Canonical Correlations for Group Symmetry Models

Jesse Crawford

Department of Mathematics
Tarleton State University
jcrawford@tarleton.edu

March 28, 2009

Outline

Outline

1. Canonical Correlations
2. Group Symmetry Models
3. Generalized Canonical Correlations

- Developed by Hotelling (1936).
- Theory covered in Chapter 12 of Anderson (1984).

Setting

- $X_1, \ldots, X_N$ are i.i.d. normally distributed random vectors with mean zero and covariance matrix $\Sigma \in \mathbf{PD}(I)$.
- $\mathbb{R}^I = \mathbb{R}^{J_1} \oplus \mathbb{R}^{J_2}$
- $$X_n = \begin{pmatrix} X_n^{(1)} \\ X_n^{(2)} \end{pmatrix}$$
- Testing problem: are $X_n^{(1)}$ and $X_n^{(2)}$ independent?

Notation

- Suppose $x \in \mathbb{R}^{J_1}$ and $y \in \mathbb{R}^{J_2}$
- Linear combinations: $x^t X_n^{(1)}$ and $y^t X_n^{(2)}$
- $\text{Cov}_\Sigma(x, y) := x^t \Sigma_{12} y$
- $\mathbb{V}_\Sigma(x) := x^t \Sigma_{11} x$
- $\mathbb{V}_\Sigma(y) := y^t \Sigma_{22} y$

- The *first canonical correlation* is
- $c_1 = \max\{\mathrm{Cov}_\Sigma(x, y) \,|\, x \in \mathbb{R}^{J_1}, y \in \mathbb{R}^{J_2}, \mathbb{V}_\Sigma(x) = \mathbb{V}_\Sigma(y) = 1\}$.
- Suppose the maximum is attained at (x_1, y_1).
- (x_1, y_1) is the *first pair of canonical covariates*.

Definitions

The *second canonical correlation* is

$$c_2 = \max\{\text{Cov}_\Sigma(x, y) \,\big|\, x \in \mathbb{R}^{J_1}, y \in \mathbb{R}^{J_2}, \mathbb{V}_\Sigma(x) = \mathbb{V}_\Sigma(y) = 1, \text{Cov}_\Sigma(x, x_1) = \text{Cov}_\Sigma(y, y_1) = 0\},$$

Max is attained at (x_2, y_2), the *second pair of canonical covariates*.

$$\vdots$$

The kth *canonical correlation* is

$$c_k = \max\{\text{Cov}_\Sigma(x, y) \,\big|\, x \in \mathbb{R}^{J_1}, y \in \mathbb{R}^{J_2}, \mathbb{V}_\Sigma(x) = \mathbb{V}_\Sigma(y) = 1, \text{Cov}_\Sigma(x, x_i) = \text{Cov}_\Sigma(y, y_i) = 0, i = 1, \ldots, k-1\},$$

Max is attained at (x_k, y_k), the kth *pair of canonical covariates*.

Results

- Canonical correlations: $c_1, \ldots, c_{J_2}$
- Canonical covariate pairs: $(x_1, y_1), \ldots, (x_{J_2}, y_{J_2})$.

Theorem

c_k is the kth largest root of

$$\begin{vmatrix} -c\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -c\Sigma_{22} \end{vmatrix} = 0,$$

and the canonical covariates satisfy

$$\begin{pmatrix} -c_k\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -c_k\Sigma_{22} \end{pmatrix} \begin{pmatrix} x_k \\ y_k \end{pmatrix} = 0.$$

Proof.

c_1 is the maximum value of $x^t\Sigma_{12}y$ subject to the constraints

- $x^t\Sigma_{11}x = 1$
- $y^t\Sigma_{22}y = 1$

- Apply Lagrange multipliers to prove results for c_1 and (x_1, y_1).
- Complete proof by inducting on J_2.

$\square$

Relation to the Maximal Invariant Statistic

- $$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \ A_1 \in \mathbf{GL}(J_1), \ A_2 \in \mathbf{GL}(J_2)$$
- Group Actions:
  - $A \cdot x = Ax$, for $x \in \mathbb{R}^{I \times N}$
  - $A \cdot \Sigma = A \Sigma A^t$, for $\Sigma \in \mathbf{PD}(I)$
- Testing problem is invariant under these actions
- The family of *empirical* canonical correlations is a maximal invariant statistic.

Empirical Canonical Correlations and Eigenvalues

- $$\hat{\Sigma} = S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$
- $$\hat{\Sigma}_0 = \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22} \end{pmatrix}$$
- Residual
$$R = \hat{\Sigma} - \hat{\Sigma}_0 = \begin{pmatrix} 0 & S_{12} \\ S_{21} & 0 \end{pmatrix}$$
- Empirical canonical correlations satisfy
$$\begin{vmatrix} -c_k S_{11} & S_{12} \\ S_{21} & -c_k S_{22} \end{vmatrix} = 0$$
- $$|R - c_k \hat{\Sigma}_0| = 0$$

Empirical Canonical Correlations and Eigenvalues

- $$|R - c_k\hat{\Sigma}_0| = 0$$
- Empirical canonical correlations are eigenvalues of R wrt. $\hat{\Sigma}_0$.
- Canonical covariates satisfy
$$(R - c_k\hat{\Sigma}_0)\begin{pmatrix} x_k \\ y_k \end{pmatrix} = 0$$
- $$u_k = \frac{1}{2}\begin{pmatrix} x_k \\ y_k \end{pmatrix} \text{ is an eigenvector of } R \text{ wrt. } \hat{\Sigma}_0 \text{ corresponding to } c_k$$
- $$v_k = \frac{1}{2}\begin{pmatrix} x_k \\ -y_k \end{pmatrix} \text{ is an eigenvector corresponding to } -c_k$$

Empirical Canonical Covariates and Eigenvectors

- $x_k = u_k + v_k$
- $y_k = u_k - v_k$

Outline

- Theory: Andersson and Madsen (1998), Appendix A
- Ten Fundamental Testing Problems: Andersson, Brøns, and Jensen (1983)

Pattern Covariance Matrices

Testing covariance structure of a multivariate normal distribution.

Example (Testing Complex Structure)

$$H_0 : \Sigma = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

Example (Testing Independence)

$$H_0 : \Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$$

Example (Bartlett's Test)

$$H_0 : \Sigma = \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix} \text{ vs. } H : \Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$$

Notation

- $G \le \mathbf{O}(I)$ is a compact group
- $\mathbf{GL}_G(I) = \{$Invertible matrices that commute with $G\}$
- $\mathbf{PD}_G(I) = \{$Positive definite matrices that commute with $G\}$

Group Symmetry Model

$X_1, X_2, \ldots, X_N$ i.i.d. Normal$(0, \Sigma)$

$$\mathrm{H}_G : \Sigma \in \mathbf{PD}_G(I)$$

Example (Testing Complex Structure)

$$G_0 = \left\{ \pm 1_I, \pm \begin{pmatrix} 0 & -1_J \\ 1_J & 0 \end{pmatrix} \right\}$$

Example (Testing Independence)

$$G_0 = \left\{ \pm 1_I, \pm \begin{pmatrix} 1_{J_1} & 0 \\ 0 & -1_{J_2} \end{pmatrix} \right\}$$

Example (Bartlett's Test)

$$G_0 = \left\langle \begin{pmatrix} 1_J & 0 \\ 0 & -1_J \end{pmatrix}, \begin{pmatrix} 0 & 1_J \\ 1_J & 0 \end{pmatrix} \right\rangle$$

Estimation

- $$\hat{\Sigma} = \Psi_G(S) = \int_G gSg^t \, dg$$
- $$\hat{\Sigma} = \frac{1}{|G|} \sum_{g \in G} gSg^t$$

Example (Testing Complex Structure)

$$\hat{\Sigma}_0 = \frac{1}{2} \begin{pmatrix} S_{11} + S_{22} & S_{12} - S_{21} \\ S_{21} - S_{12} & S_{11} + S_{22} \end{pmatrix}$$

- $\hat{\Sigma} \sim$ generalized Wishart distribution on $\mathbf{PD}_G(I)$.

Testing

- $G \leq G_0$
- $\mathbf{PD}_{G_0}(I) \subseteq \mathbf{PD}_G(I)$
- Testing problem:

$$\mathrm{H}_0 : \Sigma \in \mathbf{PD}_{G_0}(I) \text{ vs. } \mathrm{H} : \Sigma \in \mathbf{PD}_G(I)$$

- Actions of $A \in \mathbf{GL}_G(I)$
  - $A \cdot x = Ax$, for $x \in \mathbb{R}^{I \times N}$
  - $A \cdot \Sigma = A \Sigma A^t$, for $\Sigma \in \mathbf{PD}_G(I)$
- Maximal invariant: eigenvalues of $R = \hat{\Sigma} - \hat{\Sigma}_0$ wrt. $\hat{\Sigma}_0$.

Questions

- Can canonical correlations be generalized?
- Are these maximally invariant eigenvalues canonical correlations?

Outline

Setting

- $X_1, X_2, \ldots, X_N$ i.i.d. Normal$(0, \Sigma)$
- $\Theta_0 \subseteq \Theta \subseteq \mathbf{PD}(I)$
- Testing problem:

$$\mathrm{H}_0 : \Sigma \in \Theta_0 \text{ vs. } \mathrm{H} : \Sigma \in \Theta$$

- $t : \Theta \to \Theta_0$
- $\hat{\Sigma}_0 = t(\hat{\Sigma})$

Definitions

- $c_1 = \max\{\mathrm{Cov}_\Sigma(x,y) \,|\, \mathrm{Cov}_{t(\Sigma)}(x,y) = 0, \mathbb{V}_\Sigma(x) = \mathbb{V}_\Sigma(y) = 1\}$
- $c_k = \max\{\mathrm{Cov}_\Sigma(x,y) \,|\, \mathrm{Cov}_{t(\Sigma)}(x,y) = 0, \mathbb{V}_\Sigma(x) = \mathbb{V}_\Sigma(y) = 1$
  - for $i = 1, \ldots, k-1$,
  - $\mathrm{Cov}_{t(\Sigma)}(x, x_i) = \mathrm{Cov}_{t(\Sigma)}(x, y_i) = 0$,
  - $\mathrm{Cov}_{t(\Sigma)}(y, x_i) = \mathrm{Cov}_{t(\Sigma)}(y, y_i) = 0\}$

Results

- Eigenvalues of $\Sigma - t(\Sigma)$ wrt. $t(\Sigma)$:

$$\lambda_1 \geq \cdots \geq \lambda_l$$

- $$c_k = \frac{\lambda_k - \lambda_{k+1-k}}{\lambda_k + \lambda_{l+1-k} + 2}, \text{ for } k = 1, \ldots, \lfloor \tfrac{l}{2} \rfloor$$

- Covariate pairs: characterized in terms of eigenvectors

Proof.

- WLOG, $t(\Sigma) = 1_l$ and $\Sigma - t(\Sigma) = \Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_l)$
- Apply Lagrange multipliers to
- $c_1 = \max\{x^t \Lambda y \,|\, x, y \in \mathbb{R}^l, x^t(1_l + \Lambda)x = y^t(1_l + \Lambda)y = 1, x^t y = 0\}$
- Induct on l

□

Canonical Correlations for Group Symmetry Models

- $\Theta = \mathbf{PD}_G(I)$
- $\Theta_0 = \mathbf{PD}_{G_0}(I)$
- $t = \Psi_G$
- Eigenvalues of $\hat{\Sigma} - \hat{\Sigma}_0$ wrt. $\hat{\Sigma}_0$

$$\lambda_1, \lambda_2, \ldots, -\lambda_2, -\lambda_1$$

- $$c_k = \lambda_k, \text{ for } k = 1, \ldots, \lfloor \tfrac{I}{2} \rfloor$$
- Covariate pairs: characterized in terms of eigenvectors

Further Research

- Eigenvalues related to other testing problems, such as graphical models.
- Unbiasedness of likelihood ratio tests for group symmetry models.

References

- Andersson, S.A., Brøns, H.K., and Tolver Jensen, S. (1983). Distribution of Eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11** 392-415.
- Andersson, S.A. and Madsen, J. (1998). Symmetry and lattice conditional independence in a multivariate normal distribution. *Ann. Statist.* **26** 525-572.
- Hotelling, Harold (1936). Relations Between Two Sets of Variates. *Biometrika* **28** 321-377.