Week 3 Quiz

Bookmark this page

Multiple Choice

1/1 point (graded)

Assume that $y \sim N(Xw, \sigma^2 I)$ is the likelihood model for the problem we are considering. Then the MAP solution $w_{MAP} = \arg\max_w \ln p(w|y, X)$ is: (1) always the same, and (2) unbiased, for any prior $p(w)$.

- ( ) (1) TRUE, (2) FALSE
- ( ) (1) TRUE, (2) TRUE
- ( ) (1) FALSE, (2) TRUE
- (•) (1) FALSE, (2) FALSE ✔

Submit You have used 1 of 1 attempt

Checkboxes

0/1 point (graded)

Which of the following are MAP solutions of a model with likelihood $p(y|w, X)$ and prior $p(w)$?

- [x] $\arg\max_w \ln p(y, w|X)$
- [x] $\arg\max_w \ln[p(y|w, X)p(w)]$
- [ ] $\arg\max_w \ln p(w|X)$
- [ ] $\arg\max_w \ln p(y|w, X)$
- [x] $\arg\max_w \ln p(w|X, y)$

- Week 11
- Week 12
- Post-Course Survey

Submit You have used 1 of 1 attempt

Numerical Input

2.0/2.0 points (graded)

Let the vector $w \in \mathbb{R}^3$ have a Gaussian distribution $w \sim N(\mu, \Sigma)$ where $\mu = [1,\ 2,\ 3]^T$ and $\Sigma = \mathbf{diag}(1,\ 1,\ 2)$.

1. The mean of $w_1 + 2w_2 + 3w_3 = \underline{\text{enter below}}$

14

14

2. The variance of $w_1 + 2w_3 = \underline{\text{enter below}}$

9

9

This question tests a fundamental property of the Gaussian distribution that could be considered a probability prerequisite. The information is not directly from the slides, but is very easily found online.

Submit You have used 1 of 1 attempt

Multiple Choice

1/1 point (graded)

For a model with likelihood $p(y|w, X)$ and prior $p(w)$, given the training pairs (y, X) we test a new observation (y_0, x_0) by predicting y_0 given x_0. To compute this predictive distribution we need to calculate $p(y_0|w, x_0, y, X)$.

- True
- False ✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

Checkboxes

0/1 point (graded)
Active learning for linear regression can be treated as a clever way to sequentially enlarge the training data by measuring new observation pairs (y, x). Which of the following are NOT active learning strategies?

- [x] Pick x uniformly at random from the choices and measure y
- [ ] Pick x to significantly reduce uncertainty according to some measure
- [x] Pick x by asking someone (e.g., an expert) for advice and measure y
- [ ] Pick x for which we were the most incorrect in the prediction of y

✖

Submit You have used 1 of 1 attempt

✖ Incorrect (0/1 point)

Checkboxes

0/1 point (graded)
For X an $n \times d$ matrix and y an n-dimensional vector, it is possible that the linear system $y = Xw$ may have multiple solutions when

- [x] $n < d$
- [ ] $n \geq d$
- [ ] The null space of X is empty
- [x] XX^T is invertible

✖

Submit You have used 1 of 1 attempt

Multiple Choice

1/1 point (graded)
The vector w that satisfies the least squares solution of the linear system $y \approx Xw$ has the smallest ℓ_2 norm among all solutions.

- (x) True ✔
- ( ) False

Submit You have used 1 of 1 attempt

Checkboxes

1/1 point (graded)
Which of the following will likely give a sparse solution for w?

- [x] $\arg\min_w \|y - Xw\|_2^2 + \lambda\|w\|_{1/2}$
- [ ] $\arg\min_w \|y - Xw\|_1 + \lambda\|w\|_2^2$
- [ ] $\arg\min_w \|y - Xw\|_2^2 + \lambda\|w\|_3^3$
- [x] $\arg\min_w \|y - Xw\|_1 + \lambda\|w\|_{3/4}$

✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

Multiple Choice

1/1 point (graded)

For an optimization problem of the form $\arg\min_w \|y - Xw\|^2 + \lambda\|w\|_p$ the values of p for which we can NOT guarantee an optimal solution are:

- ○ $p > 2$
- ◉ $p < 1$ ✔
- ○ $1 < p < 2$

Submit You have used 1 of 1 attempt