

Homework 8

STA-360/602

Total points: 10 (reproducibility) + 10 (Q1) = 20 points. **Q2 is worth up to 8 points of extra credit on this assignment to help anyone that has had a difficult time in the course and also to help you prepare for the final exam.**

General instructions for homeworks: Please follow the uploading file instructions according to the syllabus. You will give the commands to answer each question in its own code block, which will also produce plots that will be automatically embedded in the output file. Each answer must be supported by written statements as well as any code used. Your code must be completely reproducible and must compile. Syllabus: (<https://github.com/resteorts/modern-bayes/blob/master/syllabus/syllabus-sta602-spring19.pdf>)

Advice: Start early on the homeworks and it is advised that you not wait until the day of. While the professor and the TA's check emails, they will be answered in the order they are received and last minute help will not be given unless we happen to be free.

Commenting code Code should be commented. See the Google style guide for questions regarding commenting or how to write code <https://google.github.io/styleguide/Rguide.xml>. No late homework's will be accepted.

Please look over the homework before lab this week. TA's will answer questions on the homework this week regarding these two problems below. I recommend that you work through them as much as possible before lab this week.

1. *Lab component* (10 points) Please complete **Lab 10**, parts c and d which correspond with linear regression, which can be found here: https://github.com/resteorts/modern-bayes/blob/master/labs/10-linear-regression/11-linear-regression_v2.pdf. It is highly recommend that you work through parts (a) and (b) on your own and derive these as these are excellent practice exercises for the exam. You can check your own work on this.
 - c) (5 points) Complete lab 10, part c.
 - d) (5 points) Complete lab 10, part d.

2. **Extra Credit** (8 points) *Multivariate Methods* This problem is optional for the homework, but it is worth extra credit. It is also a type of problem that is reasonable to appear on the final exam, and it's similar to what we did in class regarding the multivariate normal distribution.

Because this problem is extra credit, the TA Team will not answer questions on this except for clarification questions. These should be sent privately. You should work on your own on this problem and not with others to help you prepare for the final examination.

Consider the following hierarchical model:

$$\mathbf{y}_i \mid \boldsymbol{\theta}, \boldsymbol{\Sigma} \stackrel{i.i.d.}{\sim} MVN(\boldsymbol{\theta}_{d \times 1}, \boldsymbol{\Sigma}_{d \times d}), \quad i = 1, \dots, n,$$

and independent priors

$$\boldsymbol{\theta}_{d \times 1} \sim MVN(\boldsymbol{\mu}_{d \times 1}, \mathbf{T}_{d \times d}), \quad \boldsymbol{\Sigma}_{d \times d} \sim \text{inverseWishart}(\nu, \Psi_{d \times d}^{-1}).$$

- (a) (1 point) Show that $(\boldsymbol{\theta}^T \mathbf{T}^{-1} \boldsymbol{\mu})^T = \boldsymbol{\mu}^T \mathbf{T}^{-1} \boldsymbol{\theta}$. *Hint*: what's the transpose of a scalar?
- (b) (1 point) Use (a) to show that:

$$p(\boldsymbol{\theta}) \propto e^{-\frac{1}{2}(\boldsymbol{\theta}^T \mathbf{T}^{-1} \boldsymbol{\theta} - 2\boldsymbol{\theta}^T \mathbf{T}^{-1} \boldsymbol{\mu})}$$

- (c) (2 points) Use (b) to show that

$$p(\boldsymbol{\theta} \mid \boldsymbol{\Sigma}, \mathbf{y}) \sim MVN\{\boldsymbol{\mu}^*(\boldsymbol{\Sigma}) := \mathbf{T}^*(n\boldsymbol{\Sigma}^{-1}\bar{\mathbf{y}} + \mathbf{T}^{-1}\boldsymbol{\mu}), \mathbf{T}^* := (n\boldsymbol{\Sigma}^{-1} + \mathbf{T}^{-1})^{-1}\}.$$

- (d) (2 points) Show that:

$$\text{tr}(\Psi \boldsymbol{\Sigma}^{-1}) + \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\theta}) = \text{tr} \left\{ \left(\Psi + \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\theta})(\mathbf{y}_i - \boldsymbol{\theta})^T \right) \boldsymbol{\Sigma}^{-1} \right\}.$$

Hint: Recall that $\text{tr}(A)$ is the trace of matrix A – the sum of its diagonal entries. *Lemma:* (i) for scalar a , $\text{tr}(a) = a$; (ii) $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$.

- (e) (2 points) Use (d) to show (step-by-step) that:

$$[\boldsymbol{\Sigma} \mid \boldsymbol{\theta}, \text{data}] \sim IW(\nu^*, \Psi^*(\boldsymbol{\theta}))$$

where $\nu^* = n + \nu$ and $\Psi^*(\boldsymbol{\theta}) = (\Psi + \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\theta})(\mathbf{y}_i - \boldsymbol{\theta})^T)^{-1}$.