Assessment sub (https://swayam.gov.in) (https://swayam.gov.in/nc_details/NPTEL)
X

sourav.ganguly1429@gmail.com

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Machine Learning for Engineering and science applications (course)

Click to register for Certification exam (https://examform.nptel.ac.in/2024_01/exam_form/dashboard)

If already registered, click to check your payment status

Course outline

About NPTEL ()

How to access the portal? ()

Prerequisites Assignment ()

Matlab and Learning Modules ()

Week 1 ()

Week 2 ()

Week 3 ()

- Machine Representation of Numbers, Overflow, Underflow, Condition Number (unit?unit=17&lesson=27)
- Derivatives,Gradient,Hessian,Jacobian,Taylor Series (unit?unit=17&lesson=28)
- Matrix Calculus (Slightly Advanced) (unit?unit=17&lesson=29)
- Optimization – 1 Unconstrained Optimization (unit?unit=17&lesson=30)
- Introduction to Constrained Optimization (unit?unit=17&lesson=31)
- Introduction to Numerical Optimization Gradient Descent - 1 (unit?unit=17&lesson=32)
- Gradient Descent – 2 Proof of Steepest Descent Numerical Gradient Calculation Stopping Criteria (unit?unit=17&lesson=33)
- Introduction to Packages (unit?unit=17&lesson=34)
- Week 3 feedback Form: Machine Learning for Engineering and Science Applications (unit?unit=17&lesson=161)
- **Quiz: Week 3 : Assignment 3 (assessment?name=209)**

Assessment submitted.
X

Week 4 ()

DOWNLOAD VIDEOS ()

Text Transcripts ()

Problem Solving Session - Jan 2024 ()

Thank you for taking the Week 3 : Assignment 3.

Week 3 : Assignment 3

Your last recorded submission was on 2024-02-14, 16:59 IST **Due date: 2024-02-14, 23:59 IST.**

1) In the context of using the gradient descent algorithm for training a machine learning model, if the algorithm is observed to not consistently reduce the cost in each iteration, which of the following strategies is most appropriate to adjust the learning rate? Also, identify the behavior the algorithm is likely exhibiting from the given options: ***1 point***

- ○ Increase the learning rate significantly; the algorithm is likely converging rapidly
- ◉ Decrease the learning rate slightly; the algorithm is likely oscillating without diverging or converging
- ○ Increase the learning rate slightly; the algorithm is likely converging slowly
- ○ Leave the learning rate unchanged; the algorithm is likely diverging

2) Consider the function: ***1 point***

$$G = 2x^2 + 3y^2 - 8x + 12y + 15.$$

Determine the critical point(s) of the function.

- ◉ x = 2, y = -2
- ○ x = 2, y = -1
- ○ x = -1, y = 2
- ○ x = -2, y = 2

3) For the above function evaluate whether the identified critical point(s) represents: ***1 point***

- ◉ Local minimum
- ○ Local maximum
- ○ Saddle point
- ○ None of the Above

4) Consider the function: ***1 point***

Assessment submitted.
X

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

Assume that we start gradient descent from (x_0, y_0) = (0, 0) with a learning rate of 0.1. Find the values of x and y after three iterations of gradient descent

- (0.143, -0.976)
- (-0.823, -0.534)
- (-0.488, 0.976)
- (-0.278, 0.488)

5) For the same function given above, assume that we start gradient descent from (x_0, y_0) = (0, 0) itself but with a learning rate of 0.01. Find the values of x and y after three iterations of gradient descent. Which learning rate brings the points closer to the theoretical minimum of the function? ***1 point***

- 0.1, because it moves the point faster towards the minimum
- 0.01, because it moves the point faster towards the minimum
- Both learning rates are equally effective
- Neither learning rate is effective

6) Consider a dataset with two features x_1 and x_2, and a dependent variable y. The dataset is given as follows: ***1 point***

x_1	x_2	y
1	1	6
2	1	8
3	2	14
4	3	18

The cost function for a linear regression model is defined as:

$$J(\theta_0, \theta_1, \theta_2) = \frac{1}{2m} \sum_{i=1}^{m} \left(\theta_0 + \theta_1 x_1^{(i)} + \theta_2 x_2^{(i)} - y^{(i)} \right)^2$$

where m is the number of training examples, and θ0, θ1, θ2 are the model parameters. Assume we perform one iteration of gradient descent from the starting point $(x_0, y_0) = (0, 0)$ with a learning rate of 0.1. What will be the new values of $\theta_0, \theta_1, \theta_2$ after this iteration?

- $\theta_0 = 1.15, \theta_1 = 3.4, \theta_2 = 2.4$
- $\theta_0 = -1.15, \theta_1 = -3.4, \theta_2 = -2.4$
- $\theta_0 = 5.75, \theta_1 = 1.7, \theta_2 = 1.2$
- $\theta_0 = -5.75, \theta_1 = -1.7, \theta_2 = -1.2$

7) For the above question, what is the absolute difference in the cost function J before and after the first iteration? ***1 point***

- 88.8
- 74.2

Assessment submitted.
X

- 808.8
- 172.2

8) Which of the following is true about gradient descent? **1 point**

- The learning rate must be constant
- (selected) After the iteration we modify the vector in the direction of the negative gradient
- After the iteration we modify the vector in the direction of the positive gradient
- After the iteration we modify the vector in the direction of the input vector

9) Given two vectors, $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ what is the correct partial derivative of a with respect to b? **1 point**

- $$\frac{\partial a}{\partial b} = \begin{bmatrix} \frac{\partial a_1}{\partial b_1} \\ \frac{\partial a_2}{\partial b_2} \end{bmatrix}$$
- $$\frac{\partial a}{\partial b} = \begin{bmatrix} \frac{\partial a_1}{\partial b_1} & \frac{\partial a_2}{\partial b_1} \\ \frac{\partial a_1}{\partial b_2} & \frac{\partial a_2}{\partial b_2} \end{bmatrix}$$
- (selected) $$\frac{\partial a}{\partial b} = \begin{bmatrix} \frac{\partial a_1}{\partial b_1} & \frac{\partial a_1}{\partial b_2} \\ \frac{\partial a_2}{\partial b_1} & \frac{\partial a_2}{\partial b_2} \end{bmatrix}$$
- $$\frac{\partial a}{\partial b} = \begin{bmatrix} \frac{\partial a_1}{\partial b_2} & \frac{\partial a_2}{\partial b_2} \\ \frac{\partial a_1}{\partial b_1} & \frac{\partial a_2}{\partial b_1} \end{bmatrix}$$

10) Given the function $J(w) = 3w_1^2 + 2w_2^2 - 12w_1 + 10w_2 + 15$, determine the theoretical value of the second component of $\text{argmin}_w(J(w))$. **1 point**

- -1.5
- (selected) -2.5
- -3.5
- -4.5

You may submit any number of times before the due date. The final submission will be considered for grading.

Submit Answers