ECE C147/C247 Winter 2023 HW3

NLDL

Yaman Yucel

605704529

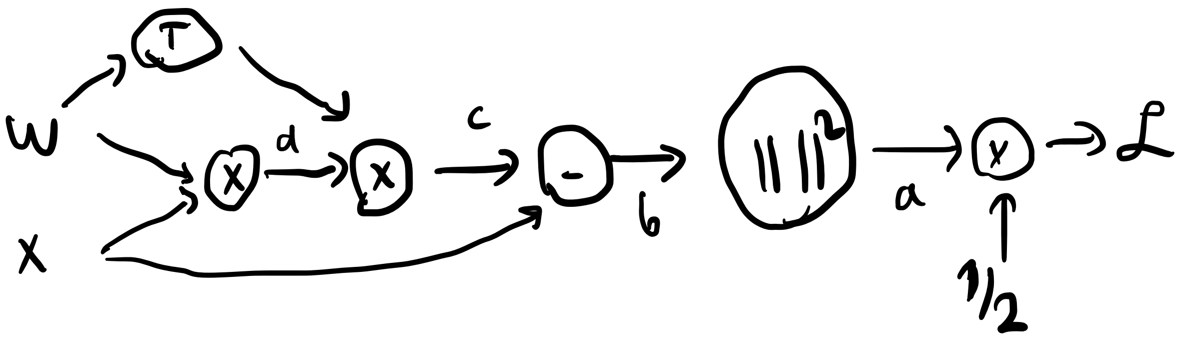
1) a)

We project vector x using W linear transformation, and in order to reconstruct projected vector, we need to multiply the projected vector Wx with W^T. Then, reconstructed vector is $W^TWx = y$. Projecting to lower dimension reduces the information of the vector. So, we try to minimize the information lost by using the loss function. At the end of training, we get W such that it preserves the information and projects data into a lower dimension.

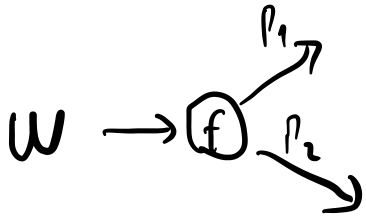
$y = W^TWx$ → reconstructed vector

minimize $\frac{1}{2}\|y - x\|_2^2$, tries to make y look like x as possible as not encoded in the first place

b) $\mathcal{L} = \frac{1}{2} \| w^T w x - x \|^2$



c)



$$\frac{\partial \mathcal{L}}{\partial w} = \frac{\partial p_1}{\partial w} \frac{\partial \mathcal{L}}{\partial p_1} + \frac{\partial p_2}{\partial w} \frac{\partial \mathcal{L}}{\partial p_2}$$

We need to sum derivatives coming from each path to find total derivative.

d)

1) $\mathcal{L} = \frac{a}{2} \Rightarrow \frac{\partial \mathcal{L}}{\partial a} = \frac{1}{2}$

2) $a = b^T b \Rightarrow \frac{\partial a}{\partial b} = 2b \Rightarrow \frac{\partial a}{\partial b} \cdot \frac{\partial \mathcal{L}}{\partial a} = 2b\frac{1}{2} = b$

3) $b = c - x \Rightarrow \frac{\partial b}{\partial c} = I \Rightarrow \frac{\partial b}{\partial c} \frac{\partial \mathcal{L}}{\partial b} = I \cdot b = b$

4) $c = w^T d \Rightarrow \frac{\partial c}{\partial d} = w \Rightarrow \frac{\partial c}{\partial d} \frac{\partial \mathcal{L}}{\partial c} = wb$

5) $c = w^T d \Rightarrow \frac{\partial c}{\partial w^T} = d^T \Rightarrow \frac{\partial c}{\partial w^T} \frac{\partial \mathcal{L}}{\partial c} = b d^T$

6) $d = wx \Rightarrow \frac{\partial d}{\partial w} = x^T \Rightarrow \frac{\partial d}{\partial w} \frac{\partial \mathcal{L}}{\partial d} = wbx^T$

$\nabla_w \mathcal{L} = 5^T + 6 = \left(\frac{\partial \mathcal{L}}{\partial w^T}\right)^T + \frac{\partial \mathcal{L}}{\partial w} = db^T + wbx^T$

$$\boxed{\Rightarrow wbx^T + wxb^T, \quad b = w^T w x - x}$$

2) $K = \alpha X X^T + \beta^{-1} I$

$$\mathcal{L} = -c - \frac{D}{2} \log|K| - \frac{1}{2} \operatorname{tr}(K^{-1} Y Y^T)$$

Note, $K^T = K$, $K^{-1} = K^{-T}$

$$\mathcal{L}_1 = -\frac{D}{2} \log|K|$$

$$\mathcal{L}_2 = -\frac{1}{2} \operatorname{tr}(K^{-1} Y Y^T)$$

a) $\mathcal{L}_1$ graph:

$$\begin{tikzcd}
 & & & & & & -D/2 \arrow[d] & \\
X \arrow[r, "e"] \arrow[d] & \times \arrow[r, "d"] & \times \arrow[r, "c"] & + \arrow[r, "K"] & |\cdot| \arrow[r, "b"] & \log \arrow[r, "a"] & \times \arrow[r] & \mathcal{L}_1 \\
X^T \arrow[ur, "f"] & & \alpha \arrow[u] & \beta^{-1} I \arrow[u] & & & &
\end{tikzcd}$$

b) $$\frac{\partial \mathcal{L}_1}{\partial a} = -\frac{D}{2}$$

$$a = \log|K|$$

$$\frac{\partial a}{\partial K} = K^{-T} \Rightarrow \frac{\partial a}{\partial K}\frac{\partial \mathcal{L}_1}{\partial a} = -\frac{D}{2} K^{-T}$$

$$K = c + \beta^{-1} I \Rightarrow \frac{\partial K}{\partial K} = I \Rightarrow \frac{\partial K}{\partial c}\frac{\partial \mathcal{L}_1}{\partial K} = -\frac{D}{2} K^{-T}$$

$$c = \alpha d \Rightarrow \frac{\partial c}{\partial d} = \alpha I \Rightarrow \frac{\partial c}{\partial d}\frac{\partial \mathcal{L}_1}{\partial c} = -\alpha \frac{D}{2} K^{-T}$$

$$d = XX^T = Xf$$

$$\frac{\partial \mathcal{L}_1}{\partial X} = \frac{\partial \mathcal{L}}{\partial d}(X^T)^T = -\alpha \frac{D}{2} K^{-T} X$$

$$\frac{\partial \mathcal{L}_1}{\partial X^T} = X^T \frac{\partial \mathcal{L}}{\partial d} = X^T\left(-\alpha \frac{D}{2} K^{-T}\right)$$

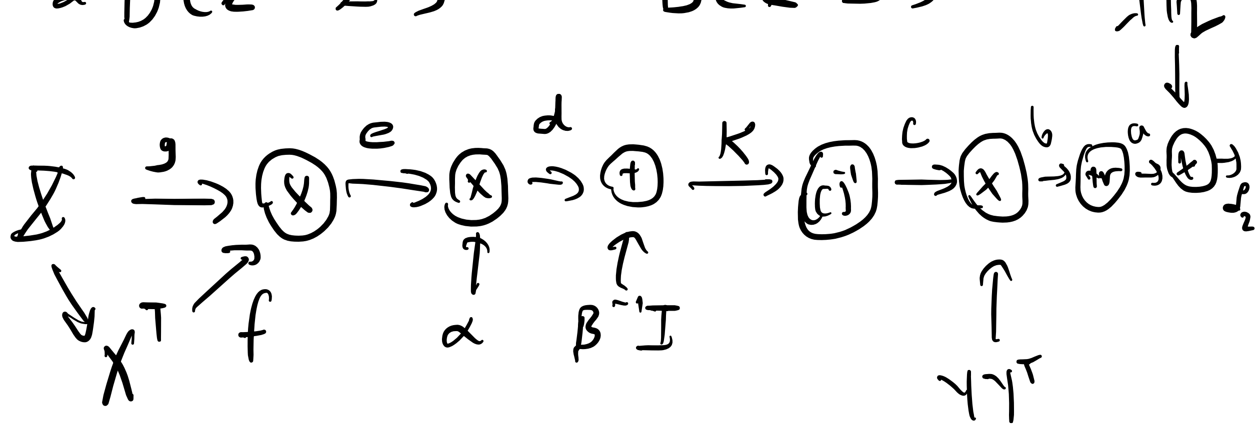
$$\frac{\partial \mathcal{L}_1}{\partial X} = -\alpha \frac{D}{2} K^{-T} X + \left(\frac{\partial \mathcal{L}_1}{\partial X^T} \right)^T$$

$$= -\alpha \frac{D}{2} \left(K^{-T} X + K^{-1} X \right)$$

↓ K symmetric

$$= -\alpha D \left(K^{-T} X \right) = -\alpha D \left(K^{-1} X \right)$$

c)



d) $$\frac{\partial \mathcal{L}_2}{\partial K} = -K^T \frac{\partial \mathcal{L}_2}{\partial K^{-1}} K^{-T}$$

$$\frac{\partial \mathcal{L}_2}{\partial a} = -\frac{1}{2} \qquad \frac{\partial b}{\partial c} = YY^T$$

$$\frac{\partial a}{\partial b} = I$$

$$\frac{\partial \mathcal{L}_2}{\partial K^{-1}} = \frac{\partial b}{\partial K^{-1}} \frac{\partial a}{\partial b} \frac{\partial \mathcal{L}_2}{\partial a} = -\frac{1}{2} y y^T$$

$\downarrow C = K^{-1}$

$$\frac{\partial \mathcal{L}_2}{\partial K} = \frac{1}{2} K^{-T} y y^T K^{-T}$$

$$\frac{\partial K}{\partial d} = I \Rightarrow \frac{\partial \mathcal{L}_2}{\partial d} = \frac{1}{2} K^{-T} y y^T K^{-T}$$

$d = \alpha e$

$$\frac{\partial \mathcal{L}_2}{\partial d} = \frac{1}{2} \alpha K^{-T} y y^T K^{-T} = \alpha \frac{\partial \mathcal{L}_2}{\partial K}$$

$e = X X^T$

$$\frac{\partial \mathcal{L}_2}{\partial X} = \frac{\partial \mathcal{L}_2}{\partial e} X = \alpha \frac{\partial \mathcal{L}_2}{\partial K} X$$

$$\frac{\partial \mathcal{L}_2}{\partial X^T} = \alpha X^T \frac{\partial \mathcal{L}_2}{\partial K}$$

$$\frac{\partial L_2}{\partial X} = \frac{\alpha}{2} K^{-T} Y Y^T K^{-T} X + \frac{\alpha}{2} K^{-1} Y Y^T K^{-1} X$$

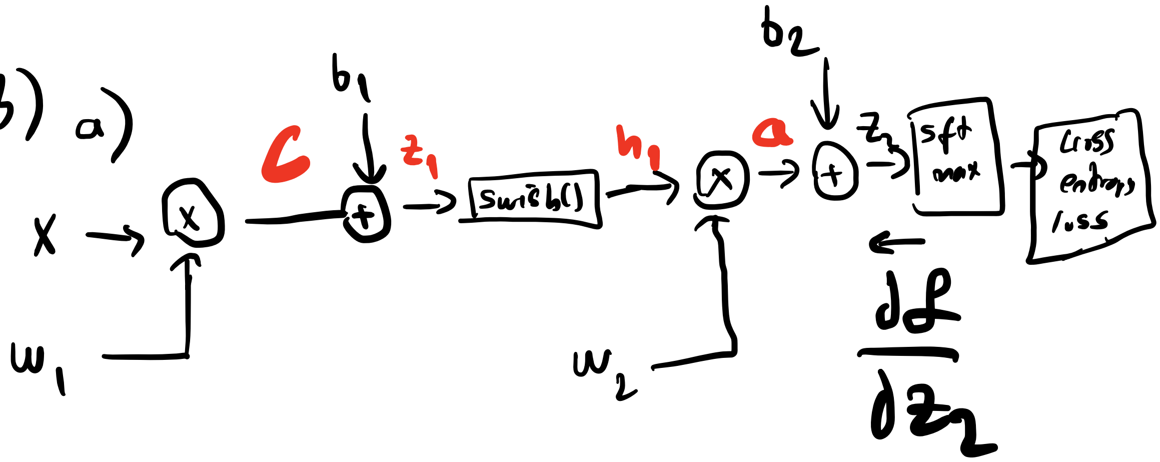
$$= \alpha K^{-1} Y Y^T K^{-1} X$$

$K^{-1} = K^{-T}$

e) sum of b) + d)

$$\Rightarrow \boxed{\begin{aligned} &= -\alpha D K^{-1} X + \alpha K^{-1} Y Y^T K^{-1} X \\ &= -\alpha D K^{-T} X + \alpha K^{-T} Y Y^T K^{-T} X \end{aligned}}$$

3) a)



Derivative of swish:

$y = swish(x)$, $y = x\,\sigma(x)$

$$y' = \sigma(x) + \sigma'(x)\,x$$
$$= \sigma(x) + \sigma(x)(1-\sigma(x))\,x$$
$$= \sigma(x)(1 + x - \sigma(x)\,x)$$
$$= \sigma(x) + \underbrace{\sigma(x)\,x}_{swish} - \sigma(x)\,\overline{\sigma(x)\,x}^{\,}_{swish}$$
$$= \sigma(x) + swish - \sigma(x)\,swish$$
$$\boxed{= swish(x) + \sigma(x)(1 - swish(x))}$$

$\frac{\partial \mathcal{L}}{\partial z_2}$, no need to compute

$$\frac{\partial \mathcal{L}}{\partial a} = \frac{\partial \mathcal{L}}{\partial z_2}$$ + sign

b)

$$\frac{\partial \mathcal{L}}{\partial b_2} = \boxed{\frac{\partial \mathcal{L}}{\partial z_2}}$$

$$\frac{\partial \mathcal{L}}{\partial W_2} = \frac{\partial a}{\partial W_2} \frac{\partial \mathcal{L}}{\partial a} = \frac{\partial \mathcal{L}}{\partial a} h_1^T = \boxed{\frac{\partial \mathcal{L}}{\partial z_2} h_1^T}$$

c)

$$\frac{\partial \mathcal{L}}{\partial h_1} = \frac{\partial a}{\partial h_1} \frac{\partial \mathcal{L}}{\partial a} = W_2^T \frac{\partial \mathcal{L}}{\partial z_2}$$

$$\frac{\partial \mathcal{L}}{\partial z_1} = \underbrace{\frac{\partial h_1}{\partial z_1}} \frac{\partial \mathcal{L}}{\partial h_1}$$

$$\mathrm{swish}(z_1) = sw(z_1)$$

$$\frac{\partial h_1}{\partial z_1} = sw(z_1) + \sigma(z_1) \odot (1 - sw(z_1))$$

$$\frac{\partial \mathcal{L}}{\partial z_1} = (sw(z_1) + \sigma(z_1) \odot (1 - sw(z_1))) \odot \left(W_2^T \frac{\partial \mathcal{L}}{\partial z_2}\right)$$

$$\boxed{\frac{\partial \mathcal{L}}{\partial b_1} = (sw(z_1) + \sigma(z_1) \odot (1 - sw(z_1))) \odot \left(W_2^T \frac{\partial \mathcal{L}}{\partial z_2}\right)}$$

$$\frac{\partial \mathcal{L}}{\partial c} = \frac{\partial \mathcal{L}}{\partial b_1}$$

$$\frac{\partial \mathcal{L}}{\partial W_1} = \frac{\partial \mathcal{L}}{\partial b_1} x^T$$

$$\boxed{= (sw(z_1) + \sigma(z_1) \odot (1 - sw(z_1))) \odot \left(W_2^T \frac{\partial \mathcal{L}}{\partial z_2}\right) x^T}$$