Home Assignment №2

Due on October 25, 2024, 23:59

Exercise 1

[3 points]. Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant rate and independently of the time since the last event. Let x has a Poisson distribution,

$$p(x = k; \lambda) = \frac{1}{k!} e^{-\lambda} \lambda^k, \tag{1}$$

where k is an occurrence number and the parameter λ is the average number of events, and the mean and variance are the same $\mathbb{E}[x] = \mathrm{Var}(x) = \lambda$.

a) [1 point]. Derive the maximum-likelihood estimate of λ, given a set of independent and identically distributed (i.i.d.) samples $\mathcal{D} = \{k^{(1)}, \ldots, k^{(M)}\}$.

b) [2 points]. The following table lists the number of intervals (maybe per minute) that are observed to have k occurrences. The total number of intervals is 230. Please calculate the maximum likelihood estimate $\lambda^\star$.

Number of occurrences (k)	0	1	2	3	4 and over
Number of intervals with k	100	81	34	9	6

Exercise 2

[3 points]. Consider the nonlinear error surface $\ell(u, v) = (ue^v - 2ve^{-u})^2$. We start at the point $(u, v) = (1, 1)$ and minimize this error using gradient descent in the u, v space. Use $\alpha = 0.1$ (*i.e.*, learning rate).

a) [1 point]. What is the partial derivative of $\ell(u, v)$ with respect to u?

b) [1 point]. How many iterations does it take for the error $\ell(u, v)$ to fall below 10^{-14} for the first time? In your programs, make sure to use double precision to get the needed accuracy.

c) [1 point]. After running enough iterations such that the error has just dropped below 10^{-14}, what is the final (u, v) you get in problem b)? Round your answer to the thousandths place.

Exercise 3

[3 points]. Suppose that $x \in \mathbb{R}^2$ $\left(x = [x_1, x_2]^T\right)$. Consider the following optimization problem:

$$\begin{aligned} \min_{x} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \\ & (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1. \end{aligned}$$

a) [1 point]. Sketch the feasible set and level sets of the objective. Find the optimal point $x^\star$ and optimal value $p^\star$.

b) [1 point]. Give the KKT conditions. Do there exist Lagrange multipliers $\lambda_1^\star$ and $\lambda_2^\star$ that prove that $x^\star$ is optimal?

c) [1 point]. Derive and solve the Lagrange dual problem. Does strong duality hold?

Exercise 4

[3 points]. In the formulation of SVM, we need to compute the margin (i.e., the distance) between an arbitrary point $x^{(i)}$ in the N-dimensional space and a hyperplane $w^T x + b = 0$, which can be formulated as the following optimization problem:

$$\begin{aligned} \min_{x} \quad & \left\| x^{(i)} - x \right\|_2 \\ \text{s.t.} \quad & w^T x + b = 0. \end{aligned}$$

a) [1 point]. Is this problem convex and why?

b) [2 points]. Using the Lagrange duality to solve for the optimal x and the distance. (Remember to form the Lagrangian and derive the Lagrange dual function).

Exercise 5

[3 points]. In the lecture notes, we have given a detailed derivation of the dual form of SVM with soft margin. With simpler arguments, derive the dual form of SVM with hard margin

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} w^T w \\ \text{s.t.} \quad & y^{(i)}(w^T x^{(i)} + b) \geq 1, \quad i = 1, \ldots, M. \end{aligned}$$

Compare the two dual forms.