1.A. Canonical Correlation Analysis :

Given the problem,

leads to an optimization problem:

$$\text{maximise : } \omega_x^T C_{xy} \omega_y ,$$

$$\text{given - } \omega_x^T C_{xx} \omega_x = 1 \ , \ \omega_y^T C_{yy} \omega_y = 1.$$

1. Formulate the Lagrangian fxn -

$$\mathcal{L} = \omega_x^T C_{xy} \omega_y - \frac{\lambda_x}{2}\left(\omega_x^T C_{xx} \omega_x - 1\right) - \frac{\lambda_y}{2}\left(\omega_y^T C_{yy} \omega_y - 1\right)$$

2. Taking differentials and solve $\nabla \mathcal{L} = 0$

$$\frac{\partial L}{\partial \omega_x} = C_{xy}\omega_y - \lambda_x C_{xx}\omega_x = 0$$

$$\frac{\partial L}{\partial \omega_y} = C_{yx}\omega_x - \lambda_y C_{yy}\omega_y = 0 .$$

3. Given the constraints.

$$C_{xy}\omega_y = \lambda_x C_{xx}\omega_x$$

$$C_{yx}\omega_x = \lambda_y C_{yy}\omega_y$$

4. $\because$ solving for maximally correlating components, $\lambda_x = \lambda_y = \lambda$.

$$\therefore \ C_{xy}\omega_y = \lambda C_{xx}\omega_x \quad , \quad C_{yx}\omega_x = \lambda C_{yy}\omega_y$$

$$\Rightarrow \ C_{xx}^{-1} C_{xy}\omega_y = \lambda \omega_x . \quad , \quad C_{yy}^{-1} C_{yx}\omega_x = \lambda \omega_y .$$

5. System of equation,

$$\begin{bmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix} = \lambda \begin{bmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{bmatrix} \underbrace{\begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix}}_{= z} .$$

$\therefore$ The solution for CCA corresponds to eigenvectors $z = \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix}$ for the above ! $\bar{c}$ eigenvalues λ.

1.B. Show

Solution to CCA is the eigenvector associated with largest eigenvalue λ, because ...

1. eigenvalue reports the maximum possible correlation between the projection of X and Y.

2. CCA for High Dimensional Data.

To show that optimal solution w_x, w_y for CCA as a linear combination of data matrices X and Y.

Note CCA finds linear projection w_x, w_y such that correlation between projected data Xw_x and Yw_y is maximum.

1. Represent w_x, w_y in span of data.

Let us prove that
$w_x = X\alpha_x$
$w_y = Y\alpha_y$ are optimal solution.

- Projection of data onto w_x, w_y are:
$$Xw_x = X(X\alpha_x) = (X^T X)\alpha_x.$$
$$Yw_y = Y(Y\alpha_y) = (Y^T Y)\alpha_y.$$
$\therefore$ Correlation can be expressed in terms of α_x, α_y.

- Optimality.

Decompose w_x, w_y into components:

- Span $\therefore$ they are in span of data matrices X & Y.
- Orthogonal: there are orthogonal to data matrices.
$$X^T w_x = 0$$
$$Y^T w_y = 0.$$

Thus,
$$w_x = w_x^{(s)} + w_x^{(n)}, \quad w_x^{(s)} = X\alpha_x$$
$$w_y = w_y^{(s)} + w_y^{(n)} \quad w_y^{(s)} = Y\alpha_y.$$

2.b.

In the original problem,

replace w_x, w_y with $X\alpha_x, Y\alpha_y$.

given $(X\alpha_x)^T C_{xy} (Y\alpha_y)$ maximize

with $(X\alpha_x)^T C_{xx} (X\alpha_x) = 1$

$(Y\alpha_y)^T C_{yy} (Y\alpha_y) = 1$

and substitute C with Q,

given $\alpha_x^T Q_{xy} \alpha_y$ maximize

with $\alpha_x^T Q_{xx} \alpha_x = 1$

$\alpha_y^T Q_{yy} \alpha_y = 1$

NOTE : same structure as original CCA ....

Following same as Exercise 1, we get the eigenvalue problem as

$$\begin{bmatrix} 0 & Q_{xy} \\ Q_{yx} & 0 \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} = \lambda \begin{bmatrix} Q_{xx} & 0 \\ 0 & Q_{yy} \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix}$$

c. Same as 1b, multiply both sides of eigenvalue by $(\alpha_x, \alpha_y)^T$ and find the pair that maximizes the objective is the one with highest eigenvalue λ.

d. If, $w_x = X\alpha_x$

$w_y = Y\alpha_y$

so we can compute w_x, w_y with α_x, α_y.