Steen A. Andersson

Canonical Correlations with Respect to a Complex Structure

Preprint
November
1978

13

Institute of Mathematical Statistics
University of Copenhagen

Steen A. Andersson

CANONICAL CORRELATIONS WITH RESPECT TO A COMPLEX STRUCTURE

Preprint 1978 No. 13

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

November 1978

1. Introduction

Let E be a vector space of dimension $2p$ over the field of real numbers $\mathbb{R}$. Let $x_1,\ldots,x_N$ $(N \geq 2p)$ be identically distributed independent observations from a normal distribution with mean value 0 and unknown covariance Σ. That is, Σ is a positive definite form on the dual space E^* to E. The maximum likelihood estimator $\hat{\Sigma}$ for Σ is well-known to be given by

$$\hat{\Sigma}\,(x_1,\ldots,x_n) = ((x^*,y^*) \to \frac{1}{N}\sum_{i=1}^{N} x^*(x_i)y^*(x_i)\ ;\ x^*,y^* \in E^*)\ .$$

The distribution of $\hat{\Sigma}$ is the Wishart distribution on the set $\rho(E^*)_r$ of positive definite forms on E^* with N degrees of freedom and parameter $\frac{1}{N}\Sigma$. Suppose now that E is also a vector space over the field $\mathbb{C}$ of complex numbers such that the restriction to the subfield of real numbers in $\mathbb{C}$ is the original vector space structure on E. The dimension of E as a vector space over $\mathbb{C}$ is then p. The vector space E^* is then also a vector space over the complex numbers under the definition $zx^* = x^* \circ \bar{z} = (x \to x^*(\bar{z}x)\ ;\ x \in E)$, $x^* \in E^*, z \in \mathbb{C}$. The set $\rho_{\mathbb{C}}(E^*)_r = \{\Sigma \in \rho(E^*)_r | \Sigma(zx^*, y^*) = \Sigma(x^*, \bar{z}\, y^*), \forall\, x^*, y^* \in E^*, \forall\, z \in \mathbb{C}\}$ defines a nulhypothesis in the statistical model described above. The condition $\Sigma(zx^*, y^*) = \Sigma(x^*, \bar{z}\, y^*)$, $\forall\, x^*, y^* \in E^*, \forall\, z \in \mathbb{C}$ is in Andersson [2] called the $\mathbb{C}$-property and in terms of matrices it has the formulation: For every basis $e_1^*,\ldots,e_p^*$ for the complex vector space E^* the matrix for a Σ with the $\mathbb{C}$-property with respect to the basis $e_1^*,\ldots,e_p^*, ie_1^*,\ldots,ie_p^*$ for the real vector space E^* has the form

$$(1.1)\qquad \begin{Bmatrix} \Pi & F \\ -F & \Pi \end{Bmatrix}$$

2. Representation of the maximal invariant

2.1. Lemma. Let Π be a positive definite form on the $\mathbb{R}$-space E. Then there exists a basis $e_1,\ldots,e_p$ for the $\mathbb{C}$-space F such that the $2p \times 2p$ real matrix for Π with respect to $e_1,\ldots,e_p$, $ie_1,\ldots,ie_p$ has the form

$$(2.1) \qquad \begin{pmatrix} I & D_\lambda \\ D_\lambda & I \end{pmatrix}$$

where I is the $p \times p$ identity matrix and

$$(2.2) \qquad D_\lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p) \quad \text{with} \quad 1 > \lambda_1 \geq \ldots \geq \lambda_p \geq 0 \;.$$

Furthermore, the matrix D_λ is uniquely determined by Π; and if $\lambda_1 > \lambda_2 > \ldots > \lambda_p > 0$, then Π also determines the basis $e_1,\ldots,e_p$ uniquely up to the sign of each basis vector.

Proof: Let $e_1',\ldots,e_p'$ be a basis for the $\mathbb{C}$-space E and let

$$\begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12}' & \Pi_{22} \end{pmatrix}$$

be the $2p \times 2p$ real matrix for Π with respect to $e_1',\ldots,e_p'$, $ie_1',\ldots,ie_p'$. The assertion is then that there exists a nonsingular complex $p \times p$ matrix $Z_1 = A + iB$ such that

$$(2.3) \qquad \begin{pmatrix} A' & B' \\ -B' & A' \end{pmatrix} \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12}' & \Pi_{22} \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \begin{pmatrix} I & D_\lambda \\ D_\lambda & I \end{pmatrix}$$

the ν'th row of Y with $\exp[-i\theta_\nu/2]$, where $d_\nu = |d_\nu|\exp[i\theta_\nu]$, $\nu = 1,\ldots,p$, and call this new matrix for Z, we obtain (2.6) with $\lambda_\nu = |d_\nu|$, $\nu = 1,\ldots,p$. Since Π is positive definite, we have $1 > \lambda_1 > \ldots \geq \lambda_p \geq 0$. The uniqueness follows from a rather elementary examination of the proof in [3] or from direct matrix calculation. Since every matrix of the form (2.1) with $1 > \lambda_1 \geq \ldots \geq \lambda_p \geq 0$ is positive definite it follows from Lemma (2.1) that the mapping from $\wp(E^*)_r$ onto $\Omega = \{(\lambda_1,\ldots,\lambda_p) \in \mathbb{R}_+^p \mid 1 > \lambda_1 \geq \ldots \geq \lambda_p \geq 0\}$ determined from Lemma 2.1 is a maximal invariant function.

3. Canonical correlations with respect to a complex structure. Interpretation.

It follows from Lemma 2.1 that there exists a basis $e_1,\ldots,e_p$ for the $\mathbb{C}$-space E such that the $2p \times 2p$ matrix for Σ with respect to $e_1^*,\ldots,e_p^*$, $ie_1^*,\ldots,ie_p^*$ has the form (2.1). In (2.1) D_λ is unique; and if $\lambda_1 > \ldots > \lambda_p > 0$, the basis $e_1^*,\ldots,e_p^*$ for the $\mathbb{C}$-space E^* is unique up to a sign for each element. λ_j is called the j-th theoretical canonical correlation of Σ with respect to the complex structure, and e_j^* is called the j-th theoretical canonical linear form of Σ with respect to the complex structure $j = 1,\ldots,p$. Let $x \in E^*$ have coordinates $(\alpha_1,\ldots,\alpha_p,\beta_1,\ldots,\beta_p)$ with respect to $e_1^*,\ldots,e_p^*$, $ie_1^*,\ldots,ie_p^*$. Then

$$\Sigma(x^*,x^*) = \sum_i \alpha_i^2 + \sum_i \beta_i^2 + 2\sum_i \lambda_i \alpha_i \beta_i \tag{3.1}$$

$$\Sigma(ix^*,ix^*) = \sum_i \alpha_i^2 + \sum_i \beta_i^2 - 2\sum_i \lambda_i \alpha_i \beta_i \tag{3.2}$$

4. The distribution of the empirical canonical correlations with respect to a complex structure.

The estimator $\hat{\Sigma}(x_1,\ldots,x_N)$ for Σ in the observations point $(x_1,\ldots,x_N)$ is given in the introduction. Suppose that $\Sigma \in \mathcal{P}_{\mathbb{C}}(E^*)_r$ and let $e_1^*,\ldots,e_p^*$ be a basis for E^* such that the $2p \times 2p$ matrix for Σ with respect to the basis $e_1^*,\ldots,e_p^*,ie_1^*,\ldots,ie_p^*$ is the $2p \times 2p$ identity matrix. The distribution of $\hat{\Sigma}$ in terms of matrices is a Wishart distribution with a representation as a density with respect to the restriction of the Lebesgue measure to all positive definite $2p \times 2p$ matrices $\mathcal{P}(\mathbb{R}^{2p})_r$ as follows

$$(4.1) \qquad c \cdot |\det \Theta|^{(N-2p-1)/2} \exp\{-\tfrac{1}{2}\operatorname{tr}(\Theta)\}d\Theta, \ \Theta \in \mathcal{P}(\mathbb{R}^{2p}) \quad .$$

The canonical correlations and linear forms (with respect to the complex structure) of $\hat{\Sigma}(x_1,\ldots,x_N)$ is called the <u>empirical canonical correlations and linear forms with respect to the complex structure.</u> The classical theory of canonical correlations is due to Hotelling [4]. We shall find the distribution of these. If we define Φ and Ψ from the $2p \times 2p$ real matrix Θ, as in formula (2.5), we have a one-to-one and onto mapping between $\mathcal{P}(\mathbb{R}^{2p})_r$ and $\mathcal{P}(\mathbb{C}^p)_r \times \mathcal{S}(\mathbb{C}^p)$, where $\mathcal{P}(\mathbb{C}^p)_r$ respectively $\mathcal{S}(\mathbb{C}^p)$ denotes the set of positive definite hermitian respectively symmetric $p \times p$ complex matrices, with Jacobian 1. Furthermore, (2.6) defines a one-to-one mapping from $GL_+(\mathbb{C}^p) \times \Omega$ into $\mathcal{P}(\mathbb{C}^p)_r \times \mathcal{S}(\mathbb{C}^p)$, where $GL_+(\mathbb{C}^p)$ is the subset of all nonsingular $p \times p$ complex matrices with a positive real part in the first row and $\Omega = \{(\lambda_1,\ldots,\lambda_p) \in \mathbb{R}^p | 1 > \lambda_1 > \ldots > \lambda_p > 0\}$.

on $GL_+(\mathbb{C}^p) \times \Omega$. Integrating over $Z \in GL_+(\mathbb{C}^p)$, we get the distribution of $f_1 = \lambda_1^2, \ldots, f_p = \lambda_p^2$:

$$(4.4) \qquad c_3 \prod_{i=1}^{p} (1 - f_i)^{(N-2p-1)/2} \prod_{i<j} (f_i - f_j)\, df_1, \ldots, df_p$$

on $\Omega = \{(f_1, \ldots, f_p) \in \mathbb{R}^p \mid 1 > f_1 > \ldots > f_p > 0\}$. Formula (13) in [1], p. 324, for $p_1 = p$, $p_2 = p+1$ and N replaced by $N+2$ gives the normings constant c_3, namely,

$$(4.5) \qquad c_3 = \Pi^{\frac{p}{2}} \prod_{i=1}^{p} \frac{\Gamma(\frac{1}{2}(N-i)+1)}{\Gamma(\frac{1}{2}(N-p+1-i))\Gamma(\frac{1}{2}(p+1-i))\Gamma(\frac{1}{2}(p-i)+1)} .$$

REFERENCES

[1] Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis, Wiley, New York.

[2] Andersson, S.A. (1975). Invariant normal models. Ann. Statist. **3**, 132-154.

[3] Bourbaki, N. (1959). Elements de Mathematique, Algebra, Chapitre 9, Hermann, Paris.

[4] Hotelling, H. (1936). Relations between two sets of variables. Biometrika **28**, 321-377.

[5] Hsu, P. L. (1939). On the distribution of the roots of certain determinantal equations. Ann. Eugen., **9**, 250-258.

PREPRINTS 1977

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Asmussen, Søren & Keiding, Niels: Martingale Central Limit Theorems and Asymptotic Estimation Theory for Multitype Branching Processes.

No. 2 Jacobsen, Martin: Stochastic Processes with Stationary Increments in Time and Space.

No. 3 Johansen, Søren: Product Integrals and Markov Processes.

No. 4 Keiding, Niels & Lauritzen, Steffen L. : Maximum likelihood estimation of the offspring mean in a simple branching process.

No. 5 Hering, Heinrich: Multitype Branching Diffusions.

No. 6 Aalen, Odd & Johansen, Søren: An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations.

No. 7 Johansen, Søren: The Product Limit Estimator as Maximum Likelihood Estimator.

No. 8 Aalen, Odd & Keiding, Niels & Thormann, Jens: Interaction Between Life History Events.

No. 9 Asmussen, Søren & Kurtz, Thomas G.: Necessary and Sufficient Conditions for Complete Convergence in the Law of Large Numbers.

No. 10 Dion, Jean-Pierre & Keiding, Niels: Statistical Inference in Branching Processes.

PREPRINTS 1978

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Tjur, Tue: Statistical Inference under the Likelihood Principle.

No. 2 Hering, Heinrich: The Non-Degenerate Limit for Supercritical Branching Diffusions.

No. 3 Henningsen, Inge: Estimation in M/G/1-Queues.

No. 4 Braun, Henry: Stochastic Stable Population Theory in Continuous Time.

No. 5 Asmussen, Søren: On some two-sex population models.

No. 6 Andersen, Per Kragh: Filtered Renewal Processes with a Two-Sided Impact Function.

No. 7 Johansen, Søren & Ramsey, Fred L.: A Bang-Bang Representation for 3x3 Embeddable Stochastic Matrix.

No. 8 Braun, Henry: A Simple Method for Testing Goodness of Fit in the Presence of Nuisance Parameters.

No. 9 Lauritzen, S.L. & Speed, T.P. & Vijayan, K. : Decomposable Graphs and Hypergraphs.

No. 10 Hald, Anders: On the Statistical Theory of Sampling Inspection by Attributes.

No. 11 Darroch, J.N. & Lauritzen, S.L. & Speed, T.P.: Markov Fields and Log-linear Interaction Models for Contingency Tables.

No. 12 Speed, T.P.: A Note on Nearest-Neighbour Gribbs and Markov Probabilities.

No. 13 Andersson, Steen A.: Canonical Correlations with Respect to a Complex Structure.