

Problem 1

Let θ be a univariate, continuous parameter. Consider the following optimization, which adds a regularization term to the log likelihood to identify a regularized MLE.

$$\hat{\theta}_{reg} = \arg \min_{\theta} (-\log p(D|\theta) + \lambda C(\theta)) \quad \lambda > 0$$

1. As an alternative, we can define the MAP estimate $\hat{\theta}_{MAP}$ for a particular choice of prior $p(\theta)$. Define a probability distribution $p(\theta)$ as a function of $C(\theta)$ such that $\hat{\theta}_{MAP} = \hat{\theta}_{reg}$. You may assume that the integral of C (and reasonable functions of C) are finite.
2. According to $p(\theta)$, what is the relative probability of two parameters $p(\theta_1)/p(\theta_2)$?
3. Qualitatively, how does $p(\theta)$ change as we decrease λ ? Increase?

Problem 2: Using Lagrange multipliers for MAP estimation for a Categorical distribution

The Categorical distribution is defined as

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

for $y \in \{1, 2, \dots, C\}$ 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 $\vec{\theta}$, $p(\vec{\theta}) \propto \exp(-\vec{\theta}^\top \vec{\theta})$.

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 3

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

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

Assignment 5

CMPT 727
Spring 2024

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 $\vec{x}^{(2)}$?

Problem 4

On average, how many hours have you spent on each assignment so far in this course?

Problem 5

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.