Partial Least Square (PLS)

Nanbo Sun

This document is about the derivation of Partial Least Squares Correlation (PLSC) which is mainly used in neuroimaging.

1 Problem Statement

Given two random vectors $\vec{x}, \vec{y}$, which has been demeaned, we want to find two transformation $\vec{w}, \vec{c}$

$$t = \vec{w}^T \vec{x}$$

$$u = \vec{c}^T \vec{y}$$

(1)

(2)

to maximize the covariance between random variable $t$ and $u$

$$\max(E[tu])^2$$

(3)

under the condition that

$$\vec{w}^T \vec{w} = 1$$

$$\vec{c}^T \vec{c} = 1$$

(4)

(5)

2 Solution

$$\max_{||\vec{w}||=1,||\vec{c}||=1} (E[zu])^2 = \max(E[\vec{w}^T \vec{x} \vec{y}^T \vec{c}])^2$$

$$= \max(\vec{w}^T C_{xy} \vec{c})^2$$

(6)

(7)

By using Lagrange multiplier, we get following equation

$$L(\vec{w}, \vec{c}) = (\vec{w}^T C_{xy} \vec{c})^2 - \lambda_w(\||\vec{w}\||-1) - \lambda_c(\||\vec{c}\||-1)$$

(8)

$$\frac{\partial L}{\partial w} = 2(\vec{w}^T C_{xy} \vec{c}) C_{xy} \vec{c} - 2\lambda_w \vec{w} = 0$$

(9)

$$\frac{\partial L}{\partial c} = 2(\vec{w}^T C_{xy} \vec{c})(\vec{w}^T C_{xy})^T - 2\lambda_c \vec{c} = 0$$

(10)
Times $\vec{w}^T$, $\vec{c}^T$ in both sides

\[ \begin{align*}
\vec{w}^T (\vec{w}^T C_{xy} \vec{c}) C_{xy} \vec{c} &= \lambda_w \vec{w}^T \vec{w} \\
\vec{c}^T (\vec{w}^T C_{xy} \vec{c}) C_{yx} \vec{w} &= \lambda_c \vec{c}^T \vec{c}
\end{align*} \] (11)

Since $\vec{w}^T \vec{w} = 1$, $\vec{c}^T \vec{c} = 1$

\[ \begin{align*}
(\vec{w}^T C_{xy} \vec{c})^2 &= \lambda_w \\
(\vec{w}^T C_{xy} \vec{c})^2 &= \lambda_c
\end{align*} \] (12)

It’s easy to see that $\lambda = \lambda_w = \lambda_c$.

We can also write the above function in this way

\[ \vec{w}^T C_{xy} \vec{c} = \sqrt{\lambda} \] (13)

Substitute above Equation into Eq. (9) and Eq. (10), we can get following

\[ \begin{align*}
C_{xy} \vec{c} &= \sqrt{\lambda} \vec{w} \\
C_{yx} \vec{w} &= \sqrt{\lambda} \vec{c}
\end{align*} \] (14)

The above equation shows that $\vec{w}, \vec{c}$ are singular vectors of $C_{xy}$.

\[ \begin{bmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{bmatrix} \begin{bmatrix} \vec{w} \\ \vec{c} \end{bmatrix} = \sqrt{\lambda} \begin{bmatrix} \vec{w} \\ \vec{c} \end{bmatrix} \] (15)

We can rewrite Eq. (17) as

\[ \vec{c} = \frac{1}{\sqrt{\lambda}} C_{yx} \vec{w} \] (16)

Then substitute Eq. (19) into Eq. (16)

\[ C_{xy} C_{yx} \vec{w} = \lambda \vec{w} \] (17)

Similarly, we can get

\[ C_{yx} C_{xy} \vec{c} = \lambda \vec{c} \] (18)

**Theorem: Singular Value Decomposition (SVD)**

If $A = U S V^T$, then $U$ and $V$ are eigenvectors of $AA^T$ and $A^T A$

\[ \begin{align*}
A &= U S V^T \\
U^T U &= V^T V = I \\
A A^T &= U S V^T V S U^T U = U S^2 \\
A^T A &= V S U^T V S U V^T V = V S^2 \\
\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} &= \begin{bmatrix} U \\ V \end{bmatrix} S
\end{align*} \] (19)
By comparing Eq. (20) and (24), we can know that \( \vec{w} \) is a singular vector of \( C_{xy} \) and \( \vec{c} \) is a singular vector of \( C_{yx} \). Therefore, by doing a simple SVD of \( C_{xy} \) we can get \( \vec{w} \) from \( U \), \( \vec{c} \) from \( V \). And, this matches the solution in the paper "Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review".