CS229 PSet #0

1a) $f(x) = \frac{1}{2}x^T A x + b^T x$

$$= \frac{1}{2}\begin{bmatrix} x_1 \\ \vdots \end{bmatrix}\begin{bmatrix} A_1 \cdots \\ \vdots \end{bmatrix}\begin{bmatrix} x_1 \cdots \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \end{bmatrix}\begin{bmatrix} x_1 \cdots \end{bmatrix}$$

$$= \frac{1}{2}\left(\sum A_{1i} x_1^2 + \sum A_{2i} x_2^2 + \dots\right) + \left(x_1 b_1 + x_2 b_2 \dots\right)$$

Look at $\frac{\partial f}{\partial x_1} = \sum A_{ij} x_1 + b_1$

$$\Rightarrow \boxed{\nabla f(x) = Ax + b}$$

1b)

$f(x) = g(h(x))$

Let's look at $h(x) = x^T x$ and $g(x) = x^2$

as an example

$$\Rightarrow f(x) = (x^T x)^2 = (x_1^2 + x_2^2 + \dots)^2$$

so $\nabla f(x) = \begin{bmatrix} 2(x_1^2 + \dots)\cdot 2x_1 \\ 2(x_1^2 + \dots)\cdot 2x_2 \\ \vdots \end{bmatrix} \Rightarrow \cancel{\nabla f(x) = g'(\nabla h(x))}$

$$\Downarrow \boxed{\nabla f(x) = g'(h(x))\nabla h(x)}$$

1c)

we saw $\nabla f(x) = Ax + b$

$$\Rightarrow \boxed{\nabla^2 f(x) = A}$$

1d)

$f(x) = g(a^T x)$

we know $\nabla f(x) = g'(h(x)) \nabla h(x) \Rightarrow g'(a^T x) a$

$$\text{so} \quad \boxed{\nabla f(x) = g'(a^T x)\, a}$$

now $\nabla^2 f(x) = g''(a^T x)\, \nabla \begin{bmatrix} 0 \\ \\ \end{bmatrix}$

$$\nabla f(x) = \begin{bmatrix} g'(a^T x) a_1 \\ \vdots \\ \end{bmatrix} \quad \text{so} \quad \nabla f(x) = \begin{bmatrix} g''(a^T x) a_1 a_1 & g''(a^T x) a_1 a_2 \\ \vdots & \vdots \\ \end{bmatrix}$$

$$\Rightarrow \boxed{\nabla f(x) = g''(a^T x)\, a a^T}$$

note $a a^T$ is outer product

i.e. $a \otimes a^T$

2a) $zz^T = \begin{bmatrix} z_1 z_1 & z_1 z_2 \\ z_2 z_1 & z_2 z_2 \end{bmatrix} = \begin{bmatrix} z_1^2 & z_1 z_2 \\ z_2 z_1 & z_2^2 \end{bmatrix}$

note $\boxed{zz^T = (zz^T)^T}$

Now $x^T zz^T x = x^T \begin{bmatrix} z_1^2 x_1 + z_1 z_2 x_2 \\ z_1 z_2 x_1 + z_2^2 x_2 \end{bmatrix} = z_1^2 x_1^2 + 2 z_1 z_2 x_1 x_2 + z_2^2 x_2^2$

$= (x_1 z_1 + x_2 z_2)^2 = (x^T z)^2 \geq 0$

~~hence~~ hence $\boxed{zz^T \succeq 0}$

2b) Building from above we need to have:

$z_1^2 x_1 + z_1 z_2 x_2 = 0$ $\Rightarrow N(x^T$

and $z_2^2 x_2 + z_1 z_2 x_1 = 0$

$\Downarrow$

$z_1(x_1 z_1 + x_2 z_2) = 0$

$z_2(x_1 z_1 + x_2 z_2) = 0$

Sine z is nonzero $\Rightarrow x_1 z_1 + x_2 z_2 = 0 \Rightarrow x^T z = 0$

$\Rightarrow$ $\boxed{\begin{gathered} N(A) = \{x \in \mathbb{R}^n : x^T z = 0\} \\ R(A) = 1 \end{gathered}}$

Sin $x^T z = 0$ is a hyperplane then nullity is $n-1$ hence rank $= 1$

2c) first show $BAB^T = (BAB^T)^T$

$A^T = A$ since $A \succeq 0$

$$BAB^T = (BA^TB^T)^T = (BAB^T)^T$$

hence $\boxed{BAB^T = (BAB^T)^T}$

now look at $\underbrace{x^TB}_{1\times m}\overset{n\times n}{A}\underbrace{B^Tx}_{m\times 1}$

say for example $y^T = x^TB$ the $y = B^Tx$

hence x^TBAB^Tx is same form as y^TAy

Since we know $A \succeq 0$ then $x^TBAB^Tx = y^TAy \geq 0$

$\Rightarrow \boxed{BAB^T \succeq 0}$

3a) $\cancel{A = T\Lambda T'}$

$\cancel{AT = TA}$ =

3a)

$$A = T\Lambda T^{-1}$$

$$At^{(i)} = T\Lambda T^{-1}t^{(i)}$$

note that $T^{-1}t^{(i)} = e_{(i)}$ basis vector

$$At^{(i)} = T\Lambda e^{(i)}$$

$$At^{(i)} = T\lambda_i e_{(i)}$$

$$At^{(i)} = \lambda_i T e^{(i)}$$

$$\Rightarrow \boxed{At^{(i)} = \lambda t^{(i)}}$$

3b) Note that proving $U^T = U^{-1}$ reduces this to 3a:

$$U^T U = I$$

$$U^T U U^{-1} = I U^{-1} \Rightarrow \boxed{U^T = U^{-1}}$$

hence: $A = U\Lambda U^T = U\Lambda U^{-1}$

and result from 3a applies

3c) consider eigenvalues λ_i and eigenvectors $t^{(i)}$ so $At^{(i)} = \lambda_i t^{(i)}$

for any $t^{(i)}$ we know $t^{(i)T} A t^{(i)} \geq 0$

$$\Downarrow$$

$$t^{(i)T} \lambda_i t^{(i)} \geq 0$$

which gurantees $\boxed{\lambda_i \geq 0}$