UNIVERSITY OF OXFORD

Machine Learning
Michaelmas Term 2021
Week 1

Problem Sheet 0

1 Linear Algebra Background

Consider the following matrix $\mathbf{X}$ and vectors $\mathbf{y}$ and $\mathbf{z}$

$$\mathbf{X} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Answer the following questions:

1. The dot (or inner) product of $\mathbf{y}$ and $\mathbf{z}$ is denoted by $\mathbf{y} \cdot \mathbf{z}$ (or sometimes $\langle \mathbf{y}, \mathbf{z} \rangle$, $\mathbf{y}^\mathsf{T}\mathbf{z}$ or $\mathbf{z}^\mathsf{T}\mathbf{y}$). What is $\mathbf{y} \cdot \mathbf{z}$?
2. What are the products $\mathbf{X}\mathbf{y}$ and $\mathbf{z}^\mathsf{T}\mathbf{X}$?
3. Does the inverse of matrix $\mathbf{X}$ exist? If so what is it?
4. What is the determinant of matrix $\mathbf{X}$?
5. What is the rank of matrix $\mathbf{X}$?

2 Calculus Background

Consider the function $f(x) = x^3 - 3x + 7$ and answer the following questions.

1. What is the derivative $\frac{\mathrm{d}f}{\mathrm{d}x}$?
2. What is the maximum value of f on the interval $[0, 2]$? How about the minimum value?
3. What are the minimum and maximum values of f on the interaval $[-2, 0]$? At what points are they attained?

Now consider the function $f(x, y) = x^3 + y^2 + xy$ and answer the following questions.

1. What is the gradient of f at the point $(2, 3)$?
2. What is the Hessian of f at the point $(0, 0)$?
3. List all the critical points of f. What can you say about each of them?

Let

$$f(x, y, z) = (x + 1)^2 + (y + 2)^2 + (z - 2)^2$$
$$g = x^2 + y^2 + z^2 - 36.$$

Using the method of Lagrange multipliers, find the critical points of $f(x, y, z)$ subject to $g(x, y, z) = 0$.

3 Probability Background

Probability Distributions

Write down the probability density/mass functions for the following distributions: Uniform over $[-1, 1]$, Univariate Gaussian, Laplace, Bernoulli, Binomial, Multivariate Gaussian. If necessary, please consult Murphy (2012, Chap 2) (or Wikipedia) and familiarize yourself with these distributions. If you are already familiar with these, go ahead and look for new interesting ones in Murphy (2012)!

Bayes' Rule

Write down Bayes' rule (look it up if necessary). Then answer the following questions.

1. Suppose an unbiased coin is tossed 5 times. Let A denote the event that the first toss resulted in heads and B the event that overall there were 3 heads and 2 tails. What is $\mathbb{P}\left[A|B\right]$?

2. Suppose your friend throws a fair (unbiased) coin and a coin with both sides having heads into a hat. She then pulls one out (assume that both have an equal chance of being chosen) tosses it and reports the outcome as heads. Conditioned on this event what is the probability that she pulled out the biased coin? What if the outcome had been tails?

4 Statistics Background

The average scores on the questions on this sheet of a group of 5 students (obviously not from this class!) were $\{6, 7, 5, 3, 8\}$ on a scale of 0 to 10. Answer the questions below. It would be a good idea to try to do the calculations in python to get used to the language if you aren't familiar with it.

1. What is the mean and median score? What is the variance and standard deviation?

2. Suppose you were asked to standardize these so that the mean is 0 and variance is 1, what transformation would you apply?

3. Is the estimator for the mean you used unbiased? How about the one for the variance?

References

Kevin P. Murphy. *Machine Learning : A Probabilistic Perspective*. MIT Press, 2012.