Ass1 Solutions

Chang Li

July 29, 2016

1 Solutions

1.1 Prob1.

If $det(A) = 0$, then $det(A^T) = 0$, therefore $det(A) = det(A^T)$.
If $det(A) \neq 0$, then matrix has inverse. So A is a full rank matrix. Since every full rank matrix can be written as a product of elementary matrices:

$$R_1 \dots R_t I_n C_1 \dots C_s$$

where R_i are row elementary matrices and C_i are column elementary matrices. We know that for all elementary matrices, $det(E) = det(E^T)$ (because column expansion or row expansion of det doesn't change the value) and $(A_1 \dots A_n)^T = A_n^T \dots A_1^T$. $det(I_n) = det(I_n^T) = 1$ Therefore,

$$\begin{aligned}
det(A^T) &= det((R_1 \dots R_t I_n C_1 \dots C_s)^T) \\
&= det(C_s^T \dots C_1^T I_n^T R_t^T \dots R_1^T) \\
&= det(C_s^T) \dots det(C_1^T) det(I_n^T) det(R_t^T) \dots det(R_1^T) \\
&= det(C_s) \dots det(C_1) det(I_n) det(R_t) \dots det(R_1) \\
&= det(R_1) \dots det(R_t) det(I_n) det(C_1) \dots det(C_s) \\
&= det(R_1 \dots R_t I_n C_1 \dots C_s) \\
&= det(A)
\end{aligned}$$

1.2 Prob 2.

For any n, an identity matrix $A \in R^{n \times n}$ is a matrix only have element equal to 1 on diagonal and all other places are 0s

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Assume that when A is an indentity matrix of $A \in R^{n \times n}$, $det(A) = \prod_{i=1}^{n} a_{ii} = 1$. We prove this using induction.

When $n = 1$, A is a scalar and $det(A) = a_{11} = 1$.

When the size is $n-1$, $det(A^{(n-1)\times(n-1)}) = \prod_{i=1}^{n-1} a_{ii} = 1$. When consider the identity matrix of size n, expanding the matrix along the first row. Since only the first element is 1 and all the other are 0s, we have:

$$\begin{aligned}
&1 \times det(A^{(n-1)\times(n-1)}) + 0 + \cdots + 0 \\
&= a_{11} \times \prod_{i=2}^{n} a_{ii} \\
&= \prod_{i=1}^{n} a_{ii} = 1
\end{aligned}$$

Therefore, for any n we have $det(A) = 1$, when A is an identity matrix and $A \in R^{n\times n}$.

1.3 Prob 3.

For any triangular matrix A of size $n \times n$, $det(A) = \prod_{i=1}^{n} a_{ii}$. Therefore, if A has eigen values then,

$$\begin{aligned}
det(\lambda I - A) &= \prod_{i=1}^{n} (\lambda - a_{ii}) && (1) \\
&= 0 && (2)
\end{aligned}$$

The equation 1 holds if and only if $\lambda_i = aii$ for some i. Therefore, the diagonal elements are equal to eigen values.

1.4 Prob 4.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{aligned}
det(\lambda I - A) &= \begin{bmatrix} \lambda & -1 \\ -1 & \lambda \end{bmatrix} \\
&= \lambda^2 - 1 = 0
\end{aligned}$$

Therefore, matrix A has two eigen values. $\lambda_1 = 1$ and $\lambda_2 = -1$. Matrix A is non-negative but $\lambda_2 = -1 < 0$.

1.5 Prob 5.

(a)

$$\frac{\partial(a^T x)}{\partial x} = \frac{\partial(x^T a)}{\partial x} = a$$
$$\nabla f(x) = a$$

(b) Suppose $y = Px$,

$$\begin{aligned}\frac{\partial(x^T Px)}{\partial x} &= \frac{\partial(x^T y)}{\partial x} + \frac{\partial y^T}{\partial x}\frac{\partial(x^T y)}{\partial y} \\ &= y + \frac{\partial(x^T P^T)}{\partial x}\frac{\partial(x^T y)}{\partial y} \\ &= y + P^T x = (P + P^T)x\end{aligned}$$

Therefore,

$$\nabla f(x) = \frac{1}{2}(P + P^T)x + q$$

(c) Because $P = P^T$,

$$\nabla f(x) = \frac{1}{2}(P + P^T)x = Px$$

(d) Suppose $g(x) = exp(a^T x + b)$, then $f(x) = f(g(x)) = \dfrac{g(x)}{1 + g(x)}$

$$\frac{\partial g(x)}{\partial x} = \frac{\partial(exp(a^T x + b))}{\partial x} = exp(a^T x + b)a$$

Therefore,

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

1.6 Prob 6.

Suppose $X^T = [X_1^T, X_2^T, \ldots X_n^T]^T$, where $X_i \in R^m$ is the ith column of X. Similarly $Y = [Y_1, Y_2, \ldots, Y_n]$ where $Y_i \in R^m$ is the ith column of Y. Therefore, let $C = X^T Y$ then,

$$C = \begin{bmatrix} X_1^T Y1 & X_1^T Y2 & \ldots & X_1^T Y_n \\ X_2^T Y1 & X_2^T Y2 & \ldots & X_2^T Y_n \\ \ldots & \ldots & & \ldots \\ X_n^T Y1 & X_n^T Y2 & \ldots & X_n^T Y_n \end{bmatrix}$$

Therefore,

$$\langle X, Y\rangle = tr(X^T Y) = \sum_{i=1}^{n} X_i^T Y_i = \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ji}y_{ji} \tag{3}$$

Property1: symmetry

$$\begin{aligned}\langle X, Y\rangle &= tr(X^T Y) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} y_{ji}x_{ji} \\ &= tr(Y^T X) \\ &= \langle Y, X\rangle\end{aligned}$$

Property2: linearity

Since $\text{tr}(X+Y) = \text{tr}(X) + \text{tr}(Y)$ and $\text{tr}(\lambda X) = \lambda \text{tr}(X)$.

$$\begin{aligned}\langle cA + B, Y\rangle &= \text{tr}((cA+B)^T Y) \\ &= \text{tr}(cA^T Y + B^T Y) \\ &= c\text{tr}(A^T Y) + \text{tr}(B^T Y) \\ &= c\langle A, Y\rangle + \langle B, Y\rangle\end{aligned}$$

Property3: definiteness

From equation 3, $\langle A, A\rangle = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji}^2$. Therefore, $\langle A, A\rangle = 0$ iff $A = 0$, otherwise $\langle A, A\rangle > 0$.

1.7 Prob 7.

Proof:

Suppose $x^T Ax = a$. Because $x \in R^n, A \in R^{n\times n}$, then $a \in R$, then $(x^T Ax)^T = a^T = a$. Because $(AA^T)^T = A^T A$,

$$\begin{aligned}(x^T Ax)^T &= x^T A^T x \\ &= x^T(-A)x \\ &= -(x^T Ax) \\ &= -a\end{aligned}$$

Therefore, $a = -a, a = 0$, namely $x^T Ax = 0$.

1.8 Prob 8.

Proof:

Because $\text{E}(X) = 0, \text{E}(Y) = 0$. Therefore, $Var(X) = E(X^2) - E(X)^2 = E(X^2)$ and similarly, $Var(Y) = E(Y^2)$. $Cov(X,Y) = E(XY) - E(X)E(Y) = E(XY)$. According to Cauchy - Schwarz inequality $|\langle X, Y\rangle|^2 \leq \langle X, X\rangle\langle Y, Y\rangle$

$$\begin{aligned}|\operatorname{Cov}(X,Y)|^2 &= |\operatorname{E}(XY)|^2 \\ &= |\langle X, Y\rangle|^2 \\ &\leq \langle X, X\rangle\langle Y, Y\rangle \\ &= \operatorname{E}(X^2)\operatorname{E}(Y^2) \\ &= \operatorname{Var}(X)\operatorname{Var}(Y)\end{aligned}$$