X to Y and I a bounded linear operator from X into a finite dimensional Hilbert space Z. Suppose C is a closed convex cone and z^0 a data vector in Z such that $I^{-1}(z^0) \cap C \neq \phi$ and*

$$C \cap \operatorname{Ker}(I) \cap \operatorname{Ker}(T) = \{0\} .$$

Then for each $\rho > 0$ there exists a $\gamma_\rho^0 = (\lambda_\rho^0, \mu_\rho^0)$ which is a saddle point of J_ρ and p_ρ is its corresponding saddle value. Moreover, the unique solution of (3.1) *is given by*

$$x_\rho^0 = G(\Pi_K(H(\gamma_\rho^0))) \quad (3.3)$$

where

$$P(\Pi_K(H(\gamma_\rho^0))) = \mu^0 \quad (3.4)$$

and

$$I(x_\rho^0) + \frac{1}{\rho} \lambda_\rho^0 = z^0 . \quad (3.5)$$

Proof : The proof is similar to the proof of Theorem 2.10 and so we shall only briefly discuss the details centering upon the differences in the proof.

Since

$$J_\rho(\lambda, \mu) = J(\lambda, \mu) - \frac{1}{2\rho} \|\lambda\|^2 + \frac{1}{2} \|z^0\|^2$$

we see that $-J_\rho(., \mu)$ for each $\mu \in \operatorname{Ker}(T)$ has no direction of recession and the functions $J_\rho(\lambda, .)$, $\lambda \in Z$ have, as with the case $\rho = 0$, no common direction of recession. Thus J_ρ has a saddle point $\gamma_\rho(\lambda_\rho^0, \mu_\rho^0) \in Z \times \operatorname{Ker}(T)$ which we will show satisfies (3.3), (3.4), and (3.5).

First, note that

$$\nabla_{(\lambda, 0)} J(\gamma_\rho^0) = \left\langle - I(G(\Pi_K(\gamma_\rho^0))) + z^0 - \frac{1}{\rho} \lambda^0, \lambda \right\rangle$$

and

$$\nabla_{(0,\mu)} J(\gamma_\rho^0) = (\mu_\rho^0 - \Pi_K(H(\gamma_\rho^0)), \mu) .$$

This proves (3.4) and (3.5) ; it remains to show (3.3) solves (3.1).

Let $\omega \in C$ (clearly $x_\rho^0 \in C$), then the equation

$$T^* T(x_\rho^0) = \Pi_K(H(\gamma_\rho^0)) - \mu_\rho^0$$

gives

$$\begin{aligned}\langle T^* T(x_\rho^0), \omega \rangle &= \langle H(\gamma_\rho^0) - \mu_\rho^0 - \Pi_{C^*}(H(\gamma_\rho^0)), \omega \rangle \\ &= \langle \lambda_\rho^0, I\omega \rangle - \langle \Pi_{C^*}(H(\gamma_\rho^0)), \omega \rangle .\end{aligned}$$

Therefore

$$\langle T(x_\rho^0), T\omega \rangle + \rho \langle I(x_\rho^0) - z^0, I\omega \rangle = - \langle \Pi_{C^*}(H(\gamma_\rho^0)), \omega \rangle \geqslant 0$$

which establishes the optimality of x_ρ^0 for (3.1). We omit the computation that verifies that the saddle value of J_ρ is p_ρ.

REFERENCES

[1] L. D. IRVINE, S. P. MARIN and P. W. SMITH, *Constrained interpolation and smoothing,* Constr. Approx., 2 (1986), 129-152.

[2] P. J. LAURENT, *Approximation et Optimization,* Herman, Paris, 1972.

[3] C. A. MICCHELLI, P. W. SMITH, J. SWETITS and J. D. WARD, *Constrained L^p approximation,* Constr. Approx., 1 (1985), 93-102.

[4] C. A. MICCHELLI and F. UTRERAS, *Smoothing and interpolation in a convex subset of a Hilbert space,* SIAM J. Sci. Statist. Comput., 9 (1988), 728-746.

[5] R. T. ROCKAFELLAR, *Convex Analysis,* Princeton University Press, Princeton, N.J., 1970.

[6] F. UTRERAS, *Smoothing noisy data under monotonicity constraints : existence, characterization and convergence, rates.* Numer. Math., 74 (1985), 611-625.

[7] F. J. BEUTLER and W. L. ROOT, *The operator pseudoinverse in control and systems identification, in Generalized Inverses and Applications,* eds. Z. Nashed, Academic Press, New York, 1973.

[8] C. K. CHUI, F. DEUTSCH and J. D. WARD, *Constrained best approximation in Hilbert space,* Constr. Approx., 6 (1990), 35-64.

[9] N. DYN and W. H. WONG, *On the characterization of non-negative volume matching surface splines,* J. Approx. Theory, 31 (1987), 1-10.