

Math 585, **Multivariate Analysis** Spring 2024. MWF 3–3:50, Neckers 156

Instructor: David Olive

Text: Olive, D.J. (2017), *Robust Multivariate Analysis*, Springer, New York, NY.

You may use the earlier online version (<http://parker.ad.siu.edu/Olive/mrun.pdf>):
Olive (2012) *Robust Multivariate Analysis*.

Course Webpage: (<http://parker.ad.siu.edu/Olive/M585.html>)

This text is also useful. Johnson, R.A., and Wichern, D.W. (1988), *Applied Multivariate Statistical Analysis*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ.

Office: Neckers 261 (wing A) *Phone:* (618)-453-6566

email: dolive@siu.edu

Office hours: MWThF 1-2:30

I am also available by appointment and on a walkin basis, especially before and after class.

The *prerequisites* for this class are a Calculus based introduction to Probability or Statistics (eg Math 483 and Linear Algebra (eg Math 221). You should be familiar with vectors, matrices, eigenvalues, eigenvectors, inverse matrix, transpose, the normal and chisquare distributions, and hypothesis testing.

This course covers robust multivariate analysis. The classical multivariate method and a robust multivariate method will be given for several common techniques. Topics include i) classical and robust estimators of multivariate location and dispersion, ii) multivariate normal and elliptically contoured distributions, iii) the DD plot for outlier detection and for determining whether the data is from a multivariate normal distribution or some other elliptically contoured distribution, iv) generalized variance, v) classical and robust analogs of the Hotelling's T^2 test, vi) MANOVA, vii) Multivariate Regression, viii) classical and robust methods of Principal Component Analysis, ix) Factor Analysis, x) classical and robust methods of Canonical Correlation Analysis, xi) Discriminant Analysis. xii) Clustering may be covered if there is time.

The free statistical software *R* will be heavily used, and *SAS* will be used. and we will meet in the Math computer lab Neckers 258 several times.

Final: Tuesday May 7, 8-10 AM in the morning in Neckers 156.

The grading and schedule below are tentative. (Drop paper work Friday March 29 with an advisor, last drop day Sunday, March 31 online.)

Except for the last week of classes, 2 homeworks may be turned in one class period late (ie on Monday) with no penalty. A third late will be accepted with 25% penalty. One or more sheets of notes will be allowed on quizzes and exams. A calculator is permitted. I sometimes give $A-$, $B+$, $B-$, and $C+$.

Grading:

HW	300		Quizzes	100	
exam1	100	exam 2	100	exam 3	100
final	300	or project		total	1000
min. grade	points	min. grade	points	min. grade	points
A	900-1000	B	800-899	C	700-799
D	550-699				

Parentheses for Johnson and Wichern (1988).

Week of	MON	WED	FRI
Jan 15	no class	2.1,(1.3)	2.2,2.4,3.1, (2.5,2.6)
Jan 22	3.1, (4.2,4.3),	3.1,3.2 (4.2,4.3,4.4), Q1	2.3,3.2,3.3 (3.2,3.3,4.5,4.6), HW1
Jan 29	3.2,3.3,3.4, (4.6,4.7)	3.4,4.3,5.1,(3.4, 3.5), Q2	5.1,5.2, HW2
Feb 5	lab	5.1,5.2, Q3	5.2, 6.1(8.2), HW3
Feb 12	6.1,(8.2)	Exam 1	6.1,6.2(8.2,8.3,8.4)
Feb 19	7.1,7.2 (10.2,10.3,10.4,10.6)	7.2, 8.2 (11.2), Q4	8.2, (11.3,11.4), HW4
Feb 26	lab	8.3, (11.5,11.6), Q5	8.3,8.7 (11.7), HW5
March 4	8.7,9.1 (5.2)	9.2, (5.3, 5.5), Q6	9.3,9.4, HW6
March 11	no class	no class	no class
March 18	9.4,10.3,	Exam 2	10.3,10.4(6.3,6.7)
March 25	10.3,10.4,11.1(6.4,6.7,9.2),	11.1,11.2 (9.4,9.5), Q7	12.1(7.2), HW7
April 1	12.2(7.3),	12.3(7.3,7.4), Q8	12.3(12.3), HW8
April 8	no class	12.3,12.7, Q9	13.1, HW9
April 15	HD	HD, Q10	HD, HW10
April 22	HD	HD, Q11	HD, HW11
April 29	HD	Exam 3	rev

Similar material on multivariate linear regression is at (<http://parker.ad.siu.edu/Olive/slch10.pdf>) and (<http://parker.ad.siu.edu/Olive/linmodch8.pdf>). Similar material on one way MANOVA is at (<http://parker.ad.siu.edu/Olive/linmodch9.pdf>).

HD is high dimensional statistics. See the following manuscripts on the course webpage.

a) Olive, D.J., and Zhang, L. (2024), One Component Partial Least Squares, High Dimensional Regression, Data Splitting, and the Multitude of Models, Communications in Statistics: Theory and Methods, to appear.

b) Olive and Alshammari (2024), Testing with the One Component Partial Least Squares and the Marginal Maximum Likelihood Estimators.

c) Haile, Zhang, and Olive (2024), Predicting Random Walks and a Data Splitting Prediction Region, Stats, 7(1), 23-33.

d) Abid and Olive (2024), Some Simple High Dimensional One and Two Sample Tests.