INFORMATION MATRIX FOR OPTIMAL DESIGN

Gérard Antille*

University of Geneva, Switzerland

Institute of Statistics Mimeo Series #1247

August 1979

* This work supported in part by the National Science Foundation under grant MCS 78-01434.

INFORMATION MATRIX FOR OPTIMAL DESIGN

by

Gérard ANTILLE

University of Geneva (Switzerland)

In the research of optimal design of experiments the information matrix has a very important part. This paper, divided into two distinct sections, gives a characterization of the structure of the information matrix for G-optimal design and for Φ-optimal design.

§ 1 . Information matrix for G-optimal design

I. Introduction and summary :

A G-optimal design is a design which minimizes the maximum of the variance of the response. In this section we prove that the G-optimal information matrix can be expressed with a discrete uniform measure defined on the space of the controllable variables.

We also give a new and short proof of one part of the equivalence theorem of Kiefer and Wolfowitz for the general and the truncated case.

II. Design and information matrix .

Let y be a random variable related to a controllable variable v by the relation of the form

$$y/v = f(v)'\theta + \varepsilon_v$$

with $f : B \longrightarrow \mathbb{R}^m$, $B \subset \mathbb{R}^k$, a continuous application

$\theta = (\theta_1, \ldots, \theta_m)'$ a vector of unknown parameters

$E[\varepsilon_v] = 0 \qquad Var[\varepsilon_v] = \sigma^2$.

The covariance between the ε_v's , for distinct v , are all zero .

If θ is estimable, it is well known, that the best linear unbiased estimator for θ is

$$\hat{\theta} = (X'X)^{-1}X'Y$$

where X is the $N \times m$ design matrix whose i^{th} row is $f(v_i)'$.

Then $\frac{1}{N} X'X$ is called the information matrix of the design

$$\frac{1}{N} X'X = \frac{1}{N} \sum_{l=1}^{N} f(v_l)f'(v_l)$$

or if we note $f(v_l)' = (x_{1l}, \ldots, x_{m\,l})$

we have $\frac{1}{N} X'X = \frac{1}{N} \sum_{l=1}^{N} x_l\, x'_l$.

Il for each v_l we have n_l observations <u>a concrete design</u> of experiment is given by $\{(v_l, \frac{n_l}{N})\}$ $l = 1 \ldots n$.

Let $D = \frac{1}{N} \begin{pmatrix} n_1 & & \\ & \ddots & \\ & & n_n \end{pmatrix}$, X'DX is called the information matrix of the design $\{(v_l, \frac{n_l}{N}\}_{l=1,\ldots,n}$, with $N = \sum n_l$

A <u>discrete design</u> is given by $\{(v_l, p_l)\}_{l=1,\dots,n}$, the information matrix of this design is

$$X' \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{pmatrix} X \quad , \text{ with } \Sigma\, p_l = 1 \quad .$$

Let Θ be a class of probability measures defined on B , with the following assumption : all the discrete measures belong to Θ .

A probability measure $\xi \in \Theta$ such that

$$m_{ij}(\xi) = \int_B f_i(v)\, f_j(v)\, d(\xi(v)) \qquad \begin{matrix} i = 1 \dots m \\ j = 1 \dots m \end{matrix}$$

exists and is finite, is a <u>continuous design</u> of experiment , and the matrix $M(\xi) = (m_{ij}(\xi))$ is called the information matrix of ξ .

By noting $\Lambda = \{\lambda | \lambda = \xi f^{-1} \; ; \; \forall\, \xi \in \Theta\}$ a class of probability measure defined on $\mathscr{X} = f(B)$, we have for the information matrix

$$(m_{ij}(\lambda)) = \int_{\mathscr{X}} xx'd(\lambda(x)) \; .$$

III. Representation and properties of the information matrix with probability measure on finite support :

The application $H : \mathscr{X} \longrightarrow \mathbb{R}^{\frac{m(m+1)}{2}}$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \longmapsto \begin{pmatrix} x_1^2 \\ \vdots \\ x_m^2 \\ \vdots \\ x_i x_j \\ \vdots \end{pmatrix} \quad i < j$$

is a one to one correspondence between

$$H(\mathscr{X}) \text{ and } \mathscr{M} = \{ xx' \mid x \in \mathscr{X} \subset \mathbb{R}^m \} .$$

If $\mathscr{X}$ is a compact set, since H is a continuous application we have $H(\mathscr{X})$ is also compact and the convex hull of $H(\mathscr{X})$ is compact, then the convex hull $\mathscr{H}(\mathscr{M})$ of $\mathscr{M}$ also .

Then $M(\lambda) \in \mathscr{H}(\mathscr{M}) \quad \forall \lambda \in \Lambda$.

Then for every matrix $M(\lambda)$ there exists a representation of the form $\sum_{i=1}^{n} \lambda_i x_i x'_i$ with $\sum \lambda_i = 1 \quad x_i \in \mathscr{X}$

and we have $n \leqslant \frac{1}{2} m(m + 1) + 1$ by a theorem of Caratheodory (Rockafellar 1970 , p.155-156).

Lemmas 1): $M(\lambda)$ is a symmetric positive-semi definite matrix (Fedorov 1972 p. 66) .

2) If Λ is the class of all the probability measures defined on $\mathscr{X}$, $\mathscr{M}_\Lambda = \{M(\lambda) \mid \lambda \in \Lambda \}$ is a convex set.

The proof is obtained by using the property of convexity of Λ .

3) Let $\lambda \in \Lambda$ we have $\int_{\mathscr{X}} x'M^{-1}(\lambda)x d(\lambda(x)) = m$ (Fedorov 1972, p. 68-69) .

4) Let $\lambda \in \Lambda$ such that $x'M^{-1}(\lambda)x \leqslant m$, $\forall x \in \mathscr{X}$

Let $\sum_{i=1}^{n} \lambda_i x_i x'_i$ a representation of $M(\lambda)$

Let $\mathscr{X}_\lambda = \{ x \in \mathscr{X} \mid x$ is an element of the given representation of $M(\lambda)\}$.

Let $\mathscr{X}_{\lambda,m} = \{ x \in \mathscr{X} \mid x'M^{-1}(\lambda)x = m \}$.

We have $\mathscr{X}_\lambda \subset \mathscr{X}_{\lambda,m}$.

<u>Proof</u> : If there exists a k such that

$$x'_k M^{-1}(\lambda)x_k < m \text{ and } x'_i M^{-1}(\lambda)x_i = m \qquad i \neq k$$

we have $\Sigma \lambda_i x'_i M^{-1}(\lambda)x_i < m$

and it is a contradiction with the preceding lemma .

<u>Proposition</u> :

Let $\lambda \in \Lambda$ such that $x'M^{-1}(\lambda)x \leqslant m \quad \forall x \in \mathscr{X}$

There exists a representation of $M(\lambda)$ of the form $\frac{1}{n} \sum^{n} x_i x'_i$.

<u>Proof</u> : we know that there exists a representation of the form $\sum^{n} \lambda_i x_i x'_i$, and for $x_i \in \mathscr{X}_\lambda$ we have

$$x'_i \, M^{-1}(\lambda) x_i = m$$

$$\sum^{n} x'_i \, M^{-1}(\lambda) \; x_i = nm$$

$$Tr(M^{-1}(\lambda) \sum^{n} \frac{1}{n} x_i \; x'_i) = m \quad .$$

A solution of this last equation is given by

$$M(\lambda) = \frac{1}{n} \Sigma \, x_i x'_i \quad . \qquad x_i \in \mathscr{X}_\lambda \subset \mathscr{X} \, .$$

Example : $\mathscr{X} = [-1, +1]^m$

If $M^{-1}(\lambda) = I_m$ we have $x'M^{-1}(\lambda)x \leqslant m$.

Hence if we choose the Hadamard matrix for the design matrix X, in the case of $n = m$, we have $\frac{1}{n} X'X = I$.

If $n > m$ the design matrix X is given by the first m columns of the Hadamard matrix of order n .

IV . Applications :

1 .G-optimal design .

A design $\lambda^* \in \Lambda$ is G-optimal if

$$\min_{\lambda\in\Lambda} \max_{x\in\mathscr{X}} x'M^{-1}(\lambda)x = \max x'M^{-1}(\lambda^*)x$$

By using lemma 4 we obtain :

$$m = \int_{\mathscr{X}} x'M^{-1}(\lambda)xd(\lambda(x)) \leqslant \max \; x'M^{-1}(\lambda)x \int_{\mathscr{X}} d(\lambda(x))$$

Then we have the information matrix of a G-optimal design defined on a compact set $X \subset \mathbb{R}^m$ that can be expressed in the form $\frac{1}{n} \sum^n x_i \, x'_i$, $x_i \in \mathscr{X}$ and $n \leqslant \frac{1}{2} m(m+1) + 1$.

When the possible choice of the controllable variables are points in an m-dimensional cube of side 2 , the points of the G-optimal design can be chosen in the set of the vertices of the cube.

2 . Remark on the Kiefer-Wolfowitz equivalence theorem .

Lemma : Let $\lambda \in \Lambda$ such that $x'M^{-1}(\lambda)x \leqslant m \quad \forall\, x \in \mathscr{X}$

We have : $\operatorname{Tr}\,[M^{-1}(\lambda)M(\xi)] \leqslant m$, $\forall\, \xi \in \Lambda$

: $\det\,[M^{-1}(\lambda)M(\xi)] \leqslant 1$, $\forall\, \xi \in \Lambda$.

Proof : $\operatorname{Tr}[M^{-1}(\lambda)M(\xi)] = \operatorname{Tr}[M^{-1}(\lambda) \int xx'd(M(\xi))]$

$$= \int x'M^{-1}(\lambda)xd(\xi(x)) \leqslant m \int d(\xi(x)) = m$$

and by using inequalities between the geometric and arithmetic mean we have

$$(\det[M^{-1}(\lambda)M(\xi)])^{1/m} \leqslant \frac{1}{m} \operatorname{Tr}[M^{-1}(\lambda)M(\xi)] \leqslant 1 \ .$$

A short proof of G-optimal implies D-optimal .

A design λ^* is D-optimal if $\min_{\lambda} \det(M^{-1}(\lambda)) = \det M(\lambda^*)$.

If we note $\tilde{N}$ the information matrix of a G-optimal design we have

$$\det(\tilde{N}^{-1}M) \leqslant 1 \qquad \forall\ M \in \mathscr{M}_{\Lambda}$$

but $$\det(\tilde{N}^{-1}M) = \det(\tilde{N}^{-1})\ \det(M)$$

Then $$\det(\tilde{N}^{-1}) \leqslant \frac{1}{\det(M)} = \det M^{-1}$$

and the matrix $\tilde{N}$ is D-optimal .

V . The truncated case :

If the experimenter is only interested in parameters in a regression with m parameters, we will speak of truncated design and truncated optimality.

In this case for $x \in \mathscr{X} \subset \mathbb{R}^m$, $s < m$, $\lambda \in \Lambda$ and if we note $x^{(1)} = (x_1, \ldots, x_s)'$, $x^{(2)} = (x_{s+1}, \ldots, x_m)$ we have

$$M(\lambda) = \int_{\mathscr{X}} xx'd(\lambda(x)) = \begin{pmatrix} M_{11}(\lambda) & M_{12}(\lambda) \\ M_{21}(\lambda) & M_{22}(\lambda) \end{pmatrix}$$

with $M_{ij}(\lambda) = \int_{\mathscr{X}} x^{(i)}x^{(j)'}d(\lambda(x))$

$M_{11}(\lambda)$ is a $s \times s$-matrix .

If $M^{-1}(\lambda)$ exists, we can make a partition of $M(\lambda)$ in the same way as for $M(\lambda)$.

Hence

$$M^{-1}(\lambda) = \begin{pmatrix} M^{11}(\lambda) & M^{12}(\lambda) \\ M^{21}(\lambda) & M^{22}(\lambda) \end{pmatrix}$$

with

$$M^{11} = (M_{11} - M_{12}M_{22}^{-1} M_{21})^{-1}$$

$$M^{22} = (M_{22} - M_{21} M_{11}^{-1} M_{12})^{-1}$$

$$M^{12} = -M_{11}^{-1} M_{12} M^{22}$$

$$M^{21} = - M_{22} M_{21} M_{11}^{-1}$$

Lemmas :

1 . Let $\tilde{x} = x^{(1)} - M_{12} M_{22}^{-1} x^{(2)}$

Let $\tilde{\mathscr{X}} = \{ \tilde{x} \mid x \in \mathscr{X}$ for a fixed $\lambda \in \Lambda \}$.

Let $M_s^{-1}(\lambda) = M^{11}(\lambda)$.

Then $M_s(\lambda) = \int_{\mathscr{X}} \tilde{x}\,\tilde{x}'d(\lambda(x))$

Proof : $\tilde{x}\tilde{x}' = x^{(1)}x^{(1)'} - x^{(1)}x^{(2)'} M_{22}^{-1} M_{12}'$

$$- M_{12} M_{22}^{-1} x^{(2)} x^{(1)'}$$

$$+ M_{12} M_{22}^{-1} x^{(2)} x^{(2)'} M_{22}^{-1} M_{12}'$$

by integration the two last terms are the same hence

$$\int_{\mathscr{X}} \tilde{x}\,\tilde{x}'d(\lambda(x)) = M_{11} - M_{12} M_{22}^{-1} M_{21} = M_s(\lambda)$$

2 . $\int_{\mathscr{X}} \tilde{x}' M_s^{-1}(\lambda)\tilde{x}\, d(\lambda(x)) = s$

Proof : $\int_{\mathscr{X}} \tilde{x}' M_s^{-1}(\lambda)\,\tilde{x}\, d(\lambda(x)) = \int_{\mathscr{X}} x' M^{-1}(\lambda)x\, d(\lambda(x))$

$$- \int_{\mathscr{X}} x^{(2)'} M_{22}^{-1}(\lambda)\, x^{(2)}d(\lambda(x))$$

$$= m - \mathrm{Tr}[M_{22}^{-1} M_{22}] =$$

$$= m - (m - s) = s$$

Corollary : $\max_{\tilde{x}} \tilde{x}' M_s^{-1}(\lambda)\,\tilde{x} \geq s$

A design λ^* is called <u>G_s-optimal</u> if

$$\min_{\lambda\in\Lambda} \max_{\tilde{x}} \tilde{x}' M_s^{-1}(\lambda)\tilde{x} = \max_{\tilde{x}} \tilde{x}' M_s^{-1}(\lambda^*)\tilde{x}$$

A design λ^* is called <u>D_s-optimal</u> if

$$\min_{\lambda} \mathrm{Det}(M_s^{-1}(\lambda)) = \mathrm{Det}(M_s^{-1}(\lambda^*)) \quad .$$

<u>A short proof of G_s-optimal implies D_s-optimal</u>

Let λ^* a G_s-optimal design, as in the general case we can prove that : $\mathrm{Tr}[M_s^{-1}(\lambda^*)\ M_s(\lambda)] \leqslant s \qquad \forall\ \lambda \in \Lambda$

$$\mathrm{Det}[M_s^{-1}(\lambda^*)\ M_s(\lambda)] \leqslant 1$$

$$\mathrm{Det}[M_s^{-1}(\lambda^*)] \leqslant \mathrm{Det}[M_s^{-1}(\lambda)]$$

Hence λ^* is D_s-optimal .

§ 2 . Information matrix for Φ-optimal design

I. Introduction and summary

A Φ-optimal design is a design which minimizes a real valued function defined on the space of the information matrix for a given regression problem.

In this section we prove that the Φ-optimal information matrix satisfies the inequality constraint $x'M^{-1}x \leqslant m$ for x belonging to the controllable variables space. It can be expressed in the form $\frac{1}{n} \Sigma x_i x_i'$.

II. Φ-optimality

Let $\Phi : \mathscr{M}_\Lambda \longrightarrow \mathbb{R}$

A design $\lambda^* \in \Lambda$ is Φ-optimal if $\Phi(M(\lambda^*)) = \min\limits_{\lambda} \Phi(M(\lambda))$. The most common examples of optimality criteria are

D-optimality : $\Phi_o(M) = \det M^{-1}$

L-optimality $\Phi_{1,C}(M) = tr(CM^{-1})$; $C \geqslant 0$

A-optimality $\Phi_1(M) = tr(M^{-1})$

E-optimality $\Phi_\infty(M) = $ maximum eigenvalue of M^{-1} .

There are particular cases or limiting cases of $\Phi_p(M) = tr(M^{-p})$ (see Kiefer 1974) .

If Λ is convex and if Φ is a convex mapping we have for

$\lambda = \alpha \lambda_1 + (1 - \alpha)\lambda_2 \qquad 0 < \alpha < 1$

$$\Phi(M(\lambda)) \leqslant \alpha \, \Phi(M(\lambda_1)) + (1 - \alpha)\Phi(M(\lambda_2)) \, .$$

If λ_1 and λ_2 are Φ-optimal we have $\Phi(M(\lambda_1)) = \Phi(M(\lambda_2))$

and
$$\Phi(M(\lambda)) \leqslant \Phi(M(\lambda_1))$$

then λ is also Φ-optimal .

Hence in the above conditions we have proved that the set of Φ -optimal measures is convex.

III. An optimisation problem .

We will note M_o the solution of the following optimisation problem :

For a given regression problem, for Φ a convex mapping and for a finite subset $\mathscr{U}$ of $\mathscr{X}$ with cardinality of $\mathscr{U} \geqslant m$, we consider :

$$\begin{cases} - \min\limits_{M \in \mathscr{M}_\Lambda} \Phi(M) \\ - x' M^{-1} x \leqslant m \, , \; \forall \, x \in \mathscr{U} \end{cases}$$

Let $L(M, \nu) = \Phi(M) + \sum\limits^{n} \nu_i (x'_i \, M^{-1} x_i - m)$ with $\nu_i > 0$ be the Lagrange function of the minimization problem. By a Kuhn-Tucker theorem (Appendix A) for convex optimisation we have $\exists \; \nu^o > 0$ such that

$$L(M_o, \nu) \leqslant L(M_o, \nu^o) \leqslant L(M, \nu^o) \quad \forall \, \nu > 0 \, , \quad \forall \, M \in \mathscr{M}_\Lambda \, .$$

Then we have

$$\Sigma\ \nu_i(x'_i\ M_o^{-1}\ x_i - m) \leqslant \Sigma\ \nu_i^o(x'_i M_o^{-1} x_i - m) \leqslant 0$$

and in the case of $\nu_i \equiv 0$ we obtain

$$\Sigma\ \nu_i^o\ (x_i'\ M_o^{-1}\ x_i - m) = 0$$

$$\mathrm{Tr}(M_o^{-1}\ \Sigma\ \lambda_i^o\ x_i x'_i) = 0 \text{ where } \lambda_i^o = \frac{\nu_i^o}{\Sigma \nu_i^o}$$

A solution of this last equation is given by $M_o = \Sigma\ \lambda_i^o\ x_i\ x_i'$

Hence, we have seen that the matrix solution of our minimization problem can be expressed as $\Sigma\ \lambda_i^o\ x_i x_i'$, $x_i \in \mathscr{U}$ so we have for this representation of $M(\lambda)$: $\mathscr{X}_\lambda = \mathscr{U}$.

<u>Proposition</u> :

A class of solution of our minimization problem is given by the set $\{ M | M = \alpha M_o + (1 - \alpha) N \}$ for all $N \in \mathscr{M}_\Lambda$ satisfying $x'\ N^{-1} x \leqslant m$, $x \in \mathscr{U}$.

Indeed : using the fact that M_o and N are positive definite matrices implies for $0 < \alpha < 1$,

$$[\alpha\ M_o + (1 - \alpha)N]^{-1} \leqslant \alpha\ M_o^{-1} + (1 - \alpha)N^{-1}$$

Appendix B

we have $x'M^{-1}x \leqslant \alpha x' M_o^{-1} x + (1 - \alpha) x' N^{-1} x \leqslant m$

and $\Phi(M) \leqslant \alpha\Phi(M_o) + (1 - \alpha)\Phi(N) = \Phi(N) + \alpha(\Phi(M_o) - \Phi(N)) \leqslant \Phi(N)$

then M is a solution of our minimization problem.

Proposition :

A solution of our minimization problem satisfies the inequality constraint over all the controllable space $\mathscr{X}$ if Φ is strongly convex.

Proof : Assume there exists x such that

$$\tilde{x}' M_o^{-1} \tilde{x} > m$$

Let $W = \mathscr{U} \cup \{\tilde{x}\}$.

Let $\check{M}$ be a solution of our minimization problem on W , we have $L(\check{M}, \check{\nu}^o) \leqslant L(M_o, \tilde{\nu}^o)$

or $\Phi(\check{M}) \leqslant \Phi(M_o)$

but $\hat{M}$ satisfies $x'\check{M}^{-1}x \leqslant m \qquad x \in \mathscr{U}$

then $\Phi(M_o) \leqslant \Phi(\check{M})$

then $\Phi(M_o) = \Phi(\check{M})$.

As Φ is strongly convex we have $M_o = \tilde{M}$ but $\tilde{x}'M_o^{-1}\tilde{x} \neq \tilde{x}'M^{-1}\tilde{x}$ which is a contradiction .

Hence $x'M_o^{-1}x \leqslant m \qquad \forall\, x \in \mathscr{X}$.

Corollary :

Let M^* be the Φ-optional matrix, Φ strongly convex . Then M^* is a solution of our minimization problem and satisfies the inequality constraint over $\mathscr{X}$.

<u>Proof</u> : Since M_o is a solution of our minimization problem we have

$$L(M_o, \nu^o) \leqslant L(M^*, \nu^o)$$

then $$\Phi(M_o) \leqslant \Phi(M^*)$$

but M^* is Φ-optimal : $\Phi(M^*) \leqslant \Phi(M_o)$

then $\Phi(M^*) = \Phi(M_o)$ as Φ is strongly convex we have $M_o = M^*$.

<u>Concluding remarks</u>:

1) In the first section we prove that there exists a representation of the form $\frac{1}{n} \sum^{n} x_i \; x_i'$ for the matrices satisfying $x'M^{-1}x \leqslant m \quad \forall \; x \in \mathscr{X}$. Now we know that the Φ-optimal information matrix satisfies $x'M^{-1}x \leqslant m \quad \forall \; x \in \mathscr{X}$. Then there exists a representation of the form

$$\frac{1}{n} \; \Sigma \; x_i \; x_i'$$

for the Φ-optimal matrix.

2) The duality theorem of Sibson (1972, 1974) is a particular case of our result by choosing $\Phi = -\ln \det$.

Thus we have a form of the general equivalence theorem of Kiefer and Wolfowitz .

3) Our result is also valid for truncated design in the following sense : if we define an application Φ_s on the set of the matrices $M_s(\lambda)$, we have : the Φ_s-optimal information matrix satisfies $x'M_s^{-1}(\lambda)x$.

Acknowledgments :

This paper is based on part of the author's Ph. D. dissertation written at the University of Geneva under the supervision of Professor F. STREIT and I.M. CHAKRAVARTI .

I wish to thank the department of Statistics of the University of North Carolina at Chapel Hill and especially Professor I.M. Chakravarti for having allowed me to spend some time there.

Appendix A :

Modified Kuhn Tucker Theorem :

M_o is a solution of our minimization problem if and only if there exists $\nu^o = (\nu_1^o , \ldots, \nu_n^o)$, $\nu_i^o \geq 0$ such that

$$L(M_o, \nu) \leq L(M_o, \nu^o) \leq L(M, \nu^o) .$$

Proof :

1) Let M_o be a solution of our minimization problem .

Let $A = \{(y_o, \ldots, y_n) \mid y_o \geq \Phi(M) , y_i \geq x_i' M^{-1} x_i - m, \quad i = 1 , \ldots, n , \text{ for at least one } M \}$

$$B = \{(y_o, \ldots, y_n) \mid y_o < \Phi(M_o), y_i < 0 \quad i = 1, \ldots, n\}$$

By the separation theorem for convex sets we have : there is $v = (v_o \ldots v_n)$, $v \neq 0$ such that

$v'a > v'b$ for $a \in A$, $b \in B$.

But this last inequality is only valid if $v \geq 0$ and is still true if

$$a_o = \Phi(M) , a_i = x_i' M^{-1} x_i - m$$

$$b_o = \Phi(M_o) , b_i = 0$$

then $v_o \Phi(M) + \Sigma v_i (x_i' M^{-1} x_i - m) \geq v_o \Phi(M_o)$ for all $M \geq 0$.

If $v_o = 0$ we obtain $\Sigma \nu_i (x_i' M^{-1} x_i - m) \geq 0$ which is in contradiction with $x_i' M^{-1} x_i - m \leq 0$.

Now set $\nu^o = \frac{1}{v^o} (v_1, \ldots, v_n)$ so $\nu^o \geq 0$

and $$\Phi(M) + \Sigma\, \nu_i^o (x_i' M^{-1} x_i - m) \geq \Phi(M_o)$$

for $M = M_o$ it follows $\Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m) \geq 0$ but from the inequality constraint we obtain

$$\Sigma\, \nu_i (x_i' M_o^{-1} x_i - m) \leq 0$$

also true for $\nu_i = \nu_i^o$.

Then $\Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m) = 0$.

And at last

$$\Phi(M_o) + \Sigma\, \nu_i (x_i' M_o^{-1} x_i - m) \leq \Phi(M_o) + \Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m)$$

$$\leq \Phi(M) + \Sigma\, \nu_i^o (x_i' M^{-1} x_i - m)$$

so $L(M_o, \nu) \leq L(M_o, \nu^o) \leq L(M, \nu^o)$.

2) Conversely : using $L(M_o, \nu) \leq L(M_o, \nu^o)$ ... we obtain

$$\Sigma\, \nu_i (x_i' M_o^{-1} x_i - m) \leq \Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m) .$$

As $\nu \geq 0$ this inequality is only true if

$$x_i' M_o^{-1} x_i - m \leq 0 .$$

Thus M_o satisfies the inequality constraint and $\Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m) \leq 0$

but for $\nu_i = 0$ we obtain $\Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m) \geq 0$ then

$\Sigma\, \nu_i^o (x_i' M_o^{-1} x_i - m) = 0$

and $\Phi(M_o) \leq \Phi(M)$.

Hence M_o is a solution of our minimization problem.

Appendix B :

Theorem (cited in Fedorov p. 19) .

If A and B are positive-definite matrices then

$$\alpha A^{-1} + (1 - \alpha)B^{-1} \geq [\alpha A + (1 - \alpha)B]^{-1} \quad 0 < \alpha < 1.$$

Proof :

$$A \geq 0 \text{ and } A^{-1} \geq 0 \text{ then } \alpha(A + A^{-1}) \geq 0$$

and

$$\alpha A \geq - \alpha A^{-1}$$

$$(1 - \alpha)B \geq -(1 - \alpha)B^{-1}$$

$$\alpha A + (1 - \alpha)B \geq - (\alpha A^{-1} + (1 - \alpha)B^{-1}) .$$

But using the fact $M \geq N$ implies $-M^{-1} \geq -N^{-1}$ for positive definite matrices we have

$$- (\alpha A + (1 - \alpha)B)^{-1} \geq (\alpha A^{-1} + (1 - \alpha)B^{-1})^{-1} .$$

On the other hand

$$(\alpha A^{-1} + (1 - \alpha)B^{-1})^{-1} + \alpha A^{-1} + (1 - \alpha)B^{-1} \geq 0$$

$$(\alpha A^{-1} + (1 - \alpha)B^{-1})^{-1} \geq - (\alpha A^{-1} + (1 - \alpha)B^{-1})$$

Then $-(\alpha A + (1 - \alpha)B)^{-1} \geq - (\alpha A^{-1} + (1 - \alpha)B^{-1})$

and $\alpha A^{-1} + (1 - \alpha)B^{-1} \geq (\alpha A + (1 - \alpha)B)^{-1}$.

B I B L I O G R A P H Y

FEDOROV V.V. Theory of optimal experiment . Academic Press, N.Y.(1972) .

KIEFER J . General equivalence Theory for Optimum designs. Ann. Stat. 2 (1974) p. 849-878.

ROCKAFELLAR Convex Analysis. Princeton University Press (1970).

SIBSON, R. Discussion on the papers by Wynn and Laycock. J.R. Stat. Soc. Series B, 34 (1972), p. 181-183.

-- D_A-optimality and duality . Colloquia Mathematica Societatis Janos Bolyai. 9. European Meeting of Statisticians, Budapest, 1972 (1974).