Statistics and Probability: Essays in Honor of C. R. Rao

Edited by

G. KALLIANPUR
P. R. KRISHNAIAH
J. K. GHOSH

Statistics and Probability: Essays in Honor of C. R. Rao

Edited by

G. KALLIANPUR
P. R. KRISHNAIAH
J. K. GHOSH

NORTH-HOLLAND PUBLISHING COMPANY—AMSTERDAM · NEW YORK · OXFORD

ISBN 0 444 86130 0

Published by:
NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM · NEW YORK · OXFORD

Sole distributors for the U.S.A. and Canada:
Elsevier Science Publishing Company, Inc.
52 Vanderbilt Avenue
New York, N.Y. 10017

Library of Congress Cataloging in Publication Data

Main entry under title:

Statistics and probability.

"Publications of C.R. Rao": p.
1. Mathematical statistics--Addresses, essays, lectures. 2. Probabilities--Addresses, essays, lectures. 3. Rao, C. Radhakrishna (Calyampudi Radhakrishna), 1920- . I. Rao, C. Radhakrishna (Calyampudi Radhakrishna), 1920- . II. Kallianpur, G. III. Krishnaiah, Paruchuri R. IV. Ghosh, J. K.
QA276.16.S843 519.5 82-2169
ISBN 0-444-86130-0 AACR2

PRINTED IN THE NETHERLANDS

PREFACE

In presenting this volume of essays in Statistics and Probability we wish to honor a distinguished scientist and one of the most eminent statisticians of our time.

The impact of C. R. Rao's work in the development of the modern theory of Statistics is so well recognized that it does not require comment in this Preface. However, scholars outside India might not be fully aware of the important role he has played in developing Statistics in India and more particularly, in the growth of the Indian Statistical Institute, and in promoting statistical education and training in South East Asian countries. After an initial period of study at Calcutta University in 1943 during which he received a Master's degree, Rao joined the Indian Statistical Institute which had been founded a decade or so earlier by P. C. Mahalanobis. The Institute had already made its mark in Statistical research, especially in such areas as Multivariate Analysis and the Design of Experiments. In the years that followed the Institute has produced, under Rao's direction, an entire generation of students and younger colleagues, many of whom have become eminent Statisticians and Probabilists in their own right.

C. R. Rao's career has not been confined to the field of Statistics in the narrow sense. After the Indian Statistical Institute assumed the functions of a University in the late fifties, its academic activities expanded enormously and many of the country's talented young scientists, economists, geologists, anthropologists, geneticists, and mathematicians were attracted to the Institute by Professor Rao's wide range of scientific interests. These interests are, to a large extent, represented by the essays in the volume, ranging from topics in Pure Mathematics, Probability Theory and Statistical Inference to those with a distinctly applied flavor. Even so, for reasons of space, it has not been possible to include several of the important areas of applications to which Rao himself has made significant contributions. Thus we regret that the volume contains no articles on Anthropology, Econometrics and Psychometrics (to name a few of these areas).

C. R. Rao was born on September 10, 1920 in Hadagali, Karnataka State, India. He has been up to now author or co-author of nine books and over 200 research publications, a list of which is given at the end of this volume. This volume is dedicated to C. R. Rao with admiration and affection on the occasion of his sixtieth birthday in appreciation of his many pioneering contributions to the field of statistics, which have found place in standard books on statistics.

We wish to express our appreciation to S. K. Mitra, K. R. Parthasarathy, B. Ramachandran and B. V. Rao, our colleagues on the Editorial Committee, for spending considerable amount of time and effort in preparation of this volume. We are grateful to R. F. Anderson, G. J. Babu, A. K. Basu, D. Basu, Sanjoy Bose, P. Bhimasankaram, Ratan DasGupta, S. DasGupta, R. H. Farrell, A. Hedayat, Subir

Ghosh, P. Jaganathan, V. M. Joshi, J. Jurečkova, S. Kotz, D. Majumdar, K. V. Mardia, M. G. Nadkarni, J. N. K. Rao, M. M. Rao, Y. R. Sarma, D. N. Shanbhag, Bikas K. Sinha, Bimal K. Sinha, A. C. Tamhane and S. J. Wolfe for reviewing the papers in this volume. Thanks are also due to the North-Holland Publishing Company for their excellent cooperation.

G. Kallianpur
P. R. Krishnaiah
J. K. Ghosh

CONTENTS

*Deceased 1981.

G. Kallianpur, P.R. Krishnaiah, J.K. Ghosh, eds., *Statistics and Probability: Essays in Honor of C.R. Rao*

COCHRAN'S THEOREM, RANK ADDITIVITY AND TRIPOTENT MATRICES

T.W. ANDERSON
Stanford University, Stanford, CA, U.S.A.

George P. H. STYAN
McGill University, Montreal, Que., Canada

Let $A_1,\ldots,A_k$ be symmetric matrices and $A=\Sigma A_i$. A matrix version of Cochran's theorem is that (a) $A_i^2=A_i$, $i=1,\ldots,k$, and (b) $A_iA_j=\mathbf{0}$ $\forall i\neq j$, are necessary and sufficient conditions for (d) $\Sigma\,\mathrm{rank}(A_i)=\mathrm{rank}(A)$ whenever (c) $A=I$. This paper reviews extensions of the theorem and its statistical interpretations in the literature, presents various proofs of the above theorem, and obtains some generalizations. In particular, (c) above is replaced by $A^2=A$ and the condition of symmetry is deleted. The relations with (e) $\mathrm{rank}(A_i^2)=\mathrm{rank}(A_i)$, $i=1,\ldots,k$, are explored. Another theorem covers the case of matrices not necessarily square. A is "tripotent" if $A^3=A$. Then $A_i^3=A_i$, $i=1,\ldots,k$, and (b) are necessary and sufficient conditions for $A^3=A$, (d), and one further condition such as $A_iA=A_i^2$, $i=1,\ldots,k$. Variations and statistical applications are treated. Tripotent is replaced by r-potent ($A^r=A$) for $r>3$.

1. Introduction

Let x be a $p\times 1$ random vector distributed according to a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix I_p. We will denote this by $x\sim \mathrm{N}(\mathbf{0}, I_p)$. Let $q_1,\ldots,q_k$ be quadratic forms in x with ranks $r_1,\ldots,r_k$, respectively, and suppose that $\Sigma q_i=x'x$. Then what has become well-known as Cochran's theorem is Theorem II of Cochran (1934, p. 179): *A necessary and sufficient condition that $q_1,\ldots,q_k$ be independently distributed as χ^2 is that $\Sigma r_i=p$.*

Rao (1973, Section 3b.4) gives this result with $x\sim\mathrm{N}(\mu, I)$ as the Fisher–Cochran theorem. Fisher (1925) showed that if the quadratic form q in x is distributed as χ_h^2, then $x'x-q$ is distributed as χ_{p-h}^2 independently of q, cf. James (1952).

Our purpose in this paper is partly expository; we review various extensions of Cochran's theorem in a bibliographic and historical perspective, with special emphasis on matrix-theoretic analogues. While we present over 30 references, we note that Scarowsky (1973) has a rather complete discussion and bibliography on the distribution of quadratic forms in normal random variables. See also the bibliography by Anderson et al. (1972), where 90 research papers published through 1966 are listed under subject-matter code 2.5 (distribution of quadratic and bilinear forms in normal variables).

Our first section is devoted to a survey of results summarized in Theorems 1.1 and 1.2. The proofs are given in Section 2. In the following section the extensions from idempotent to tripotent matrices are given and proved. All matrices considered in this paper will be real.

To formulate our first matrix-theoretic extension of Cochran's theorem we let $A_1, \ldots, A_k$ be $p \times p$ symmetric matrices and write $A = \Sigma A_i$. Consider the following statements:

(a) $$A_i^2 = A_i, \quad i = 1, \ldots, k,$$

(b) $$A_i A_j = 0 \quad \text{for all } i \neq j,$$

(c) $$A = I,$$

(d) $$\sum \text{rank}(A_i) = \text{rank}(A).$$

Then the matrix-theoretic analogue of Cochran's theorem is:

$$(a), (b), (c) \rightarrow (d), \tag{1.1}$$

$$(c), (d) \rightarrow (a), (b). \tag{1.2}$$

The reason that these two versions of Cochran's theorem are equivalent follows from the following two well-known results:

Lemma 1.1. *Let $x \sim \mathrm{N}(\mu, \Sigma)$, with Σ positive definite, and let A be non-random and symmetric. Then $x'Ax \sim \chi_f^2(\delta^2)$, a non-central χ^2 distribution with f degrees of freedom and non-centrality parameter δ^2, if and only if $A\Sigma A = A$, and then $f = \mathrm{tr}\, A\Sigma = \mathrm{rank}(A)$ and $\delta^2 = \mu' A \mu$.*

We write $\mathrm{tr}\, A$ for the trace of A and note that when $A = A^2$ then $\mathrm{tr}\, A = \mathrm{rank}(A)$; this result holds even when A is not symmetric [cf., e.g. Rao (1973, p. 28)].

When $\Sigma = I$, the condition in Lemma 1.1 reduces to $A^2 = A$, and this was first given by Craig (1943) with $\mu = 0$ and then by Carpenter (1950) with μ possibly non-zero. [Thus (a) is equivalent to $q_i = x'A_i x$ having a χ^2-distribution with number of degrees of freedom equal to $\mathrm{rank}(A_i)$.] Sakamoto (1944, Th. II, p. 5) gave the more general version with Σ positive definite and $\mu = 0$. Cochran (1934, Corollary 1, p. 179) took $x \sim \mathrm{N}(0, I)$ and gave Lemma 1.1 with the condition that all the non-zero eigenvalues of A be equal to 1 instead of the condition $A^2 = A$.

Lemma 1.2. *Let x and A be defined as in Lemma 1.1 and let B be non-random and symmetric. Then $x'Ax$ and $x'Bx$ are independently distributed if and only if $A\Sigma B = 0$.*

When $\Sigma = I$ the condition in Lemma 1.2 reduces to $AB = 0$, and this was first given by Craig (1943) with $\mu = 0$ and then by Carpenter (1950) with μ possibly non-zero. Again Sakamoto (1944, Th. I, p. 5) gave the more general version with Σ positive definite and $\mu = 0$. Their proofs, however, turned out to be incorrect and the first correct proof of Lemma 1.2 (with $\mu = 0$) seems to be by Ogawa (1948; 1949, cf. p. 85). Cochran (1934, Theorem III, p. 181) let $x \sim \mathrm{N}(0, I)$ and gave the condition in Lemma 1.2 as

$$|I - \mathrm{i}sA| \cdot |I - \mathrm{i}tB| = |I - \mathrm{i}sA - \mathrm{i}tB| \tag{1.3}$$

for all real s and t, where $\mathrm{i} = \sqrt{-1}$ and $|\cdot|$ denotes determinant. Ogasawara and

Takahashi (1951, Lemma 1) gave a short proof that (1.3) implies $\boldsymbol{AB}=\mathbf{0}$ when the symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are not necessarily positive semi-definite.

Cochran's theorem was first extended to $\boldsymbol{x}\sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{I}_p)$ by Madow (1940) and then to $\boldsymbol{x}\sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ positive definite, by Ogawa (1946, 1947), who also relaxed the condition (c) to $\boldsymbol{A}^2=\boldsymbol{A}$. Ogasawara and Takahashi (1951) extended Cochran's theorem to $\boldsymbol{x}\sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ positive definite, and to $\boldsymbol{x}\sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ possibly singular. Extensions to $\boldsymbol{x}\sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ possibly singular, have been given by Styan (1970, Theorem 6) and Tan (1977, Theorem 4.2); Ogasawara and Takahashi (1951) extended Lemmas 1.1 and 1.2 to $\boldsymbol{x}\sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ possibly singular.

James (1952) appears to be the first to notice that (1.1) could be extended to

$$\text{(a)},\text{(c)}\rightarrow\text{(b)},\text{(d)},$$
$$\text{(b)},\text{(c)}\rightarrow\text{(a)},\text{(d)},$$

while

$$\text{(a)},\text{(b)}\rightarrow \boldsymbol{A}^2=\boldsymbol{A} \text{ and (d)}$$

follows at once from the definition of the χ^2-distribution.

Chipman and Rao (1964) and Khatri (1968) extended the matrix analogue of Cochran's theorem to square matrices which are not necessarily symmetric:

Theorem 1.1. *Let $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k$ be square matrices, not necessarily symmetric, and let $\boldsymbol{A}=\Sigma \boldsymbol{A}_i$. Consider the following statements*:

(a) $$\boldsymbol{A}_i^2=\boldsymbol{A}_i, \quad i=1,\ldots,k,$$

(b) $$\boldsymbol{A}_i\boldsymbol{A}_j=\mathbf{0} \quad \textit{for all } i\neq j,$$

(c) $$\boldsymbol{A}^2=\boldsymbol{A},$$

(d) $$\sum \operatorname{rank}(\boldsymbol{A}_i)=\operatorname{rank}(\boldsymbol{A}),$$

(e) $$\operatorname{rank}(\boldsymbol{A}_i^2)=\operatorname{rank}(\boldsymbol{A}_i), \quad i=1,\ldots,k.$$

Then

$$\text{(a)},\text{(b)}\rightarrow\text{(c)},\text{(d)},\text{(e)}, \tag{1.4}$$
$$\text{(a)},\text{(c)}\rightarrow\text{(b)},\text{(d)},\text{(e)}, \tag{1.5}$$
$$\text{(b)},\text{(c)},\text{(e)}\rightarrow\text{(a)},\text{(d)}, \tag{1.6}$$
$$\text{(c)},\text{(d)}\rightarrow\text{(a)},\text{(b)},\text{(e)}. \tag{1.7}$$

As Rao and Mitra (1971, p. 112) point out, the extra condition (e) in (1.6) is required (to cover possible asymmetry); for if $k=2$ and

$$\boldsymbol{A}_1=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \qquad \boldsymbol{A}_2=\begin{pmatrix}0 & -1\\ 0 & 0\end{pmatrix},$$

then (b), (c) hold, but (a) and (d) do not. Banerjee and Nagase (1976) replace the extra condition (e) in (1.6) by

(f) $$\operatorname{rank}(A_i) = \operatorname{tr} A_i, \quad i = 1, \ldots, k,$$

and prove that

$$(\mathrm{b}), (\mathrm{c}), (\mathrm{f}) \rightarrow (\mathrm{a}), (\mathrm{d}); \tag{1.8}$$

however, the condition (b) is now no longer required on the left of (1.8) since

$$(\mathrm{c}), (\mathrm{f}) \rightarrow (\mathrm{a}), (\mathrm{b}), (\mathrm{d})$$

follows from

$$\operatorname{rank}(A) = \operatorname{tr} A = \operatorname{tr} \sum A_i = \sum \operatorname{tr} A_i = \sum \operatorname{rank}(A_i)$$

and (1.7).

In Section 2 we present several proofs of Theorem 1.1.

Marsaglia and Styan (1974) extended Theorem 1.1 by considering an arbitrary sum of matrices, which may now be rectangular. The analogue of Theorem 1.1 is

Theorem 1.2. *Let $A_1, \ldots, A_k$ be $p \times q$ matrices, and let $A = \Sigma A_i$. Consider the following statements*:

(a′) $$A_i A^- A_i = A_i, \quad i = 1, \ldots, k,$$

(b′) $$A_i A^- A_j = 0 \quad \textit{for all } i \neq j,$$

(c′) $$\operatorname{rank}(A_i A^- A_i) = \operatorname{rank}(A_i), \quad i = 1, \ldots, k,$$

(d′) $$\sum \operatorname{rank}(A_i) = \operatorname{rank}(A),$$

where A^- is some g-inverse of A. Then

$$(\mathrm{a}') \rightarrow (\mathrm{b}'), (\mathrm{c}'), (\mathrm{d}'), \tag{1.9}$$

$$(\mathrm{b}'), (\mathrm{c}') \rightarrow (\mathrm{a}'), (\mathrm{d}'), \tag{1.10}$$

$$(\mathrm{d}') \rightarrow (\mathrm{a}'), (\mathrm{b}'), (\mathrm{c}'). \tag{1.11}$$

If (a′) *or if* (b′) *and* (c′) *hold for some g-inverse A^-, then* (a′), (b′) *and* (c′) *hold for every g-inverse A^-.*

In Theorem 1.2 we define a g-inverse of A as any solution A^- to $AA^-A = A$, cf. Rao (1962), Rao and Mitra (1971).

The condition (c′) in Theorem 1.2 plays the role of condition (e) in Theorem 1.1.

Marsaglia and Styan (1974, Th. 13) proved (1.11), while Hartwig (1981) has established (1.9). The proposition (1.10), however, appears to be new and is proved in Section 2, where we also present several different proofs of (1.7). In Section 3 we extend Theorem 1.1 to tripotent matrices, following the work by Luther (1965), Tan (1975, 1976) and Khatri (1977). In Section 4 we discuss the applications of these algebraic theorems to statistics.

2. Some proofs

2.1. Proofs of Theorem 1.1

To prove (1.7) in Theorem 1.1 we begin by reducing condition (c) to a sum being $\boldsymbol{I}$, as in the earlier version of Cochran's theorem; then (1.7) reduces to (1.2). We may do this since if $\boldsymbol{A}$ is $p\times p$, not necessarily symmetric, then, as we shall show,

$$\boldsymbol{A}^2=\boldsymbol{A}\leftrightarrow \operatorname{rank}(\boldsymbol{I}-\boldsymbol{A})=p-\operatorname{rank}(\boldsymbol{A}). \tag{2.1}$$

(Note Fisher's 1925 result goes both ways, cf. Section 1, paragraph 2.) To prove (2.1) let $\boldsymbol{A}^2=\boldsymbol{A}$; then $(\boldsymbol{I}-\boldsymbol{A})^2=\boldsymbol{I}-\boldsymbol{A}$ and so

$$\operatorname{rank}(\boldsymbol{I}-\boldsymbol{A})=\operatorname{tr}(\boldsymbol{I}-\boldsymbol{A})=p-\operatorname{tr}\boldsymbol{A}=p-\operatorname{rank}(\boldsymbol{A}).$$

To go the other way we follow Krafft (1978, pp. 407–408) by noting that

$$\mathcal{N}(\boldsymbol{A})\subset\mathcal{C}(\boldsymbol{I}-\boldsymbol{A}), \tag{2.2}$$

where $\mathcal{N}(\boldsymbol{A})=\{\boldsymbol{x}\colon \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}$ is the null space of $\boldsymbol{A}$ and $\mathcal{C}(\boldsymbol{I}-\boldsymbol{A})=\{(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}\}$ is the column space of $\boldsymbol{I}-\boldsymbol{A}$. [If $\boldsymbol{x}\in\mathcal{N}(\boldsymbol{A})$, then $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ and $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{x}\in\mathcal{C}(\boldsymbol{I}-\boldsymbol{A})$.] If $\operatorname{rank}(\boldsymbol{I}-\boldsymbol{A})=p-\operatorname{rank}(\boldsymbol{A})$, then equality must hold in (2.2) and so $\boldsymbol{A}^2=\boldsymbol{A}$.

We are grateful to a reviewer for suggesting an alternative proof of (2.1). Consider the equations

$$\begin{pmatrix}\boldsymbol{I} & \boldsymbol{A}\\ \boldsymbol{A} & \boldsymbol{A}\end{pmatrix}=\begin{pmatrix}\boldsymbol{I} & \boldsymbol{I}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}-\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{I} & \boldsymbol{I}\end{pmatrix}$$
$$=\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{A} & \boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}-\boldsymbol{A}^2\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{A}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}.$$

Then

$$\operatorname{rank}(\boldsymbol{A}-\boldsymbol{A}^2)+p=\operatorname{rank}(\boldsymbol{I}-\boldsymbol{A})+\operatorname{rank}(\boldsymbol{A}) \tag{2.2a}$$

from which (2.1) follows at once.

The equation (2.2a) represents "rank additivity on the Schur complement", cf. Marsaglia and Styan (1974, p. 291). Let

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{E} & \boldsymbol{F}\\ \boldsymbol{G} & \boldsymbol{H}\end{pmatrix}$$

and suppose that

$$\operatorname{rank}(\boldsymbol{E},\boldsymbol{F})=\operatorname{rank}\begin{pmatrix}\boldsymbol{E}\\ \boldsymbol{G}\end{pmatrix}=\operatorname{rank}(\boldsymbol{E}),$$

$$\operatorname{rank}(\boldsymbol{G},\boldsymbol{H})=\operatorname{rank}\begin{pmatrix}\boldsymbol{F}\\ \boldsymbol{H}\end{pmatrix}=\operatorname{rank}(\boldsymbol{H}).$$

Then

$$(\boldsymbol{M}/\boldsymbol{E})=\boldsymbol{H}-\boldsymbol{G}\boldsymbol{E}^-\boldsymbol{F}$$

is the Schur complement of $\boldsymbol{E}$ in $\boldsymbol{M}$ and this is invariant over choices of $\boldsymbol{E}^-$. Similarly

$$(\boldsymbol{M}/\boldsymbol{H})=\boldsymbol{E}-\boldsymbol{F}\boldsymbol{H}^-\boldsymbol{G}$$

is the Schur complement of $\boldsymbol{H}$ in $\boldsymbol{M}$ and this is invariant over choices of $\boldsymbol{H}^{-}$. It follows that

$$\begin{aligned}\operatorname{rank}(\boldsymbol{M}) &= \operatorname{rank}(\boldsymbol{M}/\boldsymbol{E}) + \operatorname{rank}(\boldsymbol{E}) \\ &= \operatorname{rank}(\boldsymbol{M}/\boldsymbol{H}) + \operatorname{rank}(\boldsymbol{H}).\end{aligned} \tag{2.2b}$$

Setting

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{A} \end{pmatrix}$$

yields (2.2a) directly.

We now write $\boldsymbol{A}_0 = \boldsymbol{I} - \boldsymbol{A}$, and in view of (2.1) we replace (c) by $\Sigma_{i=0}^{k} \boldsymbol{A}_i = \boldsymbol{I}$, and (d) by $\Sigma_{i=0}^{k} \operatorname{rank}(\boldsymbol{A}_i) = p$.

The proof of (1.7) by Cochran (1934, p. 180), cf. also Anderson (1958, p. 164) and Rao (1973, Section 3b.4), requires that $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ be symmetric. In this event we may write

$$\boldsymbol{A}_i = \boldsymbol{P}_i \boldsymbol{P}_i' - \boldsymbol{Q}_i \boldsymbol{Q}_i', \quad i = 0, 1, \ldots, k, \tag{2.3}$$

where $\boldsymbol{P}_i$ is $p \times p_i$, $\boldsymbol{Q}_i$ is $p \times q_i$, and $\boldsymbol{A}_i$ has p_i positive and q_i negative eigenvalues, cf., e.g. Anderson (1958, p. 346). In (2.3) we assume that $\boldsymbol{P}_i$ has rank p_i, $\boldsymbol{Q}_i$ has rank q_i and $p_i + q_i = r_i$, the rank of $\boldsymbol{A}_i$. Hence

$$\begin{aligned}\boldsymbol{I}_p &= \sum_{i=0}^{k} \boldsymbol{A}_i = \sum_{i=0}^{k} \boldsymbol{P}_i \boldsymbol{P}_i' - \sum_{i=0}^{k} \boldsymbol{Q}_i \boldsymbol{Q}_i' \\ &= (\boldsymbol{P}_0, \ldots, \boldsymbol{P}_k, \boldsymbol{Q}_0, \ldots, \boldsymbol{Q}_k) \begin{pmatrix} \boldsymbol{I}_{p-q} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_q \end{pmatrix} \begin{bmatrix} \boldsymbol{P}_0' \\ \vdots \\ \boldsymbol{P}_k' \\ \boldsymbol{Q}_0' \\ \vdots \\ \boldsymbol{Q}_k' \end{bmatrix} \\ &= \boldsymbol{PJP}',\end{aligned} \tag{2.4}$$

say, where $q = \Sigma_{i=0}^{k} q_i$, since from (d) now $p = \Sigma_{i=0}^{k} r_i = \Sigma_{i=0}^{k} (p_i + q_i) = (\Sigma_{i=0}^{k} p_i) + q$. But (2.4) is positive definite and $\boldsymbol{P}$ is non-singular; hence $q = 0$ and $\boldsymbol{J} = \boldsymbol{I}_p$. Thus $q_0 = \cdots = q_k = 0$ and (2.4) reduces to

$$\boldsymbol{I}_p = (\boldsymbol{P}_0, \ldots, \boldsymbol{P}_k) \begin{pmatrix} \boldsymbol{P}_0' \\ \vdots \\ \boldsymbol{P}_k' \end{pmatrix} = \boldsymbol{PP}',$$

and so $\boldsymbol{P} = (\boldsymbol{P}_0, \ldots, \boldsymbol{P}_k)$ is an orthogonal matrix. Hence $\boldsymbol{A}_i^2 = \boldsymbol{P}_i \boldsymbol{P}_i' \boldsymbol{P}_i \boldsymbol{P}_i' = \boldsymbol{P}_i \boldsymbol{P}_i' = \boldsymbol{A}_i$ since $\boldsymbol{P}_i' \boldsymbol{P}_i = \boldsymbol{I}_{r_i}$, and $\boldsymbol{A}_i \boldsymbol{A}_j = \boldsymbol{P}_i \boldsymbol{P}_i' \boldsymbol{P}_j \boldsymbol{P}_j' = \boldsymbol{0}$ for all $i \neq j$ since then $\boldsymbol{P}_i' \boldsymbol{P}_j = \boldsymbol{0}$.

We now present four other proofs of (1.7); these four proofs do not require that $\boldsymbol{A}_0, \ldots, \boldsymbol{A}_k$ be symmetric.

Following Craig (1938, p. 49), cf. also Aitken (1950, Section 6) and Rao and Mitra (1971, pp. 111–112), we may prove (1.7) using a rank-subadditivity argument. From (2.1) with $\boldsymbol{A}_k$ replacing $\boldsymbol{A}$ we have

$$\begin{aligned} p-\operatorname{rank}(\boldsymbol{A}_k) &\leqslant \operatorname{rank}(\boldsymbol{I}_p-\boldsymbol{A}_k) \\ &= \operatorname{rank}(\boldsymbol{A}_0+\cdots+\boldsymbol{A}_{k-1}) \\ &\leqslant \operatorname{rank}(\boldsymbol{A}_0)+\cdots+\operatorname{rank}(\boldsymbol{A}_{k-1}) \\ &= p-\operatorname{rank}(\boldsymbol{A}_k) \end{aligned} \tag{2.5}$$

when (d) holds. This inequality string, therefore, collapses, and $\operatorname{rank}(\boldsymbol{I}_p-\boldsymbol{A}_k)=p-\operatorname{rank}(\boldsymbol{A}_k)$, which implies $\boldsymbol{A}_k^2=\boldsymbol{A}_k$ by (2.1); repeating the argument with $\boldsymbol{A}_{k-1}, \boldsymbol{A}_{k-2},\ldots$ yields (a). To see that this implies (b) we follow Rao and Mitra (1971, p. 112) by noting that the argument used in (2.5) implies that

$$(\boldsymbol{A}_i+\boldsymbol{A}_j)^2=\boldsymbol{A}_i+\boldsymbol{A}_j$$

and so

$$\boldsymbol{A}_i\boldsymbol{A}_j+\boldsymbol{A}_j\boldsymbol{A}_i=\boldsymbol{0}.$$

Premultiplying by $\boldsymbol{A}_i$ yields

$$\boldsymbol{A}_i\boldsymbol{A}_j+\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_i=\boldsymbol{0}, \tag{2.6}$$

while postmultiplying (2.6) by $\boldsymbol{A}_i$ yields

$$2\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_i=\boldsymbol{0}=\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_i.$$

Substituting into (2.6) yields (b).

Our next proof of (1.7) follows Chipman and Rao (1964, p. 4), cf. also Styan (1970, p. 571). We write

$$\boldsymbol{A}_i=\boldsymbol{B}_i\boldsymbol{C}_i',$$

where $\boldsymbol{B}_i$ and $\boldsymbol{C}_i$ are $p\times r_i$ of rank r_i. Then

$$\begin{aligned} \boldsymbol{I}_p=\sum \boldsymbol{A}_i &= \sum \boldsymbol{B}_i\boldsymbol{C}_i' \\ &= (\boldsymbol{B}_0,\ldots,\boldsymbol{B}_k)\begin{pmatrix} \boldsymbol{C}_0' \\ \\ \boldsymbol{C}_k' \end{pmatrix} \\ &= \boldsymbol{BC}', \end{aligned}$$

say. By (d) $\boldsymbol{B}$ and $\boldsymbol{C}$ are both non-singular and so $\boldsymbol{C}'=\boldsymbol{B}^{-1}$ and

$$\boldsymbol{C}'\boldsymbol{B}=\boldsymbol{I}_p=\begin{pmatrix} \boldsymbol{C}_0'\boldsymbol{B}_0,\ldots,\boldsymbol{C}_0'\boldsymbol{B}_k \\ \vdots \qquad\qquad \vdots \\ \boldsymbol{C}_k'\boldsymbol{B}_0,\ldots,\boldsymbol{C}_k'\boldsymbol{B}_k \end{pmatrix},$$

which implies that

$$A_i^2 = B_i C_i' B_i C_i' = B_i C_i' = A_i,$$
$$A_i A_j = B_i C_i' B_j C_j' = 0 \quad \text{for all } i \neq j.$$

Hence (1.7) is established.

Our fourth proof of (1.7) follows Loynes (1966), cf. also Searle (1971, p. 63). A rank-subadditivity argument is used similar to that used in (2.5):

$$\begin{aligned} p - \operatorname{rank}(A_k) &\leq \operatorname{rank}(I_p - A_k) \\ &\leq \operatorname{rank}(A_0, A_1, \ldots, A_{k-1}, I_p - A_k) \\ &= \operatorname{rank}(A_0, \ldots, A_{k-1}, I - A_0 - \cdots - A_{k-1} - A_k) \\ &= \operatorname{rank}(A_0, \ldots, A_{k-1}) \\ &\leq \operatorname{rank}(A_0) + \cdots + \operatorname{rank}(A_{k-1}) \\ &= p - \operatorname{rank}(A_k). \end{aligned}$$

Our fifth and final proof of (1.7) follows a suggestion made by a reviewer. Let the $pk \times p$ matrix

$$K = (I, I, \ldots, I)'$$

and let

$$D = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}.$$

Then

$$A = \sum_{i=1}^{k} A_i = K'DK$$

and

$$(c) \leftrightarrow K'DK = K'DKK'DK, \tag{2.6a}$$
$$\begin{aligned} (d) \leftrightarrow \operatorname{rank}(D) &= \operatorname{rank}(K'DK) \\ &= \operatorname{rank}(K'D) = \operatorname{rank}(DK) \end{aligned}$$

using Sylvester's Law of Nullity. Applying the rank cancellation rules Lemmas 2.1 and 2.2 below to (2.6a) yields

$$D = DKK'D$$

or

$$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix} (A_1, \ldots, A_k),$$

which is (a) and (b). See also Marsaglia and Styan (1974, p. 285) and Srivastava and Khatri (1979, p. 14).

The rest of Theorem 1.1 is easily proved. To prove (1.4) we see that (a), (b) →

$$A^2=\left(\sum A_i\right)^2=\sum A_i^2+\sum_{i\neq j} A_iA_j=\sum A_i=A,$$

while

$$\sum \operatorname{rank}(A_i)=\sum \operatorname{tr} A_i=\operatorname{tr}\sum A_i=\operatorname{tr} A=\operatorname{rank}(A), \tag{2.7}$$

and so (1.4) is established.

To prove (1.5) we see that (a), (c) → (d) from (2.7) and so (1.5) follows from (1.7).

To prove (1.6) we see that (b), (c) →

$$A^2=\left(\sum A_i\right)^2=\sum A_i^2+\sum_{i\neq j} A_iA_j=\sum A_i^2=\sum A_i=A;$$

multiplying through by A_i yields

$$A_i^3=A_i^2 \tag{2.8}$$

using (b). To see that (2.8) → (a) we use the rank cancellation rule Lemma 2.1 below, cf. (2.13) in Marsaglia and Styan (1974, p. 271); this rule will also be useful later on.

Lemma 2.1 (Right-hand rank cancellation rule). *If*

$$LAX=MAX \text{ and } \operatorname{rank}(AX)=\operatorname{rank}(A) \tag{2.9}$$

for some conformable matrices A, L, M *and* X, *then*

$$LA=MA. \tag{2.10}$$

Thus (2.8) → (a) by replacing L, A and X in (2.9) by A_i and M by I. Then (2.9) becomes (2.8) and (e), while (2.10) becomes (a). [We note that the two matrices A_1 and A_2 displayed right after Theorem 1.1 satisfy (2.8) but not (e).] Then (d) follows from (1.4) or (1.5).

Proof of Lemma 2.1. Let $A=BC'$, where B and C have r columns and $r=\operatorname{rank}(A)=\operatorname{rank}(B)=\operatorname{rank}(C)$. Then $\operatorname{rank}(AX)=\operatorname{rank}(BC'X)=\operatorname{rank}(C'X)=\operatorname{rank}(A)$, and so $C'X$ has full row rank. Thus $LAX=MAX$ equals $LBC'X=MBC'X \rightarrow LB=MB \rightarrow LBC'=MBC'$, which is (2.10).

Transposing the matrices in Lemma 2.1 yields:

Lemma 2.2 (Left-hand rank cancellation rule). *If*

$$LAX=LAY \text{ and } \operatorname{rank}(LA)=\operatorname{rank}(A)$$

for some conformable matrices A, L, X *and* Y, *then*

$$AX=AY.$$