MachineLearning Home Work-1

Sayantan Saha

29^{th} September 2019

1

1.1

E[x]=μ
$Var[x] = E[(X - \mu)^2]$
$= E[X^2 + \mu^2 - 2X\mu]$
$= E[X^2] + E[\mu^2] - 2\mu E[X]$
$= E[X^2] + \mu^2 - 2\mu^2$
$= E[X^2] - \mu^2$
$= E[X^2] - E[X]^2$

1.2

Cov(X,Y) = E[(X-E[X])(X-E[X])^T] $\quad whereE[X] = \mu E[X]^T = \mu^T$
$= E[(X - \mu)(X^T - \mu^T)]$
$= E[XX^T - X\mu^T - \mu X^T + \mu\mu^T]$
$= E[XX^T] - E[X]\mu^T - \mu E[X^T] + \mu\mu^T$
$= E[XX^T] - \mu\mu^T - \mu E[X^T] + \mu\mu^T$ $\quad E[X^T] = E[X]^T = \mu^T$
$= E[XX^T] - \mu\mu^T$

2

2.1

The normalizing constant is the factor that changes a probability function to a probability density function such that the probability density function can integrate to unit area.

p(x) = probability density function
f(x) = probability function

p(x) = C f(x) where c = normalizing constant

We know that $\int_{-\infty}^{\infty} p(x)dx = 1$

Given $p(X|L = l) = Ce^{\frac{-|x-a_l|}{b_l}}$ where l $\in 1, 2$
We will replace the above equation inplace of p(x) to get

$\int_{-\infty}^{\infty} e^{\frac{-|x-a_l|}{b_l}} dx = \frac{1}{C}$

For $x \geq a_l \rightarrow |x - a_l| = x - a_l$
For $x <$a$_l \rightarrow |x - a_l| = -(x - a_l)$
We can now divide the integrals into two parts of $-\infty$ to a_l and a_l to ∞

$\int_{-\infty}^{a_l} e^{\frac{x-a_l}{b_l}} dx + \int_{a_l}^{\infty} e^{\frac{-(x-a_l)}{b_l}} dx = \frac{1}{C}$

$e^{-\frac{a_l}{b_l}} \int_{-\infty}^{a_l} e^{\frac{x}{b_l}} dx + e^{\frac{a_l}{b_l}} \int_{a_l}^{\infty} e^{-\frac{x}{b_l}} dx = \frac{1}{C}$

We are converting the lower bound limit for the 1^{st} integral in terms of limit l_1 and upper bound limit for the 2^{nd} integral in terms of limits l_2

e$^{-\frac{a_l}{b_l}}$ $\lim_{l_1 \rightarrow -\infty} \int_{-l_1}^{a_l} e^{\frac{x}{b_l}} dx + e^{\frac{a_l}{b_l}} \lim_{l_2 \rightarrow \infty} \int_{a_l}^{l_2} e^{-\frac{x}{b_l}} dx = \frac{1}{C}$

$b_l e^{-\frac{a_l}{b_l}} \lim_{l_1 \rightarrow -\infty} \left[e^{\frac{x}{b_l}}\right]_{l_1}^{a_l} - b_l e^{\frac{a_l}{b_l}} \lim_{l_2 \rightarrow \infty} \left[e^{-\frac{x}{b_l}}\right]_{a_l}^{l_2} = \frac{1}{C}$

$b_l e^{-\frac{a_l}{b_l}} \lim_{l_1 \rightarrow -\infty} [e^{\frac{a_l}{b_l}} - e^{\frac{l_1}{b_l}}] - b_l e^{\frac{a_l}{b_l}} \lim_{l_2 \rightarrow \infty} [e^{-\frac{l_2}{b_l}} - e^{\frac{-a_l}{b_l}}] = \frac{1}{C}$
$b_l e^{-\frac{a_l}{b_l}} e^{\frac{a_l}{b_l}} - b_l e^{-\frac{a_l}{b_l}} \lim_{l_1 \rightarrow -\infty} e^{\frac{l_1}{b_l}} - b_l e^{\frac{a_l}{b_l}} \lim_{l_2 \rightarrow \infty} e^{-\frac{l_2}{b_l}} + be^{\frac{a}{b}} e^{\frac{a_l}{b_l}} = \frac{1}{C}$

As $\lim_{l_1 \rightarrow -\infty} e^{\frac{l_1}{b}} = 0$ and $\lim_{l_2 \rightarrow \infty} e^{-\frac{l_2}{b}} = 0$ we get

$b_l e^{-\frac{a_l}{b_l}} e^{\frac{a_l}{b_l}} + b_l e^{\frac{a_l}{b_l}} e^{-\frac{a_l}{b_l}} = \frac{1}{C}$

$\therefore C = \frac{1}{2b_l}$
$p(X|L = 1) = \frac{1}{2b_1} e^{\frac{-|x-a_1|}{b_1}}$

$p(X|L=2) = \frac{1}{2b_2}e^{\frac{-|x-a_2|}{b_2}}$

2.2

$l(x) = ln(p(X|L=1)) - ln(p(X|L=2))$

$ln(p(X|L=1)) = \frac{-|x-a_1|}{b_1} - ln(2b_1)$

$ln(p(X|L=2)) = \frac{-|x-a_2|}{b_2} - ln(2b_2)$

$l(x) = \frac{-|x-a_1|}{b_1} - ln(2b_1) - \frac{-|x-a_2|}{b_2} + ln(2b_2)$

$l(x) = \frac{|x-a_2|}{b_2} - \frac{|x-a_1|}{b_1} + ln\left[\frac{b_2}{b_1}\right]$

2.3

Given $a_1 = 0, b_1 = 1, a_2 = 1, b_2 = 2$ Substitute the following in $l(X)$

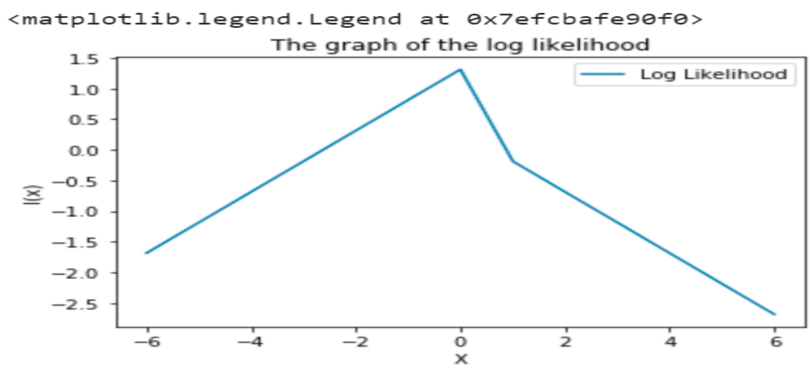

Figure 1: l(x) vs x

3

We know $P(\omega_1) = P(\omega_2) = 0.5$
$P(X|\omega_1) = \frac{1}{b-a}$
$P(X|\omega_2) = \frac{1}{t-r}$
$\frac{P(X|\omega_1)}{P(X|\omega_2)} = \frac{t-r}{b-a}$
We consider 0 - 1 Loss:
$\lambda_{11} = 0, \lambda_{12} = 1, \lambda_{21} = 1, \lambda_{22} = 0$

Minimum Probability of error classification Rule is given by:
$\frac{P(X|\omega_1)}{P(X|\omega_2)} \ \frac{>^{\omega 1}}{<_{\omega 2}} \ \frac{P(\omega_2)}{P(\omega_1)}$

$\frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}}$

substitute values that we know in the given equation

$\frac{t-r}{b-a} \ \frac{>^{\omega 1}}{<_{\omega 2}} \ 1$
$\frac{1}{b-a} \ \frac{>^{\omega 1}}{<_{\omega 2}} \ \frac{1}{t-r}$

if $\frac{1}{b-a}$ is greater than $\frac{1}{t-r}$ we will choose in favour of ω_1 and vice versa.

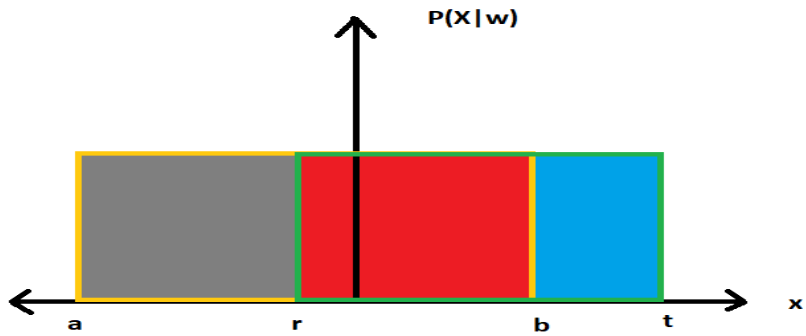

Figure 2: $P(x|\omega)$ vs x

In the above image the area in grey between a to r marks the region in which we select ω_1, the region in blue is where we select ω_2 between b and t. The region of indecision lies between r and b the red region. The size of the red region is dependent on the values of r and b.

Overall Risk (R)= $\int R(\alpha_i|X)P(X)dx$

$R(\alpha_i|X) = \sum_{j=1}^{2} \lambda_{ij} P(\omega_j|X)$

$$R = \int_{R_1} \lambda_{12} P(\omega_2) P(X|\omega_2) + \int_{R_2} \lambda_{21} P(\omega_1) P(X|\omega_1) \qquad\qquad as\ \lambda_{11} = \lambda_{22} = 0$$

4

Given : $P(X|\omega_1) = N(0,1) = \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}$ and $P(X|\omega_2) = N(\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-1}{2}\frac{(x-\mu)^2}{\sigma^2}}$

4.1

$\frac{P(X|\omega_1)}{P(X|\omega_2)} \gtrless^{\omega 1}_{\omega 2} \frac{P(\omega_2)}{P(\omega_1)}\frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}}$

The RHS of that equation is 1

$\frac{\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}}{\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-1}{2}\frac{(x-\mu)^2}{\sigma^2}}} \gtrless^{\omega 1}_{\omega 2} 1$

$\sigma e^{\frac{-x^2}{2}} \gtrless^{\omega 1}_{\omega 2} e^{\frac{-1}{2}\frac{(x-\mu)^2}{\sigma^2}}$ taking ln on both sides.

$ln(\sigma) - \frac{x^2}{2} \gtrless^{\omega 1}_{\omega 2} \frac{-1}{2}\frac{(x-\mu)^2}{\sigma^2}$ Bring the RHS to the LHS and make all the denominators the same

$ln(\sigma) - \frac{x^2}{2} + \frac{1}{2}\frac{(x-\mu)^2}{\sigma^2} \gtrless^{\omega 1}_{\omega 2} 0$

$2\sigma^2 ln(\sigma) - x^2\sigma^2 + x^2 + \mu^2 - 2x\mu \gtrless^{\omega 1}_{\omega 2} 0$ The common denominator $2\sigma^2$ is taken to the RHS and multiplied with zero.

We group terms so that we can get it in the quadratic form.

$x^2(1-\sigma^2) - 2x\mu + \mu^2 + 2\sigma^2 ln(\sigma) \gtrless^{\omega 1}_{\omega 2} 0$

Thus we derive the nice form.

4.2

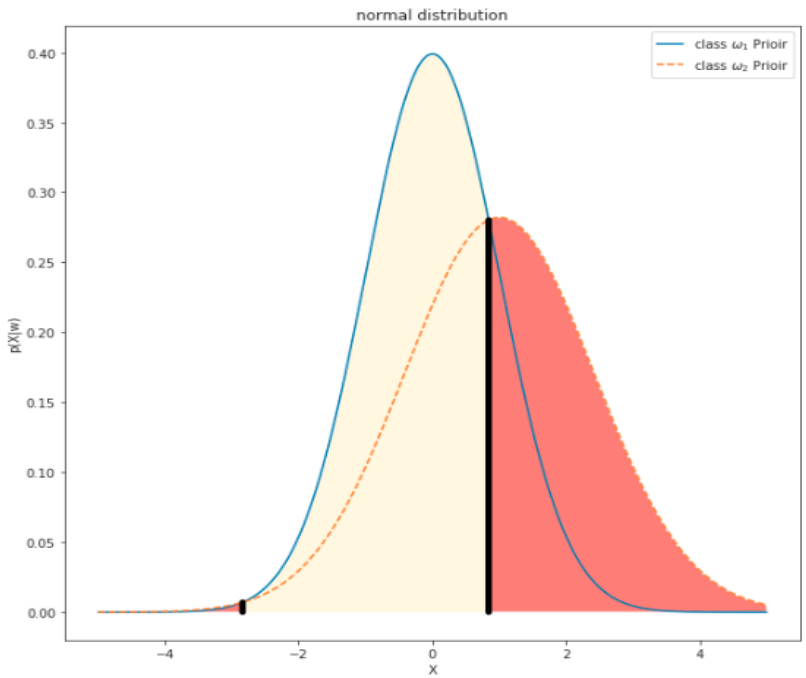

Figure 3: $P(X|w)$ vs x in prior terms

In Figure 3 the area shaded in light yellow signifies the area where we select ω_1, the region in salmon colour signifies the are where we will select ω_2 . The two black lines mark the decision boundary for the problem.

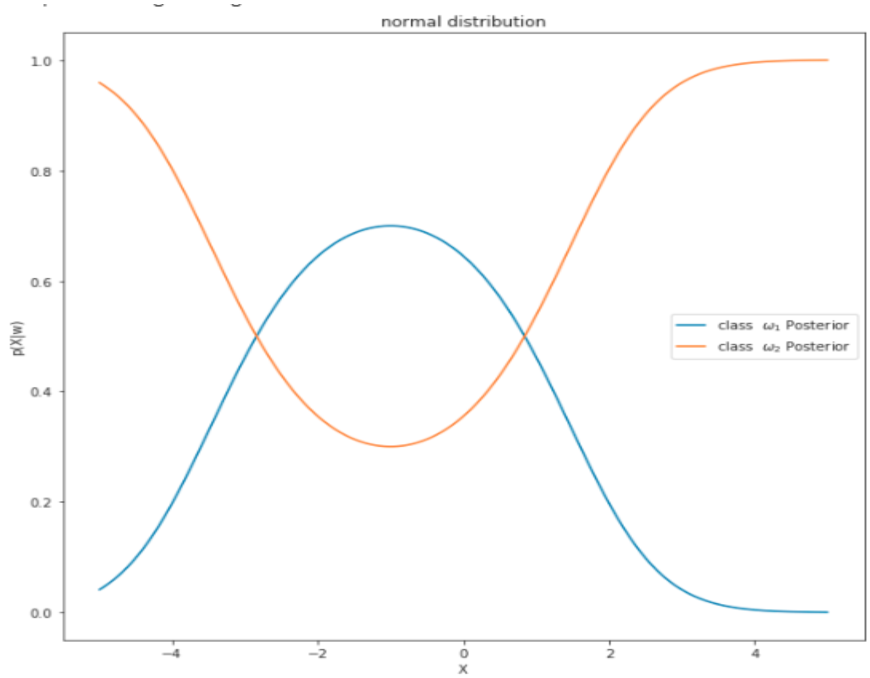

Figure 4: $P(X|w)$ vs x in posterior terms

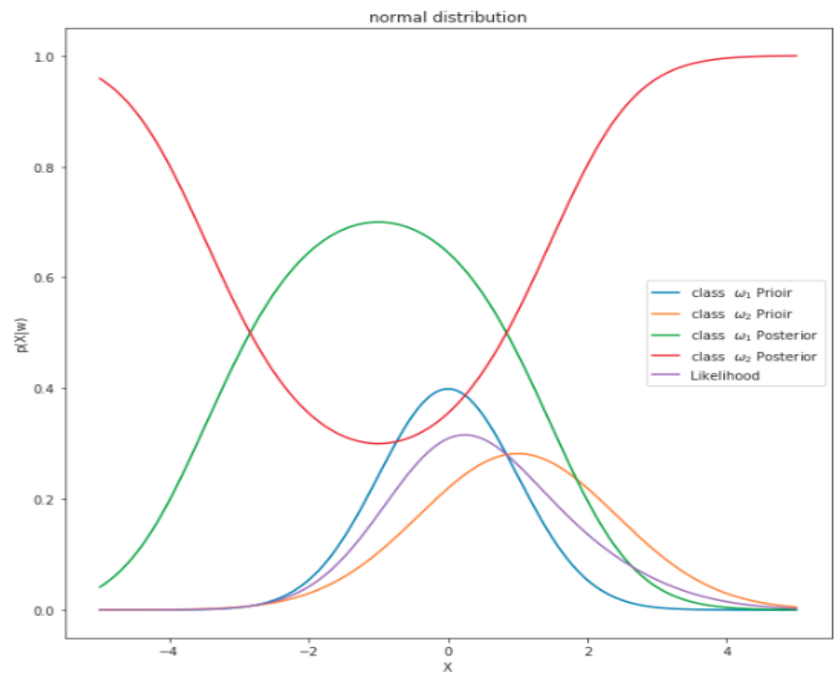

Figure 5: $P(X|w)$ vs x in prior and posterior terms

4.3

Minimum probability of error: $R = \int_{R_1} \lambda_{12} P(\omega_2) P(X|\omega_2) + \int_{R_2} \lambda_{21} P(\omega_1) P(X|\omega_1)$
as $\lambda_{11} = \lambda_{22} = 0$
See Figure 3 for reference. Within the three given decision regions we select the prior prob with the minimum probability to calculate the error.

$$= \frac{1}{2} \int_{-\infty}^{-2.84} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} + \frac{1}{2} \int_{0.84}^{\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} + \frac{1}{2} \int_{-2.84}^{0.84} \frac{e^{-\frac{x-1^2}{4}}}{\sqrt{4\pi}}$$

$$= \frac{1}{2} \left[\frac{1}{2} erf(\frac{x}{\sqrt{2}}) + C_1 \right]_{-\infty}^{-2.84} + \frac{1}{2} \left[\frac{1}{2} erf(\frac{x}{\sqrt{2}}) + C_2 \right]_{0.84}^{\infty} + \frac{1}{2} \left[\frac{1}{2} erf(\frac{x-1}{2}) + C_3 \right]_{-2.84}^{0.84}$$

$$= \frac{1}{2}(0.00225 + 0.2004 + 0.4518)$$

$$= 0.326$$

4.4

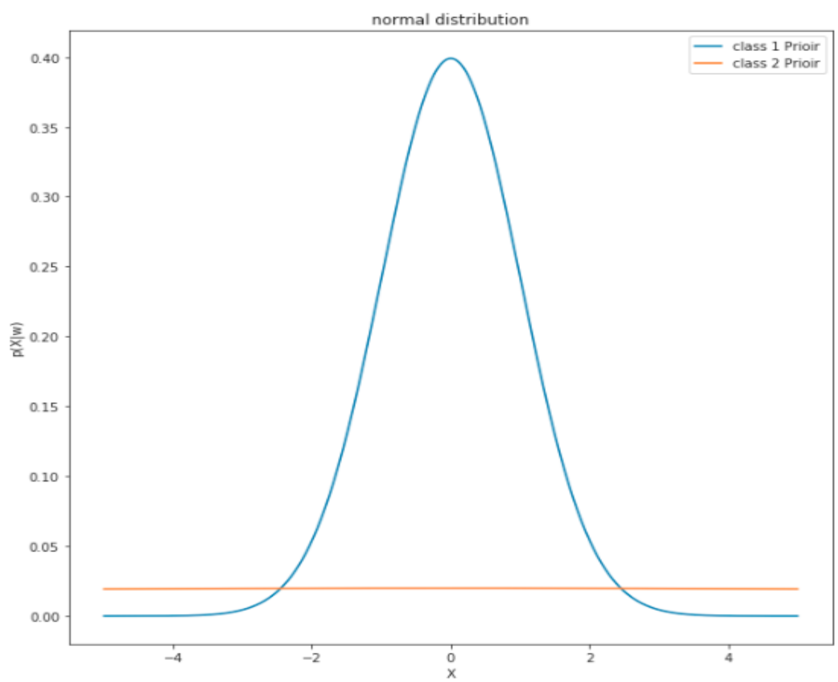

Figure 6: $P(X|w)$ vs x in prior terms

The case where $\mu = 0$ and $\sigma >> 1$ arises in a case of White gaussian noise.

5

5.1

Given :$f(\xi) = \frac{e^{(\frac{-1}{2}(\xi-\mu)^T(\xi-\mu)\frac{1}{\Sigma})}}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}$ ....Equation 1

We know $z \sim N(0, I)$ Assuming A is full rank we need to transform z to x such that $x = Az + b$.

$f(z) = \frac{e^{((\frac{-1}{2})z^T Iz)}}{(2\pi)^{\frac{n}{2}}}$ as $|\Sigma| = |I| = 1$, $\sqrt{1} = 1$ and $\Sigma^{-1} = I^{-1} = I$

Lets calculate the probability of $P(z < K)$.
$P(z < K) = \int_{-\infty}^{K} f(z)dz$ Substituting the value of the function from top.
$P(z < K) = \int_{-\infty}^{K} \frac{e^{(\frac{-1}{2}(z)^T I(z))}}{(2\pi)^{\frac{n}{2}}} dz$
Say we had to calculate $P(x < K) \to P(Az + b < K) = P(z < \frac{K-b}{A})$

$P(z < \frac{K-b}{a}) = \int_{-\infty}^{\frac{K-b}{A}} \frac{e^{(\frac{-1}{2}(z)^T I(z))}}{(2\pi)^{\frac{n}{2}}} dz$ equation for new limits

We are now performing variable transformation.
To calculate the probability in terms of x we will have to modify the integral limits, substitute $z = \frac{x-b}{A}$ and change $\frac{dx}{A} = dz$.

We know $x = Az + b$
$dx = Adz$
$\therefore dz = \frac{dx}{A}$
We have the equation for $P(z < \frac{k-b}{A})$, we are expressing the same probability but now in terms of x, we get $P(\frac{X-b}{A} < \frac{k-b}{A})$, which on simplification gives us $P(x < K)$. Thus the integrals will change back to $-\infty$ to K.
$P(x < K) = \int_{-\infty}^{K} \frac{e^{(\frac{-1}{2}(\frac{x-b}{A})^T I(\frac{x-b}{A}))}}{(2\pi)^{\frac{n}{2}}} \frac{dx}{A}$
$= \int_{-\infty}^{K} \frac{e^{(\frac{-1}{2}(\frac{x^T-b^T}{A^T})I(\frac{x-b}{A}))}}{(2\pi)^{\frac{n}{2}}A} dx$
$= \int_{-\infty}^{K} \frac{e^{(\frac{-1}{2AA^T}(x^T-b^T)I(x-b))}}{(2\pi)^{\frac{n}{2}}A} dx$
$= \int_{-\infty}^{K} \frac{e^{(\frac{-1}{2}((x-b)^T(x-b))\frac{1}{AA^T})}}{(2\pi)^{\frac{n}{2}}A} dx$ ....Equation 2

Comparing Equation 1 and Equation 2 we get the following results :
$\mu = b$
$\Sigma = AA^T$
$A = |\Sigma|^{\frac{1}{2}}$

We can now say $x \sim N(b, AA^T)$ on the linear transformation of z

5.2

From the conclusion above we get:
$b = \mu$
$A = |\Sigma|^{\frac{1}{2}}$
$\Sigma = AA^T$

5.3

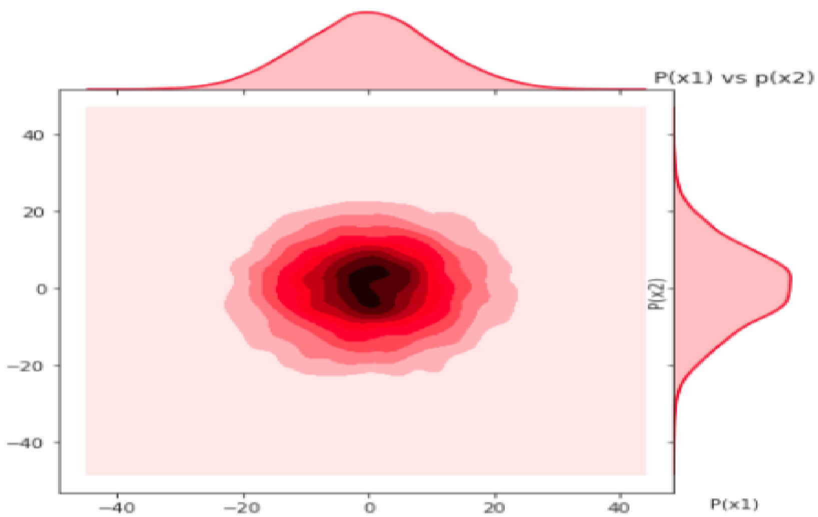

Figure 7: multivariate normal distribution

Please check the link below for code.

6

The following link leads to he github gist containing the code and graphs for the homework:
https://gist.github.com/sayan3710/ebdcddadfce96f852bff7a604f78069a