A DUAL APPROACH TO MULTIDIMENSIONAL L_p SPECTRAL ESTIMATION PROBLEMS*

A. BEN-TAL[†], J. M. BORWEIN[‡], AND M. TEOULLE[§]

Abstract. A complete duality theory is presented for the multidimensional L_p spectral estimation problem. The authors use a new constraint qualification (BWCQ) for infinite-dimensional convex programs with linear type constraints recently introduced in [Borwein and Wolkowicz, *Math. Programming*, 35 (1986), pp. 83–96]. This allows direct derivation of the explicit optimal solution of the problem as presented in [Goodrich and Steinhardt, *SIAM J. Appl. Math.*, 46 (1986), pp. 417–426], and establishment of the existence of a simple and computationally tractable unconstrained Lagrangian dual problem. Moreover, the results illustrate that (BWCQ) is more appropriate to spectral estimation problems than the traditional Slater condition (which may only be applied after transformation of the problem into an L_p space [Goodrich and Steinhardt, *op. cit.*] and which therefore yields only necessary conditions).

Key words. spectral estimation, infinite-dimensional convex duality, constraint qualifications, moment problems

AMS (MOS) subject classifications 49, 90C, 93

1. Introduction. A basic problem in spectral estimation is the estimation of a *power spectrum*, a measure $\mu \in \mathbb{R}^n$, with a known support, given a finite collection of measured correlations. This problem arises in a wide range of settings such as geophysics, radio astronomy, sonar, and radar, to mention just a few. (See Kay and Marple [8] and the references therein.) In many of these physical problems, the power spectrum is represented by a density. Let $K \subset \mathbb{R}^n$ be a measure space with a σ -finite positive measure dk and let μ be absolutely continuous with respect to dk with density (Radon–Nikodym derivative) $s(k) = d\mu/dk$. Then the problem of estimating a power spectral density function may be stated as the following: Find a nonnegative integrable s on K which vanishes outside of K and exactly matches the measured correlations:

$$(1.1) \quad r(\delta) = \int_K e^{j\delta \cdot k} s(k) dk, \quad \delta \in \Delta \quad (j := \sqrt{-1})$$

where Δ is a finite subset of $\mathbb{R}^n$ with $0 \in \Delta$, $\Delta = -\Delta$ and $\bar{r}(\delta) = r(-\delta)$ for $\delta \in \Delta$. Even when feasible, this problem does not have a unique solution. To overcome this ambiguity, the selection of such an estimate s from the infinite number of possible choices is done by requiring the spectral estimate to optimize some convex functional.

One popular method is Burg's maximum entropy method [2]. In this method the spectral estimate is required to maximize the entropy functional

$$(1.2) \quad H(s) = \int_K \log s(k)$$

* Received by the editors November 20, 1986; accepted for publication (in revised form) October 21, 1987.

[†] Department of Industrial and Operation Engineering, University of Michigan, Ann Arbor, Michigan 48105-2117.

[‡] Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 3J5. The research of this author was partially supported by an operating grant from the National Sciences and Engineering Research Council of Canada.

[§] Department of Mathematics, University of Maryland, Baltimore County, Catonsville, Maryland 21228. The research of this author was supported by a Postdoctoral Fellowship at the Department of Mathematics, Statistics and Computing Science, Dalhousie University.

subject to the correlation-matching constraints (1.1). It can be shown [2] that if there exists a strictly positive trigonometric polynomial P such that the spectral density $1/P(k)$ satisfies the constraints (1.1), then $s_0(k) = 1/P(k)$ is the unique spectral density function solving (1.2) subject to (1.1) (see also the Appendix). In some applications [8], a strictly positive solution of the form $1/P(k)$ fails to exist. This motivated Goodrich and Steinhardt [5] to suggest in a recent paper, an alternative way for selecting a solution to problem (1.1) by formulating the following minimal L^2 norm problem; find $s^* \in L^2$ solving

$$(1.3) \quad \inf \left\{ \int_K s(k)^2 dk : \int_K e^{j\delta \cdot k} s(k) dk = r(\delta), \delta \in \Delta, S \geq 0 \right\}.$$

Under appropriate conditions, it is shown there that the optimal solution $s^*(k) = \max(0, P(k))$, where P is a trigonometric polynomial. The method of proof used in [5] is the Lagrange multiplier technique (11) applied to problem (1.3). The difficulty with this approach is that the nonnegative orthant in L_2 has an empty interior and thus the usual ‘‘Slater type’’ constraint qualification fails. To overcome this difficulty, it is shown in [5] that problem (1.3) can be reduced to a minimization problem in L_∞ by using the transformation $t(k) = (s(k))^{1/2}$. For the transformed problem the constraint qualification is satisfied and thus a Kuhn–Tucker theorem for infinite-dimensional problems (see Luenberger [13, p. 217]) is applicable.

In this paper we develop a duality theory for problem (1.3) which enables us to derive the results obtained in [5], directly from this duality framework. We prove our results using a constraint qualification recently introduced by Borwein and Wolkowicz [1], called accordingly (BWCQ). The dual problem obtained is a *finite-dimensional concave* program and the optimal dual variables are exactly the parameters of the optimal spectral density s^* . Moreover the simple unconstrained nature of the dual problem is suitable for computational purposes and may lead to the construction of reliable algorithms for computing the L^2 optimal spectral estimate s^* . In § 2 we review some basic results on the *extendibility problem* as developed by Lang and McClellan [11] and state the multidimensional spectral estimation problem (P) in an appropriate vector space. In § 3 we present the new constraint qualification (BWCQ) and obtain an explicit version of it for problems of type (P). In particular we show that the (BWCQ) is equivalent to the extendibility condition given in § 2. In § 4 a complete duality theory for the multidimensional problem is derived, including the explicit solution presented in [5]. In § 5 we specialize the duality results to the one-dimensional time series problem. In this case we prove that (BWCQ) simply reduces to the classical test of positive definiteness of a Toeplitz matrix formed from the correlation samples.

Finally, § 6 contains concluding remarks and briefly discusses how to extend our results when the constraint qualification fails.

2. Multi-dimensional spectral estimation and properties. We collect in this section some preliminary results needed in the sequel. We follow the work of Lang and McClellan [11]. Let Δ be a finite set in $\mathbb{R}^n$ of the form $\Delta = \{0, \pm\delta_1, \dots, \pm\delta_m\}$ and assume that

$$\{e^{jk \cdot \delta} : \delta \in \Delta\} \quad (j \text{ denotes } \sqrt{-1})$$

is a set of linearly independent functions on K a compact subset of $\mathbb{R}^n$. The measure correlations $r(\delta)$ are *conjugate symmetric* functions, i.e., $r(-\delta) = \bar{r}(\delta)$ for all $\delta \in \Delta$. The set Δ has $(2m+1)$ elements and so a conjugate symmetric function on Δ is characterized by $(2m+1)$ real numbers; thus we may think of $(r(\delta))_{\delta \in \Delta}$ as a $(2m+1)$ vector r in

$\mathbb{R}^{2m+1}$. Similarly we denote by ψ_k the vector in the $(2m+1)$ -dimensional Euclidean space with components $\psi_k(\delta) = e^{jk \cdot \delta}$.

Let E denote the set of *extendible correlation vectors*, that is, $r \in E$ if

$$(2.1) \quad r = \int_K \psi_k d\mu$$

for some finite positive measure μ on K . We define a real-valued Δ -polynomial by

$$P(k) = \sum_{\delta \in \Delta} p(\delta) e^{-jk \cdot \delta} \quad \text{where } p(-\delta) = \bar{p}(\delta).$$

Here we will also use the vector notation $p \in \mathbb{R}^{2m+1}$ for the coefficients $(p(\delta))_{\delta \in \Delta}$ of a Δ -polynomial $P(k)$. Define the set Q of positive Δ -polynomials $Q := \{p \in \mathbb{R}^{2m+1} : P(k) \geq 0, k \in K\}$. The inner product between a vector r of correlations samples and a vector p of polynomial coefficients is defined as

$$(r, p) = \sum_{\delta \in \Delta} \bar{r}(\delta) p(\delta).$$

Note that if $r = \int_K \psi_k d\mu$ then $(r, p) = \int_K P(k) d\mu$. The following theorem provides a characterization of extendibility.

THEOREM 2.1 [11]. *The vector r is extendible if and only if*

$$(2.2) \quad (r, p) \geq 0 \quad \text{for all } p \in Q.$$

Note that E and Q are closed convex cones and thus (2.2) simply states that $E = Q^+$, the polar cone of Q . A simple property of extendibility [11] is that $\text{int } E \neq \emptyset$ and

$$(2.3) \quad \text{int } E = \{r \in E : (r, p) > 0 \text{ for all } p \in Q, p \neq 0\}.$$

As mentioned in the Introduction, in many applications the power spectrum is not represented merely as a measure but as a density. Let K be a measure space with a σ -finite positive measure dk . For $1 < \alpha < \infty$, $L_\alpha \equiv L_\alpha(K, dk)$ will denote the Lebesgue space of α th power integrable real-valued functions on K , i.e.,

$$L_\alpha = \left\{ f : \int_K |f(k)|^\alpha dk < \infty \right\}.$$

Let μ be absolutely continuous with respect to dk with density (Radon-Nikodym derivative) $s(k) = d\mu/dk$. The constraint correlation measurements (2.1) are now

$$r = \int_K \psi_k s(k) dk, \quad s \in L_\alpha.$$

This leads to the following modified extension theorem for spectral densities.

THEOREM 2.2 [11]. *Suppose that K has a finite measure. If every neighborhood of each point in K has a strictly positive dk -measure (as holds with dk equivalent to Lebesgue measure on K), then $r \in \text{int } E$ if and only if there exists a continuous strictly positive function $s(k)$ such that*

$$r = \int_K \psi_k s(k) dk.$$

Remark 2.1. It is interesting to note that Theorems 2.1 and 2.2 remain valid when the linear independence assumption of the particular functions $\{e^{jk \cdot \delta} : \delta \in \Delta\}$ is replaced by any linear independent set of continuous functions on K .

Remark 2.2. It can be shown easily using Theorem 2.1 that in Theorem 2.2 the statement “there exists a *continuous* strictly positive function $s(k)$ ” can be replaced

by the formally weaker statement “there exists a strictly positive function in L_α .” This weaker version will in fact be sufficient for our purposes (see § 3).

We close this section by formulating our problem (1.3) in the appropriate normed vector space. We denote the nonnegative cone in L_α by $L_\alpha^+ = \{f \in L_\alpha : f \geq 0 \text{ a.e.}\}$. Consider the integral functional $I : L_\alpha \rightarrow \mathbb{R}$ defined by

$$I(s) := \|s\|_\alpha^\alpha = \int_K |s(k)|^\alpha dk$$

and the linear operator $A : L_\alpha \rightarrow \mathbb{R}^{2m+1}$ given by

$$(2.4) \quad s \rightarrow \int_K \psi_k s(k) dk.$$

Then our L_α -density estimation problem is to solve

$$(2.5) \quad (P) \quad \inf \{I(s) : As = r, s \in L_\alpha^+\}.$$

Note that for $1 < \alpha < \infty$, $I(s)$ is a convex Fréchet differentiable functional and so (P) is a linearly constrained convex differentiable optimization problem.

3. A constraint qualification for problem (P). In this section we study programs of type (P) and introduce a duality theory for them using a setting recently developed by Borwein and Wolkowicz [1]. We state a version of their results which will be appropriate for problem (P). Consider the following primal problem:

$$(A) \quad v(P) = \inf \{f(x) : Ax = b, x \in S\}$$

where $f : X \rightarrow \mathbb{R}$ is a differentiable convex functional on X , $A : X \rightarrow \mathbb{R}^N$ is a continuous linear operator, X is a normed space, $b \in \mathbb{R}^N$ and $S \subset X$ is a convex cone. We let X^* denote the continuous dual space of X and $\langle \cdot, \cdot \rangle$ the bilinear form in the duality between X and X^* . The Lagrangian dual problem (B) associated with (A) is defined as

$$(B) \quad v(D) = \sup_{y \in \mathbb{R}^N} \{y^T b + \inf_{x \in S} \{f(x) - \langle y, Ax \rangle\}\}.$$

The main result concerning the dual pair (A) and (B) is the existence of a saddle point (x^*, y^*) , for (A) or equivalently the validity of a strong duality result:

$$(3.1) \quad v(P) = v(D) \quad \text{and} \quad v(D) \text{ is attained by some } y^* \in \mathbb{R}^N.$$

Usually, under the familiar “Slater type” constraint qualification,

$$(SCQ) \quad \text{There exists } x \in \text{int } S \text{ such that } Ax = b$$

the strong duality relation is guaranteed. Unfortunately, in problem (P) (see (2.5)) (SCQ) does not hold, since there $S = L_\alpha^+$ the nonnegative cone in L_α , which has empty interior. However, we shall make use of a recent result given in [1] which guarantees the strong duality relation (3.1). We need the following definitions and notation. Given any set C in X , the *polar cone* of C is the set

$$C^+ = \{x' \in X^* : \langle x', x \rangle \geq 0 \text{ if } x \in C\}.$$

Let P be any generating set for S^+ , i.e.,

$$\text{cl } \{\text{cone } P\} = S^+$$

and define

$$(3.2) \quad P^\circ := \{s^+ \in P : Ax = b, x \in S \Rightarrow \langle s^+, x \rangle = 0\}.$$

We denote by $F = \{x : Ax = b, x \in S\}$ the feasible set of (A) and we assume $F \neq \emptyset$.

THEOREM 3.1 [1]. Assume that $v(P)$ is finite. If (BWCQ): $P^\circ \subset \{0\}$ holds. Then, $v(P) = v(D)$ and $v(D)$ is attained by some $y^* \in \mathbb{R}^N$. $\square$

Remark 3.1. It is interesting to mention that (BWCQ) is equivalent to the condition $\text{cl cone}(F - S) = X$ (see [1]).

We specialize (BWCQ) to a problem of type (P). Let K be a measure space with σ -finite positive measure dk . Let $X = L_\alpha(K, dk) = L_\alpha$, $1 < \alpha < \infty$, and $S := L_\alpha^+$ be the nonnegative orthant in L_α . Further, the bilinear form

$$\langle s, x \rangle = \int_K s(k)x(k) dk$$

defines the duality between L_α and L_β , where L_β is identified as usual with the dual space of L_α , and $\alpha + \beta = \alpha\beta$, ($1 < \beta < \infty$).

THEOREM 3.2. Suppose L_α is separable. For $1 \leq \alpha < \infty$ (BWCQ) holds if and only if the weak "Slater type" condition

$$(3.3) \quad \exists \hat{x} \in L_\alpha : A\hat{x} = b, \quad \hat{x} > 0 \quad (\text{a.e.})$$

is satisfied.

Proof. First, we show that if (3.3) holds then $P^\circ \subset \{0\}$. Let $\hat{x} \in L_\alpha$ such that $A\hat{x} = b$ and $\hat{x} > 0$ (a.e.). If $s^+ \in P^\circ$, then by Definition (3.2), $s^+ \in S^+$ and $\langle s^+, \hat{x} \rangle = \int_K s^+(k)\hat{x}(k) dk = 0$. Let $A_n := \{k : \hat{x}(k) \geq 1/n\}$, $n = 1, 2, \dots$; then since $s^+ \in S^+$ implies that $s^+ \in L_\beta$ with $s^+ \geq 0$ (a.e.), we have

$$0 \leq \frac{1}{n} \int_{A_n} s^+(k) dk \leq \int_K s^+(k)\hat{x}(k) dk = 0$$

so that $\int_{A_n} s^+(k) dk = 0$. But $\{k : x(k) > 0\} = \bigcup_{n=1}^\infty A_n$; hence

$$0 = \int_{\bigcup A_n} s^+(k) dk = \int_K s^+(k) dk \quad \text{implying } s^+(k) = 0 \quad (\text{a.e. } dk).$$

Next, we show the reverse implication. Since L_α is assumed separable, and since P° is trivial, there is a countable set $\{x_n\}_{n=1}^\infty$ in L_α such that

- (i) $Ax_n = b$, $x_n \in S$;
- (ii) If $0 \neq s^+ \in S^+$, $\langle s^+, x_n \rangle > 0$ for some $n \in \mathbb{N}$.

If $A \equiv 0$ there is nothing to prove; then let us assume $A \neq 0$. We distinguish between two cases.

Case 1. $b \neq 0$. Then since $F \neq \emptyset$, $Ax_n = b$ implies that $\|A\| \|x_n\| \geq \|b\|$ and thus $\|x_n\| \geq c > 0$. Hence the series $\sum_{n=1}^\infty 1/2^n \|x_n\|$ converges, say, to γ . Let $\hat{x} := 1/\gamma \sum_{n=1}^\infty x_n/2^n \|x_n\|$. Since L_α is Banach and $\sum_{n=1}^\infty \|x_n/2^n \|x_n\|\| = 1$, then $1/\gamma \sum x_n/2^n \|x_n\|$ converges. Then (i) shows that $\hat{x}$ is feasible and so satisfies (3.3) if we show $\hat{x} > 0$. Now (ii) shows that $\langle s^+, \hat{x} \rangle > 0$ if $s^+ \neq 0$, $s^+ \in S^+$. In particular, for any E in K of finite measure

$$\int_E \hat{x}(k) dk > 0$$

which easily shows that $\hat{x} > 0$ almost everywhere.

Case 2. $b = 0$. From (ii) $x_n > 0$. Thus with $\hat{x} := \sum_{n=1}^\infty x_n/2^n \|x_n\|$. Part (i) of Theorem 3.2 shows that $A\hat{x} = 0$ and with the same proof as above we also have $\hat{x} > 0$ almost everywhere. $\square$

Note that the separability assumption on L_α in Theorem 3.2 is needed only to prove the second implication.

An immediate consequence of Theorem 3.2 which associates the notion of extendibility with (BWCQ) is given in the next result.

COROLLARY 3.1. *Suppose K has a finite measure. Assume that every neighborhood of each point in K has a strictly positive dk -measure and $\{e^{jk \cdot \delta} : \delta \in \Delta\}$ is linearly independent on K . Then (BWCQ) holds if and only if $r \in \text{int } E$.*

Proof. The result follows from Theorem 3.2. Indeed, under our assumptions, Theorem 2.2 and Remark 2.2 are applicable and thus $r \in \text{int } E$ if and only if

$$(3.4) \quad \exists s \in L_\alpha : \int_K \psi_k s(k) dk = r, \quad s > 0 \quad (\text{a.e.})$$

which is exactly (3.3) with $b := r \in \mathbb{R}^{2m+1}$ and A as defined in (2.4). $\square$

4. Duality theory for problem (P). Throughout this section we assume that $\{e^{jk \cdot \delta} : \delta \in \Delta\}$ is a linearly independent set of functions on K , which has finite measure and that every neighborhood of each point in K has a strictly positive dk -measure. Recall also that $0 \in \Delta = -\Delta$ so that Δ has $2m+1$ elements, and $r(-\delta) = \bar{r}(\delta)$ so that the problem (P) is potentially feasible. For a real number a we denote $a_+ = \max(0, a)$. Our results are given in L_α with $1 < \alpha < \infty$.

THEOREM 4.1. *If $r \in \text{int } E$, then*

$$(4.1) \quad \min (P) = \max (D)$$

where the dual problem (D) of (P) is explicitly given by

$$(4.2) \quad \sup_{p \in \mathbb{R}^{2m+1}} \left\{ \sum_{\delta \in \Delta} p(\delta) \bar{r}(\delta) - c(\alpha, \beta) \int_K \left(\sum_{\delta \in \Delta} p(\delta) e^{-jk \cdot \delta} \right)_+^\beta dk \right\}.$$

Moreover if p^* solves (D) then the unique optimal solution of (P) is given by

$$(4.3) \quad s_*(k) = d(\alpha, \beta) \left(\sum_{\delta \in \Delta} p^*(\delta) e^{-jk \cdot \delta} \right)_+^{\beta-1}$$

where $c(\alpha, \beta) = \alpha^{1-\beta}/\beta > 0$ and $d(\alpha, \beta) = \alpha^{1-\beta}$.

Proof. Consider problem (P) defined in (2.5):

$$(4.4) \quad \inf I(s) \quad \text{s.t.} \quad \int_K \psi_k(\delta) s(k) dk = r(\delta), \\ \delta \in \Delta, s \in L_\alpha, \quad s(k) \geq 0 \quad \text{a.e.}$$

To construct its Lagrangian dual we associate a complex dual variable $p(\delta)$ (with $p(-\delta) = \bar{p}(\delta)$) to each complex constraint. In order to work in the real field, we split the real and imaginary parts of (4.4). Observing that for two complex numbers z_1, z_2 we have

$$(4.5) \quad \text{Re } z_1 \text{ Re } z_2 + \text{Im } z_1 \text{ Im } z_2 = \frac{1}{2}(z_1 \bar{z}_2 + \bar{z}_1 z_2).$$

Then, using relation (4.5), the Lagrangian function for (P) can be written

$$L(s, p) = I(s) + \frac{1}{2} \sum_{\delta \in \Delta} \{r(\delta) \bar{p}(\delta) + \bar{r}(\delta) p(\delta)\} \\ - \frac{1}{2} \int_K \left\{ \sum_{\delta \in \Delta} p(\delta) \bar{\psi}_k(\delta) + \sum_{\delta \in \Delta} \bar{p}(\delta) \psi_k(\delta) \right\} s(k) dk.$$

But since $r(\delta)$ and $p(\delta)$ are conjugate symmetric on Δ and $\Delta = -\Delta$, thus $L(s, p)$ is simply

$$L(s, p) = I(s) + \sum_{\delta \in \Delta} r(\delta) \bar{p}(\delta) - \int_K P(k) s(k) dk$$

where

$$P(k) := \sum_{\delta \in \Delta} p(\delta) \bar{\psi}_k(\delta) = \sum_{\delta \in \Delta} p(\delta) e^{-jk \cdot \delta}.$$

A Lagrangian dual for (P) is then

$$(4.6) \quad \sup_{p \in \mathbb{R}^{2m+1}} \left\{ \sum_{\delta \in \Delta} p(\delta) \bar{r}(\delta) + \inf_{s \in L_{\alpha}^+} \{I(s) - \langle s, P \rangle\} \right\}.$$

Since here $I(s) = \int_K |s(k)|^{\alpha} dk$ with $1 < \alpha < \infty$, $I(s)$ is a proper convex integral functional [15], and thus applying Theorem 3A of Rockafellar [16] we have

$$(4.7) \quad \inf_{s \in L_{\alpha}^+} \{I(s) - \langle s, P \rangle\} = \int_K \inf_{s \geq 0} \{s(k)^{\alpha} - P(k)s(k)\} dk.$$

Now it is easily verified that

$$(4.8) \quad \begin{aligned} \inf_{s \geq 0} \{s^{\alpha} - Ps\} &= -\frac{\alpha^{1-\beta}}{\beta} \{\max(0, P)\}^{\beta} \\ &= -c(\alpha, \beta)(P(k))_+^{\beta} = -c(\alpha, \beta) \left(\sum_{\delta \in \Delta} p(\delta) e^{-jk \cdot \delta} \right)_+^{\beta}. \end{aligned}$$

Hence, substituting (4.8) in (4.7), from (4.6) we obtain the dual problem (4.2).

From Theorem 2.2, $r \in \text{int } E$ implies that the feasible set

$$F := \{s \in L_{\alpha} : As = r, s \geq 0\}$$

is nonempty. Also note that F is a closed convex subset of the reflexive space L_{α} , so that the existence of an S^* with minimal L_{α} -norm is guaranteed, and since $I(s)$ is a strictly convex functional, S^* is unique. Moreover, by Corollary 3.1, $r \in \text{int } E$ if and only if (BWCQ) holds, thus Theorem 3.1 is applicable and (4.1) is proved.

The optimal solution s^* given in (4.3) is just the optimality condition for $s = s^*$ to solve the minimization problem (4.8). Now, the dual problem (D) in (4.2) is an unconstrained problem; its supremum is attained say at p^* only if p^* is a critical point of the supremand which is differentiable since $\beta > 1$, i.e., p^* is a solution of

$$(4.9) \quad \bar{r}(\delta) = \int_K \beta c(\alpha, \beta) \left(\sum_{\delta \in \Delta} p^*(\delta) e^{-jk \cdot \delta} \right)_+^{\beta-1} e^{-jk \cdot \delta} dk, \quad \delta \in \Delta.$$

Since $r(\delta)$ and $\psi_k(\delta) = e^{jk \cdot \delta}$ are conjugate symmetric on Δ , and $\Delta = -\Delta$, thus (4.9) is nothing else but

$$r(\delta) = \int_K s_*(k) \psi_k(\delta) dk, \quad \delta \in \Delta$$

with $s_*(k) = d(\alpha, \beta) \left(\sum_{\delta \in \Delta} p^*(\delta) e^{-jk \cdot \delta} \right)_+^{\beta-1}$, and the proof of the theorem is completed. $\square$

We notice that in the simplest case $\alpha = 2$, the dual problem (4.1) is

$$(4.10) \quad \sup_{p \in \mathbb{R}^{2m+1}} \left\{ \sum_{\delta \in \Delta} \bar{r}(\delta) p(\delta) - \frac{1}{4} \int_K \left(\sum_{\delta \in \Delta} p(\delta) e^{-jk \cdot \delta} \right)_+^2 dk \right\}$$

and when p^* solves (4.10) the optimal L_2 -estimate is given by

$$s_*(k) = \frac{1}{2} \left(\sum_{\delta \in \Delta} p^*(\delta) e^{-jk \cdot \delta} \right)_+.$$

Remark 4.1. The corresponding problem (P) in L_1 , is trivially solved by all feasible s . Indeed

$$r_0 = \int_K s(k) dk = \|s\|_1$$

since $s \geq 0$.

We have shown that we can directly derive the explicit solution of the L_α -estimation problem as presented in [5] by using convex programming duality. Moreover, Theorem 4.1 demonstrates that the multidimensional estimation problem subject to correlation matching constraints can be solved via an *unconstrained finite-dimensional* concave programming problem (D). The desired parameters p^* for the optimal spectrum estimate s^* are precisely the optimal dual variables of problem (D), which can be solved efficiently via nonlinear programming techniques [17].

Consider the case when the spectral support may be approximated by a finite number of points:

$$K = \{k_i \in \mathbb{R}^n : i = 1, \dots, N\}.$$

A measure ν on the discrete support K is completely characterized by its value $\nu(k_i)$ at each point. Thus

$$\int_K \left(\sum_{\delta \in \Delta} p(\delta) e^{-jk \cdot \delta} \right)_+^\beta dk = \sum_{i=1}^N \left(\sum_{\delta \in \Delta} p(\delta) e^{-jk_i \cdot \delta} \right)_+^\beta \nu(k_i)$$

and the dual problem (D) gives rise to the *nonsmooth* optimization problem

$$(\hat{D}) \quad \sup_{p \in \mathbb{R}^{2m+1}} \left\{ \sum_{\delta \in \Delta} r(\delta) p(\delta) + \sum_{i=1}^N \gamma_i \left(\sum_{\delta \in \Delta} p(\delta) e^{-jk_i \cdot \delta} \right) \right\}$$

where

$$\gamma_i(t_i) := -c(\alpha, \beta)(t_i)_+^\beta \nu(k_i).$$

Problems of the above type are typical in the nonsmooth optimization literature and may be solved using existing numerical schemes [9], [12]. It is also worth mentioning that the objective function of problem $(\hat{D})$ has a very special structure which can be exploited using techniques for convex composite optimization problems (see, e.g., [4], [18]).

5. Duality for the time series case. In this section we specialize our duality results derived in § 4 for the important time series case. While for the multidimensional problem we show that (BWCQ) holds if and only if $r \in \text{int } E$, a condition which is not always easy to check; in the time series case we will show that to satisfy (BWCQ) it is necessary and sufficient to test the positive definiteness of a specific Toeplitz matrix (see Theorem-5.1).

Let $K = [-\pi, \pi]$. Given a finite sequence of complex numbers $\{r_n\}_{|n| \leq m}$ with $r_{-n} = \bar{r}_n$ the problem is now to find a nonnegative finite measure μ on $[-\pi, \pi]$ satisfying

$$(5.1) \quad r_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{jkn} d\mu(k), \quad n = 0, \pm 1, \dots, \pm m.$$

This is the so-called *trigonometric moment problem* (see, e.g., Grenander and Szegő [6]). A necessary condition in order to find a nonnegative finite measure μ satisfying (5.1) is that the *Toeplitz matrix*

$$M = [r_{l-h}]_{l,h=0}^m$$

must be positive semidefinite. This condition is also sufficient (see [6, p. 19]).

We now ask the same question for problems involving densities, i.e., finding an integrable function $s \geq 0$ that satisfies

$$(5.2) \quad r_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{jkn} s(k) dk, \quad n = 0, \pm 1, \dots, \pm m.$$

A necessary condition is that M must be positive definite. In fact this condition is also sufficient as has been shown in [6]. (See also the appendix; this also provides a nice illustration of duality theory.)

For the convenience of the reader, we state and prove this result in the following theorem.

THEOREM 5.1. *The following four statements are equivalent:*

- (a) M is positive definite;
- (b) $r \in \text{int } E$;
- (c) (BWCQ) holds;
- (d) Equation (5.2) has a nonzero measurable solution.

Proof. That (b) and (c) are equivalent is direct from Corollary 3.1. It remains to show that (a) is equivalent to (b) and (d). Suppose that (d) holds. Let $d_1, \dots, d_m$ be given complex numbers, not all zero, and consider, for a feasible s , the following Toeplitz form:

$$\begin{aligned} T_m &:= \sum_{l,h=0}^m \bar{d}_l d_h r_{l-h} = \frac{1}{2\pi} \sum_{l,h=0}^m \int_{-\pi}^{\pi} \bar{d}_l d_h e^{jkl} e^{-jkh} s(k) dk \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \lambda(e^{jk}) \bar{\lambda}(e^{+jk}) s(k) dk \end{aligned}$$

where $\lambda[z] := \sum_{l=0}^m d_l z^l$, $z = e^{jk}$.

Any such nonzero polynomial gives rise to a *nonnegative* trigonometric polynomial (see [6, p. 20]):

$$(5.3) \quad p(k) = \lambda(e^{jk}) \bar{\lambda}(e^{jk}) = \lambda(z) \overline{\lambda(z)}$$

which is nonzero (except finitely) for $|z| = 1$. By (d), s is positive on a set of positive measure. Thus

$$T_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} s(k) p(k) dk > 0$$

and M , being a Hermitian matrix, is positive definite. This gives (a).

Now suppose (a) holds. Let $\mathcal{M}$ be the convex cone of positive semidefinite Toeplitz matrices, and let L be the linear operator which sends $M \in \mathcal{M}$ to the transverse diagonal $r := (\text{Re } r, \text{Im } r)$. Then, using Theorem 6.6 in [14]; $M \in \text{relint } \mathcal{M}$ implies $r = L(M) \in \text{relint } L(\mathcal{M})$ and thus if we show that $\text{relint } L(\mathcal{M}) = \text{relint } E$ then we have shown that (a) implies (b). In fact $L(\mathcal{M}) = E$ as follows from the classical moment spaces theory. That is $r \in E$ is extendible if and only if M is in $\mathcal{M}$ (see Karlin and Studden [7, Thm. 4.1, p. 184]). Finally, (b) implies (d), which follows from Theorem 2.2. $\square$

Theorem 5.1 is exactly what we need in order to derive our duality results for the time series L_α -norm estimation problem:

$$(P_1) \quad \inf_{S \in L_\alpha^+[-\pi, \pi]} \left\{ \int_{-\pi}^{\pi} s^\alpha(k) dk : \frac{1}{2\pi} \int_{-\pi}^{\pi} s(k) e^{jkn} dk = r_n, n = 0, \pm 1, \dots, \pm m \right\}.$$

The one-dimensional analogue of Theorem 4.1 is given in the following result.

THEOREM 5.2. *Suppose M is positive definite. Then*

$$(5.4) \quad \min (P_1) = \max (D_1)$$

where the dual problem (D_1) of (P_1) is explicitly given by

$$(5.5) \quad \sup_{\lambda \in \mathbb{R}^{2m+1}} \left\{ \sum_{n=-m}^m \bar{r}_n \lambda_n - \frac{\alpha^{1-\beta}}{\beta} \int_{-\pi}^{\pi} \left(\sum_{n=-m}^m \lambda_n e^{-jkn} \right)_+^{\beta} dk \right\}.$$

Moreover if λ^* solves (D_1) then the unique optimal solution of (P_1) is given by

$$(5.6) \quad s_*(k) = \alpha^{1-\beta} \left(\sum_{n=-m}^m \lambda_n^* e^{-jkn} \right)_+^{\beta-1}.$$

Proof. Since M is assumed positive definite it follows from Theorem 5.1 that (BWCQ) holds. Thus theorem 3.1 is applicable. The remaining statements in the theorem follow immediately as a special case of Theorem 4.1. $\square$

6. Conclusions. A complete duality theory for the multidimensional spectral estimation problem has been developed by using the (BWCQ). The simple unconstrained nature of the dual problem seems appropriate for computational purposes and may lead to the construction of reliable algorithms for computing the L_α optimal spectral estimate. Furthermore, the results in the paper demonstrate that the new constraint qualification of Borwein and Wolkowicz [1] is very natural and is equivalent to the extendibility condition.

We also mention that a duality theory without constraint qualification was also developed in [1]. Following [1], particularly example 5.1 therein, it can be shown that our duality framework can also be extended when the constraint qualification fails. Finally we remark that our duality framework can be used to consider problem (P) with other convex objective functionals. In particular the choice of the “entropy type” functional $I(s) := \int_K s(k) \log s(k) dk$ lead to the explicit solution $s^*(k) = e^{P(k)}$. Note that while the computational tractability of this solution seems less attractive than the L_α norm solution, it provides us automatically with a unique strictly positive optimal solution.

Appendix.

THEOREM. *If M is positive definite, then there exists a unique strictly positive trigonometric polynomial*

$$P(k) = \sum_{n=-m}^m a_n e^{-jkn}, \quad \bar{a}_n = a_{-n} \quad \text{for } |n| \leq m$$

such that $s_0(k) = 1/P(k)$ solved the maximum entropy problem

$$\max \left\{ H(s) = \int_{-\pi}^{\pi} \ln s(k) dk \right\}$$

subject to the constraints (5.2).

Proof. Consider the primal entropy maximum problem

$$p = \sup \left\{ H(s) : \int_{-\pi}^{\pi} e^{jkn} s(k) dk = 2\pi r, n = 0, \pm 1, \dots, \pm m, s \geq 0 \right\}.$$

First we note that, since $H(s) := \int_{-\pi}^{\pi} \ln s(k) dk$ is a strictly concave integral functional, then if an optimal solution exists it is unique. The Lagrangian dual of p (following the same arguments as in Theorem 4.1) is

$$d = \inf_{y \in \mathbb{R}^{2m+1}} \left\{ 2\pi \sum_{n=-m}^m \bar{r}_n y_n + \sup \left\{ H(s) - \int_{-\pi}^{\pi} \left(\sum_{n=-m}^m y_n e^{-jkn} \right) s(k) dk \right\} \right\}.$$

It is easy to verify, using again Theorem 3A of [16], that the inner supremum (where necessarily $s > 0$ a.e.) is

$$\int_{-\pi}^{\pi} \operatorname{Ln} \left(\frac{1}{\sum_{n=-m}^m y_n e^{-jkn}} \right) dk - 2\pi$$

with the optimal $s^*(k) = (1/\sum_{n=-m}^m y_n e^{-jkn})$. Hence the dual is

$$d = \inf_{y \in \mathbb{R}^{2m+1}} \left\{ 2\pi \left(\sum_{n=-m}^m \bar{r}_n y_n - 1 \right) + \int_{-\pi}^{\pi} \operatorname{Ln} \left(\frac{1}{\sum_{n=-m}^m y_n e^{-jkn}} \right) dk \right\}.$$

By Theorems 5.1 and 2.2, a suitable form of Slater's condition is met (in L_1); thus $p = d$ and d is attained. By differentiating in the last equation, we can see that this happens if and only if y^* solves

$$(A1) \quad \bar{r}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-jkn}}{\sum_{n=-m}^m y_n e^{-jkn}} dk = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-jkn} s^*(k) dk, \quad n = 0, \pm 1, \dots, \pm m.$$

However, $\bar{r}_n = r_{-n}$; thus (7.1) is equivalent to

$$(A2) \quad r_n = \frac{1}{2\pi} \int e^{jkn} s^*(k) dk, \quad n = 0, \pm 1, \dots, m.$$

It remains to show that M positive definite guarantees the existence of a *finite strictly positive* $s^*(k)$. To this end observe that when M is positive definite (A2) implies that

$$(A3) \quad \sum_{l,h=0}^m \bar{d}_l d_h r_{l-h} = \int_{-\pi}^{\pi} \frac{p(k)}{q(k)} dk > 0$$

with

$$q(k) := \sum_{n=-m}^m y_n e^{-jkn} = \frac{1}{s^*(k)}$$

for all nonzero positive trigonometric polynomials $p(k)$ defined in (5.3). Suppose $q(k) = 0$ for some k . Then using the representation (5.3) which is valid for any non-negative trigonometric polynomial, we can write

$$q(k) = \prod_{\mu=1}^m |e^{-jk} - \alpha_{\mu}|^2$$

where $\alpha_1, \dots, \alpha_m$ denote the zeros of $q(k)$. Also note that there exists at least one $h \in [1, m]$ such that $|\alpha_h| = 1$. Now since (A3) holds for any nonzero positive trigonometric polynomial $p(k)$, by choosing $p(k) = \prod_{\nu \neq h} |e^{-jk} - \alpha_{\nu}|^2$ we have

$$\infty > \int_{-\pi}^{\pi} \frac{p(k)}{q(k)} dk = \int_{-\pi}^{\pi} \frac{dk}{|e^{-jk} - \alpha_h|^2} = \int_{-\pi}^{\pi} \frac{dk}{|e^{-jk} - 1|^2} = \frac{1}{2} \int_{-\pi}^{\pi} \frac{dk}{1 - \cos k} = \infty.$$

Hence $q(k) > 0$ and so $0 < s^*(k) < \infty$. $\square$

We note that it is formally "obvious" that (A1) holds by differentiation. It is, however, technically nontrivial to justify this step. Indeed a fully satisfactory proof is somewhat extended and requires reorganizing the proof we have given.

It is also important to mention that we can solve (7.2) explicitly to find the parameters $\{y_n\}$ (see [3] for details). The maximum entropy spectrum is then given by

$$s^*(k) = \frac{M_{11}^{-1}}{|(M^{-1}\delta)^T \phi(k)|^2}$$

where $\phi(k) = (1, \dots, e^{jmk})$, $\delta = (1, 0, \dots, 0)$ and M^{-1} is the inverse of M (which exists since M is positive definite).

REFERENCES

- [1] J. BORWEIN AND H. WOLKOWICZ, *A simple constraint qualification in infinite dimensional programming*, Math. Programming, 35 (1986), pp. 83–96.
- [2] J. P. BURG, *Maximum entropy spectral analysis*, Ph.D. thesis, Stanford University, Stanford, CA, 1975.
- [3] J. A. EDWARD AND M. M. FITELSON, *Notes on maximum entropy processing*, IEEE Trans. Inform. Theory, IT-19 (1973), pp. 232–234.
- [4] R. FLETCHER, *A model algorithm for composite nondifferential optimization problems*, Math. Programming Stud., 17 (1982), pp. 67–76.
- [5] B. K. GOODRICH AND A. STEINHARDT, *L_2 spectral estimation*, SIAM J. Appl. Math., 46 (1986), pp. 417–426.
- [6] U. GRENANDER AND G. SZEGÖ, *Toeplitz Forms and Their Applications*, Univ. of California Press, Berkeley, Los Angeles, 1958.
- [7] S. KARLIN AND W. STUDDEN, *Tchebycheff Systems; with Applications in Analysis and Statistics*, Wiley-Interscience, New York, 1986.
- [8] S. M. KAY AND S. L. MARPLE, *Spectrum analysis—a modern perspective*, Proc. IEEE, 69 (1981), pp. 1380–1419.
- [9] K. C. KIWIEL, *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Math. 1133, Springer-Verlag, Berlin, New York, 1985.
- [10] S. W. LANG AND J. H. MCCLELLAN, *Multidimensional MEM spectral estimation*, IEEE Trans. Acoust. Speech Signal Process., ASSP-30 (1982), pp. 880–887.
- [11] ———, *Spectral estimation for sensor arrays*, IEEE Trans. Acoust. Speech Signal Process., ASSP-31 (1983), pp. 349–358.
- [12] C. LEMARECHAL, *Nondifferentiable Optimization*, in Nonlinear Optimization Theory and Algorithms, L. C. W. Dixon, J. Spedicato, and G. F. Szegö, eds., Birkhäuser, Boston, 1980.
- [13] D. G. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley, New York, 1969.
- [14] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [15] ———, *Integrals which are convex functions*, II, Pacific J. Math., 39 (1971), pp. 439–469.
- [16] ———, *Integral Functionals, Normal Integrands and Measurable Selections*, Lecture Notes in Math., 543, Springer-Verlag, Berlin, New York, 1976, pp. 157–207.
- [17] K. SCHITTOWSKI, *Nonlinear Programming Codes*, Lecture Notes Econom. Math. Systems, 183, Springer-Verlag, Berlin, 1980.
- [18] R. S. WOMERSLEY, *Local properties of algorithms for minimizing nonsmooth composite functions*, Math. Programming, 32 (1985), pp. 69–89.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.