DATA 140

Probability for Data Science

Fall 2025

WEEK 15 STUDY GUIDE

The Big Picture

We conclude the term with conditioning in the multivariate normal model, and inference in the standard multiple regression model.

- The regression line can be written in multiple forms, one of which extends to the case of multiple regression.
- Prediction based on multiple predictors has familiar properties: There is a general formula for the best linear predictor, which is a natural extension of the formula for simple regression; and if the underlying distribution is multivariate normal then the best linear predictor is also the best among all predictors.
- The multiple regression model with normal errors is fundamentally important in data science. Properties of the estimated parameters lead to straightforward methods of inference.

Week At a Glance

Mon 12/1	Tue 12/2	Wed 12/3	Thu 12/4	Fri 12/5
	Lecture	Section	Lecture	Mega Section
HW 14 Due at 5PM HW 15 (Due 5PM Mon 12/8)			HW 15 party 9 AM to 12 PM	HW 15 party 2 PM to 5 PM
Focus on understanding HW 14	Work through Chapter 24	Skim Section 25.4; work on HW 15	Work through Section 25.4	Work on HW 15

Reading, Practice, and Class Meetings

Book	Topic	Lectures: Professor	Sections: TAs	Optional Additional Practice
Ch 24, 25	**Towards Multiple Regression** - 24.4 writes the regression equation in multiple different ways, each one illuminating a different property and making it easier to understand the corresponding formulas in multiple regression - 25.1, 25.2, 25.3 extend the corresponding simple regression sections (24.1, 24.3, 24.4) to the multivariate case; we will just talk through these and not do the details - 25.4 introduces the multiple regression model most commonly used in data science	Tuesday 12/2 - MSE in simple regression; connection with the bivariate normal - The big picture of the multivariate case - The multiple linear regression model: understanding the assumptions	Wednesday 12/3 - Ch 24 Ex 2, 3, 4	None; focus on the homework
	Multiple Linear Regression - 25.4 continued: the estimates and their distribution under the model	Thursday 12/4 - Multiple linear regression model: parameter estimation and inference	Friday 12/5 - Ch 24 Ex 6, 7 - Multiple regression model True/False	