CCA → Linear Regression: Notes

Guohao Dou

January 30, 2019

We can use Lagrangian multiplier on CCA to find out

$$C_x^{-1}C_{xy}C_y^{-1}C_{yx}w_x = \lambda^2 w_x$$

where $C_x = XX^T, C_y = YY^T, C_{xy} = XY^T, C_{yx} = YX^T$

Then we take the top eigenvector of $C_x^{-1}C_{xy}C_y^{-1}C_{yx}$ to find the solution to CCA.

The trick is to analytically diagonalize $C_x^{-1}C_{xy}C_y^{-1}C_{yx}$

Define $H = Y^T(YY^T)^{-1/2}$, also svd $x = U\Sigma V^T = [U_1, U_2]\mathrm{diag}(\Sigma_r, 0)[V_1, V_2]^T = U_1\Sigma_r V_1^T$

$$\begin{aligned} C_x^{-1}C_{xy}C_y^{-1}C_{yx} &= (XX^T)^{-1}XHH^TX^T \\ &= (U_1\Sigma_r^{-2}U_1^T) \cdot XHH^TX^T \\ &= U_1\Sigma_r^{-1} \cdot (\Sigma_r^{-1}U_1^TXH)H^TX^T \end{aligned}$$

Define $A = \Sigma_r^{-1}U_1^TXH$, also svd $A = P\Sigma_A Q^T$

$$\begin{aligned} C_x^{-1}C_{xy}C_y^{-1}C_{yx} &= U_1\Sigma_r^{-1} \cdot A \cdot H^TX^TUU^T \\ &= U\begin{bmatrix} I_r \\ 0 \end{bmatrix}\Sigma_r^{-1}A[H^TX^TU_1, H^TX^TU_2]U^T \\ &= U\begin{bmatrix} I_r \\ 0 \end{bmatrix}\Sigma_r^{-1}A[H^TX^TU_1, 0]U^T \\ &= U\begin{bmatrix} I_r \\ 0 \end{bmatrix}\Sigma_r^{-1}A[A^T\Sigma_r, 0]U^T \\ &= U\begin{bmatrix} \Sigma_r^{-1}AA^T\Sigma_r & 0 \\ 0 & 0 \end{bmatrix}U^T \\ &= U\begin{bmatrix} \Sigma_r^{-1}P\Sigma_A^2P^T\Sigma_r & 0 \\ 0 & 0 \end{bmatrix}U^T \\ &= U\begin{bmatrix} \Sigma_r^{-1}P & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} \Sigma_A^2 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} P^T\Sigma_r & 0 \\ 0 & I \end{bmatrix}U^T \end{aligned}$$

Note that

$$U\begin{bmatrix} \Sigma_r^{-1}P & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} P^T\Sigma_r & 0 \\ 0 & I \end{bmatrix}U^T = I$$

which means $C_x^{-1}C_{xy}C_y^{-1}C_{yx}$ is similar to $\begin{bmatrix}\Sigma_A^2 & 0\\ 0 & 0\end{bmatrix}$ and $U_1\Sigma_r P$ are the top eigenvectors.

Now consider the linear regression problem $\left(X,\tilde{Y}\right)$ where $\tilde{Y} = H^T = (YY^T)^{-1/2}Y$

The solution is the well-known

$$\begin{aligned}(XX^T)^{-1}X\tilde{Y}^T &= (XX^T)^{-1}XH\\ &= U_1\Sigma_r^{-2}U_1^TXH\\ &= U_1\Sigma_r^{-1}(\Sigma_r^{-1}U_1^TXH)\\ &= U_1\Sigma_r^{-1}A\\ &= U_1\Sigma_r^{-1}P\Sigma_AQ^T\end{aligned}$$

It can be shown that $\Sigma_A = I$.

Then

$$(XX^T)^{-1}X\tilde{Y}^T = (U_1\Sigma_r^{-1}P)Q^T$$

The main result: suppose we have a CCA problem (X, Y) with a solution W_{CCA} and a linear regression problem $(X, (YY^T)^{-1/2}Y)$ with a solution W_{LS}, then $W_{LS} = W_{CCA}Q^T$. That is, the two differ by an orthogonal matrix.