HW4 CS 4501 - SLGM Rishi Shah

① a) $K = \mathrm{cov}(x) = \mathrm{cov}(x, x) = E[xx^T] - E[x]E[x^T]$

$\therefore K^T = (E[xx^T] - E[x]E[x^T])^T = E[xx^T]^T - (E[x]E[x^T])^T$

$= E[(xx^T)^T] - E[x^T]^T E[x]^T = E[(x^T)^T x^T] - E[(x^T)^T]E[x^T]$

$= E[xx^T] - E[x]E[x^T] = \mathrm{cov}(x, x) = \mathrm{cov}(x) = K$

$\Rightarrow K = K^T \quad \therefore K$ is symmetric.

b) $v^T K v = v^T(E[xx^T] - E[x]E[x^T])v$

$= v^T E[xx^T]v - v^T E[x] E[x^T v] = E[v^T x x^T v] - E[v^T x]E[x^T v]$

$= E[(v^T x)(v^T x)^T] - E[v^T x]E[(v^T x)^T] = E[(v^T x)^2] - E[v^T x]^2$

since $v^T x$ is a scalar.

$= \mathrm{Var}(v^T x) \geq 0$

since variance is always nonnegative.

$\therefore K$ is positive semi-definite.

c) i. $\hat{K}^T = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^T\right)^T$

$= \frac{1}{n}\sum_{i=1}^{n}((x_i - \mu)(x_i - \mu)^T)^T = \frac{1}{n}\sum_{i=1}^{n}((x_i - \mu)^T)^T(x_i - \mu)^T$

$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^T = \hat{K} \Rightarrow \hat{K}^T = \hat{K}$

$\therefore \hat{K}$ is symmetric.

ii. $v^T K v = v^T\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^T\right)v$

$= \frac{1}{n}\sum_{i=1}^{n} v^T(x_i - \mu)(x_i - \mu)^T v$

$= \frac{1}{n}\sum_{i=1}^{n} v^T(x_i - \mu)(v^T(x_i - \mu))^T$

$= \frac{1}{n}\sum (v^T(x_i - \mu))^2 \geq 0$ since $v^T(x_i - \mu)$ is a scalar.

$\therefore \hat{K}$ is positive semidefinite.

② Find $\hat{K}_{MLE}$: we know $l(\mu, K) = \ln(p(D;\mu,K)) = \sum_{i=1}^{n} \ln((p(x_i;\mu,K)))$

$$= \frac{n}{2}\ln|K^{-1}| - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^T K^{-1}(x_i-\mu)$$

$$\frac{\partial l(\mu,K)}{\partial(K)} = \frac{n}{2}K^T - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)(x_i-\mu)^T = \frac{n}{2}K^T - \frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})$$

Given that $\frac{\partial}{\partial A}x^TAx = xx^T$, $\frac{\partial}{\partial A}\ln|A| = A^{-T}$, and $\mu_{MLE} = \bar{x}$.

$$0 = \frac{n}{2}K^T - \frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^T \Rightarrow K^T = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^T = K$$

since we know K is symmetric.

$$\therefore \hat{K}_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^T$$

③ $\text{Bias}(\hat{\mu}) = E[(\hat{\mu}-\mu)] = E[(\frac{1}{n}\sum_{i=1}^{n}x_i)-\mu] = E[\frac{1}{n}\sum_{i=1}^{n}x_i] - \mu$

$$= \frac{1}{n}\sum_{i=1}^{n}E[x_i] - \mu = \frac{1}{n}(n\mu) - \mu = \mu - \mu = \underline{0}$$

$$\text{Var}(\hat{\mu}) = E[(\hat{\mu}-\mu)(\hat{\mu}-\mu)^T] = \frac{1}{n^2}E[n(\hat{\mu}-\mu)n(\hat{\mu}-\mu)^T]$$

$$= \frac{1}{n^2}E[\sum_{i=1}^{n}(x_i-\mu)\sum_{j=1}^{n}(x_j-\mu)^T] = \frac{1}{n^2}\sum_{i,j=1}^{n}E[(x_i-\mu)(x_j-\mu)^T]$$

$$= \frac{1}{n^2}\left(\sum_{i,j=1,\, i\neq j}^{n}E[(x_i-\mu)(x_j-\mu)^T] + \sum_{i,j=1,\, i=j}^{n}E[(x_i-\mu)(x_j-\mu)^T]\right)$$

$$= \frac{1}{n^2}\left(0 + \sum_{i=1}^{n}E[(x_i-\mu)(x_i-\mu)^T]\right)$$

$$= \frac{1}{n^2}(nK) = \frac{K}{n}$$

$$\boxed{\text{Bias}(\hat{\mu}) = 0, \quad \text{Var}(\hat{\mu}) = \frac{K}{n}}$$