Quiz 5

Problem 1

1/1 point (graded)

When gradient descent is used to solve a minimization problem, it is guaranteed to find a local minimum (that may or may not be the global minimum).

- (x) True
- ( ) False

Submit

Problem 2

1/1 point (graded)

You are trying to find the global minimum for a convex function of one variable, $F(w)$. At the current point $w = w_0$, you find that the derivative dF/dw is equal to 2.3. Based on this information, how should you update w?

- ( ) Choose $w > w_0$
- (x) Choose $w < w_0$
- ( ) Choose $w > -w_0$

More information is needed

Submit

Problem 3

1/1 point (graded)

What is the derivative, $\nabla F(\mathbf{w})$, of the function $F(\mathbf{w}) = (3\mathbf{w} \cdot \mathbf{x})$?

- $\nabla F(\mathbf{w}) = \mathbf{x}$
- $\nabla F(\mathbf{w}) = \mathbf{w}$
- $\nabla F(\mathbf{w}) = 3\mathbf{x}$
- $\nabla F(\mathbf{w}) = 3\mathbf{w}$

Submit

Problem 4

1/1 point (graded)

In the equation $\mathbf{w}_{t+1} = \mathbf{w}_t - \eta_t \nabla L(\mathbf{w}_t)$, what does η_t represent?

- The direction in which to adjust $\mathbf{w}$ to find a minimum
- The dimension of the vector $\mathbf{w}$

- ○ The approximate number of iterations the optimization algorithm has run
- ● The size of the adjustment made to $\mathbf{w}$

✔

Submit

Problem 5

1/1 point (graded)

True or false: An adjustment to $\mathbf{w}$ in the direction of the gradient is guaranteed to result in a vector of lower cost.

- ○ True
- ● False

✔

Submit

Problem 6

1/1 point (graded)

Given a function $L(\mathbf{x}) = 3x_2x_3 + 2x_1x_3 + 2x_1x_2$, compute the gradient $\nabla L(\mathbf{x})$.

- ○ $\nabla L(\mathbf{x}) = (4x_1, 5x_2, 5x_3)$
- ● $\nabla L(\mathbf{x}) = (2x_3 + 2x_2, 3x_3 + 2x_1, 3x_2 + 2x_1)$
- ○ $\nabla L(\mathbf{x}) = (2x_1x_3 + 2x_1x_2, 3x_2x_3 + 2x_1x_2, 3x_2x_3 + 2x_1x_3)$

- ◯ $\nabla L(\mathbf{x}) = (4x_2x_3, 6x_3x_1, 6x_1x_2)$

✔

Submit

Problem 7

1/1 point (graded)
Stochastic gradient descent is a better alternative to gradient descent in which of the following cases?

- [ ] There are multiple local minima in a function
- [x] There are a large number of data points
- [ ] The function contains more than 3 variables
- [ ] The function is discontinuous in at least one location

✔

Submit

Problem 8

1/1 point (graded)
A key difference between gradient descent and stochastic gradient descent is:

- ◯ Stochastic gradient descent takes longer to perform than gradient descent, but can be used on very large data sets

- ○ Stochastic gradient descent replaces gradient descent when the loss function contains a large number of variables
- ● Each move made by gradient descent is based on the entire data set, while each move made by stochastic gradient descent is based on a single data point.
- ○ Gradient descent only makes one pass through the training set, while stochastic gradient descent makes numerous passes before convergence

✔

Submit

Problem 9

1/1 point (graded)

Using mini-batch stochastic gradient descent, a group of data points are used to make adjustments to $\mathbf{w}$. Why might this be preferable to stochastic gradient descent based on a single point?

- ○ It takes less time to compute adjustments to $\mathbf{w}$
- ○ It results in a larger adjustment
- ● The batch-based gradient calculation is a closer approximation to the actual gradient
- ○ Fewer passes over the training set are required to find a minimum

✔

Submit

Problem 10

1/1 point (graded)

True or false: The negation of any convex function is a concave function.

- (x) True
- ( ) False

✔

Submit

Problem 11

1/1 point (graded)

Given a convex function $f(x)$ and two points in the domain, a and b, which of the following must be true? Select all that apply.

- [x] The line segment connecting $(a, f(a))$ and $(b, f(b))$ must lie above the function at every point on the line connecting a and b
- [ ] $f(x)$ must be monotonically increasing along the line segment joining a and b
- [ ] $f(a) > f(b)$ when $a < b$, and $f(b) > f(a)$ when $b < a$
- [ ] $f(x)$ must have a global minimum between a and b

✔

Submit

Problem 12

1/1 point (graded)

Which of the following functions are convex? Select all that apply.

- [x] $y = e^{-x}$
- [x] $y = x^2$
- [x] $y = 2x$
- [ ] $y = \sin(x),\ x \in [0, \pi]$

Submit

Problem 13

1/1 point (graded)

True or false: A function whose 2nd derivative is always negative is a convex function.

- ( ) True
- (•) False

Submit

Problem 14

1/1 point (graded)

The matrix $M = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$ has a positive determinant.

- ( ) Yes

- [x] No

Submit

Problem 15

1/1 point (graded)

Given matrix $M = \begin{pmatrix} 4 & 1 \\ k & 1 \end{pmatrix}$, what value of k results in a singular matrix?

- [ ] $k = -4$
- [ ] $k = -1$
- [ ] $k = 1$
- [x] $k = 4$

Submit

Problem 16

1/1 point (graded)

All matrices of the form $M = UU^T$ are always positive semidefinite

- [x] True
- [ ] False

Submit

Problem 17

1/1 point (graded)

The matrix $\begin{pmatrix} 14 & 7 \\ 7 & 6 \end{pmatrix}$ is positive semidefinite and follows the form $M = UU^T$. Which of the following matrices U satisfies this equation?

- $U = \begin{pmatrix} 1 & 4 \\ 1 & 6 \end{pmatrix}$
- $U = \begin{pmatrix} 2 & 7 \\ 3 & 1 \end{pmatrix}$
- $U = \begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \end{pmatrix}$
- $U = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \end{pmatrix}$

Submit

Problem 18

1/1 point (graded)

A function, $F(\mathbf{z})$, is convex if which of the following statements hold true?

- [x] The Hessian, $H(\mathbf{z})$, is positive semidefinite at all $\mathbf{z}$
- [ ] $F(\mathbf{z}) \geq 0, \forall \mathbf{z}$
- [ ] The gradient, $\nabla F(\mathbf{z})$, is monotonically decreasing
- [ ] The Hessian, $H(\mathbf{z})$, is symmetric

Submit

Problem 19

1/1 point (graded)

Is the identity matrix positive semidefinite?

- (•) Yes
- ( ) No

Submit