Lecture Notes on Canonical Correlation Analysis

John C. Chao

Econ 721 Lecture Notes

October 11, 2021

Canonical Correlation Analysis

- Consider the random vector

$$\underset{m\times 1}{X} = \begin{pmatrix} \underset{m_1\times 1}{U} \\ \underset{m_2\times 1}{V} \end{pmatrix}$$

where

$$\begin{aligned} E[X] &= 0, \\ V(X) &= \Sigma = \begin{pmatrix} \underset{m_1\times m_1}{\Sigma_{11}} & \underset{m_1\times m_2}{\Sigma_{12}} \\ \underset{m_2\times m_1}{\Sigma_{21}} & \underset{m_2\times m_2}{\Sigma_{22}} \end{pmatrix} \end{aligned}$$

with $\Sigma_{12} = \Sigma_{21}'$ and, without loss of generality, we assume that $m_1 \leq m_2$.

Canonical Correlation Analysis

- Now, consider a linear combination

$$Y = \alpha' U$$

of the components of U and a linear combination

$$Z = \gamma' V$$

of the components of V.

- The objective is to find the linear combinations of U and V which give the maximum correlation.

- To proceed, note first that correlations are unique only up to a scalar multiplication since

$$
\begin{aligned}
|Corr(Y,Z)| &= \left| \frac{Cov(Y,Z)}{\sqrt{Var(Y)}\sqrt{Var(Z)}} \right| \\
&= \left| \frac{ab}{|ab|} \frac{Cov(Y,Z)}{\sqrt{Var(Y)}\sqrt{Var(Z)}} \right| \\
&= \left| \frac{Cov(aY,bZ)}{\sqrt{Var(aY)}\sqrt{Var(bZ)}} \right| \\
&= |Corr(aY,bZ)|
\end{aligned}
$$

for $a, b \in \mathbb{R}$ such that $a \neq 0$ and $b \neq 0$.

- It follows that, without loss of generality, we can normalize the variance of Y and Z as follows:

$$\begin{aligned} 1 &= Var\left(Y\right) = E\left[Y^2\right] = \alpha' E\left[UU'\right]\alpha = \alpha'\Sigma_{11}\alpha, \\ 1 &= Var\left(Z\right) = E\left[Z^2\right] = \gamma' E\left[VV'\right]\gamma = \gamma'\Sigma_{22}\gamma, \end{aligned}$$

where $Var\left(Y\right) = E\left[Y^2\right]$ and $Var\left(Z\right) = E\left[Z^2\right]$ follow from the fact that $E\left[Y\right] = \alpha' E\left[U\right] = 0$ and $E\left[Z\right] = \gamma' E\left[V\right] = 0$.

- Note that the above normalization is tantamount to setting

$$a = \frac{1}{\sqrt{Var\left(Y\right)}} \text{ and } b = \frac{1}{\sqrt{Var\left(Z\right)}}.$$

- Under this normalization, correlation between Y and Z is given by the covariance

$$E\left[YZ\right] = \alpha' E\left[UV\right]\gamma = \alpha'\Sigma_{12}\gamma.$$

Canonical Correlation Analysis

- To maximize the correlation, we set up the Lagrangian

$$\mathcal{L}_1 = \alpha'\Sigma_{12}\gamma - \frac{1}{2}\lambda\left(\alpha'\Sigma_{11}\alpha - 1\right) - \frac{1}{2}\mu\left(\gamma'\Sigma_{22}\gamma - 1\right)$$

- The first-order conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}_1}{\partial \alpha} &= \Sigma_{12}\gamma - \lambda\Sigma_{11}\alpha = 0 \qquad (1) \\ \frac{\partial \mathcal{L}_1}{\partial \gamma} &= \Sigma_{21}\alpha - \mu\Sigma_{22}\gamma = 0 \qquad (2) \end{aligned}$$

- Premultiply (1) by α' and (2) by γ', we get

$$\begin{aligned} \alpha'\Sigma_{12}\gamma - \lambda\alpha'\Sigma_{11}\alpha &= 0, \\ \gamma'\Sigma_{21}\alpha - \mu\gamma'\Sigma_{22}\gamma &= 0. \end{aligned}$$

- Next, differentiating $\mathcal{L}_1$

$$\mathcal{L}_1 = \alpha' \Sigma_{12} \gamma - \frac{1}{2} \lambda \left(\alpha' \Sigma_{11} \alpha - 1 \right) - \frac{1}{2} \mu \left(\gamma' \Sigma_{22} \gamma - 1 \right)$$

with respect to λ and μ and setting the partial derivatives equal to zero, we get

$$\begin{aligned} \frac{\partial \mathcal{L}_1}{\partial \lambda} &= -\frac{1}{2} \left(\alpha' \Sigma_{11} \alpha - 1 \right) = 0 \text{ implying that } \alpha' \Sigma_{11} \alpha = 1 \\ \frac{\partial \mathcal{L}_1}{\partial \mu} &= -\frac{1}{2} \left(\gamma' \Sigma_{22} \gamma - 1 \right) = 0 \text{ implying that } \gamma' \Sigma_{22} \gamma = 1 \end{aligned}$$

- It follows that

$$\begin{aligned} \alpha'\Sigma_{12}\gamma - \lambda\alpha'\Sigma_{11}\alpha &= \alpha'\Sigma_{12}\gamma - \lambda = 0, \\ \gamma'\Sigma_{21}\alpha - \mu\gamma'\Sigma_{22}\gamma &= \gamma'\Sigma_{21}\alpha - \mu = 0. \end{aligned}$$

Moreover, because $\alpha'\Sigma_{12}\gamma = \gamma'\Sigma'_{12}\alpha = \gamma'\Sigma_{21}\alpha$, we further deduce that

$$\lambda = \mu = \alpha'\Sigma_{12}\gamma.$$

- Putting everything together, we can rewrite the first-order conditions

$$\begin{aligned} \frac{\partial \mathcal{L}_1}{\partial \alpha} &= \Sigma_{12}\gamma - \lambda\Sigma_{11}\alpha = 0 \\ \frac{\partial \mathcal{L}_1}{\partial \gamma} &= \Sigma_{21}\alpha - \mu\Sigma_{22}\gamma = 0 \end{aligned}$$

as

$$\begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = 0 \tag{3}$$

Canonical Correlation Analysis

- Now, for the system of equations

$$\begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = 0$$

  to have a nontrivial solution, i.e., $(\alpha', \gamma')' \neq 0$, it must be that the matrix on the left-hand side is singular, so that

$$\det \begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} = 0 \qquad (4)$$

- Since our objective is to maximize the correlation $\lambda = \alpha'\Sigma_{12}\gamma$, we choose λ_1 to be the largest root of the determinantal equation above and let $(\alpha_1', \gamma_1')'$ be the associated eigenvector.

Canonical Correlation Analysis

- Next, we construct

$$Y_1 = \alpha_1' U \text{ and } Z_1 = \gamma_1' V$$

so that Y_1 and Z_1 are the normalized linear combinations of U and V which give the maximum correlation, with λ_1 being the maximal (canonical) correlation.

- The analysis can then proceed further with finding a second linear combination of U, say $Y_2 = \alpha_2' U$, and a second linear combination of V, say $Z_2 = \gamma_2' V$, which have maximum correlation among all pairs not correlated with Y_1 and Z_1.

- To proceed, we again set up a Lagrangian

$$\mathcal{L}_2 = \alpha'\Sigma_{12}\gamma - \frac{1}{2}\lambda\left(\alpha'\Sigma_{11}\alpha - 1\right) - \frac{1}{2}\mu\left(\gamma'\Sigma_{22}\gamma - 1\right) + \nu\alpha'\Sigma_{11}\alpha_1 + \theta\gamma'\Sigma_{22}\gamma_1$$

- In looking at the above Lagrangian, you may wonder why there are constraints only for $Cov\left(Y_2, Y_1\right) = \alpha_2'\Sigma_{11}\alpha_1 = 0$ and $Cov\left(Z_2, Z_1\right) = \gamma_2'\Sigma_{22}\gamma_1 = 0$ and not for either $Cov\left(Y_2, Z_1\right) = \alpha_2'\Sigma_{12}\gamma_1 = 0$ and $Cov\left(Z_2, Y_1\right) = \gamma_2'\Sigma_{21}\alpha_1 = 0$. It turns out that if the former two constraints are satisfied, then the latter two are automatically satisfied.

- To see this, recall the first-order conditions from step 1 above

$$\frac{\partial \mathcal{L}_1}{\partial \alpha} = \Sigma_{12}\gamma_1 - \lambda_1\Sigma_{11}\alpha_1 = 0$$
$$\frac{\partial \mathcal{L}_1}{\partial \gamma} = \Sigma_{21}\alpha_1 - \lambda_1\Sigma_{22}\gamma_1 = 0$$

Given these relations, note that, if $Cov\left(Y_2, Y_1\right) = \alpha_2'\Sigma_{11}\alpha_1 = 0$, then

$$Cov\left(Y_2, Z_1\right) = \alpha_2'\Sigma_{12}\gamma_1 = \lambda_1\alpha_2'\Sigma_{11}\alpha_1 = 0.$$

In addition, if $Cov\left(Z_2, Z_1\right) = \gamma_2'\Sigma_{22}\gamma_1 = 0$, then

$$Cov\left(Y_1, Z_2\right) = \gamma_2'\Sigma_{21}\alpha_1 = \lambda_1\gamma_2'\Sigma_{22}\gamma_1 = 0.$$

- Next, note that the first-order conditions associated with $\mathcal{L}_2$ are

$$\begin{aligned}\frac{\partial \mathcal{L}_2}{\partial \alpha} &= \Sigma_{12}\gamma_2 - \lambda_2 \Sigma_{11}\alpha_2 + \nu_2 \Sigma_{11}\alpha_1 = 0 \qquad (5)\\ \frac{\partial \mathcal{L}_2}{\partial \gamma} &= \Sigma_{21}\alpha_2 - \mu_2 \Sigma_{22}\gamma_2 + \theta_2 \Sigma_{22}\gamma_1 = 0 \qquad (6)\end{aligned}$$

Premultiply (5) by α_1' and (6) by γ_1', we get

$$\begin{aligned}\alpha_1'\Sigma_{12}\gamma_2 - \lambda_2\alpha_1'\Sigma_{11}\alpha_2 + \nu_2\alpha_1'\Sigma_{11}\alpha_1 &= \nu_2 = 0,\\ \gamma_1'\Sigma_{21}\alpha_2 - \mu_2\gamma_1'\Sigma_{22}\gamma_2 + \theta_2\gamma_1'\Sigma_{22}\gamma_1 &= \theta_2 = 0.\end{aligned}$$

- Hence, the last two constraints in $\mathcal{L}_2$ are not binding.

Canonical Correlation Analysis

- It follows finding this second linear combination pair involves solving the same eigenvalue problem that was given by

$$\begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = 0$$

  and

$$\det \begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} = 0$$

  In particular, let λ_2 be the second largest root of the determinantal equation above and let $(\alpha_2', \gamma_2')'$ be the associated eigenvector; then,

$$\begin{aligned} Y_2 &= \alpha_2' U, \\ Z_2 &= \gamma_2' V \end{aligned}$$

  are the normalized linear combination pair which yields the highest correlation among all pairs not correlated with Y_1 and Z_1 and λ_2 gives the value of their correlation.

- We can proceed in this way to obtain

Canonical Correlations (Eigenvalues)	Weights (Eigenvalues)	Y	Z
λ_1	$(\alpha_1', \gamma_1')'$	$Y_1 = \alpha_1' U$	$Z_1 = \gamma_1' V$
λ_2	$(\alpha_2', \gamma_2')'$	$Y_2 = \alpha_2' U$	$Z_2 = \gamma_2' V$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
λ_{m_1}	$(\alpha_{m_1}', \gamma_{m_1}')'$	$Y_{m_1} = \alpha_{m_1}' U$	$Z_{m_1} = \gamma_{m_1}' V$

- As a final remark, we note that the equations

$$\begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = 0$$

and

$$\det \begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} = 0$$

can be written in a different form which may be more convenient for cointegration analysis.

- In particular, note that, since $\lambda = \mu$, we can rewrite the first-order conditions given by equations (1) and (2) as

$$\Sigma_{12}\gamma - \lambda\Sigma_{11}\alpha = 0 \quad (7)$$
$$\Sigma_{21}\alpha - \lambda\Sigma_{22}\gamma = 0 \quad (8)$$

- Premultiplying expression (8) above by $\Sigma_{12}\Sigma_{22}^{-1}$ yields

$$\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\alpha - \lambda\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22}\gamma = \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\alpha - \lambda\Sigma_{12}\gamma = 0$$

or

$$\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\alpha = \lambda\Sigma_{12}\gamma = \lambda^2\Sigma_{11}\alpha$$

where the last equality above follows from (7).

- It follows that, alternatively, we can obtain the squared canonical correlations and the eigenvectors α_i $(i = 1, ..., m_1)$ as solutions of the system of equations

$$\left(\lambda^2 \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}\right) \alpha = 0$$

where λ_i^2 $(i = 1, ..., m_1)$, i.e., the squared canonical correlations, are the roots of the determinantal equation

$$\det\left(\lambda^2 \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}\right) = 0.$$

- Similarly, it can be shown that λ_i^2 and γ_i $(i = 1, ..., m_1)$ can be obtained as the solutions of the system of equations

$$\left(\lambda^2 \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}\right) \gamma = 0$$

with λ_i^2 $(i = 1, ..., m_1)$ being the m_1 largest roots of

$$\det\left(\lambda^2 \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}\right) = 0.$$