Week 8 PYQs
Machine Learning Foundations
May 22 - Sep 25

Note

1. If no questions are listed in that term Quiz, it indicates that the corresponding quiz did not include questions from that week.

2. The answer key may contain minor inaccuracies, as answers were taken directly from the original quiz/exam papers. Some of these might have been corrected later.

3. For any issues or suggestions regarding this document, feel free to contact:
   `23f3004032@ds.study.iitm.ac.in`
   or reach out on Discourse: `@23f3004032`

Quiz 1 May 22

Quiz 2 May 22

1. We need to find optimal value of

$$f(x, y, z) = \log(xyz) + xyz$$

Starting with

$$X_0 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad X_1 = \begin{bmatrix} i \\ j \\ k \end{bmatrix}$$

using gradient descent with $\eta = 1$, find $|i + j + k|$.

[Ans: 7.5 to 8.5]

2. It is known that for some function, $f(0) = 1$, $f'(x) = -\cos(x)e^{\sin(x)}$, and $f''(x) = \sin(x)e^{\sin(x)} + (\cos(x))^2 e^{\sin(x)}$. What is the value of $f(1)$ using Taylor series expansion. Use second order approximation (up to $f''(x)$) starting with $x = 0$ and $h = 1$.

[Ans:0.5]

3. We need to find cheapest cylindrical container to hold $1000m^3$ water. Side cost = 1000 per m^2, top and bottom = 2000 each.

Find optimal radius.

[Ans: 3.1 to 3.8]

ET May 22

1. Equation of plane $r \cdot n = d$, where $n = \frac{6i+4j-2k}{\sqrt{56}}$ and $d = \frac{40}{\sqrt{56}}$. Then the Minimum distance of the plane from point (1,2, 1) is: *NOTE: Enter your answer in one decimal place.* [Ans: 3.5 to 4.0]

2. Using gradient descent for linear regression with $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $X = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, and $w^t = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. If $w^{t+1} = \begin{bmatrix} i \\ j \end{bmatrix}$ in the next iteration with $\eta = 1$, what is the absolute value $|i+j|$? *NOTE: Enter your answer to the nearest integer.* [Ans: 5 to 7]

Quiz 1 Sep 22

Quiz 2 Sep 22

1. Find the height of a cylinder of maximum volume that can be inscribed in a sphere of radius $a = \sqrt{3}$. [Ans:2]

2. You are told to make an open box with a square base from a wooden piece of $27m^2$ area such that the volume is maximum. If dimensions are l, b, h, find l.

   [Ans: 3]

3. Given $f(x, y) = 10x^2 + 10xy + 5y^2$, then point $(0, 0)$ is a _____.

   (a) maxima
   (b) minima
   (c) saddle point
   (d) None of these

   [Ans:b]

4. A furniture dealer deals in tables and chairs. Max investment = 10000, storage = 60 pieces. Table cost = 500, chair cost = 200. Profit = 50 and 15.

   (a) Objective function:
      i. min $50x + 15y$
      ii. max $50x + 15y$
      iii. min $50x - 15y$
      iv. max $50x - 15y$

   [Ans: b]

   (b) The constraint function for the problem is
      i. $x \geq 0,\ y \geq 0$
      ii. $5x + 2y \leq 100$
      iii. $5x + 2y \geq 100$
      iv. $x + y \leq 60$

   [Ans:a,b,d]

ET Sep 22

1. What is value of the function $f(x) = x_1^3 + x_2^3 + x_3^3 - x_1x_2 - x_2x_3 - x_1x_3$ with an initial guess of (1, 1, 1) and a step size of 0.5 after one iteration using gradient descent? Enter the value up to two decimal points accuracy. [Ans: -0.39 to -0.36]

2. While solving for the optimal weight vector (w) for a linear regression problem using gradient descent, we have $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $X = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$. $w^t = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ after t iterations. If the value of $w^{t+1} = \begin{bmatrix} i \\ j \end{bmatrix}$ in next iteration using gradient descent method, then what is the absolute value of $|i + j|$? Assume $\eta = 1$. [Ans: 5 to 7]

Quiz 1 Jan 23

Quiz 2 Jan 23

1. Let $f(x, y) = x^2y^2 - 2x - 2y$. Which statements are correct?

   (a) (0,0) is stationary

   (b) (1,1) is stationary

   (c) minimum at (0,0)

   (d) minimum at (1,1)

   [Ans: B]

2. If $f(20) = 1$, $f'(20) = 10$, and $f''(20) = 5$, find the second order approximation of $f(10)$.

   [Ans: 151]

3. What is the maximum area of a rectangle that can be inscribed in the ellipse

$$\frac{x^2}{2} + y^2 = 1$$

   (correct to 2 decimals) [Ans: 2.83]

4. $$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_2x_3 - 2x_3x_1$$

   After one step of gradient descent from $(1, 1, 1)$ with $\eta = 1$, find the value of f.

   [Ans: -3]

ET Jan 23

1. Find the maximum value of the function $f(x, y) = 4x^2 - 4xy + y^2$ on the circle $x^2 + y^2 = 25$. Enter the answer as integer. [Ans: 125]

2. By using first order taylor series approximation, what would be the value of $f([3, 2, 1]^T)$, if $f([1, 1, 1]^T) = 10$ and $\nabla f([1, 1, 1]^T) = [2, 1, 3]^T$? Enter the answer as integer. [Ans: 15]

Quiz 1 May 23

Quiz 2 May 23

1. If $f([1,2,3]^T) = 10$ and $\nabla f([1,2,3]^T) = [1,5,7]^T$, then find the value of $f([2,2,2]^T)$ using first order Taylor series expansion.

   [Ans: 4]

2. What is the maximum area of a circle that can be inscribed in the region formed by $y = 2 - x^2$ and $y = x^2 - 2$? (center at origin)

   [Ans: 5.5]

3. Consider the dataset

   $$D = \{(-1,1),(0,1),(1,1)\}$$

   What is the first principal component?

   (a) $\begin{bmatrix}1\\0\end{bmatrix}$

   (b) $\begin{bmatrix}0\\1\end{bmatrix}$

   (c) $\begin{bmatrix}1\\1\end{bmatrix}$

   (d) $\begin{bmatrix}-1\\0\end{bmatrix}$

   [Ans: 1]

4. If the dataset is changed to

   $$D' = \{(-1,1),(0,0),(1,1)\}$$

   what is the first principal component?

   (a) $\begin{bmatrix}1\\0\end{bmatrix}$

   (b) $\begin{bmatrix}0\\1\end{bmatrix}$

   (c) $\begin{bmatrix}1\\1\end{bmatrix}$

   (d) $\begin{bmatrix}-1\\0\end{bmatrix}$

   [Ans: 1]

5. For the dataset D', let $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ be projections onto the first principal component. Which is correct?

(a)
$$\tilde{x}_1 = \begin{bmatrix} -1 \\ 2/3 \end{bmatrix},\ \tilde{x}_2 = \begin{bmatrix} 1 \\ 2/3 \end{bmatrix},\ \tilde{x}_3 = \begin{bmatrix} 1/2 \\ 2/3 \end{bmatrix}$$

(b)
$$\tilde{x}_1 = \begin{bmatrix} -1 \\ 2/3 \end{bmatrix},\ \tilde{x}_2 = \begin{bmatrix} 0 \\ 2/3 \end{bmatrix},\ \tilde{x}_3 = \begin{bmatrix} 1 \\ 2/3 \end{bmatrix}$$

(c)
$$\tilde{x}_1 = \begin{bmatrix} 1 \\ 2/3 \end{bmatrix},\ \tilde{x}_2 = \begin{bmatrix} 0 \\ 2/3 \end{bmatrix},\ \tilde{x}_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

(d)
$$\tilde{x}_1 = \begin{bmatrix} 1 \\ 2/3 \end{bmatrix},\ \tilde{x}_2 = \begin{bmatrix} 1 \\ 2/3 \end{bmatrix},\ \tilde{x}_3 = \begin{bmatrix} -1 \\ 2/3 \end{bmatrix}$$

[Ans: 2]

6. What is the reconstruction error after projecting D' onto the first principal component?

[Ans: 0.22]

ET May 23

1. What will be the value of $f(x) = 4x^2 + 2$ with an initial guess of 2, after two iterations of gradient descent algorithm? Take step size as 0.1. Enter the answer correct to three decimal places. [Ans: 2.022 to 2.028]

2. Which of the following statements regarding gradient descent is/are correct?

   (a) If the step size is chosen as 1, we may "not always" arrive as close to the optimal solution even after many number of iterations.

   (b) Solution obtained for optimization problem is the local minimum of its objective function.

   (c) Solution obtained for optimization problem is the global minimum of its objective function.

   (d) Gradient descent converges to the global optimum in the case of convex functions.

   (e) Gradient descent does not converge to the global optimum in the case of convex functions.

   [Ans: 1, 2, 4]

Quiz 1 Sep 23

Quiz 2 Sep 23

1. Consider a function f(x, y) = $x^2 + y^2$ 4x 2y + 5. Using gradient descent algorithm, with an initial guess of (3, 2) and learning rate of 0.1, what will be the value of (x, y) after one iteration?

   (a) x = 3.2, y = 2.2

   (b) x = 3, y = 2

   (c) x = 1.8,y = 0.8

   (d) x = 2.8, y = 1.8

   [Ans:d]

2. Find the area of the largest rectangle that can be inscribed in a semicircle $y = \sqrt{2 - x^2}$.

   [Ans: 2]

3. A firm produces two goods A and B. Let x and y denote their prices. Given $x + y = 60$ and profit $= 2x^2y$, find the optimal prices.

   (a) $x = 40,\ y = 20$

   (b) $x = 20,\ y = 40$

   (c) $x = 30,\ y = 30$

   (d) $x = 0,\ y = 60$

   (e) None of these

   [Ans: 1]

ET Sep 23

1. There once lived a queen named Queen Dido who was exiled from her kingdom, and so she fled to another nearby kingdom. She begged the king of the nearby kingdom to give her some land. The king agreed to give her a rectangular piece of land, which she could enclose with a rope of length 96 units. How should Queen Dido choose the sides of the rectangle so that she gets as much land as possible? Hint: Perimeter of the rectangle with sides x and y is $2(x + y)$.

   (a) $x = 24, y = 24$

   (b) $x = 30, y = 18$

   (c) $x = 48, y = 0$

   (d) $x = 48, y = 48$

[Ans: 1]

Quiz 1 Jan 24

Quiz 2 Jan 24

1. The value of a function at point 5 is 10. The values of the function's first and second order derivatives at this point are 5 and 2 respectively. What will be the function's approximate value at the point 5.1? (Enter the answer correct up to two decimal places).

   [Ans: 10.2 to 10.8]

2. Consider

   $$f(x, y) = x^2 - 4x + 2y + 1$$

   Using gradient descent with initial point $(3, 2)$ and learning rate $\frac{1}{t+1}$, find (x, y) after two iterations.

   (a) $x = 0.5,\ y = 0$

   (b) $x = 2.512,\ y = 1.612$

   (c) $x = 0,\ y = 0$

   (d) $x = 2,\ y = 0.33$

   [Ans: 4]

3. Find the point(s) on the circle $y^2 = 4 + x^2$ closest to $(3, 0)$.

   (a) $(-1.5, -2.5)$

   (b) $(1.5, 2.5)$

   (c) $(1.5, -2.5)$

   (d) $(-1.5, 2.5)$

   [Ans: 2]

ET Jan 24

Quiz 1 May 24

Quiz 2 May 24

1. Which of the following statements about choosing the learning rate in gradient descent is incorrect?

   (a) A small learning rate can lead to slow convergence

   (b) A large learning rate can cause oscillations around minimum

   (c) A large learning rate can cause divergence

   (d) A small learning rate always causes very fast convergence

   [Ans: 4]

2. Consider

   $$f(x, y) = (x - 1)^2 + (y - 2)^2$$

   Using gradient descent with initial point $(0, 0)$ and learning rate 0.1, find the absolute difference between function values after first and second iteration.

   [Ans: 1.152]

3. The value of a function at $x = 25$ is 5. The first and second derivatives are 0.1 and -0.002. Find approximate value at $x = 25.1$.

   [Ans: 5.009]

ET May 24

1. A farmer has 120 meters of fencing and wants to fence off a rectangular field that borders a straight river. The farmer does not need to fence along the river. Find the maximum possible fenced area of the field. [Ans: 1800]

2. Suppose we have 10 data points randomly distributed in space, $\mathbb{R}^3$ given by $D = \{x_1, x_2, ...x_{10}\}$ Let $g(p) = \sum_{i=1}^{10} ||p - x_i||^2$ be a function defined to calculate the sum of the square of distances of data points from a fixed point, say $p \in \mathbb{R}^3$. If $g(p)$ attains the minimum at $q = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}$, then which of the following options is true?

   (a) $x_1 + x_2 + ... + x_{10} = \begin{bmatrix} 30 \\ 0 \\ 40 \end{bmatrix}$

   (b) $x_1 + x_2 + ... + x_{10} = \begin{bmatrix} 40 \\ 0 \\ 30 \end{bmatrix}$

   (c) The distance of point $(x_1 + x_2 + ... + x_{10})$ from the origin is 50.

   (d) The distance of point $(x_1 + x_2 + ... + x_{10})$ from the origin is 40.

   [Ans: 2, 3]

Quiz 1 Sep 24

Quiz 2 Sep 24

1. Consider
$$f(x, y) = 3x^2 - 6xy + y^2 + 4$$
Using gradient descent with initial point $(1, 2)$ and learning rate $\frac{1}{t^2}$, find (x, y) after two iterations.

   (a) $(7, 4)$
   (b) $(2.5, 12.5)$
   (c) $(4, 6)$
   (d) $(3.5, 10)$

   [Ans: 2]

2. Find the point on line $y = 2x + 1$ closest to $(4, 0)$.

   (a) $(3, 7)$
   (b) $(0.4, 1.8)$
   (c) $(0, 0)$
   (d) $(-1, 0)$

   [Ans: 2]

ET Sep 24

1. A company is optimizing a model that is given by its production cost function $f(x, y) = x^2 + y^2 - 6x - 8y$, where x is the number of units produced and y is the number of hours spent on labor. The company has decided to use the gradient descent algorithm to minimize the production cost, starting from the initial point (1, 2) with a learning rate of $\eta = 0.2$. Find the value of $20x + 20y$ after one iteration. [Ans: 92]

Quiz 1 Jan 25

Quiz 2 Jan 25

1. Which of the following represents the Taylor's series expansions up to the fourth degree term of $f(x) = \frac{1}{x}$ around $x = 1$?

   (a) $1 + (x-1) - (x-1)^2 + (x-1)^3 - (x-1)^4$

   (b) $1 - (x+1) + (x+1)^2 - (x+1)^3 + (x+1)^4$

   (c) $1 - (x-1) + (x-1)^2 - (x-1)^3 + (x-1)^4$

   (d) $1 + (x-1) + (x-1)^2 + (x-1)^3 + (x-1)^4$

   [Ans: 3]

2. For the function: $f(x, y) = x^2 + 3y^2 - 4xy + 5$, using the initial values $x_0 = 2$ and $y_0 = 3$, perform one iteration of gradient descent with a learning rate of 0.1. What is the value of $f(x, y)$ after this iteration? Enter the answer correct to two decimal places. [Ans: 2.41 to 2.47]

3. If the minimum value of the function $f(x) = e^{x^2-8x+15}$ is e^{α}, then find the value of α.

   [Ans: -1]

ET Jan 25

1. Consider the function $f(x, y) = x^2 + 2y^2 - 6x + 8y$. We apply gradient descent with an initial point $(x_0, y_0) = (1, -2)$ and a learning rate of $\eta = 0.2$. Compute the updated values of x_1 and y_1 after one iteration of gradient descent, and enter the value of $x_1 + y_1$. [Ans: -0.2]

Quiz 1 May 25

Quiz 2 May 25

1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be an times continuously differentiable function. Suppose that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) > 0$. *Note:* $\nabla^2 f(x^*) > 0$ *means hessian matrix is positive definite.* What can we conclude about x^*?

   (a) x^* is a local maximum

   (b) x^* is a saddle point

   (c) x^* is a local minimum

   (d) Nothing can be concluded

   [Ans: 3]

2. Consider the function $f(x, y) = (x-1)^2+(y+2)^2$. Using the gradient descent algorithm with an initial guess $(x_0, y_0) = (3, -1)$ and learning rate $\eta_t = \frac{1}{t+1}$ where $t = 0, 1, 2, ...$, what will be the value of (x, y) after two iterations?

   (a) $x = 1$, $y = -2$

   (b) $x = 2$, $y = -2$

   (c) $x = 1.5$, $y = -2.5$

   (d) $x = 1.625$, $y = -2.375$

   [Ans: 4]

3. Find the point(s) on the parabola $y = x^2$ that is closest to the point (0,2). If (x, y) is the closest point, find the value of $x^2 + y^2$. Give your answer correct to two decimal places. [Ans: 3.45 to 4.05]

4. Consider the matrix $A = \begin{pmatrix} 3 & 2 & -2 \\ 2 & 3 & 2 \end{pmatrix}$. Let the singular value decomposition (SVD) form of A be written as $A = Q_1 \Sigma Q_2^T$. Which of the following options is/are true for Q_1?

   (a) $Q_1 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

   (b) $Q_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

   (c) $Q_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

   (d) $Q_1 = \sqrt{2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$

   [Ans: 3]

5. Consider the matrix $A = \begin{pmatrix} 3 & 2 & -2 \\ 2 & 3 & 2 \end{pmatrix}$. Which of the following options is true for Σ?

(a) $\Sigma = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$

(b) $\Sigma = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$

(c) $\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$

(d) $\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

[Ans: 3]

ET May 25

1. A company manufactures x units of oil A and y units of another oil B. The total cost in rupees, C, of producing these two items is approximated by the function $C = 5x^2 + 2xy + 3y^2 + 500$. If the total production for both oil (both types combined) is 39 units, find the quantity of the oil B in the case of minimum production cost.

[Ans: 24]

Quiz 1 Sep 25

Quiz 2 Sep 25

1. Suppose $x + y = 16$. What is the value of $x + y$ when $x^3 + y^3$ is minimum? [Ans: 63.9 to 64.1]

2. Consider the task of minimizing the function $f(x, y) = x^2 + y^2$. Starting at (x_0, y_0), run two steps of gradient descent with learning rate $\eta = 0.2$ to obtain (x_2, y_2). If $(x_2, y_2) = \alpha \cdot (x_0, y_0)$, enter the value of α correct to two places after the decimal. [Ans: 0.35 to 0.37]

3. For what value of η will gradient descent reach the minimum after exactly one step for $f(x, y) = x^2 + y^2$? Enter your answer correct to one decimal place. [Ans: 0.49 to 0.51]

4. Select all values of η for which gradient descent converges to the minimum for $f(x, y) = x^2 + y^2$.

   (a) 0.3

   (b) 0.9

   (c) 1

   (d) 1.3

   [Ans: 1, 2]

ET Sep 25

1. Consider the function $f(x) = 2x^3 + 3x^2 - 12x + 5$. Which of the following is true?

   (a) $x = 1$ is a local maximum.

   (b) $x = -2$ is a local minimum.

   (c) $x = 1$ is a local minimum.

   (d) $x = -2$ is a local maximum.

[Ans: 3, 4]