SIT787 -Mathematics for AI

Trimester 2, 2021

Due: no later than the end of Week 11, Sunday 3 October 2021, 8:00pm AEST

Note:

- A proper way of presenting your solutions is part of the assessment. Please follow the order of questions in your submission.

- Your submission can be handwritten but it must be legible. Please write neatly. If I cannot read your solution, I cannot mark it.

- Provide the way you solve the questions. All steps (workings) to arrive at the answer must be clearly shown. I need to see your thoughts.

- For a final answer without a proper justification no score will be given.

- Only (scanned) electronic submission would be accepted via the unit site (Deakin Sync).

- Your submission must be in ONE pdf file. Multiple files and/or in different file format, e.g. .jpg, will NOT be accepted. If you need to change the format of your submission, it will be subject to the late submission penalty.

Question 1) Probability, Distributions: Let X be a discrete random variable that takes values in $\{-2, -1, 0, 1, 2\}$ each with the probability of $\frac{1}{5}$. Also, Y is another discrete random variable defined as $Y = X^2$.

    (i) Construct the joint probability distribution table.

    (ii) Are X and Y independent? Justify.

    (iii) Find $\text{Corr}(X, Y)$.

    (iv) Based on your answer to part (b), can you explain the result in part (c)?

[3+3+3+3= 12 marks]

Question 2) Probability, Bayes' Theorem: There is a corona virus test with 90% accuracy. Given that only 2% of the population is infected, answer the following questions.

    (i) What does it mean that the accuracy of the test is 90%?

    (ii) If a person tests positive, what is the probability that the person actually has the disease.

    (iii) What is the probability that a patient is missclassified?

    (iv) If the person decides to repeat the test again (independent of the first one) what is the probability of being infected given that both tests are positive?

    (v) If the person decides instead of repeating the same test, he takes independently a cheaper test with accuracy of 80%. What is the probability of being infected given that both tests are positive?

[3+3+3+3+4=16 marks]

Question 3) Probability, Distributions: Let X be a continuous random variable with a probability density function given by

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

    (i) Find the value of k.

    (ii) Calculate $P(X \leq \frac{1}{2})$.

    (iii) Find the cumulative distribution of X.

    (iv) Compute $P(X \leq \frac{1}{2})$ using the cumulative distribution function you found in part (c). If there is any difference, can you explain the reason?

[2+2+2+2=8 marks]

Question 4) **Multivariate functions and Optimisation**. The idea of local and global minimum and maximum exists in studying functions with several variables. For a function of two variables $z = f(x, y)$, if f has a local maximum or minimum at a point (a, b), and the first order partial derivatives of f exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$. You can imagine that if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, then there will be a flat tangent plane, similar to a horizontal tangent line in single variable functions.

Formally, a point (a, b) is a critical point (or stationary point) of $z = f(x, y)$ either if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, or if one of these partial derivatives does not exist. Also, we need to be careful that at a critical point, a multivariate function could have a local maximum, a local minimum or neither.

Second derivative test for a local max or min for functions with 2 variables: Consider $z = f(x, y)$. We need to be able to determine whether a function has an extreme value at a critical point. For a critical point (a, b), let

$$D = D(a, b) = f_{xx}(a, b) f_{yy}(a, b) - [f_{xy}(a, b)]^2$$

be the discriminant.

- If $D > 0$ and $f_{xx}(a, b) > 0$, then f has a local minimum at (a, b).

- If $D > 0$ and $f_{xx}(a, b) < 0$, then f has a local maximum at (a, b).

- If $D < 0$, then f does not have a local max or a local min at (a, b). It has a saddle point there.

- If $D = 0$, the test gives no information. We need to use other mathematical techniques to check the situation.

To find a local minimum, a local maximum, or a saddle point of multivariate functions with more than 2 variables, we need to compute its Hessian matrix at all the critical points. Consider the multivariate function $w = f(x, y, z)$. You see that the number of independent variables is more than 2. The Hessian of this matrix is defined as

$$H(x, y, z) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\ \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

Generally, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, $\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x}$, and $\frac{\partial^2 f}{\partial z \partial y} = \frac{\partial^2 f}{\partial y \partial z}$, hence, the Hessian matrix is always symmetric. Also, note that the Hessian is a function of x, y, and z. The critical points of this function is either $\nabla f = \vec{0}$, or the partial derivative does not exist. To find a local minimum, a local maximum, or a saddle point, we plug the critical points in the Hessian matrix, compute its eigenvalues, and we make a decision based on signs of the eigenvalues, as described in the second derivative test.

Second derivative test: Find all critical points, plug them in the Hessian matrix, and compute its eigenvalues.

- If **all** eigenvalues are strictly positive, then the critical point is a local minimum.

- If **all** the eigenvalues are strictly negative, the critical point is a local maximum point.

- If eigenvalues do not have the same sign and **all** are non-zero, then the critical point is a saddle point.

- If any of the eigenvalues is zero, then no conclusion can be drawn without further information.

It is important to note that by this extension, you can find local maximum, local minimum, or saddle points of a function with any number of variables. In its most general form, to find critical points of a multivariate function with n variables $y = f(x_1, x_2, \ldots, x_n)$, you need to solve the system $\nabla f = \overrightarrow{\mathbf{0}}$, or find the points that any of the partial derivatives does not exist. Then, you need to from the Hessian matrix at the critical points, find eigenvalues, and make a decision based on the signs of the eigenvalues.

So far what we had was unconstrained optimisation. In other words, we wanted to find a local maximum or minimum, or a saddle point of a multivariate function without any restriction or constraints. However, sometimes we want to find a maximum or a minimum of $f(x, y)$ subject to a constraint $g(x, y) = k$, where k is a constant. To solve this problem, we use the Lagrange Multiplier technique. To find maximum and minimum values of $f(x, y)$ subject to the constraint $g(x, y) = k$ (assuming that these max and min values exist and $\nabla g \neq 0$ on $g(x, y) = k$):

(a) Find all values of x, y, and λ such that

$$\nabla f(x, y) = \lambda \nabla g(x, y) \text{ and}$$

$$g(x, y) = k$$

(b) Evaluate f at all the points (x, y) that result from step (a). The largest of these values is the global maximum value and the smallest is the global minimum value of f.

[**Note:** Please pay attention to the way the λ notation is used here. It is not an eigenvalue. Here, it is a nonzero multiplier – we are not interested in its value, but it will facilitate the finding of the critical points.]

This technique could be applied for functions having more than two variables. Imagine, you want to find the maximum and minimum of $w = f(x, y, z)$ subject to constraint $g(x, y, z) = 0$. Then you need to solve the following system:

$$\begin{cases} \nabla f(x, y, z) = \lambda \nabla g(x, y, z) \\ g(x, y, z) = 0 \end{cases}$$

This system has four equations and 4 unknowns. However, generally it is not linear. Sometimes, you may use some mathematical tricks to find the solution of the system.

A further extension is when there are two constraints (you can see that we can easily generalise the idea to many constraints). To find the maximum and minimum of a function $w = f(x, y, z)$

subject to two constraints $g(x, y, z) = 0$ and $h(x, y, z) = 0$, we need to introduce two Lagrange multipliers λ and μ to obtain the following system of equations:

$$\begin{cases} \nabla f(x, y, z) = \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z) \\ g(x, y, z) = 0 \\ h(x, y, z) = 0 \end{cases}$$

To summarise, you need to be familiar with the following optimisation problems, and how we can solve them. Also, you need to be extremely careful about using local or global adjectives in your communications.

- Optimisation problems
    - Unconstrained Optimisation
        * Single-variable function: (covered in assignment 1)
        * Multivariate functions
            · Multivariate with two variables (use the discriminant)
            · Multivariate with more than two variables (use eigenvalues)
    - Constraint Optimisation
        * Global max and min of a single variable function in an interval (covered in assignment 1)
        * Global max and min of a multivariate functions with some constraints (Lagrange multipliers)

Now, let's solve some optimisation problems.

(i) Find and classify all the critical points of the function

$$f(x, y) = x^3 + y^3 + 3x^2 - 3y^2 - 8.$$

(ii) Find all values for k so that $f(x, y) = x^2 + kxy + y^2$ has a local minimum at $(0, 0)$. Give your answer in the form of an interval.

Now solve the constrained optimisation problems below using Lagrange multipliers.

(iii) Find the minimum, maximum, both, or neither, value of $f(x, y) = 4x + 4y - x^2 - y^2$ subject to the constraint $x^2 + y^2 = 2$.

(iv) Find the extreme value of $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraints $2x + 2y - 4z = 0$ and $4x - 2y + 2z = 0$, and determine whether the extreme point is a maximum or minimum point.

[7+10+7+10=34 marks]

Question 5) Graph Theory. Consider the following directed Graph $G(V, E)$.
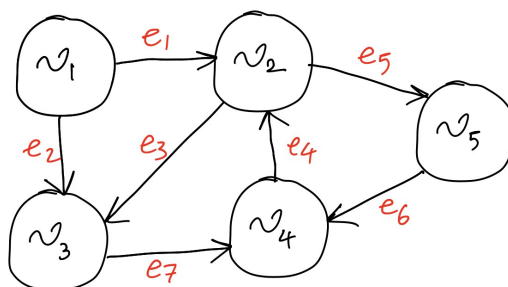
(i) Find the adjacency matrix (call it A) and the incidence matrix (call it B) of this graph.

(ii) Remove the direction of the arcs in the graph, and find the adjacency matrix of the undirected version of this graph. Call the matrix C.

(iii) Find $\text{rank}(B)$.

(iv) Show that $B^T \vec{1} = \vec{0}$, where $\vec{1}$ is the (column) vector whose components are all equal to 1, and $\vec{0}$ is the (column) vector whose components are all equal to 0. Based on $B^T \vec{1} = \vec{0}$, what can you conclude about eigenvalues and eigenvectors of BB^T?

(v) Set $L = BB^T$. Show that $L = D - C$, where $D = \text{diag}(\deg(v_1), \deg(v_2), \dots, \deg(v_5))$, namely, the degrees of vertices are on the main diagonal of D.

(vi) Find $\det(L)$ without direct calculation.

(vii) Without direct calculation, show that all the eigenvalues of L are nonnegative.

(viii) Without direct calculation, show that one of the eigenvalues of L is zero and the corresponding eigenvector is $\vec{1}$.

[4+3+3+4+4+4+4+4=30 marks]