CMP784 Deep Learning, Fall 2024

MATH PREREQUISITES QUIZ

Due Date: 5pm, Saturday, October 12, 2024 (No late submissions!)

Each student enrolled to CMP784 must complete and pass this quiz on prerequisite math knowledge. The purpose is to check whether you have the right background for the course. The topics covered in this problem set are very crucial so if you are having trouble with solving a problem, this indicates that you should spend a considerable amount of time to study that topic in its entirety.

Points and Vectors

1. Given two vectors $x = [a_1, a_2, a_3]$ and $y = [a_1, -a_2, a_3]$. Write down the equation for calculating the angle between x and y. When is x orthogonal to y?

Planes

2. Consider a hyperplane described by the d-dimensional normal vector $[\theta_1, \dots, \theta_d]$ and offset θ_0. Derive the equation for the *signed distance* of a point x from the hyperplane, which is defined as the perpendicular distance between x and the hyperplane, multiplied by +1 if x lies on the same side of the plane as the vector θ points and by -1 if x lies on the opposite side x from the hyperplane.

Matrices

3. Suppose that $A^T(AB - C) = 0$, where 0 is an $m \times 1$ vector of zeros, derive an expression for B. Assume that all relevant matrices needed for this calculation are invertible.

4. Find the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 13 & 5 \\ 2 & 4 \end{bmatrix}$.

Probability

5. Let

$$p(X_1 = x_1) = \alpha_1 e^{-\frac{(x_1 - \mu_1)^2}{2\sigma_1^2}}$$
$$p(X_2 = x_2 \mid X_1 = x_1) = \alpha e^{-\frac{(x_2 - x_1)^2}{2\sigma^2}}$$

where X_1 and X_2 are continuous random variables. Show that

$$p(X_2 = x_2) = \alpha_2 e^{-\frac{(x_2 - \mu_2)^2}{2\sigma_2^2}}$$

by explicitly calculating the values of α_2, μ_2 and σ_2.

MLE and MAP

6. Let p be the probability of landing head of a coin. You flip the coin 3 times and note that it landed 2 times on tails and 1 time on heads. Suppose p can only take two values: 0.3 or 0.6. Find the Maximum Likelihood Estimate of p over the set of possible values {0.3,0.6}

7. Suppose that you have the following prior on the parameter p: P(p = 0.3) = 0.3 and P(p = 0.6) = 0.7. Given that you flipped the coin 3 times with the observations described above, find the MAP estimate of p over the set {0.3, 0.6}, using the prior.

Optimization

Gradient ascent/descent methods are typical tools for maximizing/minimizing functions. Let $L(x, \theta)$ be a function of two vector arguments, $x = [x_1, x_2]^T$ and $\theta = [\theta_1, \theta_2]^T$. We would like to find the optimum value of vector θ which maximizes/minimizes $L(x, \theta)$ where x is assumed to be given.

The gradient $\nabla_\theta L(x, \theta)$ is a vector with two components corresponding to partial derivatives

$$\frac{\partial}{\partial \theta_j} L(x, \theta), \quad j = 1,2$$

8. Evaluate the gradient when $L(x, \theta) = \log\left(1 + exp(-\theta^T x)\right)$.

9. Into which direction does the gradient (viewed as a vector) point? Is the value of $L(x, \theta)$ larger or smaller if we evaluate it at $\theta' = \theta + \epsilon \nabla_\theta L(x, \theta)$ where ϵ is a small real number?