Statistics 5401
26. Inference for Canonical Correlation

Gary W. Oehlert
School of Statistics
313B Ford Hall
612-625-1557
gary@stat.umn.edu

Let $x^{(1)}$ be a random p-vector, and let $x^{(2)}$ be a random q-vector.

$$\begin{aligned} Var(x^{(1)}) &= \Sigma_{11} \\ Var(x^{(2)}) &= \Sigma_{22} \\ Cov(x^{(1)}, x^{(2)}) &= \Sigma_{12}(p \times q) \end{aligned}$$

If $\Sigma_{12} = 0$, then all linear combinations of $x^{(1)}$ and $x^{(2)}$ have correlation 0, so all canonical correlations will also be 0.

Even if $\Sigma_{12} = 0$, the sample covariance $\mathbf{S}_{12}$ will not be zero, so canonical correlation analysis of the sample will lead to nonzero sample canonical correlations.

Can we use a random sample to test whether $\Sigma_{12} = 0$, or equivalently, whether all the population canonical correlations are zero?

Yes (with some distributional assumptions).

We will examine two approaches.

The first assumes multivariate normality for the vector

$$x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix}$$

This approach is based on likelihood ratio tests, and can be extended to test whether a subset of the canonical correlations is zero.

The second approach is based on Bonferronizing pairwise tests of zero correlation between each $x_i^{(1)}$ and each $x_j^{(2)}$. This approach assumes only bivariate normality among the pairs.

We want to test $H_0 : \Sigma_{12} = 0$ versus $H_1 : \Sigma_{12} \neq 0$. This null hypothesis involves pq parameters (elements of Σ_{12}).

Under H_1, the maximum likelihood estimate of Σ is $(n-1)/n\mathbf{S}$.

Under H_0, the maximum likelihood estimate of Σ is

$$\mathbf{S}_0 = \frac{n-1}{n} \begin{bmatrix} \mathbf{S}_{11} & 0 \\ 0 & \mathbf{S}_{22} \end{bmatrix}$$

The basic form for the test is then

$$-2 \ln \Lambda = n \ln \left(\frac{|\mathbf{S}_0|}{|\mathbf{S}|} \right)$$

What's not so obvious is that

$$-2 \ln \Lambda = n \ln \left(\frac{|\mathbf{S}_0|}{|\mathbf{S}|} \right) = -n \ln \left(\prod_{i=1}^{p} (1 - \hat{\rho}_i^2) \right)$$

where the $\hat{\rho}_i$s are the sample canonical correlations.
Using a slightly improved multiplier, we compute

$$-(n-1-(p+1+1)/2)\ln\left(\prod_{i=1}^{p}(1-\hat{\rho}_i^2)\right)$$

and compare it to a χ^2_{pq} to get a p-value, rejecting for large values.
In fact, we can test that only some of the canonical correlation are zero. Test

$$H_0 : \rho_{k+1} = \rho_{k+2} = \ldots = \rho_p = 0$$

versus

$$H_1 : \rho_j \neq 0, \quad \text{for some} j > k$$

using

$$-(n-1-(p+1+1)/2)\ln\left(\prod_{i=k+1}^{p}(1-\hat{\rho}_i^2)\right)$$

compared to $\chi^2_{(p-k)(q-k)}$ to get a p-value, rejecting for large values.

```
Cmd> S <- tabs(X,covar:T)

Cmd> S11 <- S[run(5),run(5)]

Cmd> S12 <- S[run(5),run(6,8)]

Cmd> S22 <- S[run(6,8),run(6,8)]

Cmd> releigen(S12'%*%solve(S11)%*%S12,S22)$values
(1)     0.88685     0.095624     0.018179

Cmd> rhohat2 <- releigen(S12'%*%solve(S11)%*%S12,S22)$values

Cmd> dim(X)
(1)          55            8

Cmd> -(55-1-(8+1)/2)*log(prod(1-rhohat2))
(1)      113.75

Cmd> 1-cumchi(113.75,5*3)
(1)          0

Cmd> -(55-1-(8+1)/2)*log(prod(1-rhohat2[-1]))
(1)     5.8834

Cmd> 1-cumchi(5.88344,(4*2))
(1)     0.66029
```

```
Cmd> readdata("",sg,sp,nas,ct,mrt,art,mt)
Read from file "~/5401/JW5data/T9-12.DAT"

Cmd> X <- hconcat(sg,sp,nas,ct,mrt,art,mt)

Cmd>  S <- tabs(X,covar:T)

Cmd> S11 <- S[run(3),run(3)]

Cmd> S12 <- S[run(3),run(4,7)]

Cmd> S22 <- S[run(4,7),run(4,7)]

Cmd> releigen(S12%*%solve(S22)%*%S12',S11)$values
(1)        0.989      0.77107      0.14715

Cmd> rhohat2 <- releigen(S12%*%solve(S22)%*%S12',S11)$values

Cmd> dim(X)
(1)           50             7

Cmd> -(49-(3+4+1)/2)*log(prod(1-rhohat2))
(1)       276.43

Cmd> 1-cumchi(213.56,12)
(1)            0

Cmd> -(49-(3+4+1)/2)*log(prod(1-rhohat2[-1]))
(1)       73.508

Cmd> 1-cumchi(41.6,3*2)
(1)   7.816-17

Cmd> -(49-(3+4+1)/2)*log(prod(1-rhohat2[-run(2)]))
(1)      7.1629

Cmd> 1-cumchi(.98,2*1)
(1)      0.0278
```

A little algebra. Let $\mathbf{S}_{11}^{1/2}$ be a symmetric square root of $\mathbf{S}_{11}$ (symmetry isn't really needed, but it makes some presentation simpler). Similarly let $\mathbf{S}_{22}^{1/2}$ be a symmetric square root of $\mathbf{S}_{22}$. Let

$$\mathbf{S}_0^{1/2} = \frac{n-1}{n} \begin{bmatrix} \mathbf{S}_{11}^{1/2} & 0 \\ 0 & \mathbf{S}_{22}^{1/2} \end{bmatrix}$$

$$
\begin{aligned}
-2\ln\Lambda &= n\ln\left(\frac{|\mathbf{S}_0|}{|\mathbf{S}|}\right) \\
&= n\ln\left(\frac{|\mathbf{S}_0^{-1/2}||\mathbf{S}_0||\mathbf{S}_0^{-1/2}|}{|\mathbf{S}_0^{-1/2}||\mathbf{S}||\mathbf{S}_0^{-1/2}|}\right) \\
&= n\ln\left(\frac{|\mathbf{I}|}{|\mathbf{S}_0^{-1/2}\mathbf{S}\mathbf{S}_0^{-1/2}|}\right) \\
&= -n\ln\left(|\mathbf{S}_0^{-1/2}\mathbf{S}\mathbf{S}_0^{-1/2}|\right) \\
&= -n\ln\left(\left|\begin{bmatrix} \mathbf{I} & \mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2} \\ \mathbf{S}_{22}^{-1/2}\mathbf{S}_{21}\mathbf{S}_{11}^{-1/2} & \mathbf{I} \end{bmatrix}\right|\right)
\end{aligned}
$$

Let

$$
\mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2} = \mathbf{UDV}'
$$

be the SVD. $\mathbf{U}$ is $p \times p$, and $\mathbf{V}$ is $q \times p$. If $p < q$, let $\mathbf{V}_f = [\mathbf{V} : \mathbf{V}_0]$ be a completion of $\mathbf{V}$ to a full orthogonal matrix. Let

$$
\mathbf{H} = \begin{bmatrix} \mathbf{U}' & 0 \\ 0 & \mathbf{V}_f' \end{bmatrix}
$$

Note that $|\mathbf{H}|\ |\mathbf{H}'| = |\mathbf{I}| = 1$.
Then

$$
\begin{aligned}
&\left|\begin{bmatrix} \mathbf{I} & \mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2} \\ \mathbf{S}_{22}^{-1/2}\mathbf{S}_{21}\mathbf{S}_{11}^{-1/2} & \mathbf{I} \end{bmatrix}\right| \\
&\quad = |\mathbf{H}|\left|\begin{bmatrix} \mathbf{I} & \mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2} \\ \mathbf{S}_{22}^{-1/2}\mathbf{S}_{21}\mathbf{S}_{11}^{-1/2} & \mathbf{I} \end{bmatrix}\right||\mathbf{H}'| \\
&\quad = \left|\begin{bmatrix} \mathbf{I} & \mathbf{U}'\mathbf{UDV}'\mathbf{V}_f \\ \mathbf{V}_f'\mathbf{VDU}'\mathbf{U} & \mathbf{I} \end{bmatrix}\right| \\
&\quad = \left|\begin{bmatrix} \mathbf{I} & (\mathbf{D} : 0) \\ (\mathbf{D} : 0)' & \mathbf{I} \end{bmatrix}\right|
\end{aligned}
$$

where the sample canonical correlations are on the diagonal of $\mathbf{D}$. The determinant of a matrix of this special form can be expressed as $\prod(1 - \rho_i^2)$.
The second kind of test is based on pairwise correlations between the components of $x^{(1)}$ and $x^{(2)}$. Let $\mathbf{S}$ be an estimate of the covariance of (u, v) with f degrees of freedom, and let $r = \mathbf{S}_{12}/\sqrt{\mathbf{S}_{11}\mathbf{S}_{22}}$ be the sample correlation.
If u and v are each normally distributed, and if they are independent, then

$$
\sqrt{f-1}\frac{r}{\sqrt{1-r^2}} \sim t_{f-1}
$$

Note that this is really a regression statement; the F-test for regression u on v (or vice versa) is $(f-1)r^2/(1-r^2)$.

For $x^{(1)}$ and $x^{(2)}$, there are pq pairwise correlations r_{ij} between $x_i^{(1)}$ and $x_j^{(2)}$. Using Bonferroni, the test of $H_0 : \Sigma_{12} = 0$ is

$$\max_{i,j} t_{ij} = \max_{i,j} |\sqrt{f-1}\frac{r_{ij}}{\sqrt{1-r_{ij}^2}}| > t_{1-\alpha/(2pq),f-1}$$

For any i, j pair for which $|t_{ij}|$ exceeds the critical value, we can reject the null that $\Sigma_{12(ij)} = 0$.

```
Cmd> S <- tabs(X,covar:T)

Cmd> R <- dmat(diag(S)^-.5)%*%S%*%\
dmat(diag(S)^-.5)

Cmd> R12 <- R[run(5),run(6,8)];R12
(1,1)      0.61946      0.63254      0.51995
(2,1)      0.69538      0.69654      0.59618
(3,1)      0.77861       0.7872      0.70499
(4,1)      0.86359      0.86905      0.80648
(5,1)      0.92811       0.9347      0.86555

Cmd> dim(X)
(1)           55            8

Cmd> tstats <- abs( sqrt(53)*R12/sqrt(1-R12^2))

Cmd> tstats
(1,1)       5.7447       5.9455       4.4314
(2,1)       7.0444       7.0672       5.4061
(3,1)       9.0332        9.293       7.2367
(4,1)        12.47       12.788       9.9299
(5,1)       18.149       19.144       12.582

Cmd> 15*2*(1-cumstu(19.144,53))
(1)            0
```