01.112 Machine Learning
Homework 1

Joel Huang, 1002530

September 23, 2018

1 Linear Algebra and Probability Review

a. Point-Hyperplane distance

Proof of $\vec{\theta}$'s orthogonality

Consider a hyperplane P in $\mathbb{R}^d$, with equation $\vec{\theta} \cdot x + \theta_0 = 0$. We pick two points, $\vec{x_1}$ and $\vec{x_2}$, that lie on P and thus satisfy (1) and (2) respectively:

$$\vec{\theta} \cdot \vec{x_1} + \theta_0 = 0 \tag{1}$$

$$\vec{\theta} \cdot \vec{x_2} + \theta_0 = 0 \tag{2}$$

Subtracting (1) from (2), we have zero as the dot product of $\vec{\theta}$ and a vector on P.

$$\vec{\theta} \cdot \ (\vec{x_2} - \vec{x_1}) = 0 \tag{3}$$

Since $\vec{\theta}$ is orthogonal to $(\vec{x_2} - \vec{x_1})$, it is orthogonal to all vectors on P, as well as P itself.

Distance from $\vec{y}$ to the hyperplane

Consider any point $\vec{x}$ on the plane. The projection $\vec{a}$ of the point vector $\vec{x}$ on orthogonal vector $\vec{\theta}$ allows us to take its length, or ℓ^2-norm, as the distance from the hyperplane to the origin.

$$\vec{a} = \vec{x} \cdot \frac{\vec{\theta}}{|\vec{\theta}|} \tag{4}$$

Let $\vec{y}$ be the point of interest, $\vec{y} \in \mathbb{R}^n$. We want to find its distance to the hyperplane P. The projection $\vec{b}$ of point $\vec{y}$ onto $\vec{\theta}$ is:

$$\vec{b} = \vec{y} \cdot \frac{\vec{\theta}}{|\vec{\theta}|} \tag{5}$$

Hence, we may obtain the distance, d, of $\vec{y}$ from P by computing the ℓ^2-norm of the projection of $\vec{b} - \vec{a}$ on $\vec{\theta}$.

$$\vec{b} - \vec{a} = (\vec{y} - \vec{x}) \cdot \frac{\vec{\theta}}{|\theta|} \tag{6}$$

Substituting hyperplane equation $\vec{\theta} \cdot x = -\theta_0$, we can express the distance d in terms of point $\vec{y}$ and hyperplane parameters $\vec{\theta}$ and θ_0:

$$d = \|(\vec{b} - \vec{a})\|_2 = \left\| \frac{\vec{y} \cdot \vec{\theta} + \theta_0}{\theta} \right\|_2 \tag{7}$$

b. Sum of Poisson random variables

Independence

Let $Z = X + Y$ where $X = Pois(\alpha)$ and $Y = Pois(\beta)$. Since X and Y are independent, $Y = Z - X$ and:

$$P_Z(z) = \sum_x P_X(x) P_Y(z - x) \tag{8}$$

Proof that Z is Poisson

$$P_Z(z) = \sum_{x=0}^{z} \frac{\alpha^x \mathrm{e}^{-\alpha}}{x!} \frac{\beta^{(z-x)} \mathrm{e}^{-\beta}}{(z-x)!} \tag{9}$$

We can use the fractional form of the binomial coefficient $\binom{n}{k}$ to simplify (9):

$$\binom{z}{x} = \frac{z!}{(z-x)!\, x!} \tag{10}$$

$$P_Z(z) = \frac{\mathrm{e}^{-(\alpha+\beta)}}{z!} \sum_{x=0}^{z} \binom{z}{x} \alpha^x \beta^{(z-x)} \tag{11}$$

Which conveniently reduces to:

$$P_Z(z) = \frac{\mathrm{e}^{-(\alpha+\beta)}}{z!} (\alpha + \beta)^z \tag{12}$$

Therefore, Z is also Poisson with rate $\gamma = (\alpha + \beta)$.

2 Python and Theano

a. Python and Theano Versions

Python 3.6.6 (Anaconda), Theano 1.0.2