The Annals of Statistics
1987, Vol. 15, No. 2, 712–723

AN OPTIMIZATION PROBLEM WITH APPLICATIONS TO OPTIMAL DESIGN THEORY[1]

BY CHING-SHUI CHENG

University of California, Berkeley

Dedicated to the memory of Jack Kiefer

The problem of minimizing $\sum_{i=1}^n f(x_i)$ subject to the constraints $\sum_{i=1}^n x_i = A$, $\sum_{i=1}^n g(x_i) = B$ and $x_i \geq 0$ is solved. The solutions are different depending upon whether $(\operatorname{sgn} g'')f''/g''$ is an increasing or decreasing function. The result is used to show that for certain designs, if they are optimal with respect to two criteria, then they are also optimal with respect to many other criteria.

1. Introduction. The main purpose of this paper is to solve and discuss applications of the following minimization problem to the theory of optimal design:

$$\text{Minimize } \sum_{i=1}^{n} f(x_i) \text{ subject to the constraints} \tag{1.1}$$
$$\sum_{i=1}^{n} x_i = A, \quad \sum_{i=1}^{n} g(x_i) = B \quad \text{and} \quad x_i \geq 0,\ i = 1, 2, \dots, n,$$

where A, B are fixed constants, $A > 0$ and f, g are real-valued functions. In fact, one can also consider the more general problem of minimizing $\sum_{i=1}^n \tilde{f}(x_i)$ subject to $\sum_{i=1}^n \tilde{h}(x_i) = A$, $\sum_{i=1}^n \tilde{g}(x_i) = B$ and $x_i \geq 0$, $i = 1, 2, \dots, n$. If $\tilde{h}$ is a one-to-one function, then a trivial transformation reduces the latter to (1.1).

Problem (1.1) arises from, but is not restricted to, the theory of optimal experimental design. Special cases of (1.1) with $g(x) = x^2$ and x^{-1} have been studied in Cheng (1978) and Cheng, Masaro and Wong (1985), respectively, as tools for solving some problems in block and weighing designs. An application to graph theory can be found in Cheng (1981). In design applications, $x_1, x_2, \dots, x_n$ are usually the eigenvalues of a symmetric nonnegative definite matrix (the so-called information matrix), which explains the constraint $x_i \geq 0$ for all i.

The use of problem (1.1) in optimal design will be discussed after the following review of some preliminaries. Let $\mathscr{D}$ be the class of all competing designs in a certain setting; to each design d in $\mathscr{D}$, there corresponds an $n \times n$ information matrix $\mathbf{C}_d$. We are interested in finding a design d^* which minimizes $\Phi(\mathbf{C}_d)$ over $\mathscr{D}$ for some optimality functional Φ [see Kiefer (1974)]. Typical examples

Received November 1985; revised May 1986.
[1]Research supported by National Science Foundation grants MCS-82-00909 and DMS-85-02784.
AMS 1980 *subject classification*. 62K05.
Key words and phrases. A-optimality, D-optimality, E-optimality, optimal design, Φ_p-optimality.

of Φ are

(i) $$\Phi_0(\mathbf{C}_d) = \left(\prod_{i=1}^{n} \lambda_{di}^{-1}\right)^{1/n},$$

(ii) $$\Phi_p(\mathbf{C}_d) = \left(\frac{1}{n}\sum_{i=1}^{n} \lambda_{di}^{-p}\right)^{1/p}, \qquad p > 0,$$

and

(iii) $$\Phi_\infty(\mathbf{C}_d) = \max_{1 \le i \le n} \lambda_{di}^{-1},$$

where $\lambda_{d1}, \lambda_{d2}, \dots, \lambda_{dn}$ are the eigenvalues of $\mathbf{C}_d$. A design minimizing (i), (ii) or (iii) is called, respectively, D-, Φ_p- or E-optimal, and a Φ_1-optimal design is also known as A-optimal. Since $\lim_{p \to \infty}\Phi_p(\mathbf{C}_d) = \Phi_\infty(\mathbf{C}_d)$ and $\lim_{p \to 0}\Phi_p(\mathbf{C}_d) = \Phi_0(\mathbf{C}_d)$, E- and D-optimal designs are often referred to as Φ_∞- and Φ_0-optimal, respectively. Notice that one can also define the D-criterion in terms of the functional $\Phi(\mathbf{C}_d) = -\sum_{i=1}^{n}\log \lambda_{di}$. In this paper, we shall restrict our discussion to criteria of the form $\sum_{i=1}^{n} f(\lambda_{di})$, where f is a real-valued function. This covers D- and Φ_p-criteria by the choices of $f(x) = -\log x$ and x^{-p}, respectively, and the E-criterion is also included as a limiting case.

Usually the direct search for an optimal design is difficult, but some simple sufficient conditions for optimality are available in the literature. For certain well-structured designs, it is often possible to show that if they are optimal with respect to *one or two simple* criteria, then they are also optimal with respect to *many* other criteria. One notable example is Kiefer's (1975) result on universal optimality:

PROPOSITION 1.1. *If a design with $\lambda_{d1} = \lambda_{d2} = \cdots = \lambda_{dn}$ maximizes* $\operatorname{tr}\mathbf{C}_d$, *then it minimizes $\Phi(\mathbf{C}_d)$ for all nonincreasing, convex and orthogonally invariant Φ.*

The same paper also contains the following result:

PROPOSITION 1.2. *If a design with $\lambda_{d1} = \lambda_{d2} = \cdots = \lambda_{dn}$ is Φ_p-optimal for some $p \ge 0$, then it is Φ_q-optimal for all $q \ge p$.*

Proposition 1.2 in fact can be viewed as a corollary of Proposition 1.1, although it was not proved this way in Kiefer's paper. To gain some insight, let us first consider the following problem:

(1.2) $$\text{Minimize } \sum_{i=1}^{n} f(x_i) \text{ subject to } \sum_{i=1}^{n} x_i = A.$$

If f is convex, then the minimum value is $nf(A/n)$, attained at $x_1 = x_2 = \cdots = x_n = A/n$. Furthermore, if f is decreasing, then $nf(A/n)$ is also a decreasing function of A. Therefore, if a design with $\lambda_{d1} = \cdots = \lambda_{dn}$ maximizes $\operatorname{tr}\mathbf{C}_d$, then it minimizes $\sum_{i=1}^{n} f(\lambda_{di})$ for all decreasing and convex f. This proves a slightly weaker version of Proposition 1.1. Proposition 1.2 can be proven in a similar fashion. In fact, instead of minimizing $\sum_{i=1}^{n} x_i^{-q}$ subject to $\sum_{i=1}^{n} x_i^{-p} = A$, upon the transformation $y = x^{-p}$ (or $-\log x$ for $p = 0$) it is sufficient to consider the minimization of $\sum_{i=1}^{n} y_i^{q/p}$ (or $\sum_{i=1}^{n} e^{qy_i}$ for $p = 0$) subject to $\sum_{i=1}^{n} y_i = A$, which then reduces to the earlier case. This kind of argument leads us to the following more general result:

PROPOSITION 1.3. *If a design with* $\lambda_{d1} = \cdots = \lambda_{dn}$ *minimizes* $\sum_{i=1}^{n} g(\lambda_{di})$ *over* $\mathscr{D}$ *for a certain real-valued function* g, *then it minimizes* $\sum_{i=1}^{n} f(g(\lambda_{di}))$ *for all convex increasing functions* f.

Proposition 1.2 becomes a special case of Proposition 1.3 with $g(x) = x^{-p}$, $f(x) = x^{q/p}$ for $q > p > 0$ and $g(x) = -\log x$, $f(x) = e^{qx}$ for $q > p = 0$. This proof of Proposition 1.2 is different from Kiefer's original proof; it links Proposition 1.2 to Proposition 1.1 and makes clear the key role played by the convexity and increasing monotonicity of $x^{q/p}$ (when $q > p > 0$) and e^{qx} (when $q > p = 0$).

When a design with $\lambda_{d1} = \cdots = \lambda_{dn}$ does not exist, one would have to add another constraint to (1.2) and solve (1.1). Hopefully some designs optimal with respect to *two* criteria can be shown to be optimal with respect to many other criteria. Thus Cheng (1978) solved the simple case $g(x) = x^2$ and showed that a design with two distinct λ_{di}'s, the larger one having multiplicity 1, minimizes $\sum_{i=1}^{n} f(\lambda_{di})$ for a large class of functions f if it maximizes $\operatorname{tr} \mathbf{C}_d$ and also minimizes another function of $\mathbf{C}_d$. The case $g(x) = x^{-1}$ was solved by Cheng, Masaro and Wong (1985) as a tool for establishing that if a design with two distinct λ_{di}'s, the smaller one being simple, is A-optimal and maximizes $\operatorname{tr} \mathbf{C}_d$, then it is Φ_q-optimal for all $0 \le q \le 1$. These results provided useful tools for solving some problems in optimal block and weighing designs. In this paper, a generalization and unification of these optimality tools will be obtained by working out the general solution of (1.1). It turns out that whether $(f''/g'')\operatorname{sgn} g''$ is an increasing or decreasing function plays an important role in determining the solution, where $\operatorname{sgn} g''$ equals 1 when $g'' > 0$ and equals -1 when $g'' < 0$. It is interesting to point out that the same kind of conditions on $(f''/g'')\operatorname{sgn} g''$ also appeared in Galil and Kiefer's (1983, Theorem 1) work in comparing the performance of two designs tied at two criteria. When $g(x) = x^2$, $(\operatorname{sgn} g'')f''/g'' = \frac{1}{2}f''$. This leads to the distinction between type 1 and type 2 criteria according to $f''' < 0$ or $f''' > 0$, as treated in Cheng (1978).

The results in this paper have been presented at the Kiefer–Wolfowitz Memorial Statistical Research Conference held in July 1983 at Cornell University. Recently Kunert and Martin (1985) independently proved the same results, but for functions of the form $f(x) = x^p$ and $g(x) = x^q$ only. They also applied the results to study optimal treatment designs for correlated errors.

2. Solution to Problem (1.1). Throughout this section, we shall assume

(i) f and g are twice continuously differentiable on $(0, A)$ and $f(0) = \lim_{x \to 0^+} f(x) = \infty$;

(ii) g'' is of one sign on $(0, A)$;

(iii) $ng(A/n) < B < g(A) + (n-1)g(0)$ if $g'' > 0$ on $(0, A)$, and $ng(A/n) > B > g(A) + (n-1)g(0)$ if $g'' < 0$ on $(0, A)$, where $g(0) = \lim_{x \to 0^+} g(x)$ and can be ∞ or $-\infty$.

Assumption (ii) means that g is either strictly convex or strictly concave on $(0, A)$. Since $f(0) = \infty$, the minimum of $\sum_{i=1}^n f(x_i)$ cannot be attained at a point with some zero coordinates. Assumption (iii) assures that the set $S_{AB} \equiv \{\mathbf{x} = (x_1, \dots, x_n): \sum_{i=1}^n x_i = A, \sum_{i=1}^n g(x_i) = B, x_i > 0\}$ is nonempty (see Lemma 2.2) and that all the points in S_{AB} have at least two distinct coordinates, ruling out the trivial case $S_{AB} = \{(A/n, \dots, A/n)\}$.

Then we have

THEOREM 2.1. *Assume* (i), (ii) *and* (iii). *If* $(\operatorname{sgn} g'')f''/g''$ *is strictly decreasing on* $(0, A)$, *then the point* $\mathbf{x} = (x_1, \dots, x_n) \in S_{AB}$ *with* $x_1 > x_2 = \cdots = x_n$ *is a solution to Problem* (1.1). *On the other hand, if* $(\operatorname{sgn} g'')f''/g''$ *is strictly increasing on* $(0, A)$, *then the minimum of* $\sum_{i=1}^n f(x_i)$ *over* S_{AB} *is attained at* $\mathbf{x} = (x_1, \dots, x_n)$ *with* $x_1 = \cdots = x_{k^*} > x_{k^*+1} = \cdots = x_n$, *where* k^* *is the largest positive integer* $k < n$ *such that* $kx + (n-k)y = A$, $kg(x) + (n-k)g(y) = B$, $x > y > 0$, *has a solution.*

Before proving Theorem 2.1, we first consider the existence and uniqueness of the kind of $\mathbf{x}$ as described in the theorem.

LEMMA 2.2. *Assume* (i), (ii) *and* (iii). *Then*

(a) *there exists* $\mathbf{x} \in S_{AB}$ *such that* $x_1 > x_2 = \cdots = x_n$; *in particular,* S_{AB} *is nonempty*;

(b) *for any real number* k (*not necessarily an integer*) *such that* $1 \le k \le n-1$, *if there is a solution* (x, y) *to* $kx + (n-k)y = A$, $kg(x) + (n-k)g(y) = B$ *and* $x > y > 0$, *then it is unique*;

(c) *if the system of equations* $kx + (n-k)y = A$, $kg(x) + (n-k)g(y) = B$ *has a solution* $x > y > 0$, *where* $1 \le k \le n-1$, *then for any* k' *such that* $1 \le k' \le k$, $k'x' + (n-k')y' = A$ *and* $k'g(x') + (n-k')g(y') = B$ *also has a solution* $x' > y' > 0$.

PROOF. We only have to prove the case $g'' > 0$ on $(0, A)$. That of $g'' < 0$ can be treated similarly.

(a) The range of $g(x) + (n-1)g((A-x)/(n-1))$ over $x \in (A/n, A)$ is $(ng(A/n), g(A) + (n-1)g(0))$. Since $ng(A/n) < B < g(A) + (n-1)g(0)$, there exists $x^* \in (A/n, A)$ such that $g(x^*) + (n-1)g((A-x^*)/(n-1)) = B$. Obviously, $x^* > (A-x^*)/(n-1) > 0$; then (a) is proven by letting $x_1 = x^*$ and $x_2 = \cdots = x_n = (A-x^*)/(n-1)$.

(b) In order that $kx + (n-k)y = A$ and $x > y > 0$, one must have $y = (A-kx)/(n-k)$ and $A/n < x < A/k$. Then we have $kg(x) + (n-k)g(y) = kg(x) + (n-k)g((A-kx)/(n-k))$, which, by differentiation, can be seen to be a strictly increasing function of x on $(A/n, A/k)$.

(c) By the first sentence in the proof of (b), the range of $kg(x) + (n-k)g(y)$ over $\{(x, y)\colon\ kx + (n-k)y = A,\ x > y > 0\}$ is $(ng(A/n), kg(A/k) + (n-k)g(0))$. We may assume $g(0) < \infty$; otherwise the result holds trivially. Then $kg(A/k) + (n-k)g(0)$ can be seen to be a strictly decreasing function of k, by differentiation and the convexity of g. It follows that if $k' \leq k$, then $(ng(A/n), kg(A/k) + (n-k)g(0))$ is contained in $(ng(A/n), k'g(A/k') + (n-k')g(0))$. □

Now we are ready to prove Theorem 2.1.

PROOF OF THEOREM 2.1. Again, we shall only consider the case $g'' > 0$ on $(0, A)$. For $g'' < 0$, one only has to reverse inequalities at several places.

Let the minimum of $\sum_{i=1}^{n} f(x_i)$ over S_{AB} be attained at a point $\mathbf{a} = (a_1, a_2, \ldots, a_n)$. By Lagrange's theorem, there exist numbers α and β such that $a_1, a_2, \ldots, a_n$ satisfy

$$\frac{\partial}{\partial x_i}\left\{\sum_{i=1}^{n} f(x_i) + \alpha\left(\sum_{i=1}^{n} x_i - A\right) + \beta\left(\sum_{i=1}^{n} g(x_i) - B\right)\right\} = 0, \qquad i = 1, 2, \ldots, n,$$

i.e.,

$$f'(x_i) + \alpha + \beta g'(x_i) = 0, \qquad i = 1, 2, \ldots, n.$$

We claim that the equation $f'(x) + \alpha + \beta g'(x) = 0$ has at most two distinct solutions, which implies that $\mathbf{a}$ must have only two distinct coordinates. Suppose x and y are two distinct solutions of the above equation. Then

$$f'(x) + \beta g'(x) = f'(y) + \beta g'(y),$$

and hence

$$[f'(x) - f'(y)]/[g'(x) - g'(y)] = -\beta.$$

Let $\tilde{x} = g'(x)$ and $\tilde{y} = g'(y)$. Then

$$\left\{f'\left[(g')^{-1}(\tilde{x})\right] - f'\left[(g')^{-1}(\tilde{y})\right]\right\}/(\tilde{x} - \tilde{y}) = -\beta. \tag{2.1}$$

The derivative of $f'[(g')^{-1}(\tilde{x})]$ is $f''[(g')^{-1}(\tilde{x})]/g''[(g')^{-1}(\tilde{x})]$, which, by assumption, is either strictly increasing or strictly decreasing. Therefore, for fixed $\tilde{x}$, there is at most one $\tilde{y}$ satisfying (2.1). This proves the claim that $\mathbf{a}$ has only two distinct coordinates.

Now consider the equations

$$\begin{aligned} ka + (n-k)b &= A, \\ kg(a) + (n-k)g(b) &= B, \end{aligned} \qquad a > b > 0, \tag{2.2}$$

where $1 \leq k \leq n-1$, but k is not necessarily an integer. By Lemma 2.2, the solution is unique if it exists. So a and b are uniquely determined by A, B and k, and can be expressed as $a(A, B, k)$ and $b(A, B, k)$. The theorem is proved if we can show that for fixed A, B,

$$F(A, B, k) \equiv kf(a(A, B, k)) + (n-k)f(b(A, B, k))$$

is a strictly increasing (respectively, decreasing) function of k when f''/g'' is strictly decreasing (respectively, increasing) on $(0, A)$.

Now assume f''/g'' is strictly decreasing. We shall show that $(\partial F/\partial k)(A, B, k) > 0$. Differentiating $ka + (n-k)b = A$ and $kg(a) + (n-k)g(b) = B$ with respect to k, we get

$$a - b + k\frac{\partial a}{\partial k} + (n-k)\frac{\partial b}{\partial k} = 0,$$

$$g(a) - g(b) + kg'(a)\frac{\partial a}{\partial k} + (n-k)g'(b)\frac{\partial b}{\partial k} = 0.$$

Therefore,

$$\frac{\partial a}{\partial k} = k^{-1}\{g'(a) - g'(b)\}^{-1}\{g(b) - g(a) - g'(b)(b-a)\},$$

$$\frac{\partial b}{\partial k} = (n-k)^{-1}\{g'(a) - g'(b)\}^{-1}\{g(a) - g(b) + g'(a)(b-a)\}.$$

Then

$$\begin{aligned}\frac{\partial F}{\partial k}(A, B, k) &= f(a) - f(b) + kf'(a)\frac{\partial a}{\partial k} + (n-k)f'(b)\frac{\partial b}{\partial k}\\ &= \{g'(a) - g'(b)\}^{-1}\{[f(a) - f(b)][g'(a) - g'(b)]\\ &\qquad + f'(a)[g(b) - g(a) - g'(b)(b-a)]\\ &\qquad + f'(b)[g(a) - g(b) + g'(a)(b-a)]\},\end{aligned}$$

which, since $g'' > 0$, has the same sign as

$$\begin{aligned}&\{g'(a) - g'(b)\}\frac{\partial F}{\partial k}(A, B, k)\\ &\quad = \{f(a) - f(b)\}\{g'(a) - g'(b)\}\\ &\qquad + f'(a)\{g(b) - g(a) - g'(b)(b-a)\}\\ &\qquad + f'(b)\{g(a) - g(b) + g'(a)(b-a)\}.\end{aligned} \tag{2.3}$$

For fixed b, denote the right-hand side of (2.3) by $G(a)$. Since $G(b) = 0$, to show that $G(a) > 0$ for $a > b$, it is enough to prove $G'(a) > 0$ for $a > b$, i.e.,

$$\begin{aligned}&g''(a)\{f(a) - f(b) - f'(b)(a-b)\}\\ &\quad - f''(a)\{g(a) - g(b) - g'(b)(a-b)\} > 0,\end{aligned}$$

for $a > b$. This is equivalent to

$$\frac{f(a) - f(b) - f'(b)(a-b)}{g(a) - g(b) - g'(b)(a-b)} > \frac{f''(a)}{g''(a)}, \quad \text{for } a > b, \tag{2.4}$$

again due to the convexity of g. Now since f''/g'' is decreasing on $(0, A)$,

$$f''(t)/g''(t) > f''(a)/g''(a), \quad \text{for } b < t < a.$$

We have

$$f''(t) > g''(t)\frac{f''(a)}{g''(a)}, \quad \text{for } b < t < a.$$

Then

$$\int_b^a \int_b^x f''(t)\,dt\,dx > \int_b^a \int_b^x g''(t)\frac{f''(a)}{g''(a)}\,dt\,dx.$$

As a consequence,

$$\frac{\int_b^a \int_b^x f''(t)\,dt\,dx}{\int_b^a \int_b^x g''(t)\,dt\,dx} > \frac{f''(a)}{g''(a)},$$

which yields (2.4) and the proof is completed. □

REMARK 2.3. The condition $f(0) = \lim_{x\to 0^+} f(x) = \infty$ in (i) is to assure that the minimum of $\sum_{i=1}^n f(x_i)$ over $\{(x_1, \ldots, x_n): \sum_{i=1}^n x_i = A, \sum_{i=1}^n g(x_i) = B, x_i \geq 0\}$ does not occur at a point with some coordinates equal to zero. The same goal can also be achieved, for instance, by requiring $|g(0)| = \lim_{x\to 0^+}|g(x)| = \infty$. Thus in Theorem 2.1, we can replace "$f(0) = \lim_{x\to 0^+} f(x) = \infty$" with "$|g(0)| = \lim_{x\to 0^+}|g(x)| = \infty$." Another way to remove the above condition on f is to assume $B < (n-1)g(A/(n-1))$ when $g'' > 0$ and $B > (n-1)g(A/(n-1))$ when $g'' < 0$.

3. Applications. In this section, we shall use Theorem 2.1 to show that for certain designs, as long as they are optimal with respect to *two* criteria, they are also optimal with respect to many others. First, we shall see how the minimum value of $\sum_{i=1}^n f(x_i)$ over S_{AB} depends on A and B.

Let $A^* = \max_{d\in\mathscr{D}} \operatorname{tr} \mathbf{C}_d$. Throughout this section, we shall also assume that g' is of one sign on $(0, A^*)$.

By differentiating $ka + (n-k)b = A$ and $kg(a) + (n-k)g(b) = B$ with respect to A and B, and then solving for $\partial a/\partial A$, $\partial b/\partial A$, $\partial a/\partial B$ and $\partial b/\partial B$, we obtain

$$\begin{aligned}
\frac{\partial a}{\partial A} &= -g'(b)k^{-1}\{g'(a) - g'(b)\}^{-1},\\
\frac{\partial b}{\partial A} &= g'(a)(n-k)^{-1}\{g'(a) - g'(b)\}^{-1},\\
\frac{\partial a}{\partial B} &= k^{-1}\{g'(a) - g'(b)\}^{-1},\\
\frac{\partial b}{\partial B} &= -(n-k)^{-1}\{g'(a) - g'(b)\}^{-1}.
\end{aligned} \tag{3.1}$$

Then

$$\begin{aligned}
\frac{\partial F}{\partial B}(A, B, k) &= kf'(a)\frac{\partial a}{\partial B} + (n-k)f'(b)\frac{\partial b}{\partial B}\\
&= \{f'(a) - f'(b)\}/\{g'(a) - g'(b)\},
\end{aligned}$$

which is positive (negative) if $f''g'' > 0$ (< 0) on $(0, A)$. Thus, for fixed k and A, $F(A, B, k)$ is strictly increasing in B if $f''g'' > 0$ on $(0, A)$, and is strictly decreasing in B if $f''g'' < 0$ on $(0, A)$.

Now we vary A. From (3.1),

$$\begin{aligned}(3.2)\quad \frac{\partial F}{\partial A}(A, B, k) &= kf'(a)\frac{\partial a}{\partial A} + (n-k)f'(b)\frac{\partial b}{\partial A} \\ &= \{g'(a)f'(b) - g'(b)f'(a)\}/\{g'(a) - g'(b)\}.\end{aligned}$$

Since g' is of one sign on $(0, A^*)$, (3.2) can be written as

$$\frac{\partial F}{\partial A}(A, B, k) = g'(a)g'(b)\{f'(b)/g'(b) - f'(a)/g'(a)\}\Big/\{g'(a) - g'(b)\},$$

which is positive (negative) if $(\operatorname{sgn} g'')f'/g'$ is strictly decreasing (increasing) on $(0, A^*)$. Thus, for fixed k and B, $F(A, B, k)$ is strictly increasing in A if $(\operatorname{sgn} g'')f'/g'$ is a strictly decreasing function, and is strictly decreasing in A if $(\operatorname{sgn} g'')f'/g'$ is a strictly increasing function. Combining the above conclusions, Theorem 2.1 and Remark 2.3, we have

THEOREM 3.1. *Let* $A^* = \max_{d \in \mathscr{D}} \operatorname{tr} \mathbf{C}_d$ *and assume*

(i) *g is twice continuously differentiable on* $(0, A^*)$,
(ii) *g' is of one sign on* $(0, A^*)$, *and*
(iii) *g'' is of one sign on* $(0, A^*)$.

If there is a design $d^* \in \mathscr{D}$ *such that*

(a) $\mathbf{C}_{d^*}$ *has two distinct nonzero eigenvalues* $\lambda_{d^*1} > \lambda_{d^*2} = \cdots = \lambda_{d^*n}$,
(b) d^* *maximizes* $\operatorname{tr} \mathbf{C}_d$ *over* $\mathscr{D}$, *and*
(c) d^* *minimizes* $\sum_{i=1}^n g(\lambda_{di})$ *over* $\mathscr{D}$,

then d^* *minimizes* $\sum_{i=1}^n f(\lambda_{di})$ *over* $\mathscr{D}$ *for all f such that*

(α) *f is twice continuously differentiable on* $(0, A^*)$ *and* $\lim_{x \to 0^+} f(x) = f(0) = \infty$,
(β) $(\operatorname{sgn} g'')f''/g''$ *is strictly decreasing on* $(0, A^*)$,
(γ) $f''g'' > 0$ *on* $(0, A^*)$, *and*
(δ) $(\operatorname{sgn} g'')f'/g'$ *is strictly increasing on* $(0, A^*)$.

In (b), "*maximizes*" *is replaced by* "*minimizes*" *if in* (δ), "*increasing*" *is changed to* "*decreasing.*" *Also, the condition* "$f(0) = \lim_{x \to 0^+} f(x) = \infty$" *in* ($\alpha$) *can be dropped if we assume* $|g(0)| = \lim_{x \to 0^+} |g(x)| = \infty$ *instead. Furthermore, all the above results hold if* (a) *and* (β) *are replaced, respectively, by*

(a′) $\mathbf{C}_{d^*}$ *has two distinct nonzero eigenvalues* $\lambda_{d^*1} = \lambda_{d^*2} = \cdots = \lambda_{d^*, n-1} > \lambda_{d^*n}$ *and*
(β') $(\operatorname{sgn} g'')f''/g''$ *is strictly increasing on* $(0, A^*)$.

The same kind of argument that leads Proposition 1.1 to Proposition 1.3 now yields

THEOREM 3.2. *Let* $A^* = \max_{d \in \mathscr{D}} \operatorname{tr} \mathbf{C}_d$, *h be a nonnegative-valued function defined on* $(0, A^*)$ *and* $\tilde{A} \equiv \sup_{x \in (0, A^*)} h(x)$. *Assume*

(i) *g is twice continuously differentiable on* $(0, \tilde{A})$,
(ii) g' *is of one sign on* $(0, \tilde{A})$, *and*
(iii) g'' *is of one sign on* $(0, \tilde{A})$.

If there is a design $d^* \in \mathscr{D}$ *such that*

(a) $\mathbf{C}_{d^*}$ *has two distinct nonzero eigenvalues* λ *and* μ, *where* λ *is simple and* $h(\lambda) > h(\mu)$,
(b) d^* *minimizes* $\sum_{i=1}^n h(\lambda_{di})$ *over* $\mathscr{D}$, *and*
(c) d^* *minimizes* $\sum_{i=1}^n g(h(\lambda_{di}))$ *over* $\mathscr{D}$,

then d^* *minimizes* $\sum_{i=1}^n f(h(\lambda_{di}))$ *over* $\mathscr{D}$ *for all f such that*

(α) *f is twice continuously differentiable on* $(0, \tilde{A})$ *and* $\lim_{x \to 0^+} f(x) = f(0) = \infty$,
(β) $(\operatorname{sgn} g'')f''/g''$ *is strictly decreasing on* $(0, \tilde{A})$,
(γ) $f''g'' > 0$ *on* $(0, \tilde{A})$, *and*
(δ) $(\operatorname{sgn} g'')f'/g'$ *is strictly decreasing on* $(0, \tilde{A})$.

If we also assume that $|g(0)| = \lim_{x \to 0^+} |g(x)| = \infty$, *then the condition* "$f(0) = \lim_{x \to 0^+} f(x) = \infty$" *in* ($\alpha$) *can be dropped. Furthermore, all the above results hold if* (a) *and* (β) *are replaced, respectively, by*

(a′) $\mathbf{C}_{d^*}$ *has two distinct nonzero eigenvalues* λ *and* μ *such that* λ *is simple and* $h(\lambda) < h(\mu)$, *and*
(β') $(\operatorname{sgn} g'')f''/g''$ *is strictly increasing on* $(0, \tilde{A})$.

REMARK. It is assumed in Theorems 3.1 and 3.2 that at least one of $f(0)$ and $|g(0)|$ is equal to ∞. The case where both $f(0)$ and $|g(0)|$ are finite is much more complicated, but can be treated along the line of Cheng (1978) in which the case $g(x) = x^2$ was dealt with. For simplicity, we shall not pursue it further in this paper.

Theorems 3.1 and 3.2 are our main results. One can write down numerous applications of these two theorems by specializing f, g and h to various functions. For brevity, we shall only list a few applications to the Φ_p-criteria in the rest of the paper.

Now choose the $g(x)$ in Theorem 3.1 to be x^{-p} with $p > 0$; also for $0 \le q < p$, let $f_q(x) = x^{-q}$ when $q \neq 0$ and $f_q(x) = -\log x$ when $q = 0$. Then (i), (ii), (iii), (α) and (γ) are easily seen to be satisfied. We also have

$$(\operatorname{sgn} g''(x)) f_q''(x)/g''(x) = \begin{cases} p^{-1}(p+1)^{-1} q(q+1) x^{p-q}, & 0 < q < p, \\ p^{-1}(p+1)^{-1} x^p, & q = 0, \end{cases}$$

and

$$(\operatorname{sgn} g''(x))f_q'(x)/g'(x) = \begin{cases} p^{-1}qx^{p-q}, & 0 < q < p, \\ p^{-1}x^p, & q = 0. \end{cases}$$

Both of the above two functions are strictly increasing on $(0, \infty)$. Thus we have

COROLLARY 3.3. *If there is a design $d^* \in \mathscr{D}$ such that*

(a) C_{d^*} *has two distinct nonzero eigenvalues* $\lambda_{d^*1} = \lambda_{d^*2} = \cdots = \lambda_{d^*, n-1} > \lambda_{d^*n}$,

(b) d^* *maximizes* $\operatorname{tr} \mathbf{C}_d$ *over* $\mathscr{D}$, *and*

(c) d^* *is* Φ_p*-optimal for some* $p > 0$,

then d^* is Φ_q-optimal for all $0 \le q \le p$.

The extension of Corollary 3.3 to the case where p or q is negative is straightforward. The result in Corollary 3.3 can also be proved for $p = \infty$; this has been treated in Kunert (1985) and also in Jacroux (1985).

As an application of Theorem 3.2, let $h(x) = x^{-p}$, $g(x) = p^{-1}\log x$ and $f(x) = x^{q/p}$,where $0 < q < p$. Then conditions (i), (ii), (iii) and (γ) in Theorem 3.2 are easily seen to be satisfied. Condition (α) is also satisfied with "$f(0) = \lim_{x \to 0^+} f(x) = \infty$" replaced by "$g(0) = \lim_{x \to 0^+} g(x) = -\infty$." It is straightforward to see that

$$(\operatorname{sgn} g''(x))f''(x)/g''(x) = p^{-1}q(q-p)x^{q/p}$$

and

$$(\operatorname{sgn} g''(x))f'(x)/g'(x) = -qx^{q/p}.$$

Both are strictly decreasing functions on $(0, \infty)$. Since $g(h(x)) = -\log x$ and $f(h(x)) = x^{-q}$, and $h(x)$ is a decreasing function, we conclude

COROLLARY 3.4. *If there is a design $d^* \in \mathscr{D}$ such that*

(a) $\mathbf{C}_{d^*}$ *has two distinct nonzero eigenvalues* $\lambda_{d^*1} = \lambda_{d^*2} = \cdots = \lambda_{d^*, n-1} > \lambda_{d^*n}$,

(b) d^* *is D-optimal, and*

(c) d^* *is* Φ_p*-optimal for some* $p > 0$,

then d^* is Φ_q-optimal for all $0 \le q \le p$.

Is there an analogue of Corollary 3.3 for the case where $\mathbf{C}_{d^*}$ has two distinct nonzero eigenvalues with the larger one being simple? A simple application of Theorem 3.1 shows that if there is a design $d^* \in \mathscr{D}$ such that $\mathbf{C}_{d^*}$ has nonzero eigenvalues $\lambda_{d^*1} > \lambda_{d^*2} = \cdots = \lambda_{d^*n}$ and d^* is Φ_p-optimal for some $p \ge 0$, then d^* is Φ_q-optimal for all $q \ge p$ over the designs with $\operatorname{tr} \mathbf{C}_d \ge \operatorname{tr} \mathbf{C}_{d^*}$. This result says nothing about the designs with $\operatorname{tr} \mathbf{C}_d < \operatorname{tr} \mathbf{C}_{d^*}$, which, unfortunately, arise more often in applications. So such a straightforward application of Theorem 3.1 does not yield a useful result for designs with the above eigenvalue structure. We shall conclude our paper, however, with a more useful result which generalizes part (a) of Theorem 2.2 in Cheng (1978), the main result in that paper. Now assume (i), (ii), (iii), (a), (b), (α) and (β) in Theorem 3.1. For any design $d \in \mathscr{D}$, let $A_d = \operatorname{tr} \mathbf{C}_d$ and $B_d = \sum_{i=1}^n g(\lambda_{di})$, where $\lambda_{d1}, \dots, \lambda_{dn}$ are the eigenvalues of $\mathbf{C}_d$. Then by Theorem 2.1,

$$\sum_{i=1}^{n} f(\lambda_{di}) \geq f(\lambda(A_d, B_d)) + (n-1) f(\mu(A_d, B_d)), \tag{3.3}$$

and

$$\sum_{i=1}^{n} f(\lambda_{d^*i}) = f(\lambda(A_{d^*}, B_{d^*})) + (n-1) f(\mu(A_{d^*}, B_{d^*})), \tag{3.4}$$

where

$$\begin{aligned} \lambda(A_d, B_d) + (n-1)\mu(A_d, B_d) &= A_d, \\ g(\lambda(A_d, B_d)) + (n-1) g(\mu(A_d, B_d)) &= B_d, \\ \lambda(A_d, B_d) \geq \mu(A_d, B_d) &> 0. \end{aligned} \tag{3.5}$$

By assumption (b), $A_d \leq A_{d^*}$. So if $\mu(A_d, B_d) \leq \mu(A_{d^*}, B_{d^*})$, then $(\lambda(A_{d^*}, B_{d^*}), \mu(A_{d^*}, B_{d^*}), \dots, \mu(A_{d^*}, B_{d^*})) <_w (\lambda(A_d, B_d), \mu(A_d, B_d), \dots, \mu(A_d, B_d))$, where $<_w$ is upper weak majorization [see Marshall and Olkin (1979)]. By Theorem A.8 in Chapter 3 of Marshall and Olkin (1979), if f is also convex and nonincreasing, then

$$\begin{aligned} & f(\lambda(A_d, B_d)) + (n-1) f(\mu(A_d, B_d)) \\ & \quad \geq f(\lambda(A_{d^*}, B_{d^*})) + (n-1) f(\mu(A_{d^*}, B_{d^*})). \end{aligned} \tag{3.6}$$

Now (3.3), (3.4) and (3.6) together imply that $\sum_{i=1}^n f(\lambda_{di}) \geq \sum_{i=1}^n f(\lambda_{d^*i})$. This proves

THEOREM 3.5. *Assume conditions* (i), (ii), (iii), (a) *and* (b) *in Theorem* 3.1. *If* d^* *also maximizes* $\mu(A_d, B_d)$ *over* $\mathscr{D}$, *where* $\mu(A_d, B_d)$ *is defined in* (3.5), *then* d^* *minimizes* $\sum_{i=1}^n f(\lambda_{di})$ *over* $\mathscr{D}$ *for all* f *satisfying* $(\alpha), (\beta)$ *in Theorem* 3.1 *and*

(ε) *f is convex and nonincreasing on* $(0, A^*)$.

Theorem 2.2 of Cheng (1978) corresponds to the choice of $g(x) = x^2$ in Theorem 3.5, which so far has had many applications in solving various design problems. Recently Kunert and Martin (1985) proved the same result for $g(x) = x^{-p}$. They also proved a variant of this result and successfully applied it to prove the optimality of finite Williams II(a) designs [Williams (1952)] for first-order autoregressive processes. This adds another item to the list of applications of problem (1.1). The results in this paper are potentially useful to other situations and it is hoped that more applications can be reported in the future.

REFERENCES

CHENG, C. S. (1978). Optimality of certain asymmetrical experimental designs. *Ann. Statist.* **6** 1239–1261.

CHENG, C. S. (1981). Maximizing the total number of spanning trees in a graph: two related problems in graph theory and optimum design theory. *J. Combin. Theory Ser. B* **31** 240–248.

CHENG, C. S., MASARO, J. S. and WONG, C. S. (1985). Optimal weighing designs. *SIAM J. Algebraic Discrete Methods* **6** 259–267.

GALIL, Z. and KIEFER, J. (1983). Comparison of designs equivalent under one or two criteria. *J. Statist. Plann. Inference* **8** 103–116.

JACROUX, M. (1985). Some sufficient conditions for the type I optimality of block designs. *J. Statist. Plann. Inference* **11** 385–398.

KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2** 849–879.

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam.

KUNERT, J. (1985). Optimal repeated measurements designs for correlated observations and analysis by weighted least squares. *Biometrika* **72** 375–389.

KUNERT, J. and MARTIN, R. J. (1985). On the optimality of finite Williams II(a) designs. To appear in *Ann. Statist.*

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic, New York.

WILLIAMS, R. M. (1952). Experimental designs for serially correlated observations. *Biometrika* **39** 151–167.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720