BIAS AND VARIANCE
OF MULTIPLE CORRESPONDENCE ANALYSIS

JAN DE LEEUW

ABSTRACT. Results on first and second derivatives of generalized
eigenvalues and eigenvectors are used in Delta Method estimates for
bias and variance of the statistics typically computed in Multiple and
Ordinary Correspondence Analysis.

1. INTRODUCTION

Multiple Correspondence Analysis or MCA [Guttman 1941, 1950, Burt
generalized eigenvalue problem $Ax = \lambda Bx$, where $A$ is the Burt Matrix of
$m$ categorical variables and $B$ is $m$ times the diagonal of the Burt Matrix.
We normalize the solution by requiring $x' B x = 1$.

The Burt matrix can be defined in terms of the profile vectors. If the $m$
variables have $k_1, k_2, \ldots, k_m$ categories, then profile vectors are binary
vectors $z_\nu$ of length $\sum k_j$, indicating the $K = \prod k_j$ possible patterns that
can be observed. We have $A = \sum p_\nu z_\nu z_\nu'$, where summation is over all $K$
possible profiles, and $p_\nu$ is the proportion of observed profile vectors in
$n$ trials equal to $z_\nu$. Also $B = \sum p_\nu (m Z_\nu)$, where $Z_\nu = \text{diag}(z_\nu z_\nu')$.

1.1. Summation over $n$. If $m$ is at all large, then $K = \prod k_j$ will tend to
be very large, and many of the $p_\nu$ will be zero. In that case it makes
more sense to translate the formulas from all $K$ possible $z_\nu$ to only the
$n$ observed $z_i$. We
1.2. **Ordinary Correspondence Analysis.** There is a simple relationship between MCA and ordinary Correspondence Analysis (CA) of a cross table [Benzécri, 1973; Greenacre, 1984]. CA is just the special case \( m = 2 \). The relationship between the eigenvalues of MCA and the canonical correlations \( \rho \) of CA is simply \( \rho = 2\lambda - 1 \). For CA we can write the stationary equations in partitioned form as

\[
\begin{bmatrix} D & F \\ F' & E \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 2\lambda \begin{bmatrix} D & \emptyset \\ \emptyset & E \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}
\]

with \( a'Da + b'Eb = 1 \). This can also be written as

\[
Fb = \rho Da, \\
F'a = \rho Db,
\]

which implies that \( a'Da = b'Eb = \frac{1}{2} \).

2. **Delta Method**

Assume the profile proportions \( p_n \) are a realization of an asymptotically normal random vector \( p_n \). More precisely, there is a vector \( \pi \), with non-negative elements \( \pi\nu \) that add up to one, such that

\[
n^{\frac{1}{2}}(p_n - \pi) \overset{D}{\to} \mathcal{N}(0, \Pi - \pi\pi'),
\]

where \( \Pi \) is a diagonal matrix with \( \pi \) on the diagonal. By the *Delta Method* [Mann and Wald, 1943; Tiago De Olivera, 1982], if \( f \) is differentiable at \( \pi \), and \( G(\pi) = Df(\pi) \), then we have the asymptotic distribution result

\[
n^{\frac{1}{2}}(f(p_n) - f(\pi)) \overset{D}{\to} \mathcal{N}(0, G(\pi)(\Pi - \pi\pi')G(\pi)').
\]

If \( f \) is real-valued, bounded, and two times continuously differentiable at \( \pi \), with \( H(\pi) = DDf(\pi) \), then we can approximate the bias by using

\[
\lim_{n \to \infty} nE(f(p_n) - f(\pi)) = \frac{1}{2} \text{tr} H(\pi)(\Pi - \pi\pi').
\]

\(^1\)Random variables are underlined [Hemelrijk, 1966].
3. MCA Results

We now combine the general results on the Delta Method from the previous section with the general results on differentiation of eigenvalues and eigenvectors from Appendix A. This extends results from O’Neill [1978a,b] and De Leeuw [1984]. Some of these results have been used recently in the anacor package on CRAN [De Leeuw and Mair, 2009].

First we determine the dispersion matrix of the asymptotic joint distribution of the vector of eigenvalues. Define \( u_{sv}(\pi) = x_{sv}(\pi)'(z_{sv}z_{sv}' - m\lambda_{sv}(\pi)Z_{sv})x_{sv}(\pi) \). Then

\[
\lim_{n \to \infty} nE\{(\lambda_{sv}(\overline{p}_n) - \lambda_{sv}(\pi))(\lambda_{tv}(\overline{p}_n) - \lambda_{tv}(\pi))\} = \sum_{\nu = 1}^{K} \pi_{\nu} u_{sv}u_{tv}.
\]

Our second result gives the dispersion matrix of the joint distribution of two eigenvectors. Define the vectors

\[
v_{sv}(\pi) = -(A(\pi) - \lambda_{sv}(\pi)B(\pi))^{-1}(z_{sv}z_{sv}' - m\lambda_{sv}(\pi)Z_{sv})x_{sv}(\pi) - \frac{1}{2}m(x_{sv}(\pi)'Z_{sv}x_{sv}(\pi))x_{sv}(\pi).
\]

Then

\[
\lim_{n \to \infty} nE\{(x_{sv}(\overline{p}_n) - x_{sv}(\pi))(x_{tv}(\overline{p}_n) - x_{tv}(\pi))'\} = \sum_{\nu = 1}^{K} \pi_{\nu} v_{sv}(\pi)v_{tv}(\pi)' - \frac{1}{4}x_{sv}(\pi)x_{sv}(\pi)'.
\]

And finally the expected value, and thus the bias correction, for an individual eigenvalue. Define the numbers

\[
w_{sv}(\pi) = -u_{sv}(\pi)(mx_{sv}(\pi)'Z_{sv}x_{sv}(\pi)) + 2x_{sv}(\pi)'(z_{sv}z_{sv}' - m\lambda_{sv}(\pi)Z_{sv})v_{sv}(\pi).
\]

Then

\[
\lim_{n \to \infty} nE(\lambda_{sv}(\overline{p}_n) - \lambda_{sv}(\pi)) = \frac{1}{2} \sum_{\nu = 1}^{K} \pi_{\nu} w_{sv}(\pi),
\]

or, for CA,

\[
\lim_{n \to \infty} nE(\rho_{sv}(\overline{p}_n) - \rho_{sv}(\pi)) = \sum_{\nu = 1}^{K} \pi_{\nu} w_{sv}(\pi).
\]
3.1. **Summation over** $n$. If $m$ is at all large, then $K = \prod k_j$ will tend to be very large, and many of the $p_\nu$ will be zero. In that case it makes more sense to translate the formulas from all $K$ possible $z_\nu$ to only the $n$ observed $z_i$. With obvious notation, we then have formulas of the form

$$\lim_{n \to \infty} nE(\rho_s(\overline{p_n}) - \rho_s(\pi)) = \frac{1}{n} \sum_{i=1}^{n} w_{si}.$$ 

3.2. **Ordinary CA.** Note that if $m = 2$, we index the probabilities by the $I \times J$ cells of the bivariate table, and the two parts of the eigenvector are $a$ and $b$, we have

$$u_{sij} = (1 - 2\lambda_s)(a_{is}^2 + b_{js}^2) + 2a_{is}b_{js} = 2a_{is}b_{js} - \rho_s(a_{is}^2 + b_{js}^2)$$

and

$$\lim_{n \to \infty} nE[\lambda_s(\overline{p_n}) - \lambda_s(\pi)(\lambda_t(\overline{p_n}) - \lambda_t(\pi))] = \sum_{i=1}^{I} \sum_{j=1}^{J} \pi_{ij}u_{sij}u_{tij}.$$
REFERENCES


A. General Case. Suppose $A$ and $B$ are positive semi-definite and depend on a vector of parameters $\theta$. Suppose $(x, \lambda)$ is a normalized eigen-pair of $(A, B)$, i.e. $Ax = \lambda Bx$ and $x'Bx = 1$. In a neighborhood where the eigenvalue is isolated we differentiate the eigen-equations and find

$$\begin{align*}
DAx + ADx - D\lambda Bx - \lambda DBx - \lambda B Dx &= 0, \quad (1a) \\
(x'DBx + 2x'B Dx &= 0. \quad (1b)
\end{align*}$$

Premultiplying both sides of (1a) by $x'$ gives

$$D\lambda = x'(DA - \lambda DB)x. \quad (2)$$

We can write (1b) as

$$x'B Dx = -\frac{1}{2}x'DBx, \quad (3)$$

and (1a) as

$$\begin{align*}
(A - \lambda B)Dx &= -(DA - \lambda DB)x + D\lambda Bx. \quad (4)
\end{align*}$$

Suppose $X$ is a non-singular matrix satisfying $X'Bx = I$ and $X'Ax = \Lambda$. Then $x$ is one of the columns of $X$ and $\lambda$ is the corresponding diagonal element of $\Lambda$. Define, following [De Leeuw [2007]], the generalized inverse

$$(A - \lambda B)^- = X(\Lambda - \lambda I)^+ x',$$

where $(\Lambda - \lambda I)^+$ is the Moore-Penrose inverse. Then $(A - \lambda B)^-Bx = 0$. Since $x$ is the only vector in the null-space of $A - \lambda B$ we have from (4) that for some $\kappa$

$$Dx = -(A - \lambda B)^-(DA - \lambda DB)x + \kappa x.$$

Using (3) we find that

$$Dx = -(A - \lambda B)^-(DA - \lambda DB)x - \frac{1}{2}(x'Bx)x. \quad (5)$$
A.2. **Linear Case.** Suppose \( A = \sum \theta \nu A_\nu \) and \( B = \sum \theta \nu B_\nu \). Switching to functional notation we find

(6a) \[ D_\nu \lambda = x'(A_\nu - \lambda B_\nu)x, \]

as well as

(6b) \[ D_\nu x = -(A - \lambda B)^{-1}(A_\nu - \lambda B_\nu)x - \frac{1}{2} (x' B_\nu x)x. \]

If we differentiate (6a) again

(7) \[ D_\xi D_\nu \lambda = -(D_\xi \lambda)x' B_\nu x + 2x' (A_\nu - \lambda B_\nu)(D_\xi x). \]

Note that \( \sum \theta \nu (D_\nu \lambda) = 0 \) while \( \sum \theta \nu (D_\nu x) = -\frac{1}{2} x \). Moreover, from (7), \( \sum \theta \nu \theta_\xi (D_\xi D_\nu \lambda) = 0 \).

**Department of Statistics, University of California, Los Angeles, CA 90095-1554**

*E-mail address*, Jan de Leeuw: deleeuw@stat.ucla.edu

*URL*, Jan de Leeuw: http://gifi.stat.ucla.edu