

DATA 140



Spring 2024

WEEK 13 STUDY GUIDE

The Big Picture

We write familiar facts about expectation and covariance in matrix notation, and use them to study the most important joint distribution in data science. We then see how this is connected with simple regression.

- Linear algebra helps us express properties of sequences of random variables. Expectation and variance are replaced by mean vectors and covariance matrices.
- The multivariate normal distribution has a few equivalent definitions, chief among which is that multivariate normal variables can be represented as a linear transformation of i.i.d. standard normals. Linear combinations of multivariate normal random variables are normal; multiple linear combinations are multivariate normal; and pairwise uncorrelated multivariate normal variables are independent.
- Simple linear regression predicts Y as a linear function of a single X . No matter what the joint distribution of X and Y , there is always a least squares line. If X and Y are bivariate normal, this line turns out to be the best among all predictors.

Week At a Glance

Mon 4/15	Tue 4/16	Wed 4/17	Thu 4/18	Fri 4/19
	Lecture	Section	Lecture	Mega Section
Lab 8 (due Mon 4/22)			Lab 8 party 9 AM to 11 AM	
HW 12 Due HW 13 (Due Mon 4/22)				HW 13 party 2 PM to 5 PM
Skim Sections 23.1, 23.2.	Work through Sections 23.1, 23.2	Skim Section 24.1.	Work through Sections 24.1 to 24.3.	Work through Chapter 23 (apart from Section 23.3) and Chapter 24.

Reading, Practice, and Class Meetings

Book	Topic	Lectures: Prof. A.	Sections: TAs	Optional Additional Practice
Ch 23	<p>Multivariate Normal Vectors</p> <ul style="list-style-type: none"> - 23.1 derives the mean vector and covariance matrix of a linear transformation of a random vector; covariance matrices are positive semidefinite - 23.2 defines the multivariate normal as a linear transformation of iid standard normals, and covers the resulting properties; in particular, normal marginals don't imply jointly multivariate normal - 23.3 examines the multivariate normal joint density - 23.4 shows that for multivariate normal variables, being pairwise uncorrelated is equivalent to independence 	<p>Tuesday 4/16</p> <ul style="list-style-type: none"> - Random vectors and linear transformations - Multivariate normal and properties 	<p>Wednesday 4/17</p> <ul style="list-style-type: none"> - Ch 23 Ex 1, 2, 3 	<p>None; focus on the homework and lab. We'll try to cover all the chapter exercises in section and homework in Weeks 13 and 14.</p>
Ch 24	<p>Simple Regression</p> <ul style="list-style-type: none"> - 24.1 derives the equation of the regression line - 24.2 constructs bivariate normal random variables so that the relation between can be expressed in terms of "linear signal plus noise" - 24.3 looks at least-squares prediction in the context of the bivariate normal, and the connection with linear regression - 24.4 writes the regression equation in multiple different ways, each one illuminating a different property 	<p>Thursday 4/18</p> <ul style="list-style-type: none"> - Simple regression: general case - Bivariate normal - Regression and the bivariate normal 	<p>Friday 4/19</p> <ul style="list-style-type: none"> - Ch 24 Ex 2, 3, 4, 5 	