Quiz_ Quiz 3_ Linear...

Quiz 3: Linear Regression

Started: Sep 30 at 2:22am

Quiz Instructions

Question 1 1 pts

Which one is not a term of degree 2 polynomial features?

- $x_1 x_2$
- $2x_1 x_2$
- x_2^2
- ✓ $x_1 x_2^2$

Question 2 1 pts

Which of the following statements is not true regarding linear regression?

- It figures out the best fitting line/plane that describes two or more variables.
- It quantifies a relationship between two continuous variables.
- Random disturbance in the relationship between the independent variables and the dependent variable) is the same across all values of the independent variables.
- ✓ It can predict the outcome of a binary variable with continuous variables. F

Question 3 1 pts

$\frac{df}{dw} f(w) = (Xw - y)w + \lambda w$

$\frac{wf}{dw^2} f(w) = X^T X + \lambda I$

Let $\mathbf{X} \in \mathbb{R}^{n \times d}$, $\mathbf{y}, \mathbf{w} \in \mathbb{R}^d$.

Let $f(\mathbf{w}) := \frac{1}{2}\|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2 + \frac{\lambda}{2}\|\mathbf{w}\|_2^2$

What is $\frac{\partial f}{\partial^2 \mathbf{w}} f(\mathbf{w})$?

- $\mathbf{X}^\top \mathbf{X}$
- $2\mathbf{X}^\top \mathbf{X} + 2\lambda I$
- ✓ $\mathbf{X}^\top \mathbf{X} + \lambda I$
- $2\mathbf{X}^\top \mathbf{X}$

Question 4 1 pts

When can we say that a model generalizes?

- When generalization gap is small
- When the training loss is low
- ✓ When the test loss is low and training loss is low
- When the test lost is higher than the train loss

Question 5 1 pts

Suppose X is invertible and let $U\Sigma V^\top$ be its singular value decomposition. Which of the following is the pseudoinverse of X?

τ

- $X^\dagger = V^\top \Sigma^{-1} U^\top$

- ✓ $X^\dagger = V\Sigma^{-1} U^\top$
- $X^\dagger = U\Sigma V^\top$
- $X^\dagger = V^\top \Sigma^{-1} U$

Question 6 1 pts

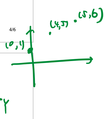
Which of the following is not true about overfitting?

- Testing loss is much higher than training loss.
- ✗ Both training and testing loss are high.
- The model is fitted too well to the data.
- The model is too complex.

Question 7 1 pts

Given a dataset $\mathbf{D} = \{(0,1), (4,5), (5,6)\}$, where the first number in each tuple is the input x_i and the second number is the label y_i. Consider the regression model $\hat{y}_i = b$ (i.e.: the

$\vec{w} = (X^T X)^{-1} X^T y$



predicted label is always the same, but b is a model parameter and can be optimized). Suppose the loss function $L(b) = \Sigma_{i=1}^N (y_i - \hat{y}_i)^2$. What is the optimal parameter ?

- ✓ 4
- 1
- 0
- 5

Question 8 1 pts

Given a dataset $D = (x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ and assume both x and y are scalars, we apply a univariate linear regression model $\hat{y}_i = wx_i + b$. Let and b^* denote the optimal model parameters , b under minimization of the square loss $L(w, b) = \Sigma_{i=1}^N (y_i - \hat{y}_i)^2$. Which of the followings is **false**?

- The loss function $L(w, b)$ is convex in both w and b T
- $b^* = \frac{1}{N}\Sigma_{i=1}^N (y_i - w^* x_i)$
- ✓ $w^* = \frac{\Sigma_{i=1}^N y_i (x_i - \bar{x})}{\Sigma_{i=1}^N x_i^2 - \bar{x}^2}$ where $\bar{x} = \frac{1}{N}\Sigma_{i=1}^N x_i$

- The optimization problem that should be solved is $L(w, b) = \Sigma_{i=1}^N (y_i - wx_i - b)^2$ T

Quiz saved at 2:22am Submit Quiz