3. (a)

unconstrained objective

$$\max_{\varphi_x \varphi_y \omega_x \omega_y} \omega_x^T C_{xy} \omega_y + \alpha . [\min(0, 1 - \omega_x^T C_{xx} \omega_x + \min(0, 1 - \omega_y^T C_{yy} \omega_y)]$$

for $\omega_x \in R^{h_1}$ and $\omega_y \in R^{h_2}$

(b) original CCA objective

$$\max_{\omega_x, \omega_y} \omega_x^T C_{xy} \omega_y \cancel{+\alpha [\min(}$$

$$\text{ST} \quad \omega_x^T C_{xx} \omega_x = 1$$

$$\omega_y^T C_{yy} \omega_y = 1$$

⇒ As, here the function $\phi_x(x;\varphi_x)$ and $\phi_y(y;\phi_y)$ is nothing but passing the input views in a stacked non-linear representation. so, on the equation given above it is simply calculating CCA over the outputs of ϕ functions.

This arrangement represents exactly the CCA form st performed linearly. and can be represented.

$$\boxed{(\varphi_x^*, \varphi_y^*, \omega_x, \omega_y) = \max_{\varphi_x, \varphi_y, \omega_x, \omega_y} \text{corr}\left(\omega_x \phi(x, \varphi_x), \omega_y \phi(y, \varphi_y)\right)}$$

~~so, to~~

To work with the constraints :-

$$\text{orig. CCA} = \max_{\omega_x, \omega_y} \omega_x^T C_{xy} \omega_y$$

$$\omega_x^T C_{xx} \omega_x = 1$$
$$\omega_y^T C_{yy} \omega_y = 1.$$

replace co-variance matrices from linear to non-linear.

$$C_{xx} = \mathbb{E}[\phi_x \phi_x^T]$$
$$C_{yy} = \mathbb{E}[\phi_y \phi_y^T]$$
$$C_{xy} = \mathbb{E}[\phi_x \phi_y^T]$$

in unconstrained environmen

$$\omega_x^T C_{xx} \omega_x - 1 = 0$$
$$\omega_y^T C_{yy} \omega_y - 1 = 0$$

$$\Rightarrow \min[0, 1 - \omega_x^T C_{xx} \omega_x] = 0 = X_{min}$$
$$\Rightarrow \min[0, 1 - \omega_y^T C_{yy} \omega_y] = 0 = Y_{min}$$

penalty

so, $$\max_{\phi_x, \phi_y, \omega_x, \omega_y} \omega_x^T C_{xy} \omega_y + \lambda * X_{min} + \lambda * Y_{min}$$

$$= \max_{\phi_x, \phi_y, \omega_x, \omega_y} \underline{\omega_x^T C_{xy} \omega_y} + \lambda [\min(0, 1 - \omega_x^T C_{xx} \omega_x) + \min(0, 1 - \omega_y^T C_{yy} \omega_y)]$$

exactly same as DCCA

or CCA in non-linear manner.

3. b

Gradient as a function of Jackobian matrix.

$$f = w_x^T C_{xy} w_y + \alpha \left[\min(0, 1 - w_x^T C_{xx} w_x) + \min(0, 1 - w_y^T C_{yy} w_y) \right]$$

gradient of f with respect to ϕ_x

$$= w_x^T \frac{C_{xy}}{d\phi_x} w_y + \alpha \left(- w_x^T \frac{C_{xx}}{d\phi_x} w_x \right) \to 0$$

$$w_x^T \frac{E[\phi_x \phi_y^T]}{d\phi_x} w_y - \alpha \left(w_x^T \frac{E(\phi_x \phi_x^T)}{d\phi_x} w_x \right) \to 0$$

$$w_x^T E\left(\frac{d\phi_x}{d\phi_x} \cdot \frac{d\phi_y}{d\phi_x} \right) w_y - \alpha \left(w_x^T E\left(\frac{d\phi_x}{d\phi_x} \frac{d\phi_x^T}{d\phi_x} \right) w_x \right) \to 0$$

ϕ_x and ϕ_y are independent

$$w_x^T E\left(\frac{d\phi_x}{d\phi_x} \right) E\left(\frac{d\phi_y}{d\phi_x} \right) w_y - \alpha \left(w_x^T E\left(2 \frac{d\phi_x}{d\phi_x} \right) w_x \right)$$

$$= 0 - 2\alpha \, w_x^T E\left(\frac{d\phi_x}{d\phi_x} \right) w_x$$

$$\Delta_{\phi_x} = -2\alpha \, w_x^T E\left(\frac{d\phi_x}{d\phi_x} \right) w_x \Rightarrow \text{solution.}$$