Assignment 6

Problem 1: Using Lagrange multipliers for MAP estimation for a Categorical distribution

The Categorical distribution is defined as

$$\text{Cat}(y|\vec{\theta}) = \prod_{c=1}^{C} \theta_c^{\mathbb{1}(y=c)},$$

for $y \in \{1, 2, \ldots, C\}$ where C is the number of labels and $C > 1$. We have observed N observations of Y, with N_k observation of each label. Suppose we decide to place a (unusual) zero-mean, identity-covariance Gaussian prior on $\vec{\theta}$, $p(\vec{\theta}) \propto \exp(-\theta^\top \theta)$.

You would like to find the MAP of $\vec{\theta}$.

1. Form the Lagrangian expression $\mathcal{L}(\vec{\theta}, \lambda)$.
2. Using your answer in Part 1., calculate the partial derivative with respect to each θ_k and λ.
3. Briefly describe how to use your answer to Part 2 to find the MAP for each θ_k. (You do not need to find an explicit solution.)

Problem 2

Consider a quadratic objective function on $\mathbf{R}^2$

$$f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2),$$

where $\gamma > 0$. We would like to apply the gradient descent method with exact line search, starting at the point $\vec{x}^{(0)} = (\gamma, 1)$.

1. Derive the exact line search update.
2. Suppose $\gamma = 10$, what is the value of $\vec{x}^{(3)}$?

Problem 3

Please write one thing from this course you found confusing, a topic you would like to hear more about, or something you found particularly interesting.