Homework W2

Nguyen Ba Dang Khoi

September 2021

1 Multivariate Gaussian Distribution

The distribution is normalized if the area under the curve is equal to 1.

Assume

$$\mathrm{I} = \int_{-\infty}^{+\infty} p(x|\mu, \sigma^2) = \int_{-\infty}^{+\infty} \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \exp^{(-\frac{1}{2}(x-\mu)^\intercal \Sigma^{-1}(x-\mu))}\, dx$$

Set

$$\begin{aligned}
\Delta^2 &= -\frac{1}{2}(x-\mu)^\intercal \Sigma^{-1}(x-\mu) \\
&= -\frac{1}{2}(x^\intercal - \mu^\intercal)\Sigma^{-1}(x-\mu) \\
&= -\frac{1}{2}x^\intercal \Sigma^{-1} x + \frac{1}{2}x^\intercal \Sigma^{-1}\mu + \frac{1}{2}\mu^\intercal \Sigma^{-1} x - \frac{1}{2}\mu^\intercal \Sigma^{-1}\mu \qquad (1)
\end{aligned}$$

With

$$-\frac{1}{2}\mu^\intercal \Sigma^{-1}\mu$$

is a constant, so we will mark it as "const."

Now, we proof that:

$$\frac{1}{2}x^\intercal \Sigma^{-1}\mu = \frac{1}{2}\mu^\intercal \Sigma^{-1} x$$

Firstly, we have

$$\frac{1}{2}x^\intercal \Sigma^{-1}\mu \in R => (\frac{1}{2}x^\intercal \Sigma^{-1}\mu)^\intercal = \frac{1}{2}x^\intercal \Sigma^{-1}\mu \qquad (2)$$

Secondly, we have

$$(\frac{1}{2}x^{\intercal}\Sigma^{-1}\mu = \frac{1}{2}\mu^{\intercal}\Sigma^{-1}x \text{ (because } (a^{\intercal}b)^{\intercal} = b^{\intercal}a) \tag{3}$$

From (2), (3)

$$=> \frac{1}{2}x^{\intercal}\Sigma^{-1}\mu = \frac{1}{2}\mu^{\intercal}\Sigma^{-1}x$$

Continue from (1), we have

$$\Delta^2 = -\frac{1}{2}x^{\intercal}\Sigma^{-1}x + x^{\intercal}\Sigma^{-1}\mu + const. \tag{4}$$

Equation (4) will be used later for conditional Gaussian distribution proof.

Consider eigenvalues λ and eigenvectors u of Σ:

$$\Sigma u_i = \lambda u_i$$

PROOF: If A is a symmetric matrix then its eigenvectors are orthorgonal
For any real matrix A and any vectors x and y, we have

$$< Ax, y >=< x, A^{\intercal}, y >$$

Now assume that A is symmetric, and x and y are eigenvectors of A corresponding to distinct eigenvalues λ and μ, then

$$\lambda < x, y >=< \lambda x, y >=< Ax, y >=< x, A^{\intercal}y >=< x, Ay >=< x, \mu y >= \mu < x, y > .$$

Therefore, $(\lambda - \mu) < x, y >= 0$. Since $\lambda - \mu \neq 0$, then ¡x,y¿ = 0, i.e, $x \perp y$.

Σ can be written as follows

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

PROOF (HAVE NOT BEEN ABLE TO PROOF IT YET):
If

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

then

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

With Σ and Σ^{-1}, we have the new Δ^2

$$\begin{aligned}
\Delta^2 &= -\frac{1}{2}(x-\mu)^\intercal \Sigma^{-1}(x-\mu) \\
&= -\frac{1}{2}\sum_{i=1}^{D}\frac{1}{\lambda_i}(x-\mu)^\intercal u_i u_i^\intercal (x-\mu) \\
&= -\frac{1}{2}\sum_{i=1}^{D}\frac{y_i^2}{\lambda_i}, \text{ with } y_i = u_i^\intercal (x-\mu)
\end{aligned} \tag{5}$$

And

$$\mid \Sigma \mid^{\frac{1}{2}} = (\prod_{i=1}^{D}\lambda_i)^{\frac{1}{2}} \tag{6}$$

We have

$$p(x) = \frac{1}{(2\pi)^{\frac{D}{2}} \mid \Sigma \mid^{\frac{1}{2}}} \exp^{(-\frac{1}{2}(x-\mu)^\intercal \Sigma^{-1}(x-\mu))} dx$$

From p(x), (5) and (6), we have the new p(y)

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

Then we have the new I equals to

$$\begin{aligned}
=> I &= \int_{-\infty}^{+\infty} p(y)dy \\
&= \prod_{i=1}^{D}\int_{-\infty}^{+\infty}\frac{1}{(2\pi\lambda_i)^{\frac{1}{2}}} \exp^{-\frac{y_i^2}{2\lambda_i}} dy_i
\end{aligned} \tag{7}$$

Leave equation (7) aside for a bit, we have a well-known equation

$$\int \exp^{-\frac{1}{2}at^2} = (\frac{2\pi}{a})^{\frac{1}{2}} \tag{8}$$

Going back to equation (7), we have

$$I = \prod_{i=1}^{D}\frac{1}{(2\pi\lambda_i)^{\frac{1}{2}}}\int_{-\infty}^{+\infty} \exp^{-\frac{y_i^2}{2\lambda_i}} dy_i$$

Apply equation (8) with $a = \frac{1}{\lambda_i}$

$$= \prod_{i=1}^{D} \frac{1}{(2\pi\lambda_i)^{\frac{1}{2}}} (2\pi\lambda_i)^{\frac{1}{2}}$$

$$=> I = 1$$