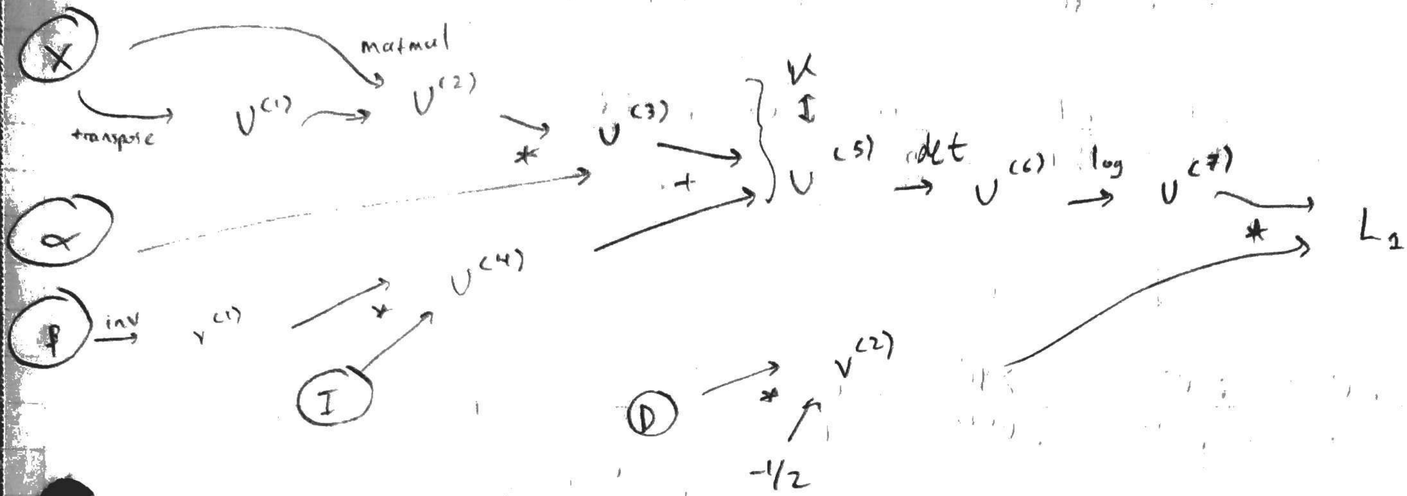
ECE C147 HW 3

Exercise 1: SKIP b/c I am a C147 student.

Exercise 2: a) I'll use the textbook convention of naming intermediate expressions, where $U^{(i)}$ represents a matrix intermediate expression, and $v^{(i)}$ represents a scalar intermediate expression.



b) We do this though backpropogation.

First, note $\frac{\partial L_1}{\partial K} = \frac{\partial}{\partial K}\left(\frac{-D}{2}\right)\log(\det K) = \frac{-D}{2}(K^{-1})^T$, as shown in the matrix cookbook.

Then, $\frac{\partial L_1}{\partial U^{(3)}} = \frac{\partial K}{\partial U^{(3)}} \frac{\partial L_1}{\partial K}$. $U^{(3)} = \alpha X X^T$ and $K = \alpha X X^T + \beta^{-1} I$.

Hence, $\frac{\partial K}{\partial U^{(3)}} = \frac{\partial(\alpha X X^T + \beta^{-1} I)}{\partial(\alpha X X^T)} = I$.

Therefore, $\frac{\partial L_1}{\partial U^{(3)}} = (I)(K^{-1})^T \overset{(-D/2)}{=} (K^{-1})^T(-D/2)$

Next, we have $\frac{\partial L_1}{\partial U^{(2)}} = \frac{\partial U^{(3)}}{\partial U^{(2)}} \frac{\partial L_1}{\partial U^{(3)}}$, where $U^{(2)} = X X^T$.

$$\frac{\partial U^{(3)}}{\partial U^{(2)}} = \frac{\partial(\alpha X X^T)}{\partial(X X^T)} = \alpha \frac{\partial X X^T}{\partial X X^T} = \alpha I.$$

So, $\frac{\partial L_1}{\partial U^{(2)}} = \alpha I \quad (K^{-1})^T \overset{(-D/2)}{=} \alpha (K^{-1})^T(-D/2)$

Finally we have $\frac{\partial L_1}{\partial X} = \frac{\partial U^{(2)}}{\partial X} \frac{\partial L_1}{\partial U^{(2)}}$.

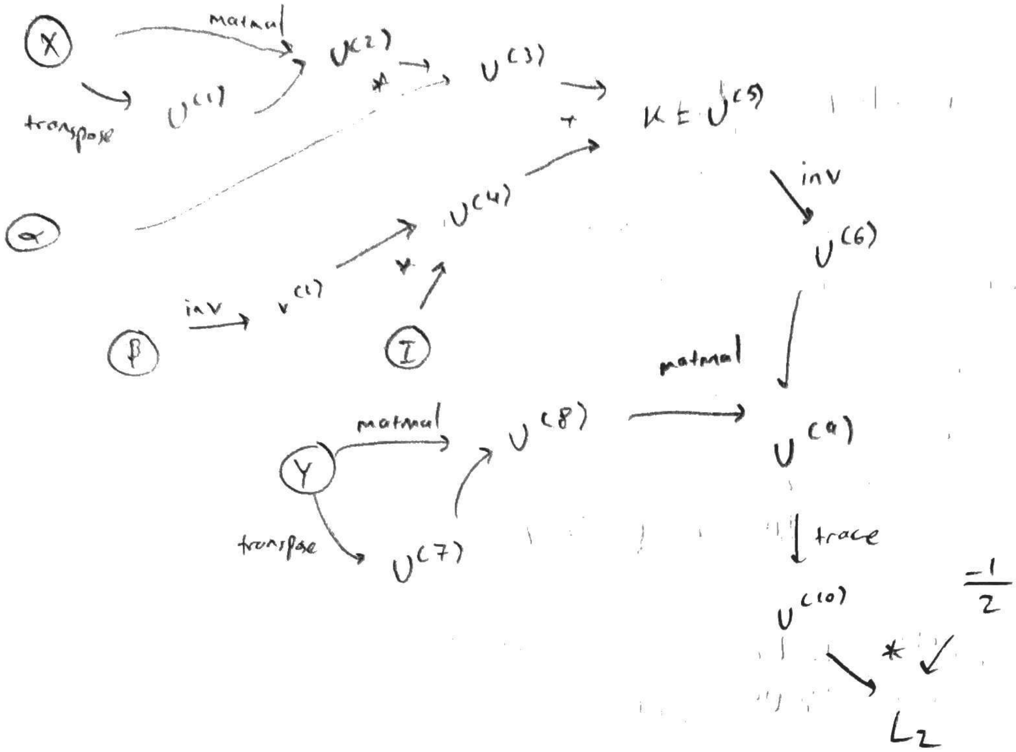
Then, $\frac{\partial U^{(2)}}{\partial X} = \frac{\partial}{\partial X}(X X^T)$. We note that if $Z = XY$, then $\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Z} Y^T$.

In this case we have $U^{(2)} = X X^T$, hence $\frac{\partial L_1}{\partial X} = \frac{\partial L_1}{\partial U^{(2)}} (X^T)^T$

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

$$\boxed{\text{Hence } \frac{\partial L_1}{\partial X} = \alpha (K^{-1})^T X (-D/2)}$$

c)


X
matmul
transpose
U(1)
U(2)
*
U(3)
+
K = U(5)
α
inv
U(4)
U(6)
β
inv
v(1)
+
I
matmul
Y
matmul
U(8)
U(9)
transpose
U(7)
trace
U(10)
-1/2
*
L2

d) Once again we use backpropagation.

First, $L_2 = -\frac{1}{2} tr(K^{-1}(YY^T)) = -\frac{1}{2} tr(K^{-1} U^{(4)})$.

From the matrix cookbook, we have $\frac{\partial}{\partial X} tr(AX^{-1}B) = -X^{-T}A^TB^TX^{-T}$.

Applying this, we get $\frac{\partial L_2}{\partial K} = \left(-\frac{1}{2}\right)\left(-K^{-T} U^{(4)T} K^{-T}\right)$

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

Next, $\frac{\partial L_2}{\partial U^{(3)}} = \frac{\partial K}{\partial U^{(3)}} \frac{\partial L_2}{\partial K}$. From (b), we know $\frac{\partial K}{\partial U^{(3)}} = I$.

In fact, the rest of the backprop is the same as (b). So we have

$$\frac{\partial L_2}{\partial X} = \frac{\partial U^{(2)}}{\partial X}\left(\frac{\partial U^{(3)}}{\partial U^{(2)}} \frac{\partial K}{\partial U^{(3)}} \frac{\partial L_2}{\partial K}\right)$$

$$= \left(\frac{\partial U^{(3)}}{\partial U^{(2)}} \frac{\partial K}{\partial U^{(3)}} \frac{\partial L_2}{\partial K}\right) X$$

Same last step as (b), since if $Z = XY$, then $\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Z} Y^T$.

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

$$\boxed{= \frac{1}{2} \alpha K^{-T} YY^T K^{-T} X}$$

e) We have $L = -c + L_1 + L_2$. Since c is a constant,

$$\frac{\partial L}{\partial X} = \frac{\partial L_1}{\partial X} + \frac{\partial L_2}{\partial X} = \alpha (K^{-1})^T X (-D/2) + \frac{1}{2} \alpha K^{-T} YY^T K^{-T} X$$

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

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