GENERALIZED HADAMARD'S INEQUALITIES AND THEIR APPLICATIONS TO STATISTICS

Nishi, Akihiro
Department of Mathematics, Faculty of Education, Saga University

https://doi.org/10.5109/13129

出版情報：統計数理研究. 18 (3/4), pp.45-49, 1979-03. 統計科学研究会
バージョン：
権利関係：

GENERALIZED HADAMARD'S INEQUALITIES AND THEIR APPLICATIONS TO STATISTICS

By

Akihiro NISHI*

(Received October 20, 1977)

Summary

Algebraic proofs are given to some inequalities involving canonical correlation coefficients which, however, seem quite natural from statistical intuition. Though all of these inequalities can be easily verified from the well-known Courant-Fischer min-max theorem (c.f. Bellman [2], pp. 115-117), our method in this note seems to have some interests notwithstanding.

Let $A=(a_{ij})$ be a p.d. (positive definite real symmetric) $N\times N$ matrix. Then the most famous determinantal inequality is due to Hadamard:

$$|A|\leqq\prod_{i=1}^{N}a_{ii}, \tag{1}$$

where the equality holds iff $a_{ij}=0(i\neq j)$.

Among various extensions of Hadamard's inequality, the extension due to Fischer seems most straightforward. Let us partition A as

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{pmatrix},\text{ where } A_{ii}(1\leqq i\leqq k) \text{ is the square matrix.}$$

Fischer's inequality is

$$|A|\leqq\prod_{i=1}^{k}|A_{ii}|, \tag{2}$$

where the equality holds iff $A_{ij}=0$ $(i\neq j)$.

A number of inequalities in multivariate statistical analysis can be derived by the inequality (2). We shall give here only two examples.

EXAMPLE 1. Let $\Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12} \\ \sigma_{21} & \Sigma_{22}\end{pmatrix}$ $(p\times p$, p.d.) be the covariance matrix of a random vector $X'=(X_1, X_2, \cdots, X_p)$. Then the square of multiple correlation coefficient between X_1 and $(X_2, \cdots, X_p)$ is given by

* Department of Mathematics, Faculty of Education, Saga University, Saga 840, Japan.

$$\rho^2_{1.(2,\dots,p)}=1-\frac{\begin{vmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \Sigma_{22}\end{vmatrix}}{\sigma_{11}|\Sigma_{22}|}. \tag{3}$$

Thus the inequality (2) shows that $0\leqq\rho^2_{1.(2,\dots,p)}<1$ and $\rho^2_{1.(2,\dots,p)}=0$ iff $\sigma_{12}=0$.

EXAMPLE 2. Let the p-dimensional random vector X be distributed according to $N_p(\mu, \Sigma)$. We partition Σ as

$$\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1k}\\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2k}\\ \vdots & \vdots & & \vdots\\ \Sigma_{k1} & \Sigma_{k2} & \cdots & \Sigma_{kk}\end{pmatrix},$$ where $\Sigma_{ii}(1\leqq i\leqq k)$ is the square matrix of order $p_i\times p_i$.

Let S be the matrix of the sum of cross-products, $S=\sum_{\alpha=1}^{n}(x_\alpha-\bar{x})(x_\alpha-\bar{x})'$, which is partitioned correspondingly as $S=(S_{ij})$, where $S_{ii}(1\leqq i\leqq k)$ is the square matrix of order p_i.

Then the likelihood ratio statistic λ for testing the hypothesis H: $\Sigma_{ij}=0$ $(i\neq j)$ (c.f. Anderson [1], Ch. 9) becomes

$$\lambda^{2/n}=\frac{|S|}{\prod_{i=1}^{k}|S_{ii}|}. \tag{4}$$

The inequality (2) implies that $0<\lambda\leqq 1$, and it is interesting to note $\lambda=1$ iff $S_{ij}=0$, $i\neq j$.

There is a refinement of Fischer's inequality (2) due to Faguet (Smirnov [6], p. 70):

$$|A|\leqq\frac{\begin{vmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{vmatrix}\begin{vmatrix}A_{22} & A_{23}\\ A_{32} & A_{33}\end{vmatrix}}{|A_{22}|}, \tag{5}$$

where the equality holds iff $A_{13}-A_{12}A_{22}^{-1}A_{23}=0$.

We shall give here elementary proofs of (2) and (5). For this purpose we use the following well-known facts in linear algebra: For square matrices A and D, we have

$$\begin{vmatrix}A & B\\ C & D\end{vmatrix}=\begin{cases}|A|\,|D-CA^{-1}B|, & \text{when } A \text{ is non-singular}\\ |D|\,|A-BD^{-1}C|, & \text{when } D \text{ is non-singular.}\end{cases} \tag{6}$$

Let $A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}$ be p.d., where A_{11} and A_{22} are square matrices. Then it follows that A^{-1}, $A_{11.2}$ and $A_{22.1}$ are p.d. and

$$\begin{aligned}\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}^{-1}&=\begin{pmatrix}A_{11.2}^{-1}, & -A_{11.2}^{-1}A_{12}A_{22}^{-1}\\ -A_{22}^{-1}A_{21}A_{11.2}^{-1}, & A_{22}^{-1}+A_{22}^{-1}A_{21}A_{11.2}^{-1}A_{12}A_{22}^{-1}\end{pmatrix}\\ &=\begin{pmatrix}A_{11}^{-1}+A_{11}^{-1}A_{12}A_{22.1}^{-1}A_{21}A_{11}^{-1}, & -A_{11}A_{12}A_{22.1}^{-1}\\ -A_{22.1}^{-1}A_{21}A_{11}^{-1}, & A_{22.1}^{-1}\end{pmatrix},\end{aligned} \tag{7}$$

where $A_{ij\cdot k}=A_{ij}-A_{ik}A_{kk}^{-1}A_{kj}$.

Though the following lemma is well-known (for instance in Lemma 5.1. in Kullback [4]), the following concise analytical proof will be of some interest.

LEMMA. *Let A and B are p. d. matrices of order N.*
(i) *If* $x'Ax \geqq x'Bx$ *for any* x, *then* $|A| \geqq |B|$.
(ii) *In addition to the condition in* (i), *if* $x'_0Ax_0 > x'_0Bx_0$ *for some* x_0, *then* $|A| > |B|$.

REMARK. This lemma is still valid for Hermite positive definite matrices because of $\int_{\{\binom{u}{v} | x^*Ax \leqq 1\}} du dv = \frac{V_{2N}}{|A|}$, where $x = u + \sqrt{-1}\, v$ (u and v are real N-vectors) and $V_{2N} = \frac{\pi^N}{\Gamma(N+1)}$.

The proofs of the inequalities (2) and (5) are based on this lemma, and hence they are valid for Hermite positive definite matrices.

PROOF. It is easily seen that the conditions of (i) and (ii) yield the following (i)′ and (ii)′, respectively.
(i)′ $\{x | x'Ax \leqq 1\} \subset \{x | x'Bx \leqq 1\}$.
(ii)′ $\{x | x'Ax \leqq 1\} \subset \{x | x'Bx \leqq 1\}$ and the difference set $\{x | x'Bx \leqq 1\} - \{x | x'Ax \leqq 1\}$ has positive Lebesgue measure. Considering the well-known definite integrals

$$\int_{\{x | x'Ax \leqq 1\}} dx = \frac{V_N}{|A|^{1/2}}, \qquad \int_{\{x | x'Bx \leqq 1\}} dx = \frac{V_N}{|B|^{1/2}},$$

where $V_N = \frac{\pi^{N/2}}{\Gamma(N/2+1)}$ is the volume of N-dimensional unit hypersphere, we have conclusions.

PROOF OF FISCHER'S INEQUALITY (2). It is enough to prove (2) for $k=2$, because we can obtain (2) for any k inductively. The inequality (2) is equivalent to $|A_{22} - A_{21}A_{11}^{-1}A_{12}| \leqq |A_{22}|$, since $\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = |A_{11}| \cdot |A_{22} - A_{21}A_{11}^{-1}A_{12}|$ and $|A_{11}| > 0$. Considering Lemma (i), (ii) and positive definiteness of A_{11}^{-1} and $A_{22} - A_{21}A_{11}^{-1}A_{12}$, we have (2).

PROOF OF FAGUET'S INEQUALITY (5). Since

$$|A| = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \times \left| A_{33} - [A_{31} A_{32}] \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} A_{13} \\ A_{23} \end{pmatrix} \right|$$

and

$$\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} = |A_{22}|\, |A_{33} - A_{32}A_{22}^{-1}A_{23}|,$$

the inequality (5) is equivalent to

$$|A_{33} - A_{32}A_{22}^{-1}A_{23}| \geqq \left| A_{33} - [A_{31} A_{32}] \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} A_{13} \\ A_{23} \end{pmatrix} \right|, \tag{8}$$

where the equality holds iff $A_{13} - A_{12}A_{22}^{-1}A_{23} = 0$. After some calculations we have

$$[A_{31} A_{32}] \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} A_{13} \\ A_{23} \end{pmatrix} = A_{32}A_{22}^{-1}A_{23} + (A_{13} - A_{12}A_{22}^{-1}A_{23})' A_{11\cdot 2}^{-1} (A_{13} - A_{12}A_{22}^{-1}A_{23}).$$

Thus we arrive at (5) by Lemma (i), (ii).

Faguet's inequality has a nice version in canonical correlation analysis. Let $X'=(\underbrace{X_1'}_{p_1}, \underbrace{X_2'}_{p_2}, \underbrace{X_3'}_{p_3})$ be a random vector with mean vector $\mu'=(\mu_1', \mu_2', \mu_3')$ and covariance matrix $\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12} & \Sigma_{13}\\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23}\\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33}\end{pmatrix}$. Let $\lambda_1^{(2)}\geqq\lambda_2^{(2)}\geqq\cdots\geqq\lambda_{p_1}^{(2)}$ be the squares of canonical correlation coefficients between X_1 and X_2. Similarly $\lambda_1^{(23)}\geqq\lambda_2^{(23)}\geqq\cdots\lambda_{p_1}^{(23)}$ are those between X_1 and $\binom{X_2}{X_3}$. Since $\lambda_i^{(2)}(1\leqq i\leqq p_1)$ and $\lambda_j^{(23)}(1\leqq j\leqq p_1)$ are the roots of the characteristic equations

$$|\lambda\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|=0 \quad\text{and}\quad \left|\lambda\Sigma_{11}-[\Sigma_{12}\Sigma_{13}]\begin{pmatrix}\Sigma_{22} & \Sigma_{23}\\ \Sigma_{32} & \Sigma_{33}\end{pmatrix}^{-1}\begin{pmatrix}\Sigma_{21}\\ \Sigma_{31}\end{pmatrix}\right|=0,$$

respectively (c. f. Anderson [1], Ch. 12), it follows that

$$\prod_{i=1}^{p_1}(1-\lambda_i^{(2)})=\frac{|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|}{|\Sigma_{11}|}=\frac{\begin{vmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{vmatrix}}{|\Sigma_{11}||\Sigma_{22}|}$$

and

$$\prod_{i=1}^{p_1}(1-\lambda_i^{(23)})=\frac{\left|\Sigma_{11}-[\Sigma_{12}\Sigma_{13}]\begin{pmatrix}\Sigma_{22} & \Sigma_{23}\\ \Sigma_{32} & \Sigma_{33}\end{pmatrix}^{-1}\begin{pmatrix}\Sigma_{21}\\ \Sigma_{31}\end{pmatrix}\right|}{|\Sigma_{11}|}$$

$$=\frac{|\Sigma|}{|\Sigma_{11}|\begin{vmatrix}\Sigma_{22} & \Sigma_{23}\\ \Sigma_{32} & \Sigma_{33}\end{vmatrix}}.$$

Thus the inequality

$$\prod_{i=1}^{p_1}(1-\lambda_i^{(2)})\geqq\prod_{i=1}^{p_1}(1-\lambda_i^{(23)}) \tag{9}$$

is equivalent to Faguet's inequality (5).

Similarly $\prod_{i=1}^{p_1}\lambda_i^{(2)}=\dfrac{|\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|}{|\Sigma_{11}|}$ and

$$\prod_{i=1}^{p_1}\lambda_i^{(23)}=\frac{\left|[\Sigma_{12}\Sigma_{13}]\begin{pmatrix}\Sigma_{22} & \Sigma_{23}\\ \Sigma_{32} & \Sigma_{33}\end{pmatrix}^{-1}\begin{pmatrix}\Sigma_{21}\\ \Sigma_{31}\end{pmatrix}\right|}{|\Sigma_{11}|}$$

$$=\frac{|\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}+(\Sigma_{31}-\Sigma_{32}\Sigma_{22}^{-1}\Sigma_{21})'\Sigma_{33.2}^{-1}(\Sigma_{31}-\Sigma_{32}\Sigma_{22}^{-1}\Sigma_{21})|}{|\Sigma_{11}|},$$

therefore Lemma (i), (ii) shows that

$$\prod_{i=1}^{p_1}\lambda_i^{(2)}\leqq\prod_{i=1}^{p_1}\lambda_i^{(23)}, \tag{10}$$

where the equality holds iff $\Sigma_{13}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{23}=0$ provided that Σ_{12} is the matrix of full rank.

We note that $\left(\prod_{i=1}^{p_1}(1-\lambda_i^{(2)})\right)^{1/2}$ is called the (vector) alienation coefficient between

X_1 and X_2 (Rozeboom [5]). Since

$$|\lambda\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|=|\Sigma_{11}^{1/2}|\,|\lambda I-\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}|\,|\Sigma_{11}^{1/2}|,$$

we have

$$\sum_{i=1}^{p_1}\lambda_i^{(2)}=\mathrm{tr}\,\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$$

and

$$\begin{aligned}\sum_{i=1}^{p_1}\lambda_i^{(23)}&=\mathrm{tr}\left(\Sigma_{11}^{-1/2}[\Sigma_{12}\Sigma_{13}]\begin{pmatrix}\Sigma_{22} & \Sigma_{23}\\ \Sigma_{32} & \Sigma_{33}\end{pmatrix}^{-1}\begin{pmatrix}\Sigma_{21}\\ \Sigma_{31}\end{pmatrix}\Sigma_{11}^{-1/2}\right)\\ &=\mathrm{tr}\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}+\mathrm{tr}(\Sigma_{11}^{-1/2}\Sigma_{13.2}\Sigma_{33.2}^{-1/2})(\Sigma_{11}^{-1/2}\Sigma_{13.2}\Sigma_{33.2}^{-1/2})'.\end{aligned}$$

Thus we have the inequality

$$\sum_{i=1}^{p_1}\lambda_i^{(2)}\leqq\sum_{i=1}^{p_1}\lambda_i^{(23)}, \tag{11}$$

where the equality holds iff $\Sigma_{13.2}=0$.

REMARK 1. Courant-Fischer min-max theorem presents the far stronger inequality than those of (9), (10) and (11):

$$\lambda_i^{(2)}\leqq\lambda_i^{(23)},\qquad 1\leqq i\leqq p_1. \tag{12}$$

REMARK 2. The partial canonical correlation coefficients (Rao [4]) between X_1 and X_3 w.r.t. X_2 are defined as the canonical correlation coefficients between $Y_1\equiv X_1-\Sigma_{12}\Sigma_{22}^{-1}(X_2-\mu_2)$ and $Y_3\equiv X_3-\Sigma_{32}\Sigma_{22}^{-1}(X_2-\mu_2)$. Since $V(Y_1)=\Sigma_{11.2}$, $V(Y_3)=\Sigma_{33.2}$ and $\mathrm{Cov}(Y_1, Y_3)=\Sigma_{13.2}$, the squares of the partial canonical correlation coefficients are the roots of the characteristic equation $|\lambda\Sigma_{11.2}-\Sigma_{13.2}\Sigma_{33.2}^{-1}\Sigma_{31.2}|=0$. Thus the equality condition $\Sigma_{13.2}=0$ on the inequalities (9), (10), (11) and (12) is equivalent to that all of the partial correlation coefficients vanish. When the joint distribution of X_1, X_2 and X_3 is multivariate normal, the condition $\Sigma_{13.2}=0$ is equivalent to conditional independence of X_1 and X_3 w.r.t. X_2, because the conditional joint distribution of X_1 and X_3 w.r.t. X_2 is multivariate normal with covariance matrix $\begin{pmatrix}\Sigma_{11.2} & \Sigma_{13.2}\\ \Sigma_{31.2} & \Sigma_{33.2}\end{pmatrix}$.

Acknowledgements

The author wishes to thank Professor A. Kudô for his suggestions. The author also wishes to Dr. H. Yanai for his guidance of the paper Rozeboom [5].

References

[1] ANDERSON, T.W., An Introduction to Multivariate Analysis. John Wiley & Sons, Inc., (1958).

[2] BELLMAN, R., An Introduction to Matrix Analysis. Second Edition. TATA McGRAW-HILL, (1974).

[3] KULLBACK, S., Information Theory and Statistics. Dover Edition, (1968).

[4] RAO, B.R., *Partial canonical correlations.* Trabajos Estadist., 20 (1969), 211-219.

[5] ROZEBOOM, W.W., *Linear correlations between sets of variables.* Psychometrika, 30 (1965), No. 1, 57-71.

[6] SMIRNOV, V.I., Linear Algebra and Group Theory. Dover Edition, (1961).