Trần Thị Thu Trang_MSV_11208164

MACHINE LEARNING

HOMEWORK 2

1. Proof that:
a) Gaussian distribution is normalized

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

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

Let:

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

$$\Rightarrow \mathrm{I}^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}y^2\right)dxdy \quad (1)$$

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

$$x = r.\cos\theta$$

$$y = r.\sin\theta$$

$\sin^2\theta + \cos^2\theta = 1$, we have $x^2 + y^2 = r^2$

Also the Jacobian of the change of variables is given by,

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.\cos^2\theta + r\sin^2\theta = r$$

Thus equation (1) can be written as:

$$I^2 = \int_0^{2\pi}\int_0^{\infty} \exp\left(\frac{-r^2}{2\sigma^2}\right)r.dr.d\theta$$

$$= 2\pi\int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right)r.dr$$

Let: $r^2 = u$

- ⇨ $2r.dr = du$
- ⇨ $rdr = \frac{1}{2}du$

$$= \ 2\pi \int_0^{\infty} \exp\left(-\frac{u}{2\sigma^2}\right).\frac{1}{2}du$$

$$= \pi \left[\exp\left(-\frac{u}{2\sigma^2}\right).(-2\sigma^2)\right]_0^{\infty}$$

$$= 2\pi\sigma^2$$

$$\Rightarrow \ I = \sqrt{2\pi\sigma^2}$$

Finally to prove that $N(x \mid \mu, \sigma^2)$ is normalized, we make the tranformation

$y = x - \mu$ so that,

$$\int_{-\infty}^{\infty} N(x|\mu,\sigma^2)dx = \frac{1}{\sqrt{2\pi\sigma^2}}.\int_{-\infty}^{\infty} \exp\left(-\frac{y}{2\sigma^2}\right)dy$$

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

$$= \frac{\sqrt{2\pi\sigma^2}}{\sqrt{2\pi\sigma^2}} \ = \ 1$$

b) Expectation of Gaussian distribution is μ (mean)

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}.\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

From the definition of the expected value of a continuous random variable:

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

So:

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

let: $\frac{x-\mu}{\sqrt{2}\,\sigma} = t$ => $x = \sqrt{2}\sigma t + \mu$

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

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

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

$$= \frac{\mu\sqrt{\pi}}{\sqrt{\pi}}$$

$$= \mu$$

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

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} . \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

From Variance as Expectation of Square minus Square of Expectation:

$$V(X) = \int_{-\infty}^{\infty} x^2 f_X(x) dx - [E(X)]^2$$

So:

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

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

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

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

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

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

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

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

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

$$= \sigma^2$$

d) Multivariate Gaussian distribution is normalized

$P(x|\mu, \sum) = \frac{1}{(2\pi)^{\frac{n}{2}} . |\sum|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right)$ $(\mu \in \mathbb{R}^n)$

Set: $\Delta^2 = (x-\mu)^T \sum^{-1}(x-\mu) = -\frac{1}{2}x^T \sum^{-1} x + x^T \sum^{-1} \mu + const$

Consider eigenvalues and eigenvectors of $\sum$

$$\sum u_i = \lambda_i u_i \quad , \; i = 1,2,3, \dots \dots , n$$

Because $\sum$ is a real, symmetric matrix, its eigenvalues will be real and itseigenvectors form an orthonormal set.

$$\sum = \sum_{i=1}^{n} \lambda_i u_i u_i^T \;\; => \;\; \sum^{-1} = \sum_{i=1}^{n} \frac{1}{\lambda_i} u_i u_i^T$$

So that: $\Delta^2 = (x-\mu)^T \sum^{-1}(x-\mu) = \sum_{i=1}^{n} \frac{1}{\lambda_i}(x-\mu)^T u_i u_i^T (x-\mu) = \sum_{i=1}^{n} \frac{y_i^2}{\lambda_i}$ $\quad with \; y_i = u_i^T(x-\mu)|$

$$|\sum|^{\frac{1}{2}} = \prod_{j=1}^{n} \lambda_i^{1/2}$$

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

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

2. Calculate:

a) The conditional of Gaussian distribution

Suppose x is a D-dimensional vector with Gaussian distribution $N(x|\mu, \sum)$ and that we partition x into two disjoint subsets x_a and x_b

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

We also define corresponding partitions of the mean vector μ given by:

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

and of the covariance matrix $\sum$ given by:

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

Σ is symmetric so Σ_{aa} and Σ_{bb} are symmetric while $\Sigma_{aa} = \Sigma_{bb}$

We are looking for conditional distribution $P(x_a|x_b)$

We have :

$$-\frac{1}{2}(x-\mu)^T(x-\mu) = -\frac{1}{2}(x-\mu)^T A(x-\mu)$$

$$= -\frac{1}{2}(x_a-\mu_a)^T A_{aa}(x_a-\mu_a) - \frac{1}{2}(x_a-\mu_a)^T A_{ab}(x_b-\mu_b) - \frac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-\mu_a)$$

$$-\frac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-\mu_b)$$

$$= -\frac{1}{2}x_a^T A_{aa}^{-1} x_a + x_a^T\left(A_{aa}\mu_a - A_{ab}(x_b-\mu_b)\right) + const$$

$$\Delta^2 = -\frac{1}{2}x^T \sum^{-1} x + x^T \sum^{-1} \mu + const$$

$\Sigma_{a|b} = A_{aa}^{-1}$

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

By using Schur complement,

$A_{aa} = \left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1}$

$A_{ab} = -\left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}$

As result :

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

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

⇨ $p(x_a|x_b) = N(x_{a|b}|\mu_{a|b}, \Sigma_{a|b})$

b) The marginal of Gaussian distribution

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

$$-\frac{1}{2}x_b^T A_{bb} x_b + x_b^T m = -\frac{1}{2}\left(x_b - A_{bb}^{-1}m\right)^T A_{bb}\left(x_b - A_{bb}^{-1}m\right) + \frac{1}{2}m^T A_{bb}^{-1} m \quad with\ m = A_{bb}\mu_b - A_{ba}(x_a-\mu_a)$$

We can integrate over unnormalized Gaussian:

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

The remaining term

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

$E[x_a] = \mu_a$

$cov[x_a] = \Sigma_{aa}$

⇒ $p(x_a) = N(x_A|\mu_A, \Sigma_{aa})$