Lecture 2

Canonical Correlation Analysis:

~~→ maximize correlation~~

→ if two datasets have vectors

$X = \{x_1, \dots, x_N\}$, $Y = \{y_1, \dots, y_N\}$

→ find linear combination of x and y that maximizes their correlation!

Assumption: Centered data $\sum_i x_i = \sum_i y_i = 0$

Cross-covariance: $C_{xy} = \frac{1}{N} X Y^T$

Auto-covariance: $C_{xx} = \frac{1}{N} X X^T$

Given: $X \in \mathbb{R}^M$, $Y \in \mathbb{R}^N$

Needed: Projections $w_x \in \mathbb{R}^M$, $w_y \in \mathbb{R}^N$ that maximize correlation

Objective:

$$\underset{w_x, w_y}{\operatorname{argmax}} \left(w_x X Y w_y \right)$$

Conditions / Constraint:

$$w_x^T X X^T w_x = 1 = w_y^T Y Y^T w_y$$

Lagrangian:

$$\mathcal{L} = w_x^T C_{xy} w_y - \frac{1}{2}\alpha (w_x^T C_{xx} w_x - 1) - \frac{1}{2}\beta (w_y^T C_{yy} w_y - 1)$$

Partial Derivative:

$$\frac{\partial \mathcal{L}}{\partial w_x^T} = C_{xy} w_y - \tilde{\alpha} C_{xx} w_x \overset{!}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial w_y^T} = C_{yx} w_x - \tilde{\beta} C_{yy} w_y \overset{!}{=} 0$$

$$\frac{\partial x^T a}{\partial x} = \frac{\partial a^T x}{\partial x} = a^T$$

$$\Rightarrow \begin{array}{ll} C_{xy} w_y = \tilde{\alpha} C_{xx} w_x \quad | \cdot w_x^T & w_x^T C_{xy} w_y = \tilde{\alpha} w_x^T C_{xx} w_x \\ & \Leftrightarrow \\ C_{yx} w_x = \tilde{\beta} C_{yy} w_y \quad | \cdot w_y^T & w_x^T C_{xy} w_y = \tilde{\beta} C_{yy} w_y \end{array}$$

$$w_x^T C_{xy} w_y = \tilde{\alpha}\, w_x^T C_{xx} w_x$$

$$w_x^T C_{xy} w_y = \tilde{\beta}\, w_y^T C_{yy} w_y$$

From constraint $w_x^T C_{xx} w_x = w_y^T C_{yy} w_y = 1$

$\Rightarrow \tilde{\alpha} = \tilde{\beta}$

<u>Partial Derivatives:</u>

$$C_{xy} w_y = \tilde{\alpha}\, C_{xx} w_x$$

$$C_{yx} w_x = \tilde{\alpha}\, C_{yy} w_y$$

<u>Matrix Form:</u>

$$\begin{pmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{pmatrix} \begin{pmatrix} w_x \\ w_y \end{pmatrix} = \tilde{\alpha} \begin{pmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{pmatrix} \begin{pmatrix} w_x \\ w_y \end{pmatrix}$$

→ Eigenvalue equation

Problem:

- covariance matrix too big
- non-linear dependencies

} → kernel CCA

- variables are coupled with delay τ
  - → tkCCA: shift variables, calc correlation for all temporal time lags τ, maximize correlation

<u>Summary</u>:

CCA: finds projections for two datasets that maximizes their correlation

kCCA: extends CCA to non-linear or high-dimensional data

temporal kCCA: extends kCCA to non-instantaneous ~~data~~ correlations, computes multivariate convolution from one modality to another (?)