

*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—1982

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

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. Kalianpur, G. III. Krishnaiiah, 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. Jurečková, 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

Preface	vii
<i>Anderson, T. W. and Styán, George P. H.:</i> Cochran's Theorem, Rank Additivity and Tripotent Matrices	1
<i>Bagchi, Somesh Chandra:</i> Invariant Subspaces of Vector Valued Function Spaces on Bohr Groups and Multivariate Prediction Theory	25
<i>Bahadur, R. R.:</i> A Note on the Effective Variance of Randomly Stopped Means	39
<i>Balakrishnan, A. V.:</i> Some Estimation Problems for Random Fields	45
<i>Barnard, G. A.:</i> Conditionality versus Similarity in the Analysis of 2×2 Tables	59
<i>Barndorff-Nielsen, O.:</i> Hyperbolic Likelihood	67
<i>Bartlett, M. S.:</i> The Development of Population Models	77
<i>Basu, D. and Tiwari, R. C.:</i> A Note on the Dirichlet Process	89
<i>Bhat, B. R.:</i> Strong Consistency of Least Squares Estimates and Conditional Consistency	105
<i>Bickel, Peter J. and Yahav, Joseph A.:</i> Asymptotic Theory of Selection Procedures and Optimality of Gupta's Rules	109
<i>Bose, R. C. and Iyer, H. K.:</i> \mathcal{S} -Invariant Designs and \mathcal{S} -Balanced Arrays with Applications	125
<i>Brown, Lawrence D.:</i> A Proof of the Central Limit Theorem Motivated by the Cramér-Rao Inequality	141
<i>Bunke, O. and Möhner, M.:</i> Minimax Estimators in Linear Models Under Normalized Quadratic Loss Functions	149
<i>Cacoullos, T. and Papageorgiou, H.:</i> Bivariate Negative Binomial-Poisson and Negative Binomial-Bernoulli Models with an Application to Accident Data	155
<i>Chakravarti, I. M. and Burton, Catherine T.:</i> Symmetries (Group of Automorphisms) of Desarguesian Finite Projective and Affine Planes and Their Role in Statistical Model Construction	169
<i>Chandra, Tapas K. and Ghosh, J. K.:</i> Deficiency for Multiparameter Testing Problems	179
<i>Chatterji, S. D.:</i> A Remark on the Cramér-Rao Inequality	193
<i>Cox, D. R.:</i> Randomization and Concomitant Variables in the Design of Experiments	197
<i>Cramér, Harald:</i> A Note on Multiplicity Theory	203

<i>Das Gupta, Somesh</i> : Classification into Two Multivariate Normal Populations with Known Means	209
<i>Dempster, A. P.</i> : Some Formulas Useful for Covariance Estimation with Gaussian Linear Component Models	213
<i>Dharmadhikari, S. W.</i> : Connectedness and Zero Variance in Sampling Designs	221
<i>Doob, J. L.</i> : A Potential Theoretic Approach to Martingale Theory	227
<i>Drygas, Hilmar</i> : Asymptotic Confidence-Intervals for the Variance in the Linear Regression Model	233
<i>Dugué, Daniel</i> : On the Way Towards a Non-parametric Multidimensional Test of Normality	241
<i>Fang, C. and Krishnaiah, P. R.</i> : Asymptotic Joint Distributions of Functions of the Elements of Sample Covariance Matrix	249
<i>Gani, J.</i> : The Early Use of Stochastic Methods: An Historical Note on McKendrick's Pioneering Papers	263
<i>Gnanadesikan, R., Kettenring, J. R. and Landwehr, J. M.</i> : Projection Plots for Displaying Clusters	269
<i>Godambe, V. P.</i> : Likelihood Principle and Randomisation	281
<i>Gupta, Shanti S. and Miescke, Klaus-J.</i> : On the Least Favorable Configurations in Certain Two-Stage Selection Procedures	295
<i>Hannan, E. J.</i> : Fitting Multivariate ARMA Models	307
<i>Heyde, C. C.</i> : Estimation in the Presence of a Threshold Theorem: Principles and Their Illustration for the Traffic Intensity	317
<i>Hodges, J. L., Jr. and Lehmann, E. L.</i> : Minimax Estimation in Simple Random Sampling	325
<i>James, A. T.</i> : Analyses of Variance Determined by Symmetry and Combinatorial Properties of Zonal Polynomials	329
<i>Johnson, Wesley and Geisser, Seymour</i> : Assessing the Predictive Influence of Observations	343
<i>Kagan, A. M., Melamed, I. A. and Zinger, A. A.</i> : A Class of Estimators of a Location Parameter in Presence of a Nuisance Scale Parameter	359
<i>Kallianpur, G.</i> : A Generalized Cameron-Feynman Integral	369
<i>Karlin, Samuel</i> : Some Results on Optimal Partitioning of Variance and Monotonicity with Truncation Level	375
<i>Kawata, Tatsuo</i> : Almost Periodic Weakly Stationary Processes	383
<i>Kemphorne, Oscar</i> : Classificatory Data Structures and Associated Linear Models	397
<i>Khatri, C. G.</i> : A Theorem on Quadratic Forms for Normal Variables	411
<i>Kiefer, J.*</i> : Optimum Rates for Non-parametric Density and Regression Estimates, Under Order Restrictions	419
<i>Kojima, Y., Morimoto, H. and Takeuchi, K.</i> : Two "Best" Unbiased Estimators of Normal Integral Mean	429

<i>Laha, R. G. and Rohatgi, V. K.</i> : Stable and Semistable Probability Measures on a Hilbert Space	443
<i>Le Cam, L.</i> : Limit Theorems for Empirical Measures and Poissonization	455
<i>Lukacs, Eugene</i> : On Some Recent Advances in the Theory of Univariate Characteristic Functions and on the Developments Which Led to Them	465
<i>Mahfoud, M. and Patil, G. P.</i> : On Weighted Distributions	479
<i>Mathew, Joseph and Nadkarni, M. G.</i> : Some Results on Cocycles and Spectra of Ergodic Flows	493
<i>Mitra, Sujit Kumar</i> : Properties of the Fundamental Bordered Matrix Used in Linear Estimation	505
<i>Moran, P. A. P.</i> : The Surface Area of an Ellipsoid	511
<i>Olkin, Ingram and Saunders, Sam C.</i> : A Model for Aerial Surveillance of Moving Objects When Errors of Observation are Multivariate Normal	519
<i>Pakshirajan, R. P. and Vasudeva, R.</i> : An Extension of Strassen's Functional Law of the Iterated Logarithm for Variables Without Variance	531
<i>Papanicolaou, George C. and Varadhan, S. R. S.</i> : Diffusions with Random Coefficients	547
<i>Parthasarathy, K. R. and Sinha, Kalyan B.</i> : A Random Trotter-Kato Product Formula	553
<i>Pathak, P. K.</i> : Asymptotic Normality of the Average of Distinct Units in Simple Random Sampling with Replacement	567
<i>Prakasa Rao, B. L. S.</i> : Maximum Probability Estimation for Diffusion Processes	575
<i>Puri, Madan L. and Ralescu, Stefan S.</i> : On the Degeneration of the Variance in the Asymptotic Normality of Signed Rank Statistics	591
<i>Ramachandran, B.</i> : An Integral Equation in Probability Theory and Its Implications	609
<i>Rao, M. M.</i> : Application and Extension of Cramér's Theorem on Distributions of Ratios	617
<i>Rice, J. and Rosenblatt, M.</i> : Boundary Effects on the Behavior of Smoothing Splines	635
<i>Rozanov, Yu. A.</i> : On Multidirectional Markov Processes	645
<i>Sen, Pranab Kumar</i> : Asymptotic Theory of Some Time-Sequential Tests Based on Progressively Censored Quantile Processes	649
<i>Shimizu, Ryoichi</i> : On the Stability of Characterizations of the Normal Distribution	661
<i>Srivastava, J. N. and Ariyaratna, W. M.</i> : Inversion of Information Matrices of Balanced 3^m Factorial Designs of Resolution V , and Optimal Designs	671
<i>Varadarajan, V. S.</i> : Some Remarks on Spherical Functions on Real Semisimple Lie Groups	689
<i>Watson, G. S.</i> : The Estimation of Palaeomagnetic Pole Positions	703
Publications of C. R. Rao	713

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, \dots, A_k be symmetric matrices and $A = \sum A_i$. A matrix version of Cochran's theorem is that (a) $A_i^2 = A_i$, $i = 1, \dots, k$, and (b) $A_i A_j = 0 \forall i \neq j$, are necessary and sufficient conditions for (d) $\sum \text{rank}(A_i) = \text{rank}(A)$ 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) $\text{rank}(A_i^2) = \text{rank}(A_i)$, $i = 1, \dots, k$, are explored. Another theorem covers the case of matrices not necessarily square. A is "tripotent" if $A^3 = A$. Then $A_i^3 = A_i$, $i = 1, \dots, k$, and (b) are necessary and sufficient conditions for $A^3 = A$, (d), and one further condition such as $A_i A = A_i^2$, $i = 1, \dots, k$. 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 0 and covariance matrix I_p . We will denote this by $x \sim N(0, I_p)$. Let q_1, \dots, q_k be quadratic forms in x with ranks r_1, \dots, r_k , respectively, and suppose that $\sum q_i = x'x$. 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, \dots, q_k be independently distributed as χ^2 is that $\sum r_i = p$.*

Rao (1973, Section 3b.4) gives this result with $x \sim N(\mu, I)$ 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, \dots, A_k be $p \times p$ symmetric matrices and write $A = \Sigma A_i$. Consider the following statements:

- (a) $A_i^2 = A_i, \quad i = 1, \dots, k,$
- (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:

$$(a), (b), (c) \rightarrow (d), \quad (1.1)$$

$$(c), (d) \rightarrow (a), (b). \quad (1.2)$$

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

Lemma 1.1. *Let $x \sim N(\mu, \Sigma)$, 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 $f = \text{tr } A\Sigma = \text{rank}(A)$ and $\delta^2 = \mu' A \mu$.*

We write $\text{tr } A$ for the trace of A and note that when $A = A^2$ then $\text{tr } A = \text{rank}(A)$; 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 $\text{rank}(A_i)$.] 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 $x \sim N(0, I)$ 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 $x \sim N(0, I)$ and gave the condition in Lemma 1.2 as

$$|I - isA| \cdot |I - itB| = |I - isA - itB| \quad (1.3)$$

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

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

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

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

$$(a), (c) \rightarrow (b), (d),$$

$$(b), (c) \rightarrow (a), (d),$$

while

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

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 A_1, \dots, A_k be square matrices, not necessarily symmetric, and let $A = \sum A_i$. Consider the following statements:*

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

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

$$(c) \quad A^2 = A,$$

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

$$(e) \quad \text{rank}(A_i^2) = \text{rank}(A_i), \quad i=1, \dots, k.$$

Then

$$(a), (b) \rightarrow (c), (d), (e), \tag{1.4}$$

$$(a), (c) \rightarrow (b), (d), (e), \tag{1.5}$$

$$(b), (c), (e) \rightarrow (a), (d), \tag{1.6}$$

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

As Rao and Mitra (1971, p. 112) point out, the extra condition (e) in (1.4) is required (to cover possible asymmetry); for if $k=2$ and

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix},$$

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

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

and prove that

$$(b), (c), (f) \rightarrow (a), (d); \quad (1.8)$$

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

$$(c), (f) \rightarrow (a), (b), (d)$$

follows from

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

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, \dots, A_k be $p \times q$ matrices, and let $A = \sum A_i$. Consider the following statements:*

- (a') $A_i A^- A_i = A_i, \quad i = 1, \dots, k,$
- (b') $A_i A^- A_j = 0 \quad \text{for all } i \neq j,$
- (c') $\text{rank}(A_i A^- A_i) = \text{rank}(A_i), \quad i = 1, \dots, k,$
- (d') $\sum \text{rank}(A_i) = \text{rank}(A),$

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

$$(a') \rightarrow (b'), (c'), (d'), \quad (1.9)$$

$$(b'), (c') \rightarrow (a'), (d'), \quad (1.10)$$

$$(d') \rightarrow (a'), (b'), (c'). \quad (1.11)$$

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 I , as in the earlier version of Cochran's theorem; then (1.7) reduces to (1.2). We may do this since if A is $p \times p$, not necessarily symmetric, then, as we shall show,

$$A^2 = A \Leftrightarrow \text{rank}(I - A) = p - \text{rank}(A). \quad (2.1)$$

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

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

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

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

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

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

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

Then

$$\text{rank}(A - A^2) + p = \text{rank}(I - A) + \text{rank}(A) \quad (2.2a)$$

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

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

and suppose that

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

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

Then

$$(M/E) = H - GE^{-1}F$$

is the Schur complement of E in M and this is invariant over choices of E^{-1} . Similarly

$$(M/H) = E - FH^{-1}G$$

is the Schur complement of H in M and this is invariant over choices of H^- . It follows that

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

Setting

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

yields (2.2a) directly.

We now write $A_0 = I - A$, and in view of (2.1) we replace (c) by $\sum_{i=0}^k A_i = I$, and (d) by $\sum_{i=0}^k \text{rank}(A_i) = p$.

The proof of (1.7) by Cochran (1934, p. 180), cf. also Anderson (1958, p. 164) and Rao (1973, Section 3b.4), requires that A_1, \dots, A_k be symmetric. In this event we may write

$$A_i = P_i P_i' - Q_i Q_i', \quad i=0, 1, \dots, k, \quad (2.3)$$

where P_i is $p \times p_i$, Q_i is $p \times q_i$, and A_i has p_i positive and q_i negative eigenvalues, cf., e.g. Anderson (1958, p. 346). In (2.3) we assume that P_i has rank p_i , Q_i has rank q_i and $p_i + q_i = r_i$, the rank of A_i . Hence

$$\begin{aligned}I_p &= \sum_{i=0}^k A_i = \sum_{i=0}^k P_i P_i' - \sum_{i=0}^k Q_i Q_i' \\ &= (P_0, \dots, P_k, Q_0, \dots, Q_k) \begin{pmatrix} I_{p-q} & 0 \\ 0 & -I_q \end{pmatrix} \begin{pmatrix} P_0' \\ \vdots \\ P_k' \\ Q_0' \\ \vdots \\ Q_k' \end{pmatrix} \\ &= PJP',\end{aligned}\quad (2.4)$$

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

$$I_p = (P_0, \dots, P_k) \begin{pmatrix} P_0' \\ \vdots \\ P_k' \end{pmatrix} = PP',$$

and so $P = (P_0, \dots, P_k)$ is an orthogonal matrix. Hence $A_i^2 = P_i P_i' P_i P_i' = P_i P_i' = A_i$ since $P_i' P_i = I_{r_i}$, and $A_i A_j = P_i P_i' P_j P_j' = 0$ for all $i \neq j$ since then $P_i' P_j = 0$.

We now present four other proofs of (1.7); these four proofs do not require that A_0, \dots, 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 A_k replacing A we have

$$\begin{aligned} p - \text{rank}(A_k) &\leq \text{rank}(I_p - A_k) \\ &= \text{rank}(A_0 + \cdots + A_{k-1}) \\ &\leq \text{rank}(A_0) + \cdots + \text{rank}(A_{k-1}) \\ &= p - \text{rank}(A_k) \end{aligned} \quad (2.5)$$

when (d) holds. This inequality string, therefore, collapses, and $\text{rank}(I_p - A_k) = p - \text{rank}(A_k)$, which implies $A_k^2 = A_k$ by (2.1); repeating the argument with A_{k-1}, A_{k-2}, \dots 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

$$(A_i + A_j)^2 = A_i + A_j$$

and so

$$A_i A_j + A_j A_i = 0.$$

Premultiplying by A_i yields

$$A_i A_j + A_i A_j A_i = 0, \quad (2.6)$$

while postmultiplying (2.6) by A_i yields

$$2 A_i A_j A_i = 0 = A_i A_j A_i.$$

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

$$A_i = B_i C_i',$$

where B_i and C_i are $p \times r_i$ of rank r_i . Then

$$\begin{aligned} I_p &= \sum A_i = \sum B_i C_i' \\ &= (B_0, \dots, B_k) \begin{pmatrix} C_0' \\ \vdots \\ C_k' \end{pmatrix} \\ &= B C', \end{aligned}$$

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

$$C' B = I_p = \begin{pmatrix} C_0' B_0, \dots, C_0' B_k \\ \vdots \\ C_k' B_0, \dots, C_k' B_k \end{pmatrix},$$

which implies that

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

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):

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

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

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

and let

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

Then

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

and

$$(c) \Leftrightarrow K' D K = K' D K K' D K, \quad (2.6a)$$

$$\begin{aligned} (d) &\Leftrightarrow \text{rank}(D) = \text{rank}(K' D K) \\ &= \text{rank}(K' D) = \text{rank}(D K) \end{aligned}$$

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

$$D = D K K' D$$

or

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

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) \rightarrow

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

while

$$\sum \text{rank}(A_i) = \sum \text{tr } A_i = \text{tr } \sum A_i = \text{tr } A = \text{rank}(A), \quad (2.7)$$

and so (1.4) is established.

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

To prove (1.6) we see that (b), (c) \rightarrow

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

multiplying through by A_i yields

$$A_i^3 = A_i^2 \quad (2.8)$$

using (b). To see that (2.8) \rightarrow (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*

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

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

$$LA = MA. \quad (2.10)$$

Thus (2.8) \rightarrow (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 $r = \text{rank}(A) = \text{rank}(B) = \text{rank}(C)$. Then $\text{rank}(AX) = \text{rank}(BC'X) = \text{rank}(C'X) = \text{rank}(A)$, 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*

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

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

$$AX = AY.$$