1) a) We are trying to minimize the loss

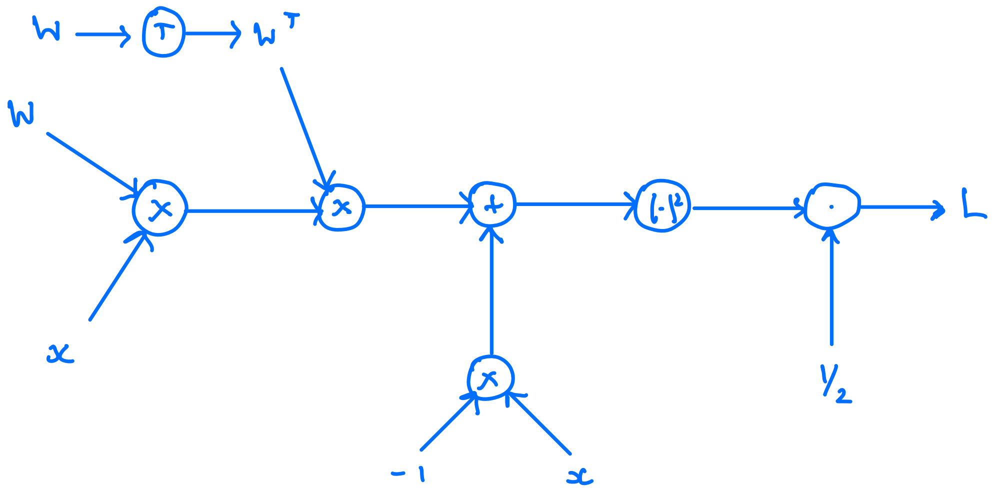
$$L = \frac{1}{2} \| W^T W x - x \|^2$$

Therefore, we try to find the W that minimizes the difference between $W^T W x$ and x.

So, we convert the given x to lower dimension form $y = Wx$, such that when we try to convert it back to the original dimensions through $W^T y$ we get an encoding of x that is very close to x (We try to minimize $W^T y - x$, i.e. $W^T W x - x$) //

Therefore, the hidden representation Wx ought to preserve information about x //

b)



c) Let's simplify all the intermediate nodes in the 2 paths and write the paths as

$$w \to a \to L$$

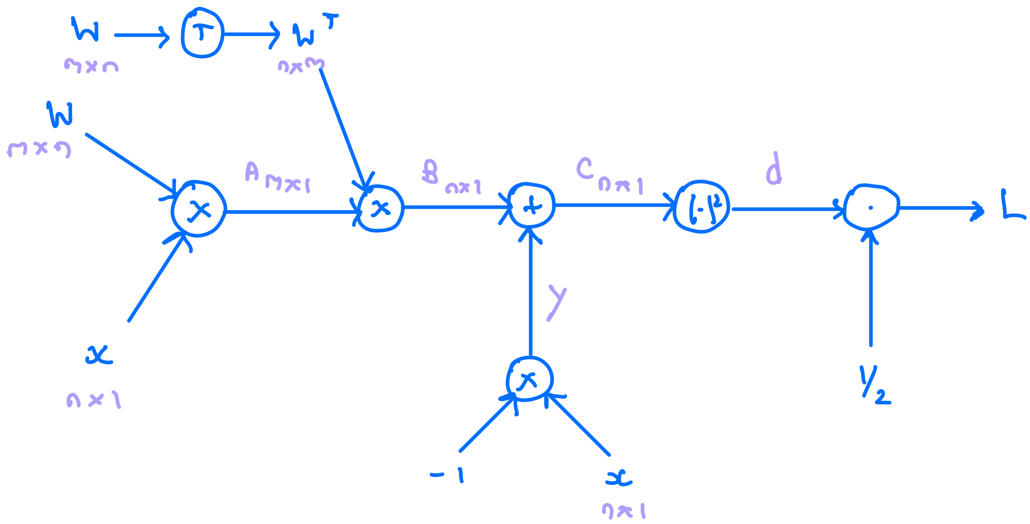
AND

$$w \to b \to L$$

BY THE RULE OF TOTAL DERIVATIVES, w affects L along 2 paths and hence the derivative is the sum of the two:

$$\frac{\partial L}{\partial w} = \frac{\partial a}{\partial w} \cdot \frac{\partial L}{\partial a} + \frac{\partial b}{\partial w} \cdot \frac{\partial L}{\partial b}$$

d)



$$L = \tfrac{1}{2} d$$

$$\rightarrow \quad \frac{\partial L}{\partial d} = \frac{1}{2}$$

$$d = \| c \|^2$$

$$\frac{\partial L}{\partial c} = \frac{\partial d}{\partial c} \cdot \frac{\partial L}{\partial d}$$

$$= 2c \cdot \frac{1}{2}$$

$$\rightarrow \quad \frac{\partial L}{\partial c} = c$$

$$c = B + y$$

$$\frac{\partial c}{\partial B} = \frac{\partial c}{\partial y} = 1$$

$$\rightarrow \frac{\partial L}{\partial B} = \frac{\partial C}{\partial B} \frac{\partial L}{\partial C} = \frac{\partial L}{\partial C} = C$$

$$\rightarrow \frac{\partial L}{\partial y} = \frac{\partial C}{\partial y} \cdot \frac{\partial L}{\partial C} = C$$

$$y = -x$$

$$\rightarrow \frac{\partial L}{\partial x} = \frac{\partial y}{\partial x} \cdot \frac{\partial L}{\partial y} \approx -\frac{\partial L}{\partial y} = -C$$

$$B = W^T A$$

$$\frac{\partial B}{\partial W^T} = A^T$$

$$\frac{\partial B}{\partial A} = W$$

$$\rightarrow \frac{\partial L}{\partial A} = \frac{\partial B}{\partial A} \frac{\partial L}{\partial B}$$

$$= WC$$

$$\rightarrow \frac{\partial L}{\partial W^T} = \frac{\partial L}{\partial B} \cdot \frac{\partial B}{\partial W^T}$$

$$= CA^T$$

$$A = Wx$$

$$\rightarrow \boxed{\frac{\partial L}{\partial W} = (CA^T)^T = AC^T}$$

$$\frac{\partial A}{\partial W} = x^T$$

$$\frac{\partial A}{\partial x} = W^T$$

$$\rightarrow \boxed{\frac{\partial L}{\partial W} = \frac{\partial L}{\partial A} \cdot \frac{\partial A}{\partial W} = WC \cdot x^T}$$

$$\rightarrow \frac{\partial L}{\partial x} = \frac{\partial A}{\partial x} \frac{\partial L}{\partial A} = W^T \cdot WC$$

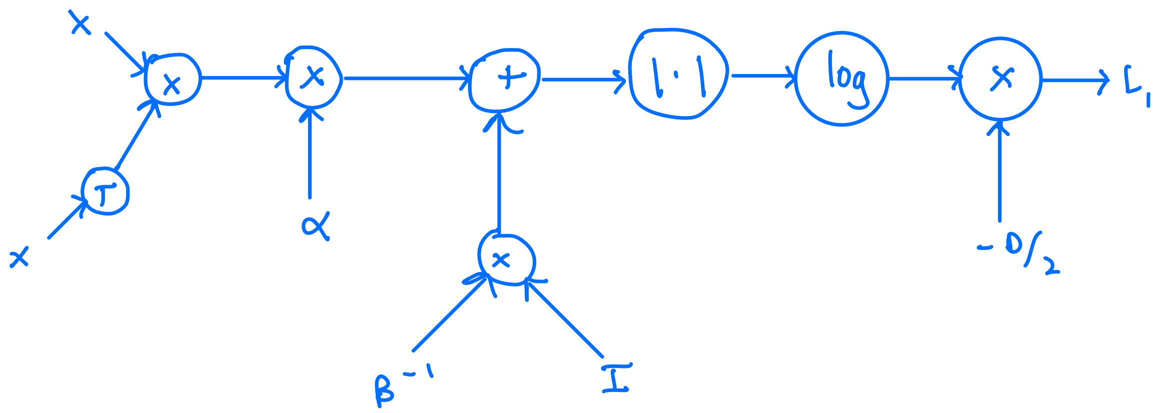
$$\frac{\partial L}{\partial W} = W C x^T + A C^T$$

$$A = Wx$$

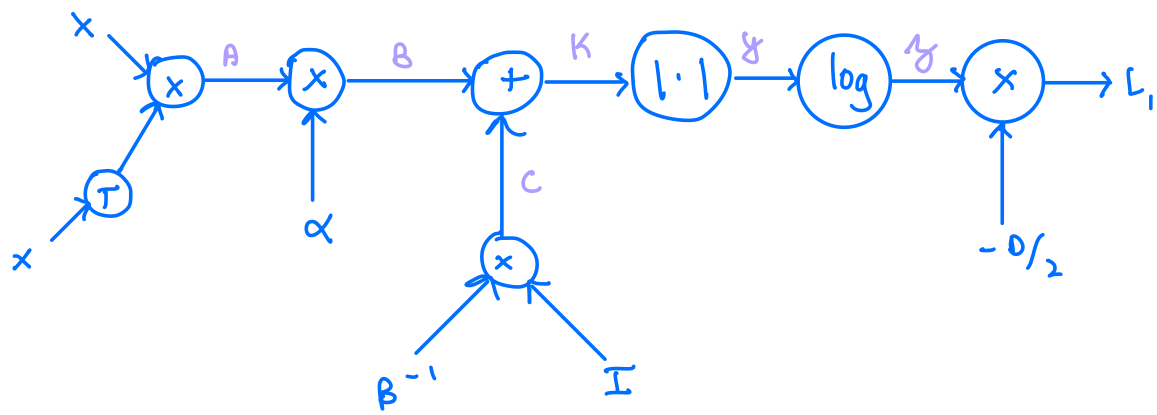
$$C = B + y = W^T A - x = W^T W x - x$$

$$\frac{\partial L}{\partial W} = W C x^T + W x C^T = W\left(W^T W x - x\right) x^T + W x \left(W^T W x - x\right)^T$$

2) a) $L_1 = -\frac{D}{2} \log |\alpha x x^T + \beta^{-1} I|$



b) $L_1 = -\frac{D}{2} \log |\alpha x x^T + \beta^{-1} I|$



$$L_1 = \frac{-D}{2} z$$

$$\frac{\partial L_1}{\partial z} = \frac{-D}{2}$$

$$z = \log y$$

$$\frac{\partial z}{\partial y} = \frac{1}{y}$$

$$\frac{\partial L_1}{\partial y} = \frac{\partial z}{\partial y} \frac{\partial L_1}{\partial z}$$

$$= \frac{1}{y} \cdot \left(\frac{-D}{2}\right)$$

$$y = |K|$$

$$\frac{\partial y}{\partial K} = |K| \left(K^{-1}\right)^T$$

$$\frac{\partial L_1}{\partial K} = \frac{\partial y}{\partial K} \frac{\partial L_1}{\partial y}$$

$$= y \left(K^{-1}\right)^T \cdot \frac{1}{y} \left(\frac{-D}{2}\right)$$

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

$$\frac{\partial L_1}{\partial K} = \left(\frac{-D}{2}\right) \left(K^T\right)^{-1}$$

$$K = B + C$$

$$\frac{\partial K}{\partial B} = 1$$

$$\frac{\partial L_1}{\partial B} = \frac{\partial L_1}{\partial K} = \left(\frac{-D}{2}\right)(K^T)^{-1}$$

$$B = \alpha A$$

$$\frac{\partial B}{\partial A} = \alpha$$

$$\frac{\partial L_1}{\partial A} = \frac{\partial B}{\partial A}\frac{\partial L_1}{\partial B} = \alpha\left(\frac{-D}{2}\right)(K^T)^{-1}$$

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

$$A = xx^T$$

$$\frac{\partial A}{\partial x} = 2x$$

$$\frac{\partial L_1}{\partial x} = \frac{\partial L_1}{\partial A} \cdot \frac{\partial A}{\partial x} = \left(\frac{-\alpha}{2}D\right)(K^T)^{-1} \cdot 2x$$

$$= (-\alpha D)(K^T)^{-1} \cdot x$$

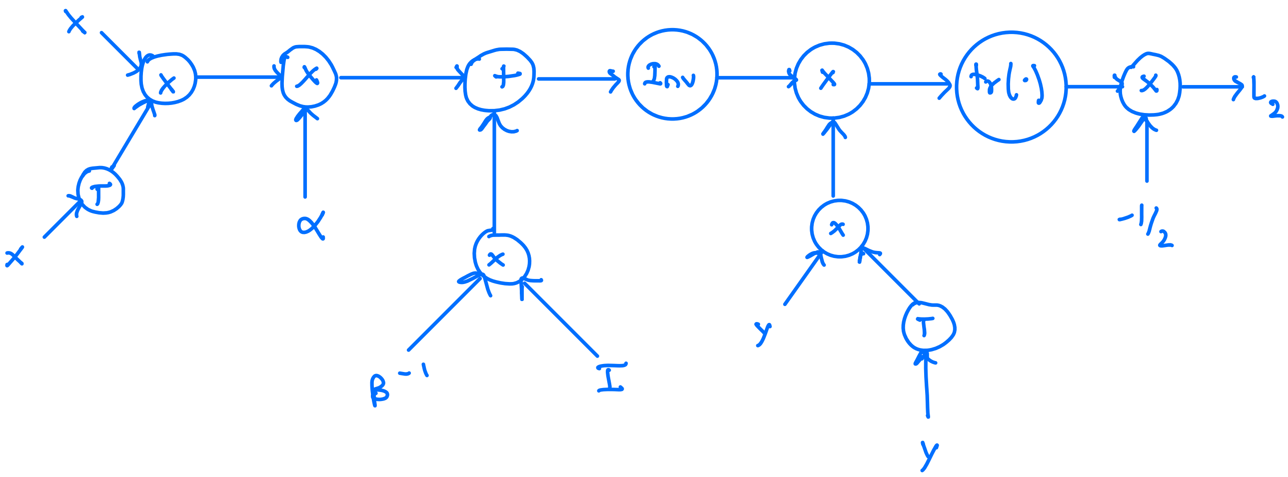
$$K = \alpha x x^T + \beta^{-1} I$$

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

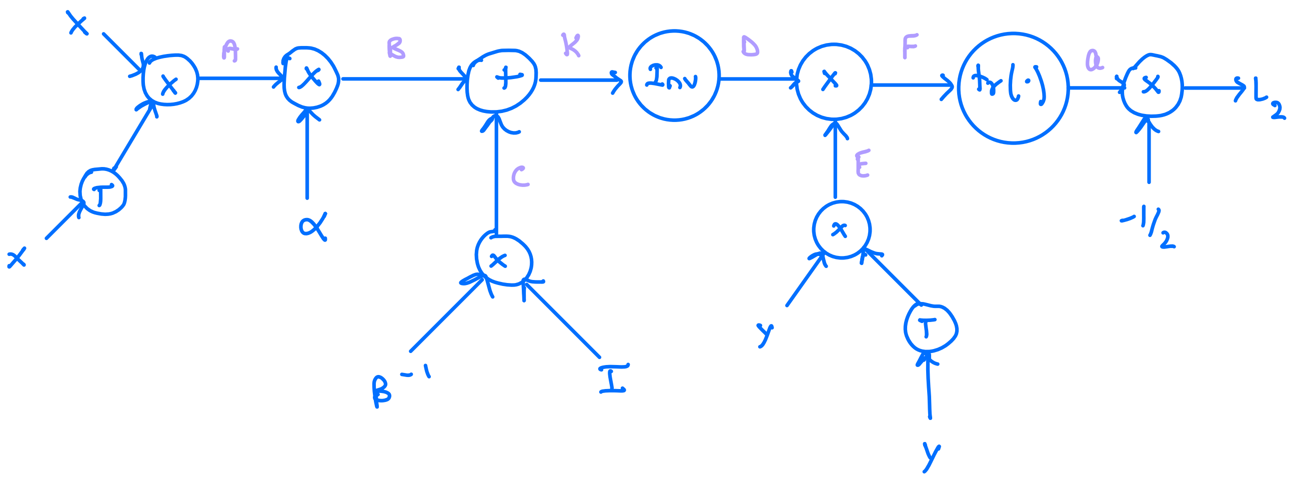
$$= \alpha x x^T + \beta^{-1} I = K$$

$$\frac{\partial L_1}{\partial x} = -\alpha D K^{-1} x$$

c) $L_2 = \frac{-1}{2} \mathrm{tr}\left((\alpha x x^T + \beta^{-1} I)^{-1} y y^T \right)$



d) $L_2 = \frac{-1}{2} \operatorname{tr}\left((\alpha x x^T + \beta^{-1} I)^{-1} y y^T\right)$



$$L_2 = -a/2$$

$$\frac{\partial L_2}{\partial a} = -1/2$$

$$a = \operatorname{tr}(F)$$

$$\frac{\partial a}{\partial F} = \operatorname{tr} \frac{\partial F}{\partial F} = I$$

$$\frac{\partial L_2}{\partial F} = \frac{\partial a}{\partial F} \frac{\partial L_2}{\partial a} = \frac{-1}{2} I$$

$$F = DE$$

$$\frac{\partial L_2}{\partial D} = \frac{\partial L_2}{\partial F} \cdot E^T = \frac{-1}{2} (y y^T)$$

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

$$\frac{\partial L_2}{\partial K} = -K^{-1} \frac{\partial L}{\partial K^{-1}} K^{-1}$$

$$= -K^{-1} \left(\frac{-1}{2} y y^T \right) K^{-1}$$

$$= \frac{1}{2} K^{-1} (y y^T) K^{-1}$$

$$K = B + C$$

$$\frac{\partial L_2}{\partial B} = \frac{\partial L_2}{\partial K}$$

$$B = \alpha A$$

$$\frac{\partial L_2}{\partial A} = \alpha \frac{\partial L_2}{\partial K}$$

$$A = xx^T$$

$$\frac{\partial A}{\partial x} = 2x$$

$$\frac{\partial L_2}{\partial x} = \frac{\partial L_2}{\partial A} \cdot \frac{\partial A}{\partial x} = \frac{\alpha}{2} K^{-1} (yy^T) K^{-1} \cdot 2x$$

$$\frac{\partial L_2}{\partial x} = \alpha K^{-1} (yy^T) K^{-1} x$$

e)

$$\frac{\partial L}{\partial x} = \frac{\partial L_1}{\partial x} + \frac{\partial L_2}{\partial x}$$

$$\frac{\partial L}{\partial x} = -\alpha D K^{-1} x \quad + \alpha K^{-1} (yy^T) K^{-1} x$$