Week 2 Quiz

Bookmark this page

Multiple Choice

1.0/1.0 point (graded)

In Lecture 3 we saw how the least squares linear regression solution from Lecture 2 could be given a probabilistic interpretation by assuming the errors to be...

- ○ independent zero-mean Gaussian random variables with different noise variances.
- ○ correlated zero-mean Gaussian random variables with different noise variances.
- ◉ independent zero-mean Gaussian random variables with shared noise variance. ✔
- ○ correlated zero-mean Gaussian random variables with shared noise variance.

Submit You have used 1 of 1 attempt

Checkboxes

1/1 point (graded)

Using the probabilistic approach to linear regression from Lecture 3, as well as the notations we have been using for the linear regression problem thus far, click all equivalent ways for generating from $p(y|X,w)$.

- ☑ $y_i = x_i^T w + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} N(0,\sigma^2), \quad \text{for} \quad i = 1,\ldots,n$
- ☑ $y_i \overset{ind}{\sim} N(x_i^T w, \sigma^2), \quad \text{for} \quad i = 1,\ldots,n$

- [x] $y \sim N(Xw, \sigma^2 I)$
- [x] $y = Xw + \vec{\epsilon}, \;\; \vec{\epsilon} \sim N(0, \sigma^2 I)$

✔

Submit You have used 1 of 1 attempt

Multiple Choice

1/1 point (graded)

Under the modeling assumption $y \sim N(Xw, \sigma^2 I)$, which of the following is true of the maximum likelihood solution for w?

- (•) $\mathbb{E}[w_{ML}] = w, \;\; Var[w_{ML}] = \sigma^2 (X^T X)^{-1}$ ✔
- ( ) $\mathbb{E}[w_{ML}] = (X^T X)^{-1} X^T y, \;\; Var[w_{ML}] = \sigma^2 (X^T X)^{-1}$
- ( ) $\mathbb{E}[w_{ML}] = (X^T X)^{-1} X^T y, \;\; Var[w_{ML}] = \sigma^2 X^T X$
- ( ) $\mathbb{E}[w_{ML}] = w, \;\; Var[w_{ML}] = \sigma^2 X^T X$

Submit You have used 1 of 1 attempt

Multiple Choice

1/1 point (graded)

Given the model $y \sim N(Xw, \sigma^2 I)$, which of the following is true about the maximum likelihood estimator w_{ML}?

- ( ) w_{ML} always has a unique solution
- ( ) w_{ML} has the smallest variance among all estimators for w
- ( ) $\mathbb{E}[w_{ML}] = (X^T X)^{-1} X^T y$
- (•) w_{ML} is an unbiased estimator of w ✔

Submit You have used 1 of 1 attempt

Multiple Choice

1/1 point (graded)

Assume $w^* = \arg\min_w \|y - Xw\|_2^2 + \lambda g(w)$, which of the following is true?

- When $g(w) = \|w\|^2$, the magnitude of values in w^* tend to increase
- The solution for w^* is analytical for arbitrary positive function $g(w)$
- When $g(w) = \|w\|^2$, the values in w^* are more stable to variations in y and X ✔
- The solution for w^* is always unique for arbitrary positive function $g(w)$

Submit You have used 1 of 1 attempt

Multiple Choice

1/1 point (graded)

The solution to ridge regression is $w_{RR} = (\lambda I + X^T X)^{-1} X^T y$. As λ increases, the value of $\|w_{RR}\|_2$

- increases
- decreases ✔

Submit You have used 1 of 1 attempt

Text Input

2.0/2.0 points (graded)

We saw how for a model with parameters θ that generates data x, Bayes rule allows us to learn about θ using:

A) prior distribution

B) likelihood distribution

C) posterior distribution

Use the letters A, B or C to identify the names of each of the distributions below.

$p(\theta|x)$

C ✔

$p(\theta)$

A ✔

$p(x|\theta)$

B ✔

Submit You have used 1 of 1 attempt

Checkboxes

2.0/2.0 points (graded)

We've discussed a few perspectives of the linear regression problem thus far. These different perspectives have led to equivalent solutions. Check all equivalent solutions below. (Please note that this problem does not have partial credit.)

- [x] $w_{LS} \Leftrightarrow w_{ML}$
- [ ] $w_{RR} \Leftrightarrow w_{LS}$
- [ ] $w_{ML} \Leftrightarrow w_{MAP}$

- [x] $w_{RR} \Leftrightarrow w_{MAP}$
- [ ] $w_{ML} \Leftrightarrow w_{RR}$
- [ ] $w_{LS} \Leftrightarrow w_{MAP}$

Submit You have used 1 of 1 attempt