Math - 789 Homework - 3

Hitesh Vaidya
hv8322@rit.edu

February 17, 2020

1 Question 2.1

$$
\begin{aligned}
I = \int_{-\infty}^{\infty} p(x)dx &= \int_{-\infty}^{\infty} \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\} \\
\text{put } \frac{x-\mu}{\sigma} &= u \\
dx &= \sigma du \\
I &= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\{-\frac{u^2}{2}\} \, \sigma dt \\
I^2 &= (\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp{(-\frac{u^2}{2})} du)(\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp{(-\frac{u^2}{2})} du) \\
\text{change of variable in second bracket} \\
&= (\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp{(-\frac{u^2}{2})} du)(\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp{(-\frac{v^2}{2})} dv) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} \exp{(-\frac{u^2+v^2}{2})} dudv \\
put \; u = r\cos\theta, v &= r\sin\theta \\
\text{calculating jacobian matrix} \\
|J| &= \frac{\partial(u,v)}{\partial(r,\theta)} \\
&= \begin{vmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \theta} \end{vmatrix} \\
&= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\
&= r \\
\text{Now,} \\
I^2 &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} \exp\{-\frac{r^2}{2}\} \, |J| dr d\theta \\
\text{Let} r^2 &= t \\
rdr &= \frac{dt}{2} \\
\text{Therefore} \\
I^2 &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} \exp\{-\frac{t}{2}\} \frac{dt}{2} d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} [\frac{1}{2} \; -2\exp\{-\frac{t}{2}\}]_0^{\infty} \; d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} -(0-1) d\theta \\
I^2 &= 1 \\
\therefore I &= \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} xp(x)dx \\
&= \int_{-\infty}^{\infty} x\frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}\ dx \\
\text{put } x-\mu &= t \\
dx &= dt \\
\mathbb{E}(X) &= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} (t+\mu)\exp\{-\frac{t^2}{2\sigma^2}\}dt \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} t\exp\{-\frac{t^2}{2\sigma^2}\}dt + \mu\int_{-\infty}^{\infty} \exp\{-\frac{t^2}{2\sigma^2}\}dt \\
\text{using B.2 from appendix B,} & \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}[\int_{-\infty}^{\infty} t\exp\{-\frac{t^2}{2\sigma^2}\}dt\ +\ \mu(2\pi\sigma^2)^{\frac{1}{2}}] \\
\text{put } t^2 &= \mu \\
tdt &= \frac{du}{2} \\
\text{Therefore,} & \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}[\int_{-\infty}^{\infty} \exp\{-\frac{u}{2\sigma^2}\}\frac{du}{2} + \mu(2\pi\sigma^2)^{\frac{1}{2}}] \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}\{[2\sigma^2\exp(-\frac{u}{2\sigma^2})]_{-\infty}^{\infty}\frac{1}{2} + \mu(2\pi\sigma^2)^{\frac{1}{2}}\} \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}[0 + \mu(2\pi\sigma^2)^{\frac{1}{2}}] \\
\mathbb{E}(X) &= \mu
\end{aligned}
$$

$$\begin{aligned}
Var(X) &= \mathbb{E}[(x-\mu)^2] \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}\int_{-\infty}^{\infty}(x-\mu)^2\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}dx \\
\text{put } \frac{x-\mu}{\sigma} &= t \\
dx &= \sigma dt \\
&= \frac{\sigma^2}{(2\pi\sigma^2)^{\frac{1}{2}}}\int_{-\infty}^{\infty}t^2\exp\{-\frac{t^2}{2}\}\sigma dt \\
\text{using integration by parts,} & \\
u = t,\ v' = t\exp\{\frac{-t^2}{2}\} \implies u' = 1, v = -\exp\{\frac{-t^2}{2}\} & \\
&= \frac{\sigma^2}{\sqrt{2\pi}}[(-t\exp\{\frac{-t^2}{2}\}|_{-\infty}^{\infty}) + \int_{-\infty}^{\infty}\exp\{\frac{-t^2}{2}\}dt]
\end{aligned}$$

The first term equals to 0.
Second term equals to 1 because it is equal to total probability integral of $\mathcal{N}(0,1)$
Therefore,

$$Var(X) = \sigma^2$$

2 Problem 2.2

consider the transformation,

$$\widetilde{X} = U^T(X - \mu)$$
$$\sum = UDU^T$$
$$\text{we know, } UU^T = 1$$
$$\therefore \Sigma^{-1} = UD^{-1}U^T$$
$$\text{also, } U\widetilde{X} = X - \mu$$
$$X = U\widetilde{X} + \mu$$

$$p(X) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right\}$$

using properties from our transformation,

$$= \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(U\widetilde{X})^T UD^{-1}U^T U\widetilde{X}\right]$$
$$= \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\widetilde{X}^T D^{-1}\widetilde{X}\right]$$
$$= \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\Sigma_{i=1}^{d}\frac{\widetilde{X}^2}{\lambda_i}\right]$$
$$= \Pi_{i=1}^{d}\frac{1}{(2\pi)^{\frac{1}{2}}(\lambda_i)^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\frac{(\widetilde{X}_i)^2}{\lambda_i}\right]$$
$$p(X) = \Pi_{i=1}^{d} p(\widetilde{X}_i)$$

where $\widetilde{X}_i$ has mean 0 and standard deviation $\sqrt{\lambda_i}$

$$\text{each } \int p(\widetilde{X}_i)d\widetilde{X}_i = 1$$

$$\int p(\widetilde{X})d\widetilde{X} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} p(\widetilde{X}_i)d\widetilde{X}_1...d\widetilde{X}_d$$
$$\text{property, } \int p(X_i)dX_i = \int |U|p(\widetilde{X}_i)d\widetilde{X}_i$$
$$\text{also we know, } |U| = 1$$
$$= \int\int...\int p(\widetilde{X}_1)p(\widetilde{X}_2)...p(\widetilde{X}_d)d\widetilde{X}_1...d\widetilde{X}_d$$
$$= \int\int ...\Pi_(i=2)^d[\int p(\widetilde{X}_1)dX1](d\widetilde{X}_2...d\widetilde{X}_d)$$
$$= 1$$

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[U\widetilde{X} + \mu] \\
&= U\mathbb{E}[\widetilde{X}] + \mu
\end{aligned}
$$

using the property of transformation,

$$
\mathbb{E} = 0 + \mu
$$

$$
\begin{aligned}
Var[X] &= Var[U\widetilde{X} + \mu] \\
&= U^T Var(\widetilde{X}) U \\
&= U^T D U \\
Var[X] &= \Sigma
\end{aligned}
$$

3 Problem 2.5

Using equation 2.30 and 2.31, we look at the exponential term of the posterior distribution $p(\mu|X)$,

$$-\frac{1}{2\sigma^2}\Sigma_{n=1}^{N}(x_n-\mu)^2-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2=\frac{1}{2\sigma_N^2}(\mu-\mu_N)^2$$

Left hand side of this equation,

first term,

$$-\frac{\Sigma_{n=1}^{N}x_N^2}{2\sigma^2}+\frac{\Sigma_{n=1}^{N}x_N\mu}{\sigma^2}-\frac{\Sigma_{n=1}^{N}\mu^2}{2\sigma^2}$$

second term,

$$-\frac{\mu^2}{2\sigma_0^2}+\frac{\mu\mu_0}{\sigma_0^2}-\frac{\mu_0^2}{2\sigma_0^2}$$

putting above two terms together gives us following,

$$-\mu^2\left(\frac{N}{2\sigma^2}+\frac{1}{2\mu_0^2}\right)+\mu\left(\frac{\Sigma_{n=1}^{N}x_N}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right)-\left(\frac{\Sigma_{n=1}^{N}x_N^2}{2\sigma^2}+\frac{\mu_0^2}{2\sigma_0^2}\right)$$

on right hand side,

$$-\frac{\mu^2}{2\sigma_N^2}+\frac{\mu\mu_N}{\sigma_N^2}-\frac{\mu_N^2}{2\sigma_N^2}$$

equating terms from both LHS and RHS,

$$-\mu^2\left(\frac{N}{2^2}+\frac{1}{2\mu_0^2}\right)=-\frac{1}{2\sigma_N^2}\mu^2,\ \mu\left(\frac{\Sigma_{n=1}^{N}x_N}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right)=\frac{\mu_N}{\sigma_N^2}\mu$$

Therefore,

$$\frac{1}{\sigma_N^2}=\frac{N}{\sigma^2}+\frac{1}{\sigma_0^2}$$

using the prior knowledge that $\Sigma_{n=1}^{N}x_N=N\cdot\overline{x}$

$$\begin{aligned}\mu_N&=\sigma_N^2\left(\frac{\Sigma_{n=1}^{N}x_N}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right)\\&=\left(\frac{1}{\sigma_0^2+\frac{N}{\sigma^2}}\right)^{-1}\cdot\left(\frac{N\overline{x}}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right)\\&=\frac{\sigma_0^2\sigma^2}{\sigma^2+N\sigma_0^2}\cdot\frac{N\overline{x}\sigma_0^2+\mu_0\sigma^2}{\sigma\sigma_0^2}\\\mu_N&=\frac{N\sigma_0^2}{N\sigma_0^2+\sigma^2}\overline{x}+\frac{\sigma^2}{N\sigma_0^2+\sigma^2}\mu_0\end{aligned}$$

4 Problem 2.9

The program for this question is submitted along with this file with the name **2.9.ipynb**.

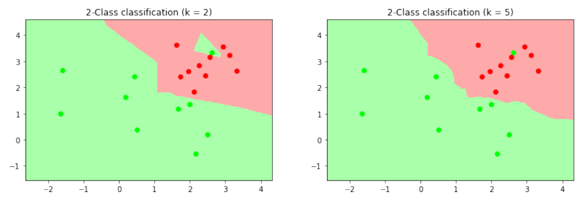

Figure 1: Graph $k = 2$ (left) and graph $k = 5$ (right).

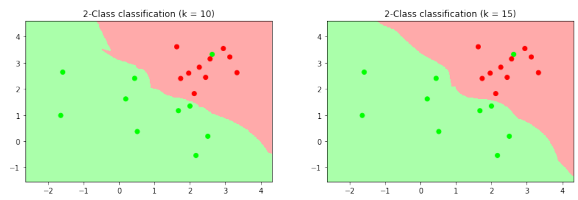

Figure 2: Graph $k = 10$ (left) and graph $k = 15$ (right).

we see that lower values of k i.e. 3 or 4 fits better as compared to other values of k.

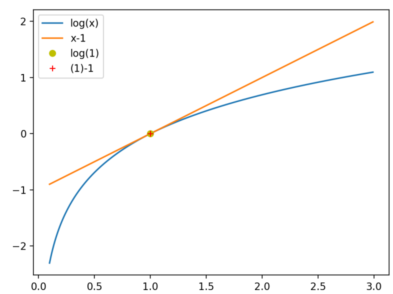

Figure 3: Graph of lnx and $x-1$

5 Problem 2.10

Differentiating $lnx - (x-1)$

$$f(x) = ln(x) - (x-1)$$
$$f'(x) = \frac{1}{x} - 1 = \frac{1-x}{x} \leq 0(x=1)$$
$$Since\ f(1) = -1 < 0\ and\ f(x)\ is\ decreasing\ for\ x > 1$$

$$\begin{aligned} L &= -\int p(x)\log\frac{p(x)}{q(x)}dx \\ &= \int p(x)\log\frac{q(x)}{p(x)}dx \\ &\leq \int p(x)\left(\frac{q(x)}{p(x)} - 1\right) \\ &= \int q(x) - \int p(x) \\ &= 1 - 1 = 0 \end{aligned}$$

The above proves true using the assumption $\int p(x) = 1\ and \int q(x) = 1$

6 Problem 2.11

The KL divergence can be given as follows,

$$L(p, q, \lambda) = \Sigma p(x) \log(\frac{q(x)}{p(x)}) - \lambda(\Sigma q_i - 1)$$

using the optimization we can obtain,

$$\log(\frac{q(x)}{p(x)}) + 1 + \lambda = 0$$

thus, we can conclude that, $p(x) = q(x) \exp \lambda - 1$

this is gaussian, hence the given constraint, $\Sigma_i q_i = 1$ is satisfied