

## A NOTE ON LEAST-SQUARES APPROXIMATION WITH RANGE CONSTRAINTS

R. J. EVANS and A. CANTONI

(Received 13 July 1978)

(Revised 15 November 1978)

### Abstract

In this note we consider various theoretical aspects of the problem of least-squares approximation subject to constraints on the range of the approximating polynomial. The problem is treated from an optimization theory viewpoint. Rice's parameter space procedure is discussed.

### 1. Introduction

In this paper we investigate least-squares approximation subject to restricted range constraints. Surprisingly, this problem does not appear to have been discussed in detail in the literature although considerable attention has been directed at the problem of range constrained uniform approximation [7, 8, 12, 13, 14]. Some general aspects of the range constrained least-squares problem were considered by Rice [11] and the special case of discrete least-squares approximation with uniform constraints was considered by Rabinowitz [10] and Krabs [6]. Here we consider the problem of least-squares approximating a square-integrable function by a function composed of a linear combination of continuous functions but with the additional requirement that range constraints are imposed on the approximating function. The problem is treated as a function space optimization problem with a continuous family of linear inequality constraints. Existence, unicity, characterization and continuity theorems are presented.

Throughout this paper we follow standard notation and denote the Hilbert space of functions square-integrable on  $[a, b]$  by  $L^2[a, b]$ , and the Banach space of continuous functions on  $[a, b]$ , equipped with a uniform norm, as  $C[a, b]$ . Whenever a norm is written without a subscript we are referring to an  $L^2$  norm.

**Problem formulation**

Restricted range least-squares approximation of a square-integrable function  $v_0 \in L^2[a, b]$  by a finite linear combination of continuous orthonormal functions  $v_i \in C[a, b], i = 1, \dots, n$ , is easily seen to be equivalent to the optimization problem:

$$\underset{g \in G}{\text{minimize}} \quad \|v_0 - g\|^2,$$

where the approximant  $g$  is given by

$$g = \sum_{i=1}^n \alpha_i v_i, \quad \alpha_i \in \mathbb{R}, \quad i = 1, \dots, n,$$

for some fixed  $n$  and the range of the approximant is constrained by the set

$$G = \{g \in C[a, b]: l(x) \leq \sum_{i=1}^n \alpha_i v_i(x) \leq u(x), x \in [a, b]\}.$$

Clearly the range of the approximating polynomial  $\sum_{i=1}^n \alpha_i v_i(x)$  is restricted to lie between two continuous functions  $u, l \in C[a, b]$ , called the range constraints. We will assume that

- (i)  $l(x) < u(x), x \in [a, b]$ ,
- (ii)  $G$  is not empty.

**REMARKS.** (i) Note that in many papers on range-constrained approximation (for example [4]) a third condition  $l(x) \leq v_0(x) \leq u(x)$ , for all  $x \in [a, b]$  is required. We make no such requirement here.

(ii) The inequality  $z(x) \leq y(x), x \in [a, b]$ , defines a partial ordering on  $C[a, b]$  and in order to simplify notation we will write this as  $z \leq y$ .

It is important to note that this partial ordering is generated by a closed positive cone

$$P = \{y \in C[a, b]: y \geq 0\}$$

and, further, since  $C[a, b]$  is equipped with a uniform norm it follows that  $P$  has a non-empty interior [9].

(iii) It will be useful in the sequel to make the following definitions

$$\begin{aligned} \text{(a)} \quad \Omega(n) &= \left\{ \alpha \in \mathbb{R}^n: l \leq \sum_{i=1}^n \alpha_i v_i \leq u \right\} \\ &= \left\{ \alpha \in \mathbb{R}^n: \sum_{i=1}^n \alpha_i v_i - u \leq 0 \right\} \cap \left\{ \alpha \in \mathbb{R}^n: - \sum_{i=1}^n \alpha_i v_i + l \leq 0 \right\}, \end{aligned}$$

and

$$(b) f: R^n \rightarrow R: f(\alpha) = \left\| \sum_{i=1}^n \alpha_i v_i - v_0 \right\|^2,$$

$$H: R^n \rightarrow C[a, b]: [H(\alpha)](x) = \sum_{i=1}^n \alpha_i v_i - u,$$

$$K: R^n \rightarrow C[a, b]: [K(\alpha)](x) = - \sum_{i=1}^n \alpha_i v_i + l.$$

(iv) We will assume that  $\Omega(n)$  is non-empty and that  $\{v_i\}_{i=1}^n$  are orthonormal on  $[a, b]$ .

It is now straightforward to see that the range constrained least-squares approximation problem can be rewritten as

$$\left. \begin{array}{l} \text{minimize } f(\alpha), \alpha \in R^n, \\ \text{subject to } H(\alpha) \leq 0, K(\alpha) \leq 0. \end{array} \right\} \tag{1}$$

There are a variety of applications of the range-constrained least-squares approximation problem. An example from control theory is the optimal pulse amplitude modulation control of a distributed parameter system with state constraints. The control is given by the vector  $\alpha$ , the desired output trajectory is the function to be approximated  $v_0, v_i, i = 1, \dots, n$ , are delayed versions of the system impulse response, and the range constraints  $u$  and  $l$  are the output state constraints. For this example, of course, it is unlikely that the  $v_i$  would be orthonormal but this is of no real consequence.

There are also applications in system theory where the problem is to approximately realize a certain system impulse response using a set of orthonormal functions. For system performance reasons it may be necessary to ensure a prescribed degree of approximation over certain portions of the response and this can be achieved easily with the range constraints.

### 2. Main results

In this section we present several results concerning existence, uniqueness and characterization of the optimal approximant.

LEMMA 1. (*Existence.*) *An optimal approximant exists under the assumptions made in Section 1.*

PROOF. Since  $H$  and  $K$  are continuous and linear, and  $P$  is closed, it follows the  $\Omega(n)$  is closed. Finally  $f(\alpha)$  is a continuous, coercive operator which is bounded below on  $\Omega(n)$  hence existence follows [3, p. 35].

LEMMA 2. (*Uniqueness.*) *The optimal approximant is unique.*

PROOF.  $f$  is strictly convex and  $H$  and  $K$  are linear [11].

In order to proceed further we need certain results from function space optimization theory. A self-contained exposition is presented in the Appendix. Applying these results to the range-constrained least-squares approximation problem formulated above (1) we have the following.

(i) The convex Lagrangian functional for (1) is given by

$$L[\alpha; \lambda_1, \lambda_2] = \left\| v_0 - \sum_{i=1}^n \alpha_i v_i \right\|^2 + \int_a^b \left\{ \sum_{i=1}^n \alpha_i v_i - u \right\} d\lambda_1 - \int_a^b \left\{ \sum_{i=1}^n \alpha_i v_i - l \right\} d\lambda_2,$$

where the Lagrange multiplier functions are positive measures which belong to the dual positive cone,  $\lambda_1, \lambda_2 \in P^*$ ,

$$P^* = \left\{ y^* \in M[a, b] : \int_a^b y dy^* \geq 0, y \in P \right\}$$

and  $M[a, b]$  is the space of regular Borel measures on  $[a, b]$  [2].

(ii) Define the set

$$\hat{\Omega}(n) = \{ \alpha \in R^n : H(\alpha) < 0 \} \cap \{ \alpha \in R^n : K(\alpha) < 0 \}.$$

(iii) The concave dual functional for (1) is given by

$$\phi(\lambda_1, \lambda_2) = \inf_{\alpha \in R^n} L[\alpha; \lambda_1, \lambda_2].$$

THEOREM 3. *Given that  $\hat{\Omega}(n)$  is non-empty, that is, there exists an  $\alpha \in R^n$  such that the range constraints can be satisfied strictly, then  $\alpha^0 \in R^n$  solves (1) if and only if there exist Lagrange multipliers  $\lambda_1^0, \lambda_2^0 \in P^*$  such that*

(a)  $L[\alpha^0; \lambda_1^0, \lambda_2^0] \leq L[\alpha; \lambda_1^0, \lambda_2^0], \alpha \in R^n,$

(b)  $\sum_{i=1}^n \alpha_i^0 v_i - u \leq 0,$

$$- \sum_{i=1}^n \alpha_i^0 v_i + l \leq 0,$$

(c)  $\int_a^b \left\{ \sum_{i=1}^n \alpha_i^0 v_i - u \right\} d\lambda_1^0 = 0,$

$$- \int_a^b \left\{ \sum_{i=1}^n \alpha_i^0 v_i - l \right\} d\lambda_2^0 = 0.$$

PROOF. This follows from a direct application of the function space version of the Kuhn–Tucker Theorem presented in the Appendix.

COROLLARY 4. *The optimal Lagrange multipliers  $\lambda_1^0, \lambda_2^0$  are mutually singular [2].*

PROOF. From conditions (b) and (c) of Theorem 3, we see that the  $d\lambda_i, i = 1, 2$ , are non-zero only where the range constraints are active. Also it is clear from the problem formulation that both range constraints cannot be active simultaneously, thus wherever  $d\lambda_1$  is non-zero,  $d\lambda_2$  must be zero and vice versa, that is the Lagrange multipliers are mutually singular.

THEOREM 5. (*Characterization.*) *Problem (1) is solved by*

$$\alpha_i^0 = \hat{\alpha}_i - \int_a^b v_i d\lambda^0, \quad i = 1, \dots, n,$$

where  $\hat{\alpha}_i = \int_a^b v_0 v_i dx, i \dots 1, n$ , solves the unconstrained least-squares approximation problem

$$\text{minimize } \left\| \sum_{i=1}^n \alpha_i v_i - v_0 \right\|^2,$$

and  $\lambda^0 \in M[a, b]$  solves the unconstrained quadratic optimization problem

$$\text{minimize } \phi(\lambda), \quad \lambda \in M[a, b],$$

where

$$\begin{aligned} \phi(\lambda) = & \left\| \sum_{i=1}^n \left( \hat{\alpha}_i - \frac{1}{2} \int_a^b v_i d\lambda \right) v_i - v_0 \right\|^2 + \int_a^b \left\{ \sum_{i=1}^n \left( \hat{\alpha}_i - \frac{1}{2} \int_a^b v_i d\lambda \right) v_i \right\} d\lambda \\ & - \frac{1}{2} \int_a^b (u+l) d\lambda - \frac{1}{2} \int_a^b (u-l) |d\lambda| - \int_a^b v_0 d\lambda. \end{aligned}$$

PROOF. Using Corollary 4, the Lagrangian can be rewritten as

$$L[\alpha; \lambda] = \left\| \sum_{i=1}^n \alpha_i v_i - v_0 \right\|^2 + \int_a^b \left\{ \sum_{i=1}^n \alpha_i v_i - \frac{1}{2}(u+l) \right\} d\lambda + \frac{1}{2} \int_a^b (u-l) |d\lambda|,$$

where  $d\lambda = d\lambda_1 - d\lambda_2$  and  $|d\lambda| = d\lambda_1 + d\lambda_2$ . Now  $L[\alpha; \lambda]$  can be minimized with respect to  $\alpha$  by setting

$$\begin{aligned} \alpha_i &= \int_a^b v_0 v_i dx - \frac{1}{2} \int_a^b v_i d\lambda \\ \hat{\alpha}_i &= -\frac{1}{2} \int_a^b v_i d\lambda. \end{aligned}$$

Next the dual functional  $\phi(\lambda)$  is found by substituting this  $\alpha_i$  into the Lagrangian, and using the orthonormality of  $\{v_i\}$ . Finally, the generalized duality theorem is applied (see Appendix).

**THEOREM 6.** (*Alternative characterization.*) *Problem (1) is solved by*

$$\alpha_i^0 = \int_a^b v_i v_0 dx - \int_a^b v_i d\lambda^0, \quad i = 1, \dots, n,$$

for any  $\lambda^0 \in M[a, b]$  satisfying

$$\sum_{i=1}^n \alpha_i^0 v_i(x) = \begin{cases} u(x) & \text{if } \lambda^0(x) > 0, \quad x \in [a, b], \\ l(x) & \text{if } \lambda^0(x) < 0, \quad x \in [a, b]. \end{cases}$$

**PROOF.** This can be obtained directly from Theorem 5. Alternatively one may use duality theory as follows. The dual functional  $\phi(\lambda)$  obtained in Theorem 5 is convex and is therefore subdifferentiable [3]. Moreover a necessary and sufficient condition for  $\lambda^0 \in M[a, b]$  to minimize  $\phi(\lambda)$  is that zero belongs to the subdifferential of  $\phi(\lambda^0)$ , that is  $0 \in \partial\phi(\lambda^0)$ . To express this condition for our problem we simply observe that the smallest element of the subgradient is given by the function which minimizes the directional Gateaux differential.

It is interesting to note that the solution to the range-constrained least-squares problem is given by the solution to the unconstrained least-squares problem plus a correction term to account for the constraints. In [11] Rice mentioned the possibility of solving constrained approximation problems by firstly solving the unconstrained problem and then finding the element in the constraint set which is closest in the parameter-space norm sense to the optimal unconstrained parameter. Using Theorem 5 above it is quite easy to show that Rice's procedure is in fact optimal for our problem.

**LEMMA 7.** *The solution to range-constrained least-squares problem can be found by solving the problem*

$$\text{minimize} \quad \sum_{i=1}^n (\hat{\alpha}_i - c_i)^2, \quad \mathbf{c} \in R^n, \tag{2}$$

$$\text{subject to} \quad H(\mathbf{c}) \leq 0, \\ K(\mathbf{c}) \leq 0.$$

where†

$$\hat{\alpha}_i = \arg \min_{\alpha \in R^n} \left\| \sum_{i=1}^n \alpha_i v_i - v_0 \right\|^2.$$

---

† The notation  $\arg \min$  is not standard but the meaning is straightforward. For example,  $\arg \min f(\alpha)$  is the value of  $\alpha$  for which  $f(\alpha)$  is minimized.

PROOF. Following the proof of Theorem 5 we see that the dual function for (2) above is the same as the dual for (1) except for a constant.

Another interesting interpretation of  $\alpha_i^0$  can be obtained by decomposing  $\lambda^0$  into an integrable part and a saltus part [2]

$$d\lambda^0 = c^0(x) dx + \sum_{j=1}^{\infty} p_j^0 \delta(x - x_j),$$

where  $c^0 \in L^1[a, b]$  is the Radon-Nikodym derivative of  $d\lambda$  and  $p_j^0 \in \mathbb{R}$ . Now

$$\alpha_i^0 = \hat{\alpha}_i - \frac{1}{2} \int_a^b v_i(x) g^0(x) dx - \frac{1}{2} \int_a^b v_i(x) \sum_{j=1}^{\infty} p_j^0 \delta(x - x_j) dx.$$

Thus  $\alpha^0$  can be thought of as the solution to the unconstrained approximation problem

$$\text{minimize } \left\| \sum_{i=1}^n \alpha_i v_i - \hat{v} \right\|^2, \quad \alpha \in \mathbb{R}^n,$$

where

$$\hat{v} = v_0 + g^0 + \sum_{j=1}^{\infty} p_j^0 \delta(x - x_j).$$

We now move on to consider the continuity properties of the best approximation operator. The best approximation operator  $A\{v_i\}, u, l$  will be written simply as  $A$  from now on, where a certain orthogonal set  $\{v_i\}$  and upper and lower bounds  $u$  and  $l$  are understood to be associated with any  $A$ , where

$$A: C[a, b] \rightarrow C[a, b]: Av_0 = \sum_{i=1}^n \alpha_i^0 v_i \quad \text{where } \alpha_i^0, i = 1, \dots, n, \text{ solves (1).}$$

THEOREM 8. (Continuity of best approximation operator.) Given two continuous functions  $v_{01}, v_{02}$  then the optimum restricted range least-squares approximations  $Av_{01}$  and  $Av_{02}$  satisfy for some  $\tau > 0$

$$\|Av_{01} - Av_{02}\| \leq \tau \|v_{01} - v_{02}\|.$$

PROOF. Assume  $Av_{01}$  and  $Av_{02}$  exist. From the continuity of unconstrained least-square approximation operator we have that there exists a  $\gamma > 0$  such that

$$\|\hat{\alpha}_1 - \hat{\alpha}_2\| \leq \gamma \|v_{01} - v_{02}\|.$$

Now applying Lemma 7 we see that continuity of  $A$  is equivalent to the existence of some  $\rho > 0$  for which

$$\|c_1^0 = c_2^0\| \leq \rho \|\hat{\alpha}_1 - \hat{\alpha}_2\|,$$

where

$$c_i^0 = \arg \min_{c \in \Omega(n)} \|\hat{\alpha}_i - c\|^2, \quad i = 1, 2.$$

Now is it clear that we can write

$$c_i^0 = \text{Proj} [\Omega(n), \hat{\alpha}_i],$$

where  $\text{Proj} [\Omega(n), \hat{\alpha}_i]$  is the projection operation which projects the element  $\hat{\alpha}_i$  onto the closed convex set  $\Omega(n)$  and it is well known that this projection operation is continuous [5]. The theorem now follows immediately.

The following definitions and lemmas will be useful in the sequel. The optimal  $n$ -dimensional range-constrained approximation of a function  $v_0$  will be denoted  $A_n v_0$  and from Theorem 5 above we have

$$A_n v_0 = U_n + C_n,$$

where

$$U_n = \sum_{i=1}^n v_i \int_a^b v_0 v_i dx$$

and

$$C_n = -\frac{1}{2} \sum_{i=1}^n v_i \int_a^b v_i d\lambda^0.$$

Note that  $\lambda^0$  which maximized  $\phi(\lambda)$  is a function of  $n$  and  $v_0$  so from now on we will write  $\lambda_n^0$ .

LEMMA 9.

$$\lim_{n \rightarrow \infty} \|U_n - U_\infty\| = 0.$$

PROOF. This follows directly from Bessel's inequality [1].

LEMMA 10. *The sequence  $\{\lambda_i^0\}_{i=1}^\infty$  is bounded in norm.*

PROOF. It is clear that the sequence of optimal primal costs  $\{f(\alpha^0)\}_{n=1}^\infty$  is a monotonically decreasing sequence which is bounded below by zero. Moreover, the dual cost  $\phi(\lambda^0)$  is equal to the primal cost for each  $n$ , thus the sequence  $\{\phi(\lambda_n^0)\}_{n=1}^\infty$  converges, that is

$$\lim_{n \rightarrow \infty} \phi(\lambda_n^0) = \phi(\lambda_\infty^0).$$

† Note that,  $\Omega(n)$  is a closed convex set for all  $l, u, n$  and  $\{v_i\}$  satisfying the assumptions stated in Section 1.

Now if  $\{\lambda_i^0\}$  is *not* bounded in norm there must exist a subsequence  $\{\lambda_k^0\}$  for which

$$\|\lambda_k^0\| \rightarrow \infty.$$

Using the assumption that  $\hat{\Omega}(n)$  is non-empty it can be shown that such a subsequence contradicts the converge of  $\phi(\lambda_i^0)$ .

LEMMA 11.

$$\lim_{n \rightarrow \infty} \left\| \int_a^b v_i d\lambda_n^0 - \int_a^b v_i d\lambda_\infty^0 \right\| = 0.$$

PROOF. From Lemma 10 we have the sequence  $\{\lambda_i^0\}_{i=1}^\infty$  is bounded in norm. Further, by Alaoglu’s Theorem [2] the unit ball in  $M[a, b]$  is weak-star compact thus there must exist a subsequence of  $\{\lambda_n^0\}$  which converges weak-star to a limit  $\hat{\lambda}$  and it can be shown that  $\hat{\lambda} = \lambda_\infty$ . From this it follows that  $\{\int_a^b v_i d\lambda_n^0\}$  converges for every  $i$  to  $\int_a^b v_i d\lambda_\infty^0$ . Moreover this is a continuous function of  $\lambda_n^0$  so  $\{\int_a^b v_i d\lambda_n^0\}$  is bounded and the Lebesgue Dominated Convergence Theorem [2] implies mean-square convergence.

THEOREM 12.

$$\lim_{n \rightarrow \infty} \|A_n v_0 - A_\infty v_0\| = 0.$$

PROOF. Clearly

$$\|A_n v_0 - A_\infty v_0\| \leq \|U_n - U_\infty\| + \|C_n - C_\infty\|.$$

The first term converges by Lemma 9 and using Lemma 11 we can show that the second term will also converge. Defining

$$\xi_i^n = \int_a^b v_i d\lambda_n^0$$

we see that

$$\|C_n - C_\infty\| \leq \left\| \sum_{i=1}^n \xi_i^\infty v_i - \sum_{i=1}^\infty \xi_i^\infty v_i \right\| + \left\| \sum_{i=1}^n \xi_i^n v_i - \sum_{i=1}^n \xi_i^\infty v_i \right\|.$$

The second term on the right-hand side converges by Lemma 11, and the first term converges because  $\sum_{i=1}^\infty |\xi_i^\infty|^2 < \infty$ .

### 3. Computational aspects and conclusions

In this paper the range-constrained least-squares approximation problem has been investigated using the techniques of function space optimization theory and

convex analysis. A variety of theorems concerning the existence and uniqueness of the optimal approximant and its characterization and properties have also been presented. One aspect which has not been discussed, however, is the availability of algorithms for finding the optimal approximant. The major difficulty of course lies in the infinite dimensionality of the constraints, making some type of discretization necessary before computational algorithms can be applied. It may be possible to discretize the primal problem directly as is done in [16], for the maximum norm problem and in [15] for an infinite dimensional quadratic programming problem. Alternatively it is possible to discretize the dual problem by approximating the Lagrange Multiplier  $\lambda \in M[a, b]$  so that

$$\int_a^b s(x) d\lambda_j(x) = \sum_{j=1}^m \lambda_j \int_{x_j}^{x_{j+1}} s(x) dx,$$

where  $\{x_j\}_{j=1}^m$  is some partition of  $[a, b]$ , that is  $d\lambda$  has been approximated by rectangular pulses of height  $\lambda_j$  on each interval  $[x_j, x_{j+1}, j = 1, \dots, m-1$  [4].

Now by solving the *unconstrained finite dimensional convex* (but non-differentiable) dual problem for increasingly finer partitions,  $m, m+1, \dots$ , we obtain a sequence of approximating functions and it can be shown using techniques similar to Lemma 1 that this sequence convergences in mean-square to the optimal approximant which solves (1).

### Acknowledgements

The authors are grateful to the anonymous reviewers for a number of important comments and suggestions which significantly improved this paper. This work was supported by the Australian Research Grants Committee and the Radio Research Board.

### Appendix

In this appendix some results from the theory of constrained minimization in normed linear spaces are stated. Since the results are well known [9] the presentation here is very brief and is designed to summarize some of the key theorems.

Let  $U$  and  $Z_i, i = 1, 2, \dots, n$ , be real normed linear spaces, with convex cones [9]  $P_i \subseteq Z_i, i = 1, 2, \dots, n$ . For each cone  $P_i$  generates a partial ordering [9] on  $Z_i$ , denoted  $\leq$  and defined by

$$x \leq y \Leftrightarrow y - x \in P_i, \quad x, y \in Z_i. \tag{A.1}$$

Here  $P_i$  is called the positive cone of  $Z_i$ , since  $P_i = \{z \in Z_i: z \geq 0\}$ , and the notation  $x < y$  indicates that  $y - x$  is an interior point of  $P_i$ . The dual positive cone  $P_i^*$  in the dual space  $Z_i^*$  is defined by

$$P_i^* = \{z^* \in Z_i^*: z^*z \geq 0 \text{ for all } z \in P_i\}$$

where the notation  $z^*z$  is used to indicate that  $z$  is the argument of the functional  $z^*$ .

The function  $H: U \rightarrow Z$  is said to be convex (with respect to the partial ordering  $\leq$ ) if

$$H[\gamma u + (1 - \gamma)v] \leq H(u) + (1 - \gamma)H(v), \quad u, v \in U, \quad \gamma \in [0, 1].$$

Consider the optimization problem

$$\begin{aligned} &\text{minimize } f(\alpha), \quad \alpha \in U, & \text{(A.2)} \\ &\text{subject to } H_i(\alpha) \leq 0, \quad i = 1, 2, \dots, n, \end{aligned}$$

where  $f: U \rightarrow R^1$  and  $H_i: U \rightarrow Z_i, i = 1, 2, \dots, n$ , are all convex. The Lagrangian function  $L: U \times Z_1^* \times Z_2^* \times \dots \times Z_n^* \rightarrow R^1$  for problem (A.2) is defined by

$$L(\alpha; \lambda_1, \lambda_2, \dots, \lambda_n) = f(\alpha) + \sum_{i=1}^n \lambda_i \cdot H_i(\alpha). \tag{A.3}$$

If  $\lambda_i \in P_i^*$  and  $Z_i^*, i = 1, 2, \dots, n$ , are all given then  $L(\cdot; \lambda_1, \lambda_2, \dots, \lambda_n)$  is a convex functional on  $U$ . The dual frequency  $\phi: Z_1^* \times Z_2^* \times \dots \times Z_n^* \rightarrow R^1$  for problem (A.2) is defined by

$$\phi(\lambda_1, \lambda_2, \dots, \lambda_n) = \inf_{\alpha \in U} L(\alpha; \lambda_1, \lambda_2, \dots, \lambda_n) \tag{A.4}$$

and it is easily verified [9] that  $\phi$  is a concave functional.

Two key theorems are now stated.

**THEOREM A-1.** (*Generalized Kuhn-Tucker conditions.*) For  $i = 1, 2, \dots, n$ , let  $U, Z_i$  be real normed linear spaces, as above, with each of the positive cones  $P_i \subseteq Z_i$  having a non-empty interior. Let  $f: U \rightarrow R^1$  and  $H_i: U \rightarrow Z_i$  be convex, as above and assume the existence of  $\bar{\alpha} \in U$  such that  $H_i(\bar{\alpha}) < 0$  for all  $i$ . Then  $\alpha^0 \in U$  solves problem (A.2) if and only if there exist Lagrange multipliers  $\lambda_i^0 \in P_i^*, i = 1, 2, \dots, n$ , such that

$$(i) \quad L(\alpha^0; \lambda_1^0, \lambda_2^0, \dots, \lambda_n^0) \leq L(\alpha; \lambda_1^0, \lambda_2^0, \dots, \lambda_n^0), \quad \alpha \in U, \tag{A.5}$$

$$(ii) \quad \lambda_i^0 \cdot H_i(\alpha^0) = 0, \quad i = 1, 2, \dots, n, \tag{A.6}$$

$$(iii) \quad H_i(u^0) \leq 0, \quad i = 1, \dots, n.$$

PROOF. See [9].

Condition (ii) of the above theorem expresses the fact that a Lagrange multiplier is non-zero only where the corresponding constraint is active. The following duality theorem provides a method for finding the Lagrange multipliers.

**THEOREM A-2.** (*Generalized duality.*) *The Lagrange multipliers of Theorem (A-1),  $\lambda_i^0 \in P_i^*$ ,  $i = 1, 2, \dots, n$ , satisfy*

$$(i) \quad \phi(\lambda_1^0, \lambda_2^0, \dots, \lambda_n^0) \geq \phi(\lambda_1, \lambda_2, \dots, \lambda_n), \lambda_i^0, \lambda_i \geq 0, \quad (A.7)$$

$$(ii) \quad \min_{H_i(u) \leq 0} f(\alpha) = \max_{\lambda_i \leq 0} \phi(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Conditions (i) and (ii) of Theorem A-1 hold for any  $\lambda_1, \lambda_2, \dots, \lambda_n$  which satisfy (i) above.

PROOF. See [9].

REMARKS. (i) There are many other approaches to optimization in function spaces but the summary presented here is adequate for our needs.

(ii) The requirement that the cone  $P \subseteq Z$  has a non-empty interior can, under certain circumstances, be quite restrictive as it dominates the choice of spaces allowed for the range of  $H$ . Several recent results allow certain relaxations of this requirement.

(iii) The requirement that a strictly feasible solution for the constraints must exist is in fact a constraint qualification. Several other constraint qualifications are available.

### References

- [1] E. Cheney, *Introduction to approximation theory* (New York: McGraw-Hill, 1966).
- [2] N. Dunford and I. Schwartz, *Linear operators. Part I* (New York: Interscience, 1958).
- [3] I. Eklund and R. Terman, *Convex analysis and variational problems* (Berlin: Springer-Verlag, 1976).
- [4] T. E. Fortmann and R. J. Evans, "Optimal filter design subject to output sidelobe constraints", *J. Opt. Theory Applics* 14 (1974), 271–290.
- [5] L. Kantorovich and G. Akilov, *Functional analysis in normed spaces* (New York: Pergamon, 1964).
- [6] W. Krabs, "Fehlerquadrat-Approximation als Mittel zur Losung des diskreten linearen Tschebyscheff-Problems", *Z. Angew. Math. Mech.* 44 (1964), 42–45.
- [7] J. Lewis, "Approximation with convex constraints", *SIAM Review* 15 (1973), 193–217.
- [8] H. Loeb, D. Moursund and G. Raylor, "Uniform rational weighted approximation having restricted ranges", *J. Approx. Theory* 1 (1968), 401–411.
- [9] D. Luenberger, *Optimization by vector space methods* (Wiley, 1969).

- [10] P. Rabinowitz, "Mathematical programming", in *Approximation theory*, (ed. A. Talbot) (London: Academic Press, 1970).
- [11] J. Rice, "Approximation with convex constraints", *SIAM J. Control* 11 (1963), 15–32.
- [12] L. Schumaker and G. Taylor, "On approximation by polynomials having restricted ranges II", *SIAM J. Num. Anal.* 6 (1969), 31–36.
- [13] G. Taylor, "On approximation by polynomials having restricted ranges", *SIAM J. Num. Anal.* 5 (1968), 258–268.
- [14] G. Taylor, "Approximation by functions having restricted ranges III", *J. Math. Anal. Appl.* 27 (1969), 241–248.
- [15] G. Wahba, "On the minimization of a quadratic functional subject to a continuous family of linear inequality constraints", *SIAM J. Control* 11 (1973), 64–79.
- [16] G. A. Watson, "The calculation of best restricted approximations", *SIAM J. Num. Anal.* 11 (1974), 693–699.

Department of Electrical Engineering  
University of Newcastle  
N.S.W. 2308  
Australia