

AD-A097 035

STANFORD UNIV CA DEPT OF STATISTICS
A-OPTIMALITY FOR REGRESSION DESIGNS.(U)
JAN 81 N N CHAN

F/G 12/1

N00014-75-C-0442
NL

UNCLASSIFIED TR-49

1 of 1
20 21 22



END
DATE
FILMED
4-2-81
DTIC

AD A 097035

(14) TR-49

(12)

(6)

A-OPTIMALITY FOR REGRESSION DESIGNS.

BY

(10) N. N. CHAN

LEVEI II

(9)

TECHNICAL REPORT NO. 49

(11) JANUARY 1981

DTIC
ELECTE

MAR 30 1981

E

(15)

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)

OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

(12) 15

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA



DTIC FILE COPY

332580

ew

81 3 30 024

A-OPTIMALITY FOR REGRESSION DESIGNS

by

N. N. CHAN

The Chinese University of Hong Kong

TECHNICAL REPORT NO. 49

JANUARY 1981

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A	

PREPARED UNDER CONTRACT N000-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government.
Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Also Issued as Technical report No. 163 under National Science Foundation
Grant MCS 78-07736 - Dept. of Statistics, Stanford University.

A-OPTIMALITY FOR REGRESSION DESIGNS

N. N. Chan

1. Introduction.

Consider the linear regression model

$$y = X\beta + \epsilon ,$$

where y is an $m \times 1$ vector of observations, X is an $m \times n$ matrix to be called the design matrix, β is an $n \times 1$ vector of unknown parameters, and ϵ is an $m \times 1$ vector of random variables with mean the $m \times 1$ zero vector and known covariance matrix Λ . We assume that $m \geq n$ and denote the eigenvalues of Λ in ascending order of magnitude by

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \leq \dots \leq \lambda_m .$$

For later use denote the diagonal matrices with diagonal elements $\lambda_1, \dots, \lambda_i$ by Λ_i , $i = n$ and m .

For a given design matrix X of rank n , an unbiased estimate of the parameter β based on the observation y is the simple least squares estimate

$$(X'X)^{-1}X'y ,$$

whose covariance matrix is given by

$$(1) \quad (X'X)^{-1}X'\Lambda X(X'X)^{-1} .$$

One of the design problems is to choose X from a given experimental region such that the trace of the matrix in (1) is minimal. This is a problem in the A-optimal designs of regression experiments and was considered by Dorogovcev (1971) under the more general setting that the observations are the realization of stochastic processes. Earlier work on A-optimal designs was given by Elfving (1952) and Chernoff (1953).

In this paper the experimental region under consideration is taken to be the set H of all $m \times n$ real matrices of rank n whose i^{th} column has a Euclidean norm not exceeding c_i , $i = 1, \dots, n$, where the c_i are given positive numbers. In section 2, it is shown that for any matrix X in H the trace of the matrix in (1) has as a lower bound of

$$\left(\sum_{i=1}^n c_i^2 \right)^{-1} \left(\sum_{i=1}^n \lambda_i^{\frac{1}{2}} \right)^2.$$

In section 3, a necessary and sufficient condition for the existence of an X in H to attain the lower bound is derived. For the case in which all the c_i are equal, a partial result was given in Chan and Wong (1981). Dorogovcev (1971) obtained the lower bound for the special case $n = 2$ and $c_1 = c_2$.

It is worth noting that in the regression model if one considers the best linear unbiased estimate $(X' \Lambda^{-1} X)^{-1} X' \Lambda^{-1} y$ and its covariance matrix $(X' \Lambda^{-1} X)^{-1}$, by minimizing the trace of the latter for all X in H , the corresponding optimal design problem has a simple solution, as is given in Rao (1973, p. 236). On the other hand, if one wishes to minimize the determinant of $(X' \Lambda^{-1} X)^{-1}$, there is the so called D-optimal design

problem, of which comprehensive reviews can be found in St. John and Draper (1975) and Kiefer and Galil (1980).

2. An Inequality.

For the regression model and the set H as given in section 1, we note that in minimizing the trace of the matrix in (1) with respect to X in H , the matrix Λ in (1) can be replaced by the diagonal matrix Λ_m without loss of generality, in view of the existence of an orthogonal matrix P such that

$$\Lambda = P' \Lambda_m P$$

and the following equality

$$(X'X)^{-1} X' \Lambda X (X'X)^{-1} = (Y'Y)^{-1} Y' \Lambda_m Y (Y'Y)^{-1} ,$$

where $Y = PX$ which is again in H . The following lemma of Fan (1949) will be required in the proof of our main inequality.

Lemma 1. Let B be a real $m \times n$ matrix whose n columns form an orthonormal set. Then

$$\text{tr } B' \Lambda B \geq \text{tr } \Lambda_n ,$$

where tr represents the trace operation.

Theorem 1. For any X in H ,

$$\text{tr} \{ (X'X)^{-1} X' \Lambda X (X'X)^{-1} \} \geq \left(\sum_{i=1}^n c_i^2 \right)^{-1} \left(\sum_{i=1}^n \lambda_i^{\frac{1}{2}} \right)^2 .$$

Proof. By the Cauchy-Schwarz inequality applied to the trace inner product $\text{tr}\{X'Y\}$ between two real $m \times n$ matrices X and Y , we have

$$(2) \quad \text{tr}\{X'X\} \times \text{tr}\{(X'X)^{-1}X'\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}X(X'X)^{-1}\} \geq \text{tr}^2\{X'\Lambda^{\frac{1}{2}}X(X'X)^{-1}\}.$$

But the trace on the right-hand side is

$$(3) \quad \text{tr}\{(X'X)^{-\frac{1}{2}}X'\Lambda^{\frac{1}{2}}X(X'X)^{-\frac{1}{2}}\},$$

which is not less than $\sum_{i=1}^n \lambda_i^{\frac{1}{2}}$ by Lemma 1 on noting that the n columns of the matrix $X(X'X)^{-\frac{1}{2}}$ are orthonormal. By the definition of the set H ,

$$(4) \quad \text{tr}\{X'X\} \leq \sum_{i=1}^n c_i^2.$$

Hence the main inequality follows.

3. A-Optimal Designs.

The main result of this work is to obtain a necessary and sufficient condition on $(\lambda_1, \dots, \lambda_m)$ and $(c_1, \dots, c_n)$ for the existence of a matrix in H such that the lower bound in Theorem 1 is attained. For this we need the following lemmas.

Lemma 2. Let D be an $n \times n$ real diagonal matrix with diagonal elements $d_1 \leq d_2 \leq \dots \leq d_n$, and $a_1, \dots, a_n$ be n real numbers such that $a_1 \leq a_2 \leq \dots \leq a_n$ and

$$\sum_{i=1}^n a_i = \sum_{i=1}^n d_i .$$

Then there exists an $n \times n$ orthogonal matrix P such that the n diagonal elements of $P'DP$ are $a_1, \dots, a_n$ if and only if

$$\sum_{i=1}^k a_i \geq \sum_{i=1}^k d_i , \quad k = 1, 2, \dots, n-1 .$$

This lemma is a version of a result by Horn (1954) and a proof is given by Mirsky (1958). See also Marshall and Olkin (1979, p. 220).

Lemma 3. Let D be as in Lemma 2 and B be an $n \times k$ matrix whose k columns form an orthonormal set. Arrange the eigenvalues of the $k \times k$ matrix $B'DB$ in ascending order $b_1 \leq b_2 \leq \dots \leq b_k$. Then $b_i \geq d_i$, $i = 1, \dots, k$.

This is the Poincaré separation theorem and can be found for example in Rao (1973, p. 64).

Theorem 2. Suppose that the positive numbers c_i , $i = 1, \dots, n$, are arranged in ascending order of magnitude and that the smallest eigenvalue λ_1 of the covariance matrix Λ is positive. Then there is an X in H such that

$$\text{tr}\{(X'X)^{-1}X'\Lambda X(X'X)^{-1}\} = \left(\sum_{i=1}^n c_i^2\right)^{-1} \left(\sum_{i=1}^n \lambda_i^2\right)^2 ,$$

if and only if

$$\left(\sum_{i=1}^n c_i^2\right)^{-1} \sum_{i=1}^k c_i^2 \geq \left(\sum_{i=1}^n \lambda_i^{1/2}\right)^{-1} \sum_{i=1}^k \lambda_i^{1/2}, \quad k = 1, \dots, n-1.$$

Proof. Sufficiency: Consider the diagonal matrix $\Lambda_n^{1/2}$ whose diagonal elements are $\lambda_i^{1/2}$, $i = 1, \dots, n$. By Lemma 2 there exists an orthogonal matrix P of order n such that the i^{th} diagonal element of $P' \Lambda_n^{1/2} P$ is bc_i^2 , $i = 1, \dots, n$, where

$$b = \left(\sum_{i=1}^n c_i^2\right)^{-1} \left(\sum_{i=1}^n \lambda_i^{1/2}\right).$$

Denote by X the $m \times n$ matrix

$$b^{-1/2} \begin{bmatrix} \Lambda_n^{1/2} P \\ 0 \end{bmatrix},$$

where 0 is an $(m-n) \times n$ submatrix of zeros. Note that X is of rank n as $\lambda_1 > 0$ and that the i^{th} diagonal element of $X'X$ equals c_i^2 as we have

$$X'X = b^{-1} P' \Lambda_n^{1/2} P.$$

Hence X is a member of the set H . Moreover, for the diagonal matrix Λ_m of order m , we have

$$\begin{aligned} X' \Lambda_m X &= b^{-1} [P' \Lambda_n^{1/2} \quad 0'] \Lambda_m \begin{bmatrix} \Lambda_n^{1/2} P \\ 0 \end{bmatrix} \\ &= b^{-1} P' \Lambda_n^{1/2} \Lambda_n \Lambda_n^{1/2} P \\ &= b^{-1} P' \Lambda_n^{3/2} P, \end{aligned}$$

and so

$$\begin{aligned} \text{tr}\{(X'X)^{-1}X'\Lambda_m X(X'X)^{-1}\} &= \text{tr}\{b(P'\Lambda_n^{-1/2}P)P'\Lambda_n^{3/2}P(P'\Lambda_n^{-1/2}P)\} \\ &= b \text{tr}\{P'\Lambda_n^{1/2}P\} \\ &= \left(\sum_{i=1}^n c_i^2\right)^{-1} \left(\sum_{i=1}^n \lambda_i^{1/2}\right)^2 . \end{aligned}$$

The proof for sufficiency is completed by replacing Λ_m by Λ as remarked at the beginning of Section 2.

Necessity. Suppose that X , a member of H , is such that the inequality in Theorem 1 becomes an equality. Then the three inequalities in the proof of Theorem 1 reduce to equalities. First, note that the i^{th} diagonal element of the matrix $X'X$ equals c_i^2 , $i = 1, \dots, n$, because it cannot exceed c_i^2 (as X is in H) and from (4)

$$\text{tr}\{X'X\} = \sum_{i=1}^n c_i^2 .$$

By Lemma 2, it is then enough to show that $\lambda_1^{1/2}, \dots, \lambda_n^{1/2}$ are the eigenvalues of the $n \times n$ matrix $bX'X$. For this, note that the Cauchy-Schwarz inequality (2) becoming an equality implies that there is a nonzero real number d such that

$$X = d\Lambda^{1/2}X(X'X)^{-1} .$$

So we have

$$X'X = dX'\Lambda^{1/2}X(X'X)^{-1} .$$

The equality corresponding to (3) then implies that

$$\begin{aligned}
 \sum_{i=1}^n \lambda_i^{\frac{1}{2}} &= \text{tr}\{X' \Lambda^{\frac{1}{2}} X (X'X)^{-1}\} \\
 (5) \qquad \qquad &= d^{-1} \text{tr}\{X'X\} \\
 &= d^{-1} \sum_{i=1}^n c_i^2.
 \end{aligned}$$

Therefore, $d = b^{-1}$, and so

$$bX'X = X' \Lambda^{\frac{1}{2}} X (X'X)^{-1}.$$

It remains to show that the $n \times n$ matrix

$$(6) \qquad (X'X)^{-\frac{1}{2}} X' \Lambda^{\frac{1}{2}} X (X'X)^{-\frac{1}{2}}$$

has $\lambda_1^{\frac{1}{2}}, \dots, \lambda_n^{\frac{1}{2}}$ as its eigenvalues. In fact, by replacing Λ by Λ_m and using Lemma 3, we see that the i^{th} smallest eigenvalue of the matrix in (6) is not less than $\lambda_i^{\frac{1}{2}}$, $i = 1, \dots, n$, and, in view of the first equality in (5), must be equal to $\lambda_i^{\frac{1}{2}}$, completing the proof.

Acknowledgement. I am thankful to Professor T. W. Anderson for helpful discussions and suggestions.

References

- Chan, N. N., and Wong, C. S. (1981), Existence of an A-optimal model for a regression experiment, J. Math. Anal. and Applications (to appear).
- Chernoff, H. (1953), Locally optimum designs for estimating parameters, Ann. Math. Statist., 24, 586-602.
- Dorogovcev, A. Ja. (1971), Problems of optimal control of a regression experiment, Selected Transl. in Math. Statist. and Probability, 10, 35-41.
- Elfving, G. (1952), Optimum allocation in linear regression theory, Ann. Math. Statist., 23, 255-262.
- Fan, K. (1949), On a theorem of Weyl concerning eigenvalues of linear transformations I, Proc. Nat. Acad. Sci. U.S.A., 35, 652-655.
- Horn, A. (1954), On the eigenvalues of a matrix with prescribed singular values, Proc. Amer. Math. Soc., 5, 4-7.
- Kiefer, J., and Galil, Z. (1980), Optimum weighing designs, Recent Developments in Statistical Inference and Data Analysis, North-Holland, Amsterdam, 183-190.
- Marshall, A. W. and Olkin, I. (1979), Inequalities: Theory of Majorization and Its Applications, Academic Press, New York.
- Mirsky, L. (1958), Matrices with prescribed characteristic roots and diagonal elements, J. London Math. Soc., 33, 14-21.
- Rao, C. R. (1973), Linear Statistical Inference and Its Applications, 2nd ed., Wiley, New York.
- St. John, R. C., and Draper, N. R. (1975), D-optimality for regression designs: A review, Technometrics, 17, 15-23.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.
2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.
3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.
4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.
5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.
6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.
7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.
8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.
9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.
10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.
11. "Population Mixing Models and Clustering Algorithms," Stanley L. Selove, February 1, 1973.
12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.
13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.
14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.
15. "Maximum Likelihood Estimation of Parameter of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.
16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations,"
Steffen L. Lauritzen, May, 1974.
18. "On the Interrelationships among Sufficiency, Total Sufficiency and
Some Related Concepts," Steffen L. Lauritzen, June, 1974.
19. "Statistical Inference for Multiply Truncated Power Series Distributions,"
T. Cacoullis, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models
in the Time and Frequency Domains," T. W. Anderson, June 1975.
21. "Asymptotic Properties of Some Estimators in Moving Average Models,"
Raul Pedro Mentz, September 8, 1975.
22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"
Mituaki Huzii, February 1976.
23. "Estimating Means when Some Observations are Classified by Linear
Discriminant Function," Chien-Pai Han, April 1976.
24. "Panels and Time Series Analysis: Markov Chains and Autoregressive
Processes," T. W. Anderson, July 1976.
25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson,
September 1976.
26. "The Recurrence Classification of Risk and Storage Processes,"
J. Michael Harrison and Sidney I. Resnick, September 1976.
27. "The Generalized Variance of a Stationary Autoregressive Process,"
T. W. Anderson and Raul P. Mentz, October 1976.
28. "Estimation of the Parameters of Finite Location and Scale Mixtures,"
Javad Behboodian, October 1976.
29. "Identification of Parameters by the Distribution of a Maximum
Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.
30. "Discrimination Between Stationary Gaussian Processes, Large Sample
Results," Will Gersch, January 1977.
31. "Principal Components in the Nonnormal Case: The Test for Sphericity,"
Christine M. Waternaux, October 1977.
32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a
Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

TECHNICAL REPORTS (continued)

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.
34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.
35. "Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.
36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.
37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.
38. "Maximum Likelihood Estimation of the Parameters of a Multivariate Normal Distribution," T.W. Anderson and I. Olkin, July 1979.
39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models," Fereydoon Ahrabi, August 1979.
40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike, August, 1979.
41. "Approximating Conditional Moments of the Multivariate Normal Distribution," Joseph G. Deken, December 1979.
42. "Methods and Applications of Time Series Analysis - Part I: Regression, Trends, Smoothing, and Differencing," T.W. Anderson and N.D. Singpurwalla, July 1980.
43. "Cochran's Theorem, Rank Additivity, and Tripotent Matrices." T.W. Anderson and George P.H. Styan, August, 1980.
44. "On Generalizations of Cochran's Theorem and Projection Matrices," Akimichi Takemura, August, 1980.
45. "Existence of Maximum Likelihood Estimators in Autoregressive and Moving Average Models," T.W. Anderson and Raúl P. Mentz, Oct. 1980.
46. "Generalized Correlations in the Singular Case," Ashis Sen Gupta, November 1980.
47. "Updating a Discriminant Function on the Basis of Unclassified Data," G.J. McLachlan and S. Ganesalingam, November 1980.
48. "A New Proof of Admissibility of Tests in the Multivariate Analysis of Variance." T.W. Anderson and Akimichi Takemura, January, 1981.
49. "A-Optimality for Regression Designs," N. N. Chan, January 1981.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 49	2. GOVT ACCESSION NO. AD-A097035	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A-Optimality for Regression Designs		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) N. N. Chan		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE JANUARY 1981
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 9
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Also issued as Technical Report No. 163 - NATIONAL SCIENCE FOUNDATION GRANT MCS 78-07736 - Dept. of Statistics - Stanford University.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) A-optimal design, eigenvalues, inequalities, linear regression		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Use is made of a result of Horn (1954) on the existence of a symmetric matrix with prescribed diagonal elements and eigenvalues. A necessary and sufficient condition is then given for the existence of an A-optimal design for a regression experiment in the Dorogovcev (1971) setting. ↑		

DD FORM 1473

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

