ECE C247 Homework 03

Chih-En Lin

1.

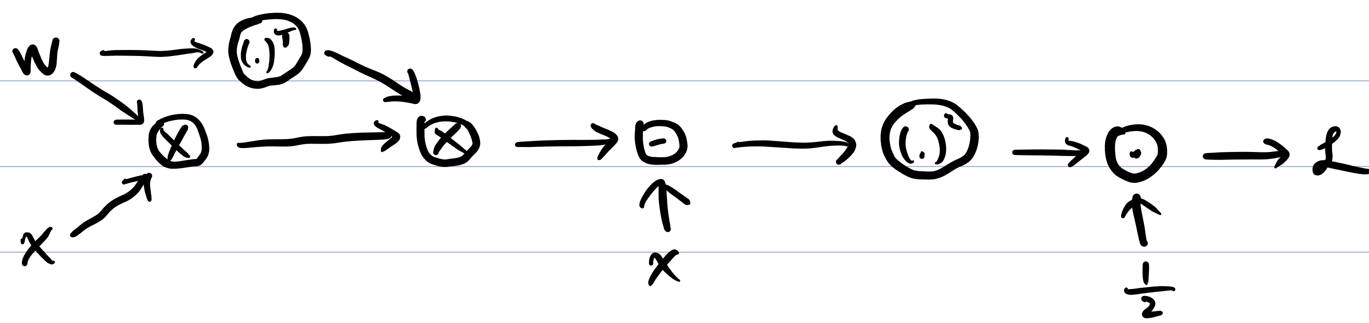
(a) Wx is to encode the information of x.

$W^T W x$ is to decode the information of x.

With the loss $\mathcal{L}$ be minimized, the difference between the reconstructed $W^T W x$ and the original x will be minimized.

Thus, Wx will preserve the information about x. #

(b)



#

(c)

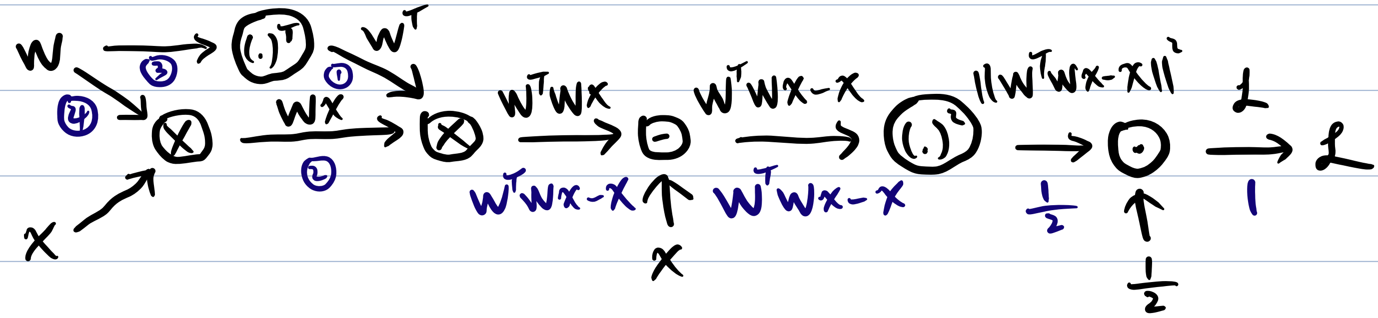
Let two paths are L_1 and L_2.

if $\begin{cases} L_1: & a \to b \to d \\ L_2: & a \to c \to d \end{cases}$

$$\nabla_W \mathcal{L} = \nabla_W \mathcal{L}_1 + \nabla_W \mathcal{L}_2 = \frac{\partial b}{\partial a} \cdot \frac{\partial d}{\partial b} + \frac{\partial c}{\partial a} \cdot \frac{\partial d}{\partial c}$$

#

(d) Use the computational graph from (b), Calculate the backpropragation



① $\dfrac{\partial \mathcal{L}}{\partial W^T} = (W^TWx - x)(Wx)^T$

② $\dfrac{\partial \mathcal{L}}{\partial Wx} = W(W^TWx - x)$

③ : ① Backpropragate to W : $(Wx)(W^TWx - x)^T$

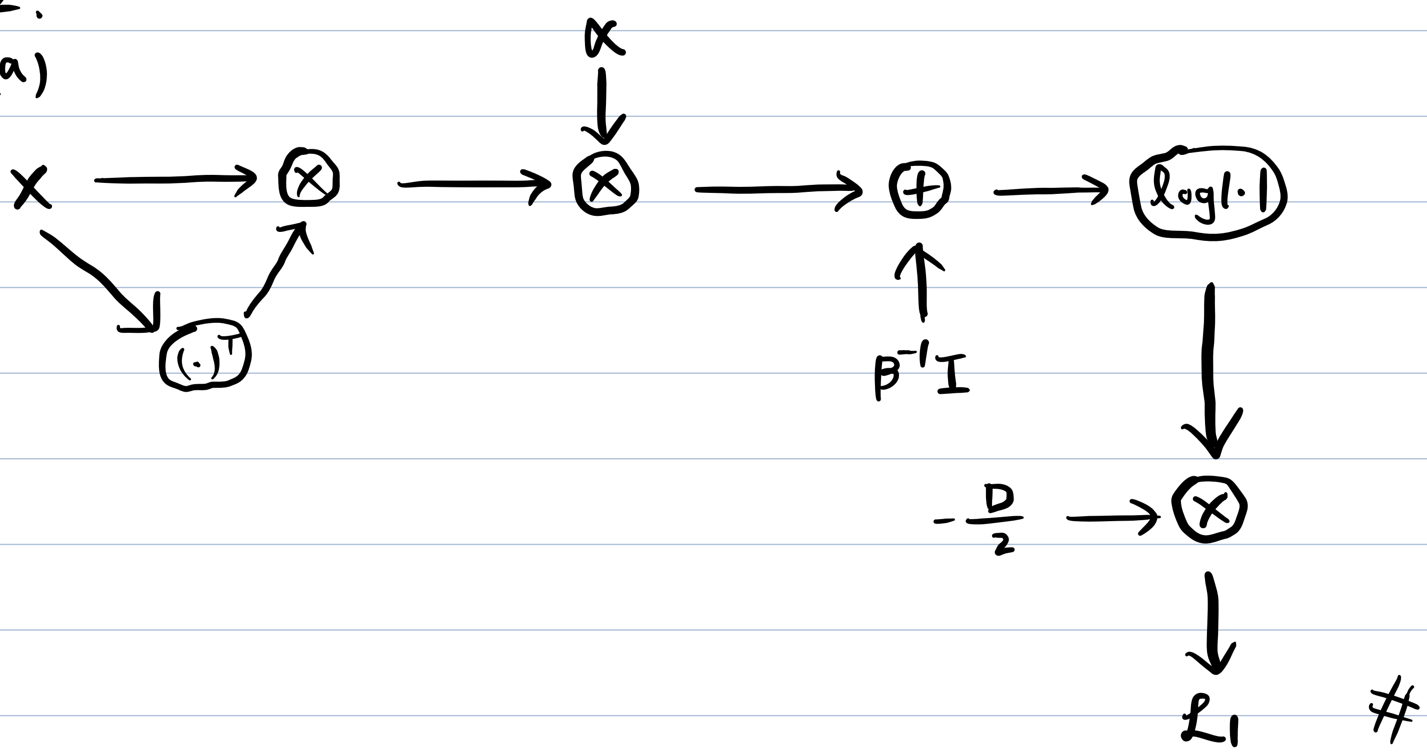
④ : ② Backpropragate to W : $W(W^TWx - x)x^T$

$$\nabla_W \mathcal{L} = ③ + ④ = (Wx)(W^TWx - x)^T + W(W^TWx - x)x^T$$

#

2.

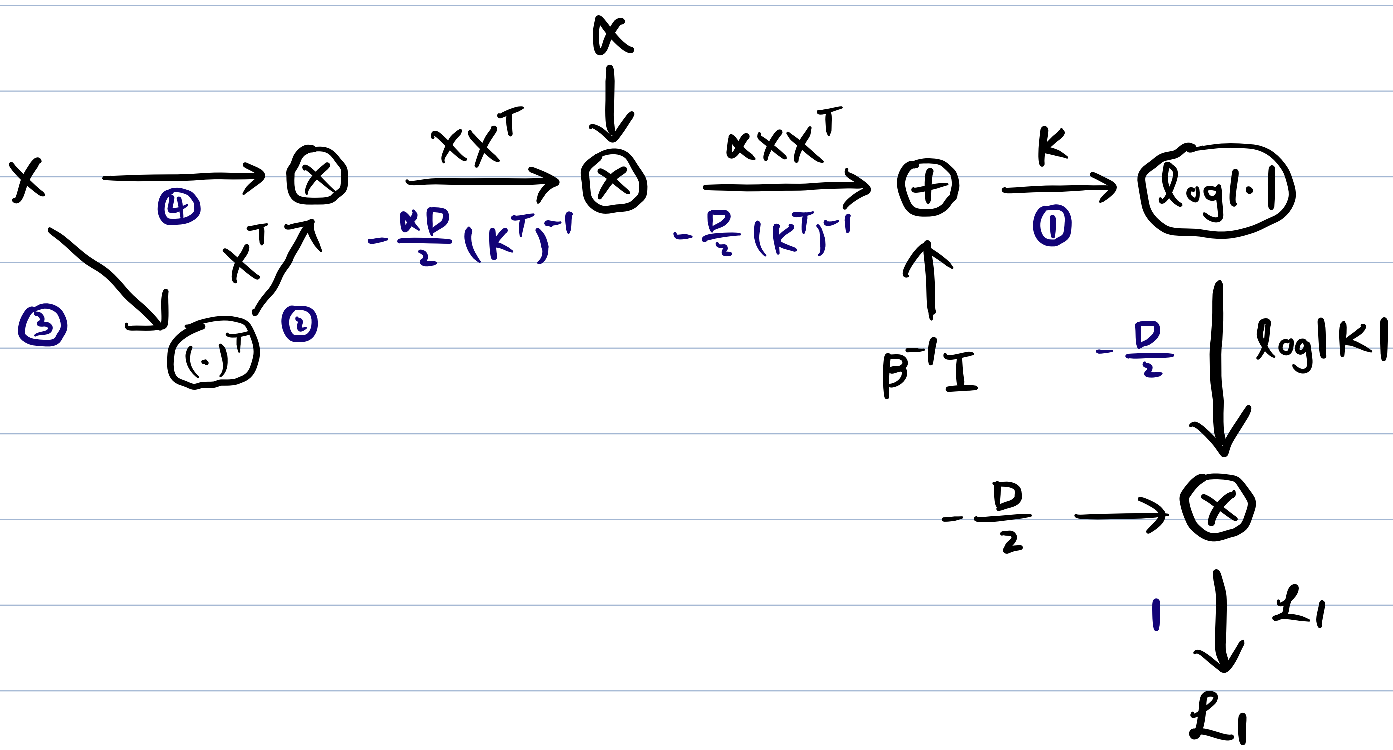
(a)



#

(b)

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



$$① = \frac{\partial \mathcal{L}_1}{\partial K} = -\frac{D}{2}(K^T)^{-1}$$

$$② = \frac{\partial \mathcal{L}_1}{\partial X^T} = X^T\left(\frac{-\alpha D}{2}(K^T)^{-1}\right)$$

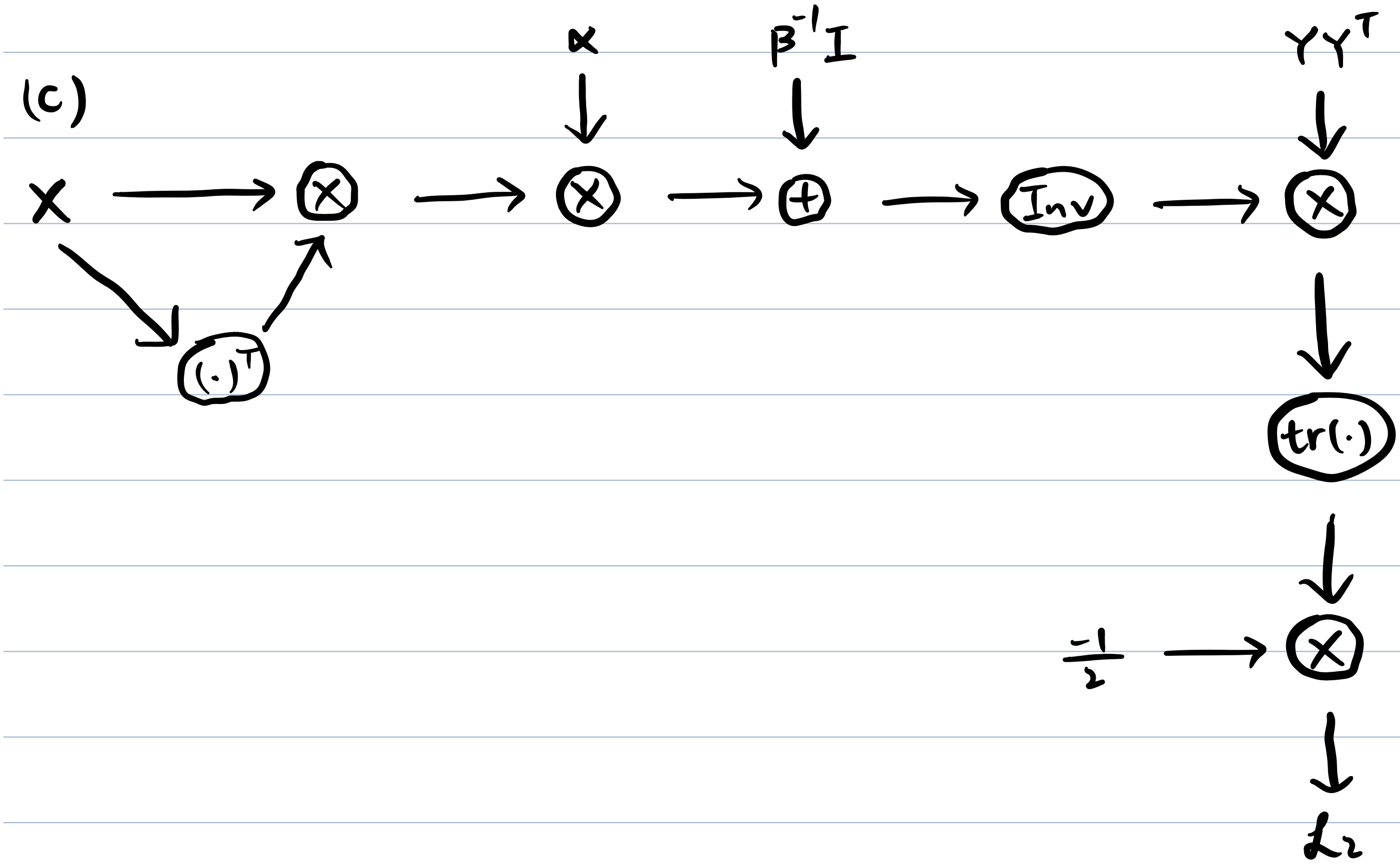
③ $= \frac{-\alpha D}{2} K^{-1} X$

④ $= \frac{-\alpha D}{2} (K^T)^{-1} X$ $\quad$ (K is symmetric)

$$\frac{\partial \mathcal{L}_1}{\partial X} = -\alpha D (K^T)^{-1} X = -\alpha D K^{-1} X$$ #

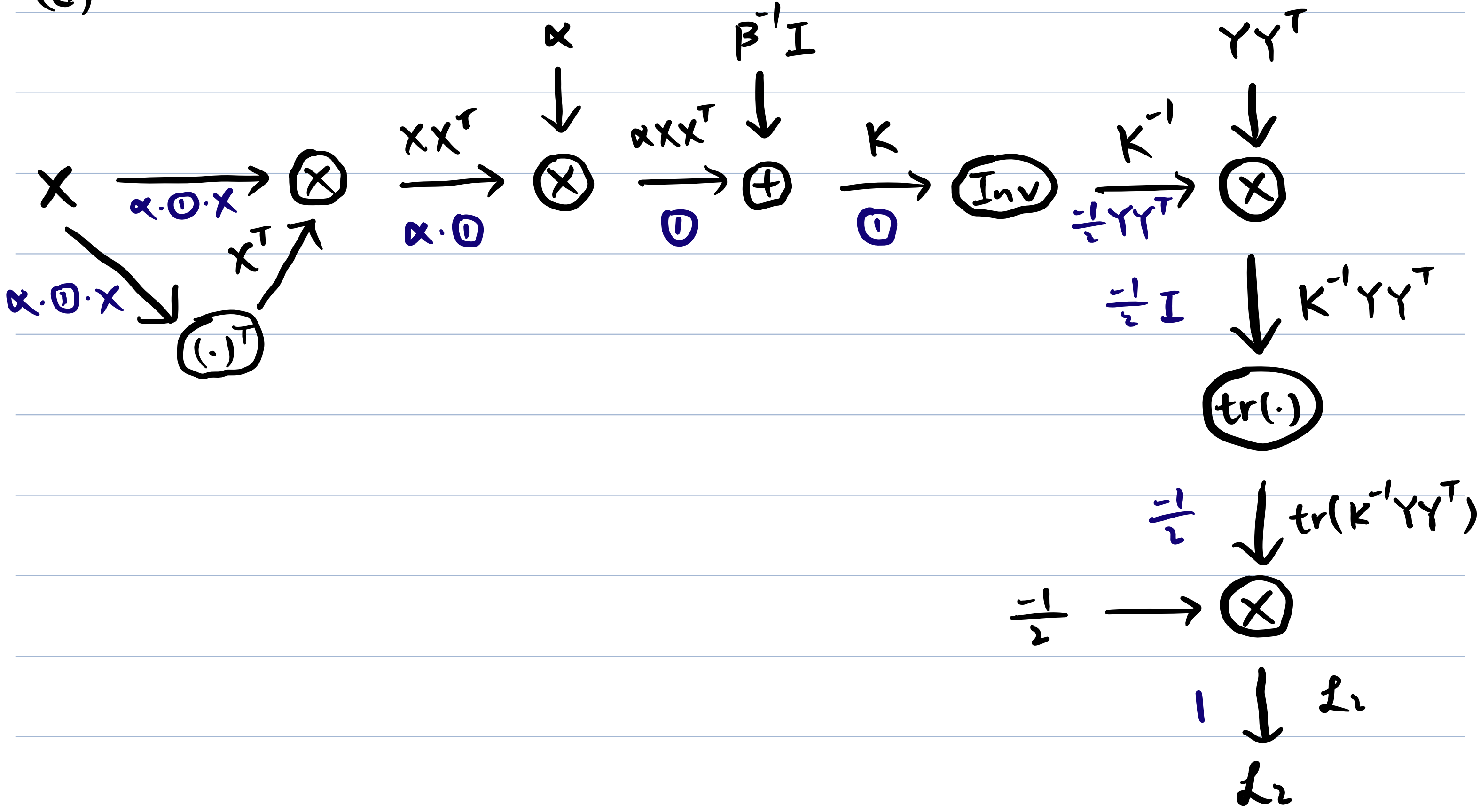
$$(K = \alpha X X^T + \beta^{-1} I)$$

(c)



#

(d)



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

$$① = \frac{\partial \mathcal{L}_2}{\partial K} = -(K^T)^{-1}\left(\frac{-1}{2}YY^T\right)(K^T)^{-1} = \frac{1}{2}K^{-1}YY^TK^{-1}$$

$$\frac{\partial \mathcal{L}_2}{\partial X} = 2\alpha \cdot \left(\frac{1}{2}K^{-1}YY^TK^{-1}\right) \cdot X = \alpha K^{-1}YY^TK^{-1}X$$

#

$(K = \alpha XX^T + \beta^{-1} I)$

(e)

$$\frac{\partial \mathcal{L}}{\partial X} = \frac{\partial \mathcal{L}_1}{\partial X} + \frac{\partial \mathcal{L}_2}{\partial X}$$

$$= \alpha K^{-1}YY^TK^{-1}X - \alpha D K^{-1} X$$

#

$(K = \alpha XX^T + \beta^{-1} I)$