mlw2

Nguyen Tuan Duy

September 9, 2022

1 problem 1

1)

a)

$$n(x|\mu,\sigma^2) = \frac{1}{((2\pi\sigma^2)^{1/2}} exp(\frac{-1}{2\sigma^2}(x-\mu)^2)$$

the distribution is normalized when

$$\int (x|\mu,\sigma^2) = 1$$

$$I = \int exp(\frac{-1}{2\sigma^2}(x-\mu)^2)dz$$

Making the transformation from Cartesian coordinates (x, y) to polar coordiante(r, θ)

$$x = rcos(\theta) \tag{1}$$
$$y = rsin(\theta) \tag{2}$$

$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial(x)}{\partial(r)} & \frac{\partial(x)}{\partial(\theta)} \\ \frac{\partial(y)}{\partial(r} & \frac{\partial(y)}{\partial(\theta)} \end{vmatrix} = \begin{vmatrix} cos(\theta) & -rsin(\theta) \\ sin(\theta) & rcos(\theta) \end{vmatrix} = r$

$x^2 + y^2 = r^2cos^2(\theta) + r^2sin^2(\theta) = r^2$

$I^2 = \int\int exp(\frac{-1}{2\sigma^2}(x^2+y^2))dydx = \int_0^{2\pi}\int_0^{\infty} exp(\frac{-1}{2\sigma^2}r^2)d\theta dr$

Let u = r^2, $du = 2rdr$

$$I^2 = 1/2\int_0^{2\pi}\int_0^{\infty} exp(\frac{-1}{2\sigma^2}u)dud\theta = 1/2\int_0^{2\pi} 2\sigma^2 d\theta = 2\pi\sigma^2 \tag{3}$$

Now we can calculate:

$$n(x|\mu,\sigma^2) = n(x|\mu,\sigma^2) = \frac{1}{((2\pi\sigma^2)^{1/2}} exp(\frac{-1}{2\sigma^2}(x-\mu)^2) \tag{4}$$

let u = (x - μ), du = dx

$$n(x|\mu, \sigma^2) = \frac{1}{((2\pi\sigma^2)^{1/2}} exp(\frac{-1}{2\sigma^2}(u)^2) \tag{5}$$

$$= \frac{1}{((2\pi\sigma^2)^{1/2}} \sqrt{2\pi\sigma^2} = 1 \tag{6}$$

the gaussian distribution is normalized

b)

$$E(x) = \int n(x|\mu, \sigma^2) x dx$$
$$= \int \frac{1}{((2\pi\sigma^2)^{1/2}} exp(\frac{-1}{2\sigma^2}(x-\mu)^2) x dx$$

Let u = x - μ, du = dx

$$E(x) = \int \frac{1}{((2\pi\sigma^2)^{1/2}} exp(\frac{-1}{2\sigma^2} u)(u+\mu)$$
$$= \frac{1}{((2\pi\sigma^2)^{1/2}} [((2\pi\sigma^2)^{1/2}(u+\mu) - u((2\pi\sigma^2)^{1/2}]$$
$$= u + \mu - u$$
$$= \mu$$

c) we have from part a:

$$\int exp(\frac{-1}{2\sigma^2}(x-\mu)^2) dz = (2\pi\sigma^2)^{1/2}$$

by differentiating both side of this equation:

$$\frac{(x-\mu)^2}{\sigma^3} \int \exp \frac{-1}{2\sigma^2}(x-\mu)^2 dx = \sqrt{2\pi}$$
$$\frac{1}{\sqrt{2\pi\sigma}} \int exp(\frac{-1}{2\sigma^2}(x-\mu)^2) dx (x-\mu) = \sigma^2$$
$$E[(x-\mu)^2] = \sigma^2$$
$$E[x^2 - 2x\mu + \mu^2] = \sigma^2$$
$$E[x^2] - 2\mu^2 + \mu^2 = \sigma^2$$
$$E[x^2] = \sigma^2 + \mu^2$$
$$E[x^2] - E[x] = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

d)

$$N(x|\mu, \Sigma) = \int \frac{1}{(2\pi)^D} \frac{1}{|\Sigma|^{1/2}} exp(\frac{-1}{2}(x-\mu)^T \Sigma^- 1 (x-\mu)) dx$$

let's consider

$$\Delta^2 = (x-\mu)^T \Sigma^- 1(x-\mu)$$

eigenvector equation for the covariance matrix:

$$\Sigma u_i = \lambda_i u_i$$

because Σ is a real, symmetric matrix so:

$$u_i^T u_j = I_{ij}$$

$$\Sigma = \sum_{i=1}^{D} \lambda_i u_i u_i^T$$

$$\Sigma^{-1} = \sum_{i=1}^{D} \frac{1}{\lambda_i} u_i u_i^T$$

Let

$$y_i = \frac{u_i^T (x-\mu)}{\lambda_i}$$

then

$$\begin{aligned}\Delta^2 &= (x-\mu)^T \Sigma^- 1(x-\mu) \\ &= \sum_{i=1}^{D} \frac{u_i^T (x-\mu)}{\lambda_i} \\ &= \sum_{i=1}^{D} \frac{y_i^2}{\lambda_i}\end{aligned}$$

$$|\Sigma| = \prod_{i=1}^{D} \lambda_i$$

$$P(y) = \prod_{i=1}^{D} \frac{1}{(2\pi\lambda_i)^{1/2}} exp(\frac{y_i^2}{2\lambda_i})$$

$$\int P(y) = \int \prod_{i=1}^{D} \frac{1}{(2\pi\lambda_i)^{1/2}} exp(\frac{y_i^2}{2\lambda_i}) = \prod_{i=1}^{D} \frac{1}{(2\pi\lambda_i)^{1/2}} (2\pi\lambda_i)^{1/2} = 1$$

2 problem 2

a)

$$x = \begin{bmatrix} x_a \\ x_b \end{bmatrix} \tag{7}$$

$$\mu = \begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix} \tag{8}$$

covariance matrix:

$$\Sigma = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix} \tag{9}$$

$$\Lambda \equiv \Sigma^{-1} \tag{10}$$

$$\Lambda = \begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{bmatrix} \tag{11}$$

$$P(a|b) = \frac{p(a,b)}{p(b)} \tag{12}$$

$$= \frac{\frac{1}{\sqrt{2\pi^D |\Sigma|}} exp(\frac{-1}{2}(\begin{bmatrix} x_a - \mu_a \\ x_b - \mu_b \end{bmatrix})^T \Sigma^{-1} \begin{bmatrix} x_a - \mu_a \\ x_b - \mu_b \end{bmatrix}}{\frac{1}{\sqrt{2\pi^D |\Sigma_{bb}|}} exp(\frac{-1}{2}(\begin{bmatrix} x_b - \mu_b \end{bmatrix})^T \Sigma_{bb}^{-1} \begin{bmatrix} x_b - \mu_b \end{bmatrix}} \tag{13}$$

$$= \frac{1}{\sqrt{2\pi^{D-D_b}}} \frac{|\Sigma_{bb}|}{|\Sigma|} exp(\frac{-1}{2}(\begin{bmatrix} x_a - \mu_a \\ x_b - \mu_b \end{bmatrix})^T \Sigma^{-1} \begin{bmatrix} x_a - \mu_a \\ x_b - \mu_b \end{bmatrix} + \frac{1}{2}(\begin{bmatrix} x_b - \mu_b \end{bmatrix})^T \Sigma_{bb}^{-1} \begin{bmatrix} x_b - \mu_b \end{bmatrix} \tag{14}$$

we can find the covariance matrix:

$$|\Sigma| = \begin{vmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{vmatrix} = |\Sigma_{aa}\Sigma_{bb} - \Sigma_{ba}\Sigma ab| = |\Sigma_{bb}||\Sigma_{aa} - \Sigma_{ba}\Sigma_{bb}^{-1}\Sigma_{ba}| \tag{15}$$

$$\frac{|\Sigma_{bb}|}{|\Sigma|} = \frac{1}{|\Sigma_{aa} - \Sigma_{ba}\Sigma_{bb}^{-1}\Sigma_{ba}|} \tag{16}$$

the inverse of a block is:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1}CMBD^{-1} \end{bmatrix} \tag{17}$$

$$\begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}^{-1} = \begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{bmatrix} \Lambda = (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1} \tag{18}$$

$$\Sigma_{a|b} = \Lambda_{aa}^{-1} \tag{19}$$

the exponential part of (14) can be written as:

$$
\begin{aligned}
&\exp\left[-\frac{1}{2}\big((x_a-\mu_a)^{\mathrm{T}}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}(x_a-\mu_a)\ - \right.\\
&\qquad 2(x_a-\mu_a)^{\mathrm{T}}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}(x_b-\mu_b)+\\
&\qquad (x_b-\mu_b)^{\mathrm{T}}\left[\Sigma_{bb}^{-1}+\Sigma_{bb}^{-1}\Sigma_{ba}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}\right](x_b-\mu_b)\big)\\
&\qquad \left.+\frac{1}{2}\big((x_b-\mu_b)^{\mathrm{T}}\Sigma_{bb}^{-1}(x_b-\mu_b)\big)\right]\\
&=\exp[-\frac{1}{2}(x_a-\mu_a)^{\mathrm{T}}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}(x_a-\mu_a)\\
&-2(x_a-\mu_a)^{T}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}(x_b-\mu_b)\\
&(x_b-\mu_b)^{\mathrm{T}}\Sigma_{bb}^{-1}\Sigma_{ba}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}(x_b-\mu_b)]\\
&=exp[-\frac{1}{2}(x_a-(\mu_a+\Sigma_{ab}\Sigma_{bb}^{-1})(x_b-\mu_b))^{T}(\Sigma_{aa}-\Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}(x_a-(\mu_a+\Sigma_{ab}\Sigma_{bb}^{-1})(x_b-\mu_b))];
\end{aligned}
\tag{20}
$$

Comparing (20) with the exponential part of the multivariate gaussian distribution:

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

along with (19) we have:

$$p(x_a|x_b)=N(x|\mu_{a|b},\Lambda_{aa}^{-1})$$

b)

$$p(x_a)=\int p(x_a,x_b|\mu,\Sigma)dx_b \tag{21}$$

$$=\frac{1}{(2\pi)^{D/2}|\Sigma^{-1}|^{1/2}|}\int exp(-\frac{1}{2}[x-\mu]^T\Sigma^{-1}[x-\mu])dx \tag{22}$$

$$=\frac{1}{Z}\int exp(-\frac{1}{2}[x_a-\mu_a]^T\Lambda_{aa}^{-1}[x_a-\mu_a] \tag{23}$$

$$+\frac{1}{2}[x_a-\mu_a]^T\Lambda_{ab}^{-1}[x_b-\mu_b] \tag{24}$$

$$+\frac{1}{2}[x_b-\mu_b]^T\Lambda_{ba}^{-1}[x_a-\mu_a] \tag{25}$$

$$+\frac{1}{2}[x_b-\mu_b]^T\Lambda_{bb}^{-1}[x_b-\mu_b])dx_b \tag{26}$$

Using coletion of squares:

$$\frac{1}{2}z^TAz+b^Tz=\frac{1}{2}(z+A^{-1}b)^TA(z+A^{-1}b)+c-\frac{1}{2}b^TA^{-1}b$$

let:

$$
\begin{aligned}
z = x_b - \mu_b \\
A &= \Lambda_{bb} \\
b &= \Lambda_{ba}[x_a - \mu_a] \\
c &= -\frac{1}{2}[x_a - \mu_a]^T \Lambda_{aa}^{-1}[x_a - \mu_a]
\end{aligned}
$$

$$
\begin{aligned}
p(x_a) &= \frac{1}{z}\int exp[-[\frac{1}{2}(x_b - \mu_b + \Lambda_{bb}^{-1}\Lambda_{ab}(x_a - \mu_a)]^T \Lambda_{bb}[(x_b - \mu_b + \Lambda_{bb}^{-1}\Lambda_{ab}(x_a - \mu_a)] \\
&+ \frac{1}{2}[x_a - \mu_a]^T \Lambda_{aa}^{-1}[x_a - \mu_a] - \frac{1}{2}[x_a - \mu_a]^T \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba}(x_a - \mu_a)]dx_b \\
&= exp[-\frac{1}{2}(x_a - \mu_a)^T \Lambda_{aa}(x_a - \mu_a) + \frac{1}{2}(x_a - \mu_a)^T \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba}(x_a - \mu_a)] \\
&+ \frac{1}{z}\int exp(-\frac{1}{2}[x_b - \mu_b + \Lambda_{bb}^{-1}\Lambda_{ba}(x_a - \mu_a)]^T \Lambda_{bb}[x_b - \mu_b + \Lambda_{bb}^{-1}\Lambda_{ba}(x_a - \mu_a)])dx_b \\
&= \frac{1}{z}exp[-\frac{1}{2}(x_a - \mu_a)^T(\Lambda_{aa} + \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})(x_a - \mu_a)]](2\pi)^{-D/2}|\Lambda_{bb}|^{-1/2}
\end{aligned}
$$

Here, we notice that $\Lambda_{aa} + \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba} = \Sigma_{aa}$
we can compare this to the multivariate gaussian distribution to see that the mean vector is μ_a and the covariance matrix is Σ_{aa}
Therefore we can conclude:

$$
p(x_a) = N(x|\mu_a, \Sigma_{aa}) \tag{27}
$$