SCALING AND DIMENSIONAL ANALYSIS

Course Objectives: Consider the three terms that are combined in the title of this course: “Measurement” is an operation that is fundamental to scientific research; however, its implications and consequences are often poorly understood. Greater appreciation for the nature of measurement is useful for discerning and exploiting systematic structure in empirical data. Next, the terms “scaling” and “dimensional analysis” refer to a wide variety of research strategies and procedures. The common element among them is that they all seek to provide quantitative and/or geometric representations of the internal structure in a set of multivariate data. Researchers apply these techniques for three main reasons: (1) Simple data reduction—summarizing a large set of variables with a smaller number of composite measures; (2) examining dimensionality—testing the underlying sources of variation in a dataset; and (3) measurement—obtaining empirical representations of the underlying (and usually unobservable) dimensions, which can be employed as analytic variables in other statistical procedures. On a less formal note, researchers will often find that dimensional analysis is very beneficial for conceptualizing the contents of their data. In addition, these techniques usually provide visual displays that are very useful for presenting analytical results to other people. Thus, for a variety of reasons, scaling and dimensional analysis are useful additions to the social scientist’s “repertoire” of research strategies.

Course Prerequisites: This course assumes that participants are familiar, and comfortable, with basic descriptive statistics and the multiple regression model. Some prior exposure to matrix algebra would also be helpful, but it is not absolutely required. However, we will encounter certain mathematical operations which are undefined outside the context of matrices (i.e., the singular value decomposition). Therefore, anyone who has never worked with matrix algebra should be sure to attend the ICPSR Summer Program course, “Introductory Lectures on Matrix Algebra.”

Software Considerations: With very few exceptions, the methods covered in this workshop are computationally intensive. Therefore, appropriate software is required to perform most of the analyses. We will rely primarily upon the R statistical computing environment. Most (but, unfortunately, not all) of the procedures also can be carried out in STATA. In addition, the other major statistical packages (e.g., SPSS, SAS, etc.) contain routines for most of the techniques that we will cover in this course. From time to time, we may use some of these other packages, as well as special-purpose software, to carry out certain analyses.

Course Web Site: The home page for this course is located at the following URL:

http://www.polisci.msu.edu/jacoby/icpsr/scaling

The contents of this website will evolve and expand as the course proceeds through the subject matter. You should regard the site as an information resource. It will contain the syllabus, copies of handouts, datasets, assignments, computing and software resources, lecture outlines, and links to other interesting and useful internet sites.
Reading Material: Unfortunately, there is no single textbook that covers all of the topics in this course. In addition, many of the texts that are available have certain drawbacks that limit their usefulness for our purposes: They tend to be very expensive; they usually assume a high level of mathematical sophistication; they often contain sections that are out of date. Because of these considerations, we will rely primarily on several shorter works taken from the Sage series on Quantitative Applications in the Social Science (i.e., the “little green books”):


The following textbooks are also quite good. Some participants may want to supplement or replace the Sage Papers with entries from this list:

Armstrong II, David A.; Ryan Bakker; Royce Carroll; Christopher Hare; Keith T. Poole; Howard Rosenthal. (2014) *Analyzing Spatial Models of Choice and Judgment with R*. CRC Press.


TOPICS AND READING ASSIGNMENTS

I. Introductory Concepts

A. Data Theory
   Jacoby (1991), Chapter 3.

B. Dimensionality
   Jacoby (1991), Chapter 4.

C. Measurement Theory and Its Implications
   Jacoby (1991), Chapters 1 and 2
   Weller and Romney (1990), Chapter 1.
   Lattin et al. (2003), Chapter 1.
   Young, Forrest W. (1981) “Quantitative Analysis of Qualitative Data.” Psychome-
   trika 46: 357-388.
   De Veaux, Richard D. (1990) Finding Transformations for Regression Using the
   ACE Algorithm.” In John Fox and J. Scott Long (Editors), Modern Methods of
   Data Analysis. Sage.

II. Exploiting Metric Information in Models of Data

A. The Unidimensional Unfolding Model
   Coombs, Clyde H. (1950) “Psychological Scaling Without a Unit of Measurement.”
   Psychological Review 57: 145-158.
   Technique: Interval Values from Ordinal Data.” In H.M. Blalock (Editor) Measure-

B. The Summated Rating Model


Sijtsma, Klaas (2009) “On the Use, Misuse, and the Very Limited Usefulness of Cronbach’s Alpha.” *Psychometrika* 74: 107-120. Also see the responses and commentaries that follow this article.

### III. Preparation for Multidimensional Models

A. Vector Geometry and Linear Models
   - Wickens (1995), Chapters 1-5.
   - Lattin et al. (2003), pp. 19-32.

B. Singular Value Decomposition and the Basic Structure of a Matrix
   - Weller and Romney (1990), Chapter 2.
   - Lattin et al. (2003), pp. 32-36.
   - Borg and Groenen (2005), pp. 146-163.

### IV. Dimension Reduction and Summarizing Multivariate Data

A. The Biplot

B. Principal Components Analysis
   - Dunteman (1989), Chapters 1-6, 8.
   - Weller and Romney (1990), Chapter 3.
   - Bartholomew et al. (2008), Chapter 5.
   - Lattin et al. (2003), Chapter 4.
   - Borg and Groenen, pages 519-526.
V. Factor Analysis

A. The Common Factor Model
   Gorsuch (1983), Chapters 1-4.
   Mulaik (2010), Chapter 6.
   Bartholomew et al. (2008), pages 175-183.

B. Estimating the Factor Model
   Gorsuch (1983), Chapters 6 and 8.
   Mulaik (2010), Chapters 7 and 8.
   Bartholomew et al. (2008), pages 183-188.
   Lattin et al. (2003), pp. 131-153

C. Rotation
   Gorsuch (1983), Chapters 9-10.
   Mulaik (2010), Chapters 10-12.
   Bartholomew et al. (2008), pages 188-192.
   Lattin et al. (2003), pp. 153-156.

D. Constructing Factor Scales
   Gorsuch (1983), Chapter 12.
   Mulaik (2010), Chapter 13.
Bartholomew et al. (2008), pages 192-207.
Lattin et al. (2003), pp. 156-166.

E. Introduction to Confirmatory Factor Analysis (if time permits)
Mulaik (2010), Chapter 15.
Bartholomew et al. (2008), pp. 289-301.
Lattin et al. (2003), Chapter 6.

VI. Multidimensional Scaling

A. Overview and Spatial Distance Models
Borg and Groenen (2003), Chapters 1, 17-19.
Lattin et al. (2003), pp. 206-211.
Armstrong II et al. (2014), Chapter 1.

B. Classical Multidimensional Scaling (Metric and Nonmetric)
Bartholomew et al. (2008), Chapter 3.
Lattin et al. (2003), pp. 211-235.
Borg and Groenen (2005), Chapters 2, 3, 8, 9, 11-13.
Armstrong II et al. (2014), pp. 103-128.
C. Weighted Multidimensional Scaling
   Borg and Groenen (2005), Chapter 22.
   Armstrong II et al. (2014), pp. 132-143.

D. Data for Multidimensional Scaling Analyses
   Kruskal and Wish (1978), pp. 73-82.
   Borg and Groenen (2005), Chapter 6.

E. Statistical Inference and Multidimensional Scaling
   Ramsay, J. O. (1977) “Maximum Likelihood Estimation in Multidimensional Scal-
   ing.” *Psychometrika* 42: 241-266.
   Bakker, Ryan and Keith T. Poole. (2013) “Bayesian Metric Multidimensional Scal-
   ing.” *Political Analysis* 21: 125-140.
   Jacoby, William G. and David A. Armstrong II. (2013) “Bootstrap Confidence Re-
   gions for Multidimensional Scaling Solutions.” *American Journal of Political Sci-
   ence* 58: 264-278.
   Armstrong II et al. (2014), pp. 128-132.

VII. Multidimensional Unfolding and Preference Mapping
   Armstrong II et al. (2014), Chapter 5.
   Weller and Romney, PP. 44-54.
   Lattin et al. (2003), pp. 244-252.
   Borg and Groenen (2005), Chapters 14-16.
   Carroll, J.D. (1972) “Individual Differences and Multidimensional Scaling.” In Roger N.
VIII. Correspondence Analysis

Weller and Romney (1990), Chapters 5-8.
Bartholomew et al. (2008), Chapter 4.
Borg and Groenen (2005), Chapter 24.

IX. Some Final Considerations: Comparison of Scaling Strategies