Linear Algebra Review Assignment 1

Simon Kafader, 15-114-945

October 11, 2017

1.a) $f : \mathbb{R}^n \to \mathbb{R}^m,\ f(x) = Ax$. Show that $Df(x) = A$.

$$f(x) = Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

$$\Rightarrow \frac{\partial f_i(x)}{\partial x_j} = a_{ij}$$

since when you derivate with respect to $x_j$only the terms containing $x_j$will not fall off as constants what will result in the vector $\begin{pmatrix} a_{1j} & \dots & a_{mj} \end{pmatrix}^T$. Now the i-th component of this vector is obviously a_{ij}.

$$\Rightarrow Df(x) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = A$$

b.)

$$g(x) = x^T x = x_1^2 + \cdots + x_n^2$$

$$Df(x) = \left[\frac{\partial x^\top x}{\partial x_1}, \dots, \frac{\partial x^\top x}{\partial x_n} \right]^\top = \begin{bmatrix} 2x_1 & \dots & 2x_n \end{bmatrix}^T$$

c.)

$$Dh(x) = D(x^\top) \cdot Qx + D(Qx) \cdot x^\top = Qx + x^\top Q$$

Now since Q is orthogonal we know that $Qx = x^\top Q \Rightarrow Dh(x) = 2Qx$

d)

$$f(x) = ||Ax - b||^2 = g(h(x))$$

where $h(x) = Ax - b$ and $g(x) = ||x||^2$.

$$Dg(x) = D < x, x >= D(x_1^2 + \cdots + x_m^2) = (2x_1 \dots\ 2x_m)^\top = 2x$$

$$Dh(x) = D(Ax - b) = A$$

$$Df(x) = Dg(z)|_{z=h(x)} Dh(x) = 2(Ax - b) \cdot A$$

2.)

$$tr(A^2) = (a_{11}^2 + \cdots + a_{1m}a_{m1}) + (a_{21}a_{12} + a_{22}^2 + \cdots + a_{2m}a_{m2}) + \cdots + (a_{m1}a_{1m} + \cdots + a_{mm}^2)$$

$$= \sum_{i=1}^{m} a_{i1}a_{1i} + \cdots + a_{im}a_{mi} = \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}a_{ji}$$

$$tr(A^T A) = (a_{11}^2 + \cdots + a_{1m}^2) + \cdots + (a_{m1}^2 + \ldots + a_{mm}^2) = \sum_{i=1}^{m} a_{i1}^2 + \cdots + a_{im}^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}^2$$

Remind, that $\forall a, b \in \mathbb{R}:\ 2ab \leq a^2 + b^2$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}a_{ji} \leq \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}^2 \Rightarrow tr(A^2) \leq tr(A^T A)$$

3.a)

$$tr\left(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}\begin{pmatrix} x_{11} & \dots & x_{m1} \\ \vdots & \ddots & \vdots \\ x_{1n} & \dots & x_{mn} \end{pmatrix}\right) = (a_{11}x_{11} + \cdots + a_{1n}x_{1n}) + \cdots + (a_{m1}x_{m1} + \cdots + a_{mn}x_{mn})$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_{ij}$$

Now obvoiusly every x_{ij}shows up exactly one time in this sum. So if we deriviate with respect to $x_{ji}$the only term left is a_{ji}. So obvoiusly $\nabla_X tr(AX^T) = A$

b)

As we've shown in the last review series it is known that $tr(AB) = tr(BA)$. So

$$\nabla_X tr(AX^T) = \nabla_X tr(X^T A) = A$$

c)

We define $C = BA$ and use knowledge from the script section 2.1 and a)

$$\nabla_X tr(AX^T B) = \nabla_X tr(BAX^T) = \nabla_X tr(CX^T) = C = BA$$

4)

$$D\left(\frac{x^\top Ax}{x^\top x}\right) = \frac{2Ax}{x^\top x} - \frac{-2x \cdot x^\top Ax}{\left(x^\top x\right)^2} = \frac{(2Ax)(x^\top x)}{(x^\top x)^2} - \frac{-2x \cdot x^\top Ax}{\left(x^\top x\right)^2} = \frac{(2Ax)(x^\top x) + 2x \cdot x^\top Ax}{\left(x^\top x\right)^2} = 0$$

$$(Ax)(x^\top x) + x \cdot (x^\top Ax) = 0$$

$$(Ax)(x^\top x) = -x \cdot (x^\top Ax)$$

$$(x^\top x)Ax = -x(x^\top Ax)$$