CS6501: Advanced Mahchine Learning **Spring 2017**

Problem Set 0

Handed Out: January 18^{st}, 2017 *Due: NONE*

1. [Probability] Assume that the probability of obtaining heads when tossing a coin is λ.

   a. What is the probability of obtaining the first head at the $(k+1)$-th toss?

   b. What is the expected number of tosses needed to get the first head?

2. [Probability] Assume X is a random variable.

   a. We define the variance of X as: $Var(X) = E[(X-E[X])^2]$. Prove that $Var(X) = E[X^2] - E[X]^2$.

   b. If $E[X] = 0$ and $E[X^2] = 1$, what is the variance of X? If $Y = a + bX$, what is the variance of Y?

3. [Calculus] Let $f(x,y) = 3x^2 + y^2 - xy - 11x$

   a. Find $\frac{\partial f}{\partial x}$, the partial derivative of f with respect to x. Find $\frac{\partial f}{\partial y}$.

   b. Find $(x,y) \in \mathbb{R}^2$ that minimizes f.

4. [Linear Algebra] Assume that $w \in \mathbb{R}^n$ and b is a scalar. A hyper-plane in $\mathbb{R}^n$ is the set, $\{x : x \in \mathbb{R}^n, w^T x + b = 0\}$.

   a. For $n = 2$ and 3, find two example hyper-planes (say, for $n = 2$, $w^T = \begin{bmatrix}1 & 1\end{bmatrix}$ and $b = 2$ and for $n = 3$, $w^T = \begin{bmatrix}1 & 1 & 1\end{bmatrix}$ and $b = 3$) and draw them on a paper.

   b. The distance between a point $x_0 \in \mathbb{R}^n$ and the hyperplane $w^T x + b = 0$ can be described as the solution of the following optimization problem:

   $$\begin{aligned} &\min_x \|x_0 - x\|^2 \\ &\text{s.t. } w^T x + b = 0 \end{aligned}$$

   However, it turns out that the distance between x_0 and $w^T x + b = 0$ has an analytic solution. Derive the solution. (*Hint: you may be familiar with another way of deriving this distance; try your way too*)

   c. Assume that we have two hyper-planes, $w^T x + b_1 = 0$ and $w^T x + b_2 = 0$. What is the distance between these two hyperplanes?