12011923 ~~07~~ 张旭东

Q1: (a) True.

(b) Consider the quadratic form of $x_a x_b$:

$$\Delta^2 = \frac{1}{2}\begin{pmatrix} x_a^T - \mu_a^T & x_b^T - \mu_b^T \end{pmatrix}\begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix}\begin{pmatrix} x_a - \mu_a \\ x_b - \mu_b \end{pmatrix}$$

And ~~we~~ the coefficient between x_a^T and x_a is the $\Sigma_{a|b}$, the ~~coffi~~ coefficient of x^T is $\mu_{a|b}$. so. the following expression will be got:

$$\cancel{\Sigma_{ab}}\ \Sigma_{a|b} = \cancel{A}\ \Lambda_{aa}^{-1}$$

$$\mu_{a|b} = \mu_a - \cancel{A}\Lambda_{aa}^{-1}\Lambda_{ab}(x_a - \mu_b).$$

And the inverse of Σ:

$$\begin{pmatrix} (A-BD^{-1}C)^{-1} & -(A-BD^{-1}C)^{-1}BD^{-1} \\ -DC(A-BD^{-1}C) & \cdots \end{pmatrix}.$$

Thus, the following expression ~~get~~ will be got:

$$\Sigma_{a|b} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$$

$$\mu_{a|b} = \mu_a - \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b).$$

Q2: (a). Because z follows normal distribution, x also follows the normal distribution, we need to calculate the coefficient of the quadratic form and the coefficient of x^T:

$$\Delta^2 = \frac{1}{2}\begin{pmatrix} x_a^T - \mu_a^T & x_b^T - \mu_b^T \end{pmatrix}\begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix}\begin{pmatrix} x_a - \mu_a \\ x_b - \mu_b \end{pmatrix}$$

$$= -\frac{1}{2}x_b^T\Lambda_{bb}x_b + x_b^T[\Lambda_{bb} - \mu_b - \Lambda_{ba}(x_a - \mu_a)]$$

$$= -\frac{1}{2}x_b^T\Lambda_{bb}x_b + x_b^T m \cancel{\ldots}, \quad m = \Lambda_{bb} - \mu_b - \Lambda_{ba}(x_a - \mu_a)$$

transform the above equation into quadratic form:

$$-\frac{1}{2}(x_b - \Lambda_{bb}^{-1}m)\Lambda_{bb}(x_b - \Lambda_{bb}^{-1}m)$$

then the quadratic containing x_a becomes:

$$-\frac{1}{2}x_a^T(\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})x_a + x_a^T(\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})\mu_a.$$

~~the~~ After that, the following expression will be got:

$$\begin{cases} \Sigma_a = (\Lambda_{aa} - \Lambda_{ab}\Lambda_{ab}^{-1}\Lambda_{ab})^{-1} \\ \mu_a = \Sigma_a(\cancel{A}\Lambda_{aa} - \Lambda_{ab}\Lambda_{ab}^{-1}\Lambda_{ab})^{-1}\mu_a = \mu_a. \end{cases}$$

then consider the inverse of Σ, we get

$$\begin{cases} \Sigma_a = \Sigma_{aa} = \Lambda^{-1} \\ \mu_a = \mu_a \end{cases}, \quad \text{which implies } x_a \sim \mathcal{N}(\mu_a, \Lambda^{-1}).$$

b): Firstly, get the inverse Λ of the covariance matrix $\vec{R}$;

$$\vec{R} = \begin{pmatrix} \Lambda + A^T L A & -A^T L \\ -LA & L \end{pmatrix}.$$

According to Q1; we get:

$$y|x \sim \mathcal{N}(y|\mu_{y|x}, \Sigma_{y|x}).$$

where:
$$\begin{cases} \mu_{y|x} = \mu_y - L^{-1}(-LA)(x-\mu) = A\mu + b + Ax - A\mu = Ax + b \\ \Sigma_{y|x} = L^{-1} \end{cases}$$

Q3: Derivative the function by A and let it be zero, we get the derivation form:

$$\sum_{k=1}^{N} (A^{-1}(x_n-\mu))(A^{-1}(x_n-\mu))^T - N(A^{-1})^T = 0$$

Thus: $[A^{-1}(\sum_{k=1}^{N}(x_n-\mu)(x_n-\mu)^T) - N](A^{-1})^T = 0$.

we get: $A = \frac{1}{N}\sum_{k=1}^{N}(x_n-\mu)(x_n-\mu)^T$

Q4: a) $\sigma^2_{ML}(N) = \frac{1}{N}\sum_{k=1}^{N}(x_n-\mu)^2$

$$= \frac{1}{N}[(N-1)\sigma^2_{ML}(N-1) + (x_n-\mu)^2]$$

$$= \sigma^{2\,(N)}_{ML}(N-1) + \frac{1}{N}[(x_n-\mu)^2 - \sigma^{2\,(N-1)}_{ML}]$$

Using the Robbins-Monro method, we get the expression of z:

$$\frac{\partial \ln p}{\partial(\sigma^2)} = \frac{1}{2\sigma^2}\left[\frac{(x-\mu)^2}{\sigma^2} - 1\right]$$

substitute the form into the formula, and choose $\alpha_N = \frac{2\sigma^4}{N}$:

$$\sigma^{2\,(N)}_{ML} = \sigma^{(N-1)}_{ML} + \frac{2\sigma^4}{N}\frac{1}{2\sigma^2}\left[\frac{(x-\mu)^2}{\sigma^2} - 1\right] = \frac{1}{1}\sigma^{(N-1)}_{ML} + \frac{1}{N}\left((x_n-\mu)^2 - \sigma^{2\,(N-1)}_{ML}\right].$$

b). Using the symbol ~~of~~ A of replace covariance matrix Σ.

$$\cancel{A^{(N)}}\ A^{(N)} = \frac{1}{N}\sum_{k=1}^{N}(x_k-\mu)(x_N-\mu)^T$$

$$= \frac{1}{N}[(N-1)A^{(N-1)} + (x_N-\mu)(x_N-\mu)^T].$$

$$= A^{(N-1)} + \frac{1}{N}[(x_N-\mu)(x_N-\mu)^T - A^{(N-1)}]$$

As the same: $\frac{\partial \ln(p)}{\partial A} = -\frac{A^{-1}}{2}[1-(x-\mu_{ML})(x-\mu_{ML})^T A^{-1}]$.

Then: $A^{(N)} = \cancel{A}A^{(N-1)} + \alpha_N \frac{A^{-1}}{2}\left[(x_N-\mu_{ML})(x_N-\mu_{ML})^T - A^{(N-1)}\right]$

Let $\alpha_N = \frac{2A}{N}$, we get the same result.

Q5: According to the formula given by the text:

$$\mu_N = \frac{\sigma^2}{N\sigma_0^2+\sigma^2}\mu_0 + \frac{N\sigma_0^2}{N\sigma_0^2+\sigma^2}\mu_{ML}, \quad \frac{1}{\sigma_N^2} = \frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}.$$

Thus:
$$\sigma_N = \frac{\sigma\sigma_0}{\sqrt{\sigma^2+N\sigma_0^2}}.$$

where: $\sigma = \sqrt{\Sigma}$, $\sigma_0 = \sqrt{\Sigma_0}$, $\mu_{ML} = \frac{1}{N}\sum_{n=1}^{N} x_n$. And $p(\mu|\vec{X}) = N(\mu|\mu_N, \sigma_N^2)$.

we get:
$$p(\mu_N|\vec{X}) = \frac{1}{\sqrt{2\pi}\sigma_N}\exp\left\{-\frac{(x-\mu_N)^2}{2\sigma_N^2}\right\}.$$

with $\mu_N = \frac{1}{N}\sum_{n=1}^{N} x_n$, $\sigma_N = \frac{\sigma\sigma_0}{\sqrt{\sigma^2+N\sigma_0^2}}$.