

Midterm Study Guide

Midterm - Friday Feb 8th in class

2019-02-01

The midterm is closed book.

I will provide statistical tables if you need them (so you should know how to use them).

You should bring a calculator (although arithmetic errors are generally forgiven).

Things you will **not** have to do:

- write any R code
- invert more than a 2x2 matrix
- perform matrix arithmetic on more than a 3x3 matrix

You should be able to:

- State the multiple regression model in matrix form along with the assumptions on the errors and design matrix.
- Describe the entries in the design matrix given a model and study description.
- Derive the least squares estimates.
- Define fitted values and residuals.
- Describe the difference between random errors and residuals.
- Derive the mean and variance-covariance matrix of the least squares estimates in multiple linear regression.
- State the Gauss-Markov theorem and discuss its consequences in practice.
- State the form and properties of the estimate for the variance of the errors.
- Describe why using `lm()` in R is preferable to performing the matrix algebra $(X^T X)^{-1} X^T Y$.
- State the distribution of the least squares estimates under the assumption of Normal errors.
- Identify properties of the least squares estimates (i.e. form of the estimates, mean, variance and distribution of the estimates, unbiasedness, BLUE, etc.) that rely on the Normality assumption.
- Describe the consequences of having orthogonal columns in the design matrix.
- State the null distribution of t-statistics and F-statistics in hypothesis tests relevant to multiple linear regression models.
- Construct t-based confidence intervals and hypothesis tests on individual parameters, or linear combinations of individual parameters, given either R output, or the necessary estimates, and $(X^T X)^{-1}$
- Construct prediction intervals for the mean response or a future response, given either R output, or the necessary estimates, and $(X^T X)^{-1}$
- Discuss the difference between a interval for the mean response and an interval for a future response.
- Discuss ways in which a prediction model can go wrong.
- Interpret a confidence interval or prediction interval in the context of a study.

- Comment on the conclusion a hypothesis test would reach based on the result of a confidence interval/region.
- Conduct an F-test to compare two models.
- Interpret the result of an F-test in context of a study.
- State the null and alternative hypotheses in the overall regression sum of squares F-test.
- Find the linear parameteric function test equivalent to a model test and vice versa (i.e. HW#4 Q1).