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 $\vec{w}, \vec{c}$

$$t = \vec{w}^T \vec{x} \tag{1}$$

$$u = \vec{c}^T \vec{y} \tag{2}$$

to maximize the covariance between random variable t and u

$$\max(E[tu])^2 \tag{3}$$

under the condition that

$$\vec{w}^T \vec{w} = 1 \tag{4}$$

$$\vec{c}^T \vec{c} = 1 \tag{5}$$

2 Solution

$$\max_{||\vec{w}||=1, ||\vec{c}||=1} (E[tu])^2 = \max(E[\vec{w}^T \vec{x} \vec{y}^T \vec{c}])^2 \tag{6}$$

$$= \max(\vec{w}^T C_{xy} \vec{c})^2 \tag{7}$$

By using Lagrange multiplier, we get following equation

$$L(\vec{w}, \vec{c}) = (\vec{w}^T C_{xy} \vec{c})^2 - \lambda_w(||\vec{w}|| - 1) - \lambda_c(||\vec{c}|| - 1) \tag{8}$$

$$\frac{\partial L}{\partial w} = 2(\vec{w}^T C_{xy} \vec{c}) C_{xy} \vec{c} - 2\lambda_w \vec{w} = 0 \tag{9}$$

$$\frac{\partial L}{\partial c} = 2(\vec{w}^T C_{xy} \vec{c})(\vec{w}^T C_{xy})^T - 2\lambda_c \vec{c} = 0 \tag{10}$$

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

$$\vec{w}^T(\vec{w}^T C_{xy}\vec{c})C_{xy}\vec{c} = \lambda_w \vec{w}^T\vec{w} \tag{11}$$

$$\vec{c}^T(\vec{w}^T C_{xy}\vec{c})C_{yx}\vec{w} = \lambda_c \vec{c}^T\vec{c} \tag{12}$$

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

$$(\vec{w}^T C_{xy}\vec{c})^2 = \lambda_w \tag{13}$$

$$(\vec{w}^T C_{xy}\vec{c})^2 = \lambda_c \tag{14}$$

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

We can also write the above function in this way

$$\vec{w}^T C_{xy}\vec{c} = \sqrt{\lambda} \tag{15}$$

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

$$C_{xy}\vec{c} = \sqrt{\lambda}\vec{w} \tag{16}$$

$$C_{yx}\vec{w} = \sqrt{\lambda}\vec{c} \tag{17}$$

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

$$\begin{bmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{bmatrix} \begin{bmatrix} \vec{w} \\ \vec{c} \end{bmatrix} = \sqrt{\lambda} \begin{bmatrix} \vec{w} \\ \vec{c} \end{bmatrix} \tag{18}$$

We can rewrite Eq. (17) as

$$\vec{c} = \frac{1}{\sqrt{\lambda}} C_{yx}\vec{w} \tag{19}$$

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

$$C_{xy}C_{yx}\vec{w} = \lambda\vec{w} \tag{20}$$

Similarly, we can get

$$C_{yx}C_{xy}\vec{c} = \lambda\vec{c} \tag{21}$$

Theorem: Singular Value Decomposition (SVD)

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

$$A = USV^T \tag{22}$$

$$U^TU = V^TV = I \tag{23}$$

$$AA^TU = USV^TVSU^TU = US^2 \tag{24}$$

$$A^TAV = VSU^TUSV^TV = VS^2 \tag{25}$$

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} S \tag{26}$$

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".