Steen A. Andersson

Canonical Correlations with Respect to a Complex Structure

Preprint November 1978
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 $\Sigma$. That is, $\Sigma$ is a positive definite form on the dual space $E^*$ to $E$. The maximum likelihood estimator $\hat{\Sigma}$ for $\Sigma$ is well-known to be given by

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

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 $zx^* = x^* \circ z = (x \to x^*(\bar{z}x) ; x \in E)$, $x^* \in E^*$, $z \in \mathbb{C}$. 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 null hypothesis 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 $\Sigma$ 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

$$(1.1) \begin{pmatrix} \Pi & F \\ -F & \Pi \end{pmatrix}$$
2. Representation of the maximal invariant

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

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

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

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

Furthermore, the matrix $D_\lambda$ is uniquely determined by $\Pi$; and if $\lambda_1 > \lambda_2 > \ldots > \lambda_p > 0$, then $\Pi$ 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 $C$-space $E$ and let

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

be the $2p \times 2p$ real matrix for $\Pi$ 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

$$(2.3) \quad \begin{pmatrix} A' & B' \\ -B' & A' \end{pmatrix} \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi'_{12} & \Pi_{22} \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \begin{pmatrix} I & D_\lambda \\ D_\lambda & I \end{pmatrix}$$
the $v$th row of $Y$ with $\exp[-i \theta_v/2]$, where $d_v = |d_v| \exp[i \theta_v]$, $v = 1, ..., p$, and call this new matrix for $Z$, we obtain (2.6) with $\lambda_v = |d_v|$, $v = 1, ..., p$. Since $\Pi$ is positive definite, we have $1 > \lambda_1 > ... > \lambda_p > 0$. 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 $1 > \lambda_1 > ... > \lambda_p > 0$ is positive definite it follows from Lemma (2.1) that the mapping from $p(E^*)_p$ onto $\Omega = \{ (\lambda_1, ..., \lambda_p) \in \mathbb{R}_p^+ | 1 > \lambda_1 > ... > \lambda_p > 0 \}$ 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 $e_1, ..., e_p$ for the $C$-space $E$ such that the $2p \times 2p$ matrix for $E$ with respect to $e_1^*, ..., e_p^*$, $ie_1^*, ..., ie_p^*$ has the form (2.1). In (2.1) $D_\lambda$ is unique; and if $\lambda_1 > ... > \lambda_p > 0$, the basis $e_1^*, ..., e_p^*$ for the $C$-space $E^*$ is unique up to a sign for each element. $\lambda_j$ is called the $j$-th theoretical canonical correlation of $E$ with respect to the complex structure, and $e_j^*$ is called the $j$-th theoretical canonical linear form of $E$ with respect to the complex structure $j = 1, ..., p$. Let $x \in E^*$ have coordinates $(\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_p)$ with respect to $e_1^*, ..., e_p^*$, $ie_1^*, ..., ie_p^*$. Then

(3.1) \[ \Sigma(x^*, x^*) = \sum_{i} \alpha_i^2 + \sum_{i} \beta_i^2 + 2 \sum_{i} \lambda_i \alpha_i \beta_i \]

(3.2) \[ \Sigma(ix^*, ix^*) = \sum_{i} \alpha_i^2 + \sum_{i} \beta_i^2 - 2 \sum_{i} \lambda_i \alpha_i \beta_i \]
The distribution of the empirical canonical correlations with respect to a complex structure.

The estimator $\hat{\Sigma}(x_1, \ldots, x_N)$ for $\Sigma$ in the observations point $(x_1, \ldots, x_N)$ is given in the introduction. Suppose that $\Sigma \in \mathcal{P}_c(E^*)_r$ and let $e^*_1, \ldots, e^*_p$ be a basis for $E^*$ such that the $2p \times 2p$ matrix for $\Sigma$ with respect to the basis $e^*_1, \ldots, e^*_p, i e^*_1, \ldots, i e^*_p$ 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 $\rho(\mathcal{P}^{2p}_r)$ as follows

$$
(4.1) \quad c \cdot |\det \Theta|^{(N-2p-1)/2} \exp\left(-\frac{1}{2} \text{tr}(\Theta)\right) d\Theta, \quad \Theta \in \mathcal{P}(R^{2p})
$$

The canonical correlations and linear forms (with respect to the complex structure) of $\hat{\Sigma}(x_1, \ldots, x_N)$ is called the empirical canonical correlations and linear forms with respect to the complex structure. The classical theory of canonical correlations is due to Hotelling [4]. We shall find the distribution of these. If we define $\Phi$ and $\Psi$ from the $2p \times 2p$ real matrix $\Theta$, as in formula (2.5), we have a one-to-one and onto mapping between $\mathcal{P}(R^{2p})_r$ and $\mathcal{P}(\mathcal{C}^p)_r \times \mathcal{G}(\mathcal{C}^p)_r$, where $\mathcal{P}(\mathcal{C}^p)_r$ respectively $\mathcal{G}(\mathcal{C}^p)_r$ 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 $\text{GL}^+_c(\mathcal{C}^p) \times \Omega$ into $\mathcal{P}(\mathcal{C}^p)_r \times \mathcal{G}(\mathcal{C}^p)_r$, where $\text{GL}^+_c(\mathcal{C}^p)_r$ 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 \( f_1 = \lambda_1^2, \ldots, f_p = \lambda_p^2 \):

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

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

\[
(4.5) \quad c_3 = \prod_{i=1}^{p} \prod_{j=1}^{p} \frac{\Gamma(\frac{1}{2}(N-i)+1)}{\Gamma(\frac{1}{2}(N-p+1-i))\Gamma(\frac{1}{2}(p+1-i))\Gamma(\frac{1}{2}(p-i)+1)}
\]
REFERENCES


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. 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. 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. 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.