

| CIA   | -   |
|---|---|
| Bookmarks   | Week 3 > Week 3 Quiz > Week 3 Quiz  Week 3 Quiz   |
| <ul> <li>Machine         <ul> <li>Learning</li> <li>Course: Getting</li> </ul> </li> <li>Started</li> </ul> | Multiple Choice 1/1 point (graded) Assume that $y\sim N(Xw,\sigma^2I)$ is the likelihood model for the problem we                           |
| ▶ Week 1  | are considering. Then the MAP solution $w_{MAP} = rg \max_w \ln p(w y,X)$ is: (1) always the same, and (2) unbiased, for any prior $p(w)$ . |
| ▶ Week 2  | (1) TRUE, (2) FALSE   |
| ▼ Week 3  | (1) TRUE, (2) TRUE  |
| Lecture 5 Bayesian<br>Linear Regression   | (1) FALSE, (2) TRUE   |
| Lecture 6 Sparse<br>Linear Regression   | <ul><li>(1) FALSE, (2) FALSE ✓</li></ul>  |
| <b>Week 3 Quiz</b> Quiz due Apr 11, 2017  07:30 MYT   |   |
| Week 3 Project: Linear Regression Project due Apr 11, 2017 07:30 MYT  | Submit You have used 1 of 1 attempt   |
| Week 3 Discussion<br>Questions  | Checkboxes  1/1 point (graded)  Which of the following are MAP solutions of a model with likelihood   |
| ▶ Week 4  | p(y w,X) and prior $p(w)$ ?   |
| ▶ Week 5  | $ ightharpoonup rg \max_w \ln p(y,w X)$   |
| ▶ Week 6  | $	extstyle 	ext{arg max}_w \ln[p(y w,X)p(w)]$   |
|   | $lacksquare \ rg \max_w \ln p(w X)$   |
|   | $lacksquare rg \max_w \ln p(y w,X)$   |
|   | $	extbf{	extit{@}} rg \max_w \ln p(w X,y)$  |
|   | <b>✓</b>  |

## **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  $\Sigma=\mathrm{diag}(1,~1,~2)$ .

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



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



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  $p(y_0|w,x_0,y,X)$ .



Submit

You have used 1 of 1 attempt

| (y,x)                                    | learning for linear regression can be treated as a clever way to ntially enlarge the training data by measuring new observation pairs . Which of the following are NOT active learning strategies? |
|--|--|
| ✓ F                                      | Pick $m{x}$ uniformly at random from the choices and measure $m{y}$  |
|  | Pick $m{x}$ to significantly reduce uncertainty according to some measure  |
| <b>✓</b> F                               | Pick $m{x}$ by asking someone (e.g., an expert) for advice and measure $m{y}$  |
| <b>✓</b> P                               | Pick $oldsymbol{x}$ for which we were the most incorrect in the prediction of $oldsymbol{y}$   |
| <b>~</b>                                 |  |
| Sub                                      | You have used 1 of 1 attempt   |
| 0/1 poi<br>For $oldsymbol{X}$<br>the lin | <b>kboxes</b><br>Int (graded)<br>Int $n 	imes d$ matrix and $y$ an $n$ -dimensional vector, it is possible that ear system $y = Xw$ may have multiple solutions when                               |
| · r                                      | . 1  |
|  | n < d  |
|  | $n < d$ $n \ge d$  |
|  |  |
| ОТ                                       | $n \geq d$   |
| ОТ                                       | $n \geq d$<br>The null space of $oldsymbol{X}$ is empty  |
|  | $n \geq d$<br>The null space of $X$ is empty $XX^T$ is invertible  |
| Sub                                      | $n \geq d$ The null space of $oldsymbol{X}$ is empty $oldsymbol{X} oldsymbol{X}^T$ is invertible   |

ypprox Xw has the smallest  $\ell_2$  norm among all solutions. True False You have used 1 of 1 attempt Submit Checkboxes 1/1 point (graded) Which of the following will likely give a sparse solution for w?  $extbf{ extit{ iny arg min}}_w \|y - Xw\|_2^2 + \lambda \|w\|_{1/2}$  $lacksquare rg \min_{w} \|y - Xw\|_2^2 + \lambda \|w\|_3^3$  $extit{ } extit{ } ext{arg min}_{w} \, \|y - Xw\|_{1} + \lambda \|w\|_{3/4}$ Submit You have used 1 of 1 attempt **Multiple Choice** 1/1 point (graded) For an optimization problem of the form  $rg \min_{w} \|y - Xw\|^2 + \lambda \|w\|_p$ the values of  $\boldsymbol{p}$  for which we can NOT guarantee an optimal solution are: p>2 p < 1 
✓</p> 0 1

Submit

You have used 1 of 1 attempt

© All Rights Reserved



© 2012-2017 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

















