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

$\nabla f(x) = \frac{1}{2} \cdot 2Ax + b = Ax + b$

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

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

$\nabla f(x) = g'(h(x)) \nabla h(x)$

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

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

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

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

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

$\nabla f(x) = g'(a^T x) a$

$\nabla_x^2 f(x) = g'(a^T x) a \left[\frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \cdots \quad \frac{\partial}{\partial x_n} \right]$

$$= \begin{bmatrix} g''(a^T x) a_1 a_1 & - - - - & g''(a^T x) a_n a_1 \\ \vdots & & \vdots \\ g''(a^T x) a_1 a_n & & g''(a^T x) a_n a_n \end{bmatrix}$$

$= g''(a^T x)\, a a^T$

$\boxed{\therefore \nabla f(x) = g'(a^T x) a, \quad \nabla^2 f(x) = g''(a^T x)\, a a^T}$

2. (a) $A = zz^T$, $z \in R^n$

i) $A = A^T$?

$(zz^T)^T = (z^T)^T z^T = zz^T$

$\therefore A = A^T$

ii) $x^T A x \geq 0$ for $\forall x \in R^n$

$$x^T A x = \sum_{j=1}^{n}\sum_{i=1}^{n} x_j A_{ij} x_i = \sum_{j=1}^{n}\sum_{i=1}^{n} x_j z_i z_j x_i = \sum_{i=1}^{n} x_i z_i \sum_{j=1}^{n} x_j z_j$$

$$= \cancel{\left\{ \sum_{i=1}^{n} x_i z_i \right\}} = \left\{ \sum_{i=1}^{n} x_i z_i \right\}^2 \geq 0$$

$\therefore A \in S_+^n$

(b) $N(A) = \{x \mid Ax = 0\}$

$Ax = zz^T x$

$\therefore N(A) = \{x \mid x = \vec{0}$ or $z^T x = 0\}$

$$zz^T = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \begin{bmatrix} z_1 & \cdots & z_n \end{bmatrix} = \begin{bmatrix} z_1 z_1 & z_1 z_2 & \cdots & z_1 z_n \\ \vdots & & & \\ z_n z_1 & \cdots & & z_n z_n \end{bmatrix}$$

$$\sim \begin{bmatrix} z_1 z_1 & \cdots & z_1 z_n \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{bmatrix} \qquad \therefore Rank(A) = 1$$

$$\boxed{\therefore N(A) = \{x \mid x = \vec{0} \text{ or } z^T x = 0\}, \ Rank(A) = 1}$$

(c) BAB^T

i) $(BAB^T)^T = (AB^T)^T B^T = (B^T)^T A^T B^T = BA^T B^T = BAB^T$

$(\because A = A^T)$ $\therefore BAB^T = (BAB^T)^T$

ii) $x^T BAB^T x \geq 0$ for $\forall x \in R^m$

$(B^T x)^T = x^T B$, $B^T x \in R^n$, $A \in S_+^n \Rightarrow X^T A X \geq 0$ for $\forall X (= B^T x) \in R^n$

$\therefore BAB^T$ is PSD.

3.

(a) $A = T\Lambda T^{-1}$, $TT^{-1} = I = T^{-1}T$

$AT = T\Lambda T^{-1}T = T\Lambda I = T\Lambda$

i) AT

$$A[t^1\ t^2 \cdots t^n] = [At^1\ At^2 \cdots At^n]$$

ii) $T\Lambda$

$$[t^1\ t^2 \cdots t^n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$= t^1[\lambda_1\ 0 \cdots 0] + t^2[0\ \lambda_2 \cdots 0] + \cdots + t^n[0 \cdots 0\ \lambda_n]$$

$$= \begin{bmatrix} t_1^1 \\ \vdots \\ t_n^1 \end{bmatrix}[\lambda_1\ 0 \cdots 0] + \cdots + \begin{bmatrix} t_1^n \\ \vdots \\ t_n^n \end{bmatrix}[0 \cdots 0\ \lambda_n]$$

$$= \begin{bmatrix} \lambda_1 t_1^1 & 0 & \cdots & 0 \\ \vdots & 0 & & \vdots \\ \vdots & \vdots & & \vdots \\ \lambda_1 t_n^1 & 0 & & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \cdots & \lambda_n t_1^n \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n t_n^n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 t_1^1 & \lambda_2 t_1^2 & \cdots \\ \vdots & \vdots & \\ \lambda_1 t_n^1 & \vdots & \end{bmatrix} = [\lambda_1 t^1\ \lambda_2 t^2 \cdots \lambda_n t^n]$$

$$\boxed{\therefore At^i = \lambda_i t^i}$$

(b) U is orthogonal $\Rightarrow$ $UU^T = I = U^TU$

i) $A = U\Lambda U^T$ $\Rightarrow$ $AU = U\Lambda U^TU = U\Lambda(U^TU) = U\Lambda I = U\Lambda$

$\therefore AU = U\Lambda$

ii) AU

$$= A[u^1 \cdots u^n] = [Au^1 \; Au^2 \cdots Au^n]$$

iii) $U\Lambda$

$$[u^1 \cdots u^n]\begin{bmatrix} \lambda_1 & \cdots & & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & \lambda_n \end{bmatrix}$$

$$= u^1[\lambda_1 \cdots 0] + u_2[0 \; \lambda_2 \cdots 0] + \cdots + u^n[0 \cdots \lambda_n]$$

$$= \begin{bmatrix} u_1^1\lambda_1 & \cdots & 0 \\ \vdots & & \\ u_n^1\lambda_2 & \cdots & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \cdots & \lambda_n u_1^n \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n u_n^n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 u_1^1 & \cdots & \lambda_n u_1^n \\ \vdots & & \vdots \\ \lambda_1 u_n^1 & \cdots & \lambda_n u_n^n \end{bmatrix} = [\lambda_1 u^1 \; \lambda_2 u^2 \cdots \lambda_n u^n]$$

Since $AU = U\Lambda$, each element of matrices is identical.

$$\boxed{\therefore Au^i = \lambda_i u^i}$$

(c) A is PSD $\Rightarrow x^T A x \geq 0$ for $\forall x \in R^n$.

$$Ax = \lambda x \Rightarrow x^T A x = x^T \lambda x$$

~~Where~~

$$x^T \lambda x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_j \lambda_{ij} x_i$$

$$\lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

So, when $i \neq j$, $x_j \lambda_{ij} x_i = 0$.

When $i = j$, $x_j \lambda_{ij} x_i = x_i \lambda_{ii} x_i = \{x_i\}^2 \lambda_i(A)$

Thus, $x^T \lambda x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_j \lambda_{ij} x_i = \sum_{i=1}^{n} \{x_i\}^2 \lambda_i(A) \geq 0$.

Since $\{x_i\}^2 \lambda_i(A) \geq 0$, $\lambda_i(A) \geq 0$.

$$\boxed{\therefore \lambda_i(A) \geq 0 \text{ for each } i}$$

4. (a) $Y = X_1 + \cdots + X_n$

$$E[Y] = E[X_1 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

$$= \mu_1 + \mu_2 + \cdots + \mu_n = 1^T \mu$$

$$Var[Y] = Var[X_1 + X_2 + \cdots + X_n]$$

$$= \sum_{i=1} Var(X_i) + \sum_{i,j} Cov(X_i, X_j) \quad (i \neq j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} (\Sigma)_{ij} = 1^T \Sigma 1$$

$\therefore E[Y] = 1^T \mu$, $Var[Y] = 1^T \Sigma 1$

Y is normal distribution.

(b)
$$E[X^T \Sigma^{-1} X] = E[tr(X^T \Sigma^{-1} X)]$$

$$= E[tr(\Sigma^{-1} X X^T)] = tr(E[\Sigma^{-1} X X^T])$$

$$= tr(\Sigma^{-1} E[X X^T])$$

$$= tr(\Sigma^{-1}(\Sigma + \mu\mu^T)) = tr(I + \Sigma^{-1}\mu\mu^T)$$

$$= tr(I) + tr(\Sigma^{-1}\mu\mu^T)$$

$$= n + \mu^T \Sigma^{-1} \mu$$

$$\therefore E[X^T \Sigma^{-1} X] = n + \mu^T \Sigma^{-1} \mu$$