Machine Learning 1 - Week 2

Nguyễn Thị Kiều Nhung - 11203041

August 24, 2022

Problem 1. Prove

(a) Gaussian distribution is normalized

To prove the Gaussian distribution is normalized, I will first show that it is normalized for a zero-mean Gaussian and extend that result to show that $\mathcal{N}(x|\mu, \sigma^2)$ is normalized.

The pdf of the zero-mean Gaussian distribution is given by:

$$\mathcal{N}(x|\mu = 1, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}x^2\right) \quad -\infty < x < \infty. \tag{1}$$

To prove that the above expression is normalized, we have to show that

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

Or

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

Let

$$I = \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}x^2\right) dx \tag{3}$$

Squaring the above expression,

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}y^2\right) dx\, dy \tag{4}$$

To integrate this expression we make the transformation from (x, y) coordinates to (r, θ) coordinates, which is defined by

$$x = r\, cos\, \theta \tag{5}$$
$$y = r\, sin\, \theta \tag{6}$$

We also have:

$$dx\, dy = \left|\frac{\partial\,(x, y)}{\partial\,(r, \theta)}\right| dr\, d\theta$$

and using the trigonometric identity $cos^2\theta + sin^2\theta = 1$, we have $x^2 + y^2 = r^2$. Also the Jacobian of the change of variables is given by,

$$
\begin{aligned}
&= \begin{vmatrix} \dfrac{\partial(x)}{\partial(r)} & \dfrac{\partial(x)}{\partial(\theta)} \\[2ex] \dfrac{\partial(y)}{\partial(r)} & \dfrac{\partial(y)}{\partial(\theta)} \end{vmatrix} \\[2ex]
&= \begin{vmatrix} cos\,\theta & -r\,sin\,\theta \\ sin\,\theta & r\,cos\,\theta \end{vmatrix} \\
&= r\,cos^2\theta + r\,sin^2\theta \\
&= r
\end{aligned}
$$

using the same trigonometric identity $cos^2\theta + sin^2\theta = 1$. So,

$$
\begin{aligned}
I^2 &= \int_0^{2\pi}\int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\,dr\,d\theta && (7)\\
&= 2\pi\int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\,dr && (8)\\
&= 2\pi\int_0^{\infty} \exp\left(-\frac{u}{2\sigma^2}\right)\frac{1}{2}\,du && (9)\\
&= \pi\left[\exp\left(-\frac{u}{2\sigma^2}\right)\left(-2\sigma^2\right)\right]_0^{\infty} && (10)\\
&= 2\pi\sigma^2 && (11)
\end{aligned}
$$

where from (8) to (9), we have used the change of variables $r^2 = u$. Thus

$$I = \left(2\pi\sigma^2\right)^{1/2}.$$

Finally to prove that $\mathcal{N}(\mathrm{x}|\mu,\,\sigma^2)$ is normalized, we make the tranformation $y = x - \mu$ so that,

$$
\begin{aligned}
\int_{-\infty}^{\infty}\mathcal{N}(x \mid \mu,\sigma^2)\,dx &= \frac{1}{\left(2\pi\sigma^2\right)^{1/2}}\int_{-\infty}^{\infty}\exp\left(-\frac{y^2}{2\sigma^2}\right)\,dy \\
&= \frac{I}{\left(2\pi\sigma^2\right)^{1/2}} \quad (\textit{substitute}\ \mathrm{I}) \\
&= 1
\end{aligned}
$$

as required.

(b) Expectation of Gaussian distribution is μ (mean)

Probability Density Function:

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

From the definition of the Expectation of Continuous Random Variable:

$E(X) = \int_{-\infty}^{\infty} x f(x) dx$

So:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \end{aligned} \tag{12}$$

Let

$$t = \frac{x-\mu}{\sigma\sqrt{2}}$$
$$\implies t^2 = \frac{(x-\mu)^2}{2\sigma^2}$$
$$dx = \sqrt{2}\sigma \, dt$$

Substituting,

$$E(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (t\sigma\sqrt{2}+\mu) \exp\left(-t^2\right) dt\sqrt{2}\sigma$$

$$= \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^{\infty} t\sigma\sqrt{2} \times \exp\left(-t^2\right) dt + \int_{-\infty}^{\infty} \mu \times \exp\left(-t^2\right) dt \right]$$

$$= \frac{1}{\sqrt{\pi}} \left[\sqrt{2}\sigma \int_{-\infty}^{\infty} t \times \exp\left(-t^2\right) dt + \mu \int_{-\infty}^{\infty} \exp\left(-t^2\right) dt \right]$$

$$= \frac{1}{\sqrt{\pi}} \left(\sqrt{2}\sigma \left[\frac{-1}{2} \times \exp\left(-t^2\right) \right]_{-\infty}^{\infty} + \mu\sqrt{\pi} \right)$$

$$= \frac{1}{\sqrt{\pi}} \left(\sqrt{2}\sigma(0-0) + \mu\sqrt{\pi} \right)$$
$$= \frac{\sigma\sqrt{\pi}}{\sqrt{\pi}}$$
$$= \mu$$

(c) Variance of Gaussian distribution is σ^2 (variance)

By definition:

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

Let

$$t = \frac{x - mu}{\sqrt{2}\sigma}$$

$$\implies dt = \frac{dx}{\sqrt{2}\sigma}$$

$$x = \sigma t\sqrt{2} + \mu$$

$$\implies x^2 = 2t^2\sigma^2 + t2\sqrt{2}\sigma t\mu + \mu^2$$

Substituting, we got:

$$V(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \exp{(-t^2)} dt\sqrt{2}\sigma - \mu^2$$

$$= \left[\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (2t^2\sigma^2 + t2\sqrt{2}\sigma t\mu + \mu^2) \exp{(-t^2)} dt \right] - \mu^2$$

$$= \frac{1}{\sqrt{\pi}} \left[2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp{(-t^2)} dt + 2\sqrt{2}\sigma\mu \int_{-\infty}^{\infty} t \exp{(-t^2)} dt + \mu^2 \int_{-\infty}^{\infty} \exp{(-t^2)} dt \right] - \mu^2$$

$$= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp{(-t^2)} dt + 2\sqrt{2}\sigma\mu \Big[\frac{-1}{2} \exp{(-t^2)} \Big]_{-\infty}^{\infty} + \mu^2\sqrt{\pi} \right) - \mu^2$$

$$= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp{(-t^2)} dt + 2\sqrt{2} \times 0 \right) + \mu^2 - \mu^2$$

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

Using integration ny parts:

$$u = t; \;\; dv = t \exp{(-t^2)} dt$$

$$du = dt; \;\; v = \frac{-1}{2} \exp{(-t^2)}$$

Then,

$$\int_{-\infty}^{\infty} u dv = uv - \int v du$$

$$= \frac{-1}{2} t \exp -(t^2) - \int_{-\infty}^{\infty} \frac{-1}{2} \exp(-t^2) dt$$

$$= \left[\frac{-t}{2} \exp(-t^2) \right]_{-\infty}^{\infty} + \frac{1}{2} \int \exp(-t^2)$$

Then,

$$\implies V(X) = \frac{2\sigma^2}{\sqrt{\pi}} \times \frac{1}{2} \int_{-\infty}^{\infty} \exp(-t^2) dt$$

$$= \frac{2\sigma^2 \sqrt{\pi}}{2\sqrt{\pi}}$$

$$= \sigma^2$$

(d) Multivariate Gaussian distribution is normalized

To proof that the Multivariate normal distribution is normalized, we first look at the form of the distribution:

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

Where:
μ is a D-dimensional mean vector,
Σ is a D $\times$ D covariance matrix,
$|\Sigma|$ denotes the determinant of Σ.
Quadratic form of Gaussian distribution:

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

$$= x^T \Sigma^{-1} x - x^T \Sigma^{-1} \mu - \mu^T \Sigma^{-1} x + \mu^T \Sigma^{-1} \mu$$

We can see that:

$$\left(x^T \Sigma^{-1} \mu \right)^T = \mu^T (\Sigma^{-1})^T x$$

Therefore:

$$\begin{aligned} \Delta^2 &= x^T \Sigma^{-1} x - x^T \Sigma^{-1} \mu - \mu^T \Sigma^{-1} x + \mu^T \Sigma^{-1} \mu \\ &= x^T \Sigma^{-1} x - 2x^T \Sigma^{-1} \mu + c \end{aligned} \tag{13}$$

Next, we have the eigenvalues and eigenvectors of Σ, i = 1,...,D:

$$\Sigma u_i = \lambda_i u_i$$

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

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

$$\implies \Sigma^{-1} = \sum_{n=1}^{D} \frac{1}{\lambda} u_i u_i^T$$

Next,

$$\begin{aligned} \Delta^2 = (x-\mu)^T \Sigma^{-1} (x-\mu) &= \sum_{n=1}^{D} \frac{1}{\lambda} (x-\mu)^T u_i u_i^T (x-\mu) \\ &= \sum_{n=1}^{D} \frac{{y_i}^2}{\lambda_i} \end{aligned} \tag{14}$$

with $y_i = u_i^T(x-\mu)$
We have that:

$$p(y) = \prod_{n=1}^{D} \frac{1}{2\pi\lambda}^{1/2} \exp\left(-\frac{{y_j}^2}{x\lambda_j}\right)$$

Integrating both sides:

$$\int_{-\infty}^{\infty} p(y) dy = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{2\pi\lambda}^{1/2} \exp\left(-\frac{{y_j}^2}{x\lambda_j}\right) dy_i$$

but we can see that:

$$\frac{1}{2\pi\lambda}^{1/2} \exp\left(-\frac{{y_j}^2}{x\lambda_j}\right) dy_i$$

have the total probability = 1 (part a)
So that:

$$\prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{2\pi\lambda}^{1/2} exp\left(-\frac{{y_j}^2}{x\lambda_j}\right) dy_i = \prod_{i=1}^{n} 1 = 1$$