$\mathbb{E}[z \mid x_1]$ and $\mathbb{E}[z \mid x_2]$ in pCCA have correlations equal to the canonical correlations

Akash Umakantha

January 9, 2016

Starting with the pCCA model,

$$
\begin{aligned}
z &\sim N(0, I_d) \\
x_1 \mid z &\sim N(W_1 z + \mu_1, \psi_1) \\
x_2 \mid z &\sim N(W_2 z + \mu_2, \psi_2)
\end{aligned}
$$

The joint distribution for z and x_1 is

$$
\begin{bmatrix} z \\ x_1 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \mu_1 \end{bmatrix}, \begin{bmatrix} I_d & W_1^T \\ W_1 & W_1 W_1^T + \psi_1 \end{bmatrix}\right)
$$

Conditioning z on x_1, we get

$$
\mathbb{E}[z \mid x_1] = 0 + W_1^T \left(W_1 W_1^T + \psi_1\right)^{-1} (x_1 - \mu_1)
$$

From pg. 4 of the Bach and Jordan pCCA paper, the maximum likelihood solutions for W_1 and ψ_1 are

$$
\begin{aligned}
\hat{W}_1 &= \Sigma_{11} U_{1d} P_d^{1/2} \\
\hat{\psi}_1 &= \Sigma_{11} - \hat{W}_1 \hat{W}_1^T
\end{aligned}
$$

Where Σ_{11} is the sample covariance of x_1, U_{1d} are the first d canonical vectors from CCA between x_1 and x_2, and P_d is a diagonal matrix containing the first d canonical correlations.

Plugging $\hat{W}_1$ and $\hat{\psi}_1$ into the expectation and simplifying gives

$$
\mathbb{E}[z \mid x_1] = P_d^{1/2} U_{1d}^T (x_1 - \mu_1)
$$

This is the result for the posterior expectation from pg. 4 of the pCCA paper. Similarly for $z \mid x_2$,

$$
\mathbb{E}[z \mid x_2] = P_d^{1/2} U_{2d}^T (x_2 - \mu_2)
$$

Let

$$
\begin{aligned}
A &= U_{1d}^T (x_1 - \mu_1) \\
B &= U_{2d}^T (x_2 - \mu_2)
\end{aligned}
$$

So that

$$
\begin{aligned}
\mathbb{E}[z \mid x_1] &= P_d^{1/2} A \\
\mathbb{E}[z \mid x_2] &= P_d^{1/2} B
\end{aligned}
$$

Note that A and B are just the CCA scores ($\mathbb{R}^{d\times N}$), and the row-wise correlations between them are equal to the canonical correlations. Since P_d is a diagonal matrix, the row-wise correlations between $\mathbb{E}[z \mid x_1]$ and $\mathbb{E}[z \mid x_2]$ are also exactly equal to the canonical correlations.

To see this, take for example d=2.

$$\mathbb{E}[z \mid x_1] = \begin{bmatrix} \sqrt{p_1} & 0 \\ 0 & \sqrt{p_2} \end{bmatrix} \begin{bmatrix} - - \vec{a_1} - - \\ - - \vec{a_2} - - \end{bmatrix} = \begin{bmatrix} - - \sqrt{p_1}\vec{a_1} - - \\ - - \sqrt{p_2}\vec{a_2} - - \end{bmatrix}$$

$$\mathbb{E}[z \mid x_2] = \begin{bmatrix} \sqrt{p_1} & 0 \\ 0 & \sqrt{p_2} \end{bmatrix} \begin{bmatrix} - - \vec{b_1} - - \\ - - \vec{b_2} - - \end{bmatrix} = \begin{bmatrix} - - \sqrt{p_1}\vec{b_1} - - \\ - - \sqrt{p_2}\vec{b_2} - - \end{bmatrix}$$