2020 ESSEX SUMMER SCHOOL

3L Scaling Methods for Social Science:
Measuring Patterns and Preferences in Surveys and Behavior

Royce Carroll

August 10, 2020

Course Description

This course focuses on methods to **discover, understand and visualize latent patterns in data** and is especially suited to students with projects using survey data and other forms of relational data. The course introduces students to measurement theory and methods of scaling techniques, including **Multidimensional Scaling**,**Ideal Point Estimation**, and **Item Response Theory**. The first part of the course will provide an overview of the foundations of these techniques and introduce students to the most common methods for "spatial" analysis of survey and behavior data. The course will demonstrate how to interpret, measure and visualize latent dimensions of data via a variety of scaling methods using the open-source programming language R. The course will also discuss a range of applications these methods to social science studies of relational and perception data derived from elite behaviour and surveys. The course concludes with discussions of the most recent advances in the field, including applications for text analysis, and practical advice for those seeking to use such methods in social science research, relevant to the students enrolled.

The course first covers how to analyse data from scales found in surveys, focusing on surveys that ask respondents to place themselves and / or stimuli on issue or attribute scales. The course begins with approaches to scaling to generate bias-adjusted and latent spatial data from survey responses, such as the **Aldrich-McKelvey scaling** and **'Basic Space' scaling** with Anchoring Vignettes as methods for addressing perceptual bias in the form of "Differential Item Functioning." The course next examines similarities and dissimilarities data and covers **multidimensional scaling (MDS)** with a focus on the **SMACOF optimization method** implemented in R as well as Bayesian applications to Metric Multidimensional Scaling. Next, the course covers unfolding analysis of rating scale data from surveys such as favorability scales for stimuli such as politicians or social groups. Finally, the course provides an overview of IRT and ideal point estimation, generally focused on binary choice data, which includes those used in 'roll call voting' analysis of elite behavior in parliaments and courts. Here we will cover Poole and Rosenthal's W-NOMINATE and Poole's Optimal Classification unfolding method, as well as a variety of Bayesian analysis techniques for binary and ordinal choice data using Item Response Theory (IRT). An extensive range of Bayesian techniques are discussed, including **Bayesian Aldrich-McKelvey Scaling**, and **Ordinal and Dynamic Item Response Theory (IRT)**. The course will conclude with several advanced methods that extend the core themese.

This course will enable students to derive latent spatial preference information and/or dimensional structure from various types of survey and behavior data, which is applicable to a wide range of social science applications, academic and non-academic alike. Consumers of research based on these methods will also benefit from a deeper understanding of this type of methodology, its potential and its limitations.

Schedule

Lecture / Lab: 14:15 – 17:45 Monday 10th August to Friday 21st August 2020

Packages

You will need to lastest installation of *R* (preferably version 3.6 or above) and RStudio. You also need to install the `asmcjr` package and other relevant scaling packages along with this course.

Install `asmcjr` package in R environment

You can install using the `install_github()` function from the `devtools` package. `asmcjr` requires compilation, so Windows users will have to install Rtools first. For Mac users, you need to make sure you have already installed latest GNU Fortran(gfortran 8.2) and Xcode Developer Tools. Visit `asmcjr` for further instructions.

```
install.packages("devtools", dependencies=TRUE)
library(devtools)
devtools::install_github("uniofessex/asmcjr")
```

You should also install JAGS, which will be used in several contexts, along with the package `rjags`

Note: lecture slides will be added after each lecture, reading materials will be available beforehand

The following scaling packages and relevant toolkits are used by this course:

Lectures	Packages
01 - An Introduction / R	`basicspace`
02 - Issue Scale I	`asmcjr`, `smacof`
03 - Issue Scale II	`asmcjr`, `basicspace`
04 - Issue Scale III	`asmcjr`, `smacof`, `rgenoud`,
05 - Similarities Data and Agreement Score	`asmcjr`, `rgenoud`,`smacof`
06 - Binary Ideal Point Estimation (W-NOMINATE)	`asmcjr`, `wnominate`, `anominate`
07 - Non-parametric Ideal Point Estimation (OC and OOC)	`asmcjr`, `oc`,`ooc`, `MCMCpack`
08 - IRT and Bayesian MCMC	`asmcjr`, `MCMCpack`, `pscl`
09 - Dynamic and Ordinal IRT	`asmcjr`, `ooc`
10 - Scaling Text and EM IRT	`asmcjr`, `emIRT`

Lecture Notes (slides will be added after each lecture)

(Note that if you are taking the exam, it will cover material from days 1-8)

01 - An Introduction / R

- Core Reading: CH 1 of Armstrong
- Recommended Reading: Furr and Bacharach Ch. 4

Slide: html | pdf

02 - Issue Scale I

- Core Reading: 2.1 and 6.1 of Armstrong

2.1 Aldrich-McKelvey Scaling – 2.1.4 Estimating Bootstrapped Standard Errors for Aldrich-McKelvey Scaling – 6 Bayesian Scaling Models (JAGS Intro) – 6.1 Bayesian Aldrich-McKelvey Scaling

- Recommended Reading: Zakharova and Warwick 2014 | Carroll and Kubo 2017 | Saiegh 2015

Slide: html | pdf

03 - Issue Scale II

- Core Reading: 2.2 and 2.3 of Armstrong

2.2 Basic Space Scaling: The blackbox Function – 2.3 Basic Space Scaling: The blackbox transpose Function – 2.3.1 Example 1: 2000 and 2006 Comparative Study of Electoral Systems (Mexican Modules) – 2.3.2 Estimating Bootstrapped Standard Errors for Black Box Transpose Scaling – 2.3.3 Using the blackbox transpose Function on Datasets with Large Numbers of Respondents

- Recommended Reading: Bakker et al. 2014 | Bakker, Jolly and Polk 2020

Slide: html | pdf

04 - Issue Scale III

- Core Reading: 2.2 of Armstrong 2.2 Basic Space Scaling: The blackbox Function – 2.2.1 Example 1: 2000 Convention Delegate Study – 2.2.2 Example 2: 2010 Swedish Parliamentary Candidate Survey – 2.2.3 Estimating Bootstrapped Standard Errors for Black Box Scaling
- Recommended Reading: **Slide:** html | pdf

Gross and Sigelman 1984 | Bornschier 2010

05 - Similarities and Rating Scale Data

- Core Reading: Chapters 3 and 4 of Armstrong

Slide: html | pdf

06 - Binary Ideal Point Estimation (W-NOMINATE)

- Core Reading: CH 5 of Armstrong
- Recommended Reading: Hix, Nouryand and Roland 2009

Slide: html | pdf

07 - Non-parametric Ideal Point Estimation (OC and OOC)

- Core Reading: CH 6 of Armstrong
- Recommended Reading: Rosenthal and Voeten 2004 | Hare, Liu ans Lupton 2018

Slide: html | pdf

08 - IRT and Bayesian MCMC

- Core Reading: CH 6 of Armstrong | Furr and Bacharach Ch. 13 'Item Response Theory' and Rasch Models
- Recommended Reading: Clinton and Jackman and Rivers 2004 | Armstrong and Lucas 2020

Slide: html | pdf

09 - Dynamic and Ordinal IRT

- Core Reading: CH 6 of Armstrong
- Recommended Reading: Martin and Quinn 2004

Slide: html | pdf

10 - Scaling Text and EM IRT

- Core Reading: Imai, Lo, and Olmsted 2016
- Recommended Reading: Barberá 2015 | Slapin and Proksch 2008

Slide: html | pdf

Books

Required Book

- Analyzing Spatial Models of Choice and Judgment, 2nd Edition 2020, David A. Armstrong II, Ryan Bakker, Royce Carroll, Christopher Hare, Keith T. Poole, Howard Rosenthal, CRC Press Available in PDF (please do not distribute PDF of book)