**Introduction to ML (CS771), Autumn 2018**
**Indian Institute of Technology Kanpur**
**Homework Assignment Number 2**

*Student Name:* Shashi Kant Gupta
*Roll Number:* 160645
*Date:* September 30, 2018

# QUESTION 1

According to the given problem:

$$p(\mathbf{w}) = C * exp(\frac{-\lambda}{2}\mathbf{w}^T\mathbf{w}) \tag{1}$$

$$p(\mathbf{y}|\mathbf{X},\mathbf{w}) = \prod_{n=1}^{N} \frac{1}{1 + exp(-y_n\mathbf{w}^T\mathbf{x}_n)} \tag{2}$$

Therefore, MAP estimate will be:

$$\hat{\mathbf{w}}_{MAP} = argmax_{\mathbf{w}} log(p(\mathbf{y}|\mathbf{X},\mathbf{w})) + log(p(\mathbf{w})) \tag{3}$$

$$\Rightarrow \hat{\mathbf{w}}_{MAP} = argmax_{\mathbf{w}} \sum_{n=1}^{N} -log(1 + exp(-y_n\mathbf{w}^T\mathbf{x}_n)) + \frac{-\lambda}{2}\mathbf{w}^T\mathbf{w} \tag{4}$$

$$\Rightarrow \hat{\mathbf{w}}_{MAP} = argmin_{\mathbf{w}} \sum_{n=1}^{N} log(1 + exp(-y_n\mathbf{w}^T\mathbf{x}_n)) + \frac{\lambda}{2}\mathbf{w}^T\mathbf{w} \tag{5}$$

Therefore, to minimise the negative log likelihood the partial derivatives w.r.t. to $\mathbf{w}$ yields.

$$\lambda\mathbf{w} + \sum_{n=1}^{N} \frac{-y_n\mathbf{x}_n exp(-y_n\mathbf{w}^T\mathbf{x}_n)}{1 + exp(-y_n\mathbf{w}^T\mathbf{x}_n)} = 0 \tag{6}$$

$$\Rightarrow \hat{\mathbf{w}}_{MAP} = \frac{1}{\lambda}\sum_{n=1}^{N} \frac{exp(-y_n\mathbf{w}^T\mathbf{x}_n)}{1 + exp(-y_n\mathbf{w}^T\mathbf{x}_n)} y_n\mathbf{x}_n \tag{7}$$

$$\therefore \alpha_n = \frac{1}{\lambda}(\frac{exp(-y_n\mathbf{w}^T\mathbf{x}_n)}{1 + exp(-y_n\mathbf{w}^T\mathbf{x}_n)}) \tag{8}$$

From the expression of $\alpha_n$, we can see that it specifies an scaled version of non-class probabilities for $xn$. This make sense as we can see from the probability expression for the right-class as the $\alpha_n$ increases the probability for the right-class will increase, which means if probability of wrong-class is high it will give us a new estimate of $\mathbf{w}$ such that probability of right class will increase.

**Introduction to ML (CS771), Autumn 2018**
**Indian Institute of Technology Kanpur**
**Homework Assignment Number 2**

# QUESTION 2


*Student Name:* Shashi Kant Gupta
*Roll Number:* 160645
*Date:* September 30, 2018


---

As per the problem:

$$p(y=1)=\pi \tag{9}$$

$$p(\mathbf{x}|y=1)=\prod_{d=1}^{D}\mu_{d,1}^{xd}(1-\mu_{d,1})^{1-xd} \tag{10}$$

$$p(\mathbf{x}|y=0)=\prod_{d=0}^{D}\mu_{d,0}^{xd}(1-\mu_{d,0})^{1-xd} \tag{11}$$

$$\therefore p(y=1|\mathbf{x})=\frac{\pi\prod_{d=1}^{D}\mu_{d,1}^{xd}(1-\mu_{d,1})^{1-xd}}{\pi\prod_{d=1}^{D}\mu_{d,1}^{xd}(1-\mu_{d,1})^{1-xd}+(1-\pi)\prod_{d=0}^{D}\mu_{d,0}^{xd}(1-\mu_{d,0})^{1-xd}} \tag{12}$$

$$=\frac{1}{1+\frac{1-\pi}{\pi}\prod_{d=1}^{D}[\frac{\mu_{d,0}}{\mu_{d,1}}]^{\mathbf{x}_d}[\frac{1-\mu_{d,0}}{1-\mu_{d,1}}]^{1-\mathbf{x}_d}} \tag{13}$$

$$=\frac{1}{1+f(\mathbf{x})} \tag{14}$$

where $f(\mathbf{x})=\frac{1-\pi}{\pi}\prod_{d=1}^{D}[\frac{\mu_{d,0}}{\mu_{d,1}}]^{\mathbf{x}_d}[\frac{1-\mu_{d,0}}{1-\mu_{d,1}}]^{1-\mathbf{x}_d}$
It can be easily shown that:

$$p(y=0|\mathbf{x})=\frac{1}{1+f(\mathbf{x})^{-1}} \tag{15}$$

Therefore, this makes a discriminative model with its distribution as $Bernoulli[g(f(\mathbf{x}))]$ where $g(f(\mathbf{x}))=\frac{1}{1+f(\mathbf{x})}$ and $f(\mathbf{x})=\frac{1-\pi}{\pi}\prod_{d=1}^{D}[\frac{\mu_{d,0}}{\mu_{d,1}}]^{\mathbf{x}_d}[\frac{1-\mu_{d,0}}{1-\mu_{d,1}}]^{1-\mathbf{x}_d}$

For the decision boundary we can equate $p(y=1|\mathbf{x})=p(y=0|\mathbf{x})$. Which gives:

$$\pi\prod_{d=1}^{D}\mu_{d,1}^{xd}(1-\mu_{d,1})^{1-xd}=(1-\pi)\prod_{d=0}^{D}\mu_{d,0}^{xd}(1-\mu_{d,0})^{1-xd} \tag{16}$$

Or equivalently we can write, $f(\mathbf{x})=1$ as the decision boundary! Therefore, here we get an exponential decision boundary!

**Introduction to ML (CS771), Autumn 2018**
**Indian Institute of Technology Kanpur**
**Homework Assignment Number 2**

*Student Name:* Shashi Kant Gupta
*Roll Number:* 160645
*Date:* September 30, 2018

# QUESTION 3

My solution to problem 3

**Introduction to ML (CS771), Autumn 2018**
**Indian Institute of Technology Kanpur**
**Homework Assignment Number 2**

*Student Name:* Shashi Kant Gupta
*Roll Number:* 160645
*Date:* September 30, 2018

# QUESTION 4

---

My solution to problem 4

*Student Name:* Shashi Kant Gupta
*Roll Number:* 160645
*Date:* September 30, 2018

# QUESTION 5

My solution to problem 5

**Introduction to ML (CS771), Autumn 2018**
**Indian Institute of Technology Kanpur**
**Homework Assignment Number 2**

*Student Name:* Shashi Kant Gupta
*Roll Number:* 160645
*Date:* September 30, 2018

# QUESTION 6

My solution to problem 6