Gaussian Distribution

Mai Tra My

August 2022

1 Proof that

(a) Gaussian distribution is normalized
To proof a distribution is normalized:

$$\int_{-\infty}^{\infty} p(x|\mu,\sigma^2)dx = 1$$

Gaussian distribution:

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

Therefore, we need to proof:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\{\frac{-(x-\mu)^2}{2\sigma^2}\}dx = 1 \tag{1}$$

Let $t = x - \mu$

$$dt = dx$$

$$(1) \leftrightarrow \int_{-\infty}^{\infty} \exp\{\frac{-t^2}{2\sigma^2}\}dt = \sqrt{2\pi\sigma^2} \tag{2}$$

Let $I = \int_{-\infty}^{\infty} \exp\{\frac{-t^2}{2\sigma^2}\}dt$

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{-\frac{t^2}{2\sigma^2} - \frac{y^2}{2\sigma^2}\}dtdy$$

Let $t = rcos\theta, y = rsin\theta$

$$\begin{aligned} dtdy &= \begin{vmatrix} \frac{\partial t}{\partial r} & \frac{\partial t}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} drd\theta \\ &= \begin{vmatrix} cos\theta & -rsin\theta \\ sin\theta & rcos\theta \end{vmatrix} drd\theta \\ &= r(cos^2\theta + sin^2\theta)drd\theta \\ &= rdrd\theta \end{aligned}$$

Let $u = r^2 \rightarrow du = 2rdr$

Therefore,

$$\begin{aligned}
I^2 &= \int_0^{2\pi}\int_0^{\infty} \exp\{-\frac{r^2}{2\sigma^2}\} r dr d\theta \\
&= 2\pi \int_0^{\infty} \exp\{-\frac{r^2}{2\sigma^2}\} r dr \\
&= \pi \int_0^{\infty} \exp\{-\frac{u}{2\sigma^2}\} du \\
&= -2\sigma^2\pi(e^{-\infty} - e^0) \\
&= 2\sigma^2\pi \\
I &= \sqrt{2\pi\sigma^2}
\end{aligned}$$

We have equation (2) proved.

(b) Expectation of Gaussian distribution is μ

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\{\frac{-(x-\mu)^2}{2\sigma^2}\} dx
\end{aligned}$$

Subtitute $t = \frac{x-\mu}{\sqrt{2\sigma^2}} \to dt = \frac{1}{2\sigma^2} dx$

$$\begin{aligned}
E(X) &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2} t + \mu) \exp\{-t^2\} dt \\
&= \frac{1}{\sqrt{\pi}} (\sqrt{2\sigma^2} \int_{-\infty}^{\infty} t * \exp\{-t^2\} dt + \mu \int_{-\infty}^{\infty} exp\{-t^2\} dt) \\
&= \frac{1}{\sqrt{\pi}} (\frac{-\sqrt{2\sigma^2}}{2} \exp\{-t^2\}]_{-\infty}^{\infty} + \mu \int_{-\infty}^{\infty} exp\{-t^2\} dt) \\
&= \frac{1}{\sqrt{\pi}} (0 + \mu \int_{-\infty}^{\infty} exp\{-t^2\} dt)
\end{aligned}$$

Using polar coordinates methods, we proved

$$\int_{-\infty}^{\infty} exp\{-t^2\} dt = \sqrt{\pi}$$

So

$$E(X) = \mu$$

(c) Variance of Gaussian distribution is σ^2
Considering the simplest Gaussian distribution, standardized normal distribution.

$$\begin{aligned}
var(X) &= E(X^2) - E(X)^2 \\
E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \exp\{\frac{-1}{2} x^2\} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \exp\{\frac{-1}{2} x^2\} x dx
\end{aligned}$$

Let $y = \frac{x^2}{2} \to dy = x dx$

$$
\begin{aligned}
E(X^2) &= 2\frac{1}{\sqrt{2\pi}}\int_0^\infty \sqrt{2y}\exp\{y\}dy \\
&= 2\frac{1}{\sqrt{\pi}}\int_0^\infty \sqrt{y}\exp\{y\}dy \\
&= \frac{2}{\sqrt{\pi}}\int_0^\infty y^{(0.5+1)-1}\exp\{y\}dy
\end{aligned}
$$

Using Gamma function:

$$
\begin{aligned}
\Gamma(a) &= \int_0^\infty y^{a-1}e^{-y}dy \\
&\rightarrow E(X^2) = \frac{2}{\sqrt{\pi}}\Gamma(\frac{1}{2}+1) \\
&= \frac{2}{\sqrt{\pi}}\frac{1}{2}\Gamma(\frac{1}{2}) \\
&= 1
\end{aligned}
$$

In general, N(T) is Gaussian distribution

$$
\begin{aligned}
var(T) &= var(\sigma X + \mu) \\
&= \sigma^2 var(X) \\
&= \sigma^2
\end{aligned}
$$

(d) Multivariate Gaussian distribution is normalized
For a D-dimensional vector x

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

μ is a D-dimensional mean vector, Σ is a DxD covariance matrix. Set

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

is a quadratic form of Gaussian distribution Considering eigenvalues and eigenvectors of Σ

$$
\Sigma u_i = \lambda_i u_i, i = 1, ..., D
$$

Because Σ is a real, symmetric matrix, its eigenvalues will be real and its eigenvectors form an orthonormal set.

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

So,

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

with $y_i = u_i^T(x-\mu)$

$$p(y) = \prod_{j=1}^{D} \frac{1}{2\pi\lambda_j}^{1/2} \exp\{-\frac{y_j^2}{2\lambda_j}\}$$

$$\Longrightarrow \int_{-\infty}^{\infty} p(y)dy = \prod_{j=1}^{D} \int_{-\infty}^{\infty} \frac{1}{2\pi\lambda_j}^{1/2} \exp\{-\frac{y_j^2}{2\lambda_j}\}dy_j = 1$$

2 Calculate

(a) The conditional of Gaussian distribution.
Suppose x: D-dimensional vector with Gaussian distribution
We partition x into 2 disjoint subsets x_a, x_b

$$x = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$$

We define mean vector

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$$

and covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \to A = \Sigma^{-1} = \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix}$$

Σ is symmetric $\to \Sigma_{aa}, \Sigma_{bb}$ are symmetric;
$\Sigma_{ab} = \Sigma_{ba}^T$
We have

$$\begin{aligned} -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) &= -\frac{1}{2}(x-\mu)^T A(x-\mu) \\ &= -\frac{1}{2}(x_a-\mu_a)^T A_{aa}(x_a-\mu_a) - \frac{1}{2}(x_a-\mu_a)^T A_{ab}(x_b-\mu_b) \\ &- \frac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-mu_a) - \frac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-mu_b) \\ &= -\frac{1}{2}x_a^T A_{aa}^{-1}x_a + x_a^T(A_{aa}\mu_a - A_{ab}(x_b-\mu_b)) + const \end{aligned}$$

It is quadratic form of x_a Comparing with Gaussian distribution

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

By using Schur comlement,

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} M & -MBD^{-1} \\ -D^{-1}CMD^{-1} & D^{-1}CMBD^{-1} \end{pmatrix}$$
$$M = (A - BD^{-1}C)^{-1}$$
$$\longrightarrow A_{aa} = (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1} A_{ab} = -(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}$$

Therefore,

$$\begin{aligned} \mu_{a|b} &= \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b-\mu_b) \\ \Sigma_{a|b} &= \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba} \\ \longrightarrow p(x_a|x_b) &= N(x_{a|b}|\mu_{a|b}, \Sigma_{a|b}) \end{aligned}$$

(b) The marginal of Gaussian distribution. The marginal distribution given by

$$p(x_a) = \int p(x_a; x_b)dx_b$$

We need to integrate out x_b by looking the quadratic form related to x_b

$$-\frac{1}{2}x_b^T A_{bb} x_b + x_b^T m = -\frac{1}{2}(x_b - A_{bb}^{-1}m)^T A_{bb}(x_b - A_{bb}^{-1}m) + \frac{1}{2}m^T A_{bb}^{-1} m$$

with $m = A_{bb}\mu_b - A_{ba}(x_a - \mu_a)$ We can integrate over unnormalized Gaussian

$$\int \exp\{-\frac{1}{2}(x_b - Abb^{-1}m)^T A_{bb}(x_b - A_{bb}^{-1}m)\}dx_b$$

The remaining term

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

Similarly, we have

$$E[x_a] = \mu_a cov[x_a] = \Sigma_{aa} \longrightarrow p(x_a) = N(x_a|\mu_a; \Sigma_{aa})$$