Identifying the Scale Origin in Comparative Judgment Data

Ulf Böckenholt

Email: ulf.bockenholt@mcgill.ca

Outline

- Comparative Judgments
    1. Paired Comparisons
    2. Rankings (partial, incomplete)
- **Study of individual differences**
- Loss of scale origin (same difference but different absolute levels) and interpretational problems
- Solutions and Applications:
    1. Graphical Representation
    2. Augmenting comparative judgments

Comparative Judgments

- Stimuli ($j = 1, \ldots, J$); judges ($i = 1, \ldots n$); covariates x_{ij}
- Tasks
  - Select the preferred stimulus out of the set of stimuli; rank the presented stimuli; select the best and worst.
  - Multiple (overlapping) stimulus sets
- Multiple responses by each judge
  - Boredom, carry–over effects
  - Study of individual differences

Judgment Process:

- Evaluation of stimulus j by judge i:

$$y_{ij} = \mu_{ij} + \epsilon_{ij}$$

- Pairwise judgments between stimuli j and k:

$$y_{i(jk)} = y_{ij} - y_{ik}$$

- Response:
  - Select j when $y_{i(jk)} > 0$ and, otherwise, select k
  - Rankings are multiple (simultaneous) difference judgments
  - Example: $j > k > h$ when $y_{i(jk)} > 0$ and $y_{i(kh)} > 0$

Individual Differences

- Individual scale values: $\mu_{i1}, \mu_{i2}, \ldots, \mu_{iJ}$
  - Group means of scale values: $\mu_1, \mu_2, \ldots, \mu_J$
  - Covariance matrix: $\mathbf{\Omega}$
- Variability within judge: Level-1
- Variability between judges: Level-2
- Easy estimation: *HLM 5, Mixor, Mplus, MLwiN* but difficult interpretation: Scale origin is lost.

Within– and Between–Judge Effects

- Comparative judgment of stimulus pair (jk) by person i:

$$y_{ijk} = \mu_{ij} - \mu_{ik} + \epsilon_{ijk},$$

- $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi} = \text{Diag}(\psi_1, \psi_2, \ldots, \psi_{\binom{r}{2}})$
- $\boldsymbol{\mu}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$
- Special cases for $\boldsymbol{\Psi}$ and $\boldsymbol{\Omega}$
- Comparative judgments can be written as a linear model

- **Level 1:** $y_{ijk} = \mu_{ij} - \mu_{ik} + \epsilon_{ijk}$

$$\begin{pmatrix} y_{ijk} \\ y_{ijl} \\ y_{ijm} \\ y_{ikl} \\ y_{ikm} \\ y_{ilm} \end{pmatrix} = \begin{pmatrix} \mu_{ij} - \mu_{ik} \\ \mu_{ij} - \mu_{il} \\ \mu_{ij} - \mu_{im} \\ \mu_{ik} - \mu_{il} \\ \mu_{ik} - \mu_{im} \\ \mu_{il} - \mu_{im} \end{pmatrix} + \begin{pmatrix} \epsilon_{ijk} \\ \epsilon_{ijl} \\ \epsilon_{ijm} \\ \epsilon_{ikl} \\ \epsilon_{ikm} \\ \epsilon_{ilm} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mu_{ij} \\ \mu_{ik} \\ \mu_{il} \\ \mu_{im} \end{pmatrix} + \begin{pmatrix} \epsilon_{ijk} \\ \epsilon_{ijl} \\ \epsilon_{ijm} \\ \epsilon_{ikl} \\ \epsilon_{ikm} \\ \epsilon_{ilm} \end{pmatrix}$$

$$= \mathbf{A}\, \boldsymbol{\mu}_i + \boldsymbol{\epsilon}_i \qquad ,$$

Between–Judge Effects—Level 2

1.
$$\boldsymbol{\mu}_i = \boldsymbol{\mu} + \boldsymbol{\nu}_i$$

2.
$$\boldsymbol{\mu}_i = \boldsymbol{\beta}\mathbf{X}_i + \boldsymbol{\nu}_i,$$
where $\mathbf{X}_i$ contains covariates about judges.

3.
$$\boldsymbol{\mu}_i = \mathbf{Z}\boldsymbol{\zeta}_i + \boldsymbol{\nu}_i,$$
where $\mathbf{Z}$ contains covariates about stimuli.

Interpretation: It is all relative

- $(\mu_{ij} - c_i) - (\mu_{ik} - c_i) = \mu_{ij} - \mu_{ik}$.
- Set one of the scale values to 0. As a result,
  - only $(J-1)$ scale values,

$$\upsilon_{ij} = \mu_{ij} - \mu_{iJ},$$

  - and the corresponding $(J-1) \times (J-1)$ covariance matrix $\mathbf{\Sigma}$ can be estimated,

$$\sigma_{jj} = \omega_{jj} + \omega_{JJ} - 2\,\omega_{jJ}.$$

- Identification of covariate effects has the same limitations.

- It is not possible to infer $\boldsymbol{\mu}_i$ from $\boldsymbol{v}_i$:

$$\boldsymbol{\mu}_i = (4, 2, 1); \;\; \boldsymbol{v}_i = (3, 1).$$

- It is not possible to infer $\boldsymbol{\Omega}$ from $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Omega} = \begin{pmatrix} .75 & -.125 & 0 \\ -.125 & 1 & .125 \\ 0 & .125 & 1.25 \end{pmatrix}, \; \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

- Two problems:

  1. How do we interpret $\boldsymbol{\Sigma}$?
  2. Can we recover the scale origin and estimate both $\boldsymbol{\mu}_i$ and $\boldsymbol{\Omega}$?

1. Interpretation of Σ

- Interpret distances between stimulus j and k

$$\delta_{jk} = \omega_{jj} + \omega_{kk} - 2\,\omega_{jk} = \sigma_{jj} + \sigma_{kk} - 2\,\sigma_{jk} \qquad (1)$$

- Distance matrix can be analyzed with multidimensional scaling software

Four Examples

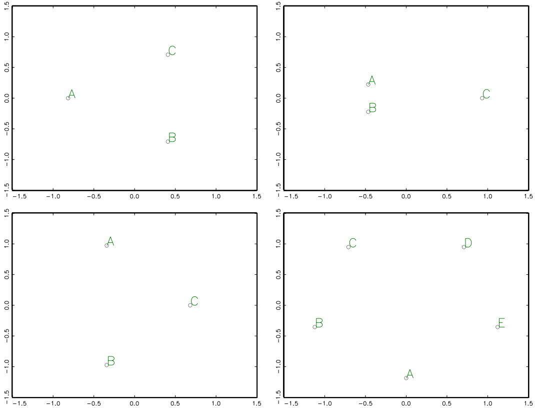
$$\mathbf{\Omega}_{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \mathbf{\Omega}_{(2)} = \begin{bmatrix} 1 & .9 & 0 \\ .9 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \mathbf{\Omega}_{(3)} = \begin{bmatrix} 1 & -.9 & 0 \\ -.9 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Inter-stimulus distance matrices:

$$\mathbf{\Delta}_{(1)} = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}; \mathbf{\Delta}_{(2)} = \begin{bmatrix} 0 & .2 & 2 \\ .2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}; \mathbf{\Delta}_{(3)} = \begin{bmatrix} 0 & 3.8 & 2 \\ 3.8 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$$

$\mathbf{\Omega}_{(4)} = \mathbf{\Lambda}\mathbf{\Lambda}' + \sigma^2\mathbf{I}$, where $\sigma^2 = 1$ and $\mathbf{\Lambda}$ is specified as

$$\mathbf{\Lambda} = \begin{bmatrix} 0.00 & 1.00 \\ 0.95 & 0.30 \\ 0.60 & -0.80 \\ -0.60 & -0.80 \\ -0.95 & 0.30 \end{bmatrix}.$$



A
B
C
A
B
C
A
B
C
A
B
C
D
E

Application: Preferences for beer brands

- $n=$ 241, complete paired comparison of 8 beer brands
- *Old Milwaukee, Meister Brau, Stroh's, Old Milwaukee Light, Budweiser, Heineken, Miller Light, Coors Light*
- Light–heavy and budget–premium beverages.
- Analysis with HLM 5, multivariate Laplace approximation to estimate scale values $\boldsymbol{v}$ and their covariance matrix $\boldsymbol{\Sigma}$.

Table 1: Estimated covariance matrix $\hat{\Sigma}$

	OM–CL	MB–CL	S–CL	OML–CL	B–CL	H–CL	ML–CL
OM–CL	34.6						
MB–CL	27.6	27.8					
S–CL	20.4	18.0	16.0				
OML–CL	17.8	15.8	12.0	15.7			
B–CL	16.8	16.5	11.8	6.6	17.7		
H–CL	1.9	3.1	3.4	-2.1	9.4	13.2	
ML–CL	7.2	7.2	5.1	5.8	3.5	0.9	5.3

Table 2: Distance matrices
(Upper triangular matrix: 2d–approximation.)

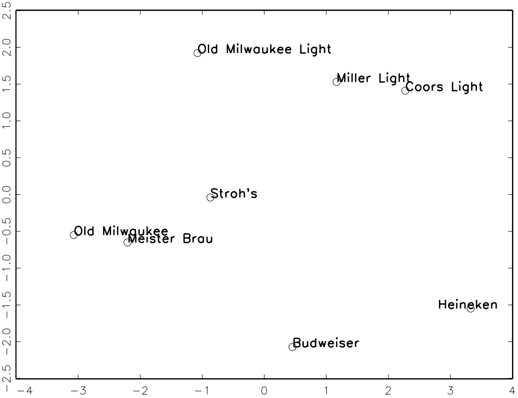
	OM	MB	S	OML	B	H	ML	CL	$\hat{\mu}$
OM		0.8	5.1	10.1	14.7	41.9	22.3	32.4	-2.37
MB	7.3		2.2	7.9	9.1	31.4	16.1	24.3	-1.47
S	9.8	7.8		3.9	5.9	19.9	6.6	12.0	.37
OML	14.7	11.9	7.7		18.3	31.4	5.2	11.5	-3.54
B	18.8	12.4	10.2	20.2		8.5	13.5	15.4	4.85
H	44.1	34.7	22.5	33.0	12.1		14.1	9.9	-.65
ML	25.6	18.8	11.2	9.5	16.1	16.7		1.2	1.26
CL	34.6	17.8	16.0	15.7	17.7	13.2	5.3		1.55

OM = Old Milwaukee, MB = Meister Brau, S = Stroh's,

OML = Old Milwaukee Light, B = Budweiser, H = Heineken,

ML = Miller Light, CL = Coors Light

- Two major dimensions which account for 55% and 25% of the total variance
- First factor contrasts light and heavy beer brands (*Budweiser* and *Heineken* versus *Coors Light* and *Old Milwaukee Light*).
- Second factor contrast less expensive brands (*Old Milwaukee*, *Meister Brau*), with more expensive ones (*Heineken*).



Old Milwaukee Light
Miller Light
Coors Light
Stroh's
Old Milwaukee
Meister Brau
Heineken
Budweiser
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−4
−3
−2
−1
0
1
2
3
4

Recovery of scale origin: Estimating μ_i and Ω

- Extension of basic comparative judgment task
- Identification of scale origin is a testable hypothesis
- Some examples:
  1. Ratings of single items: Absolute judgments are equal to comparative judgments
  2. A priori zero point: Risky evaluations are equal to non–risky evaluations
  3. Stimulus–bundle comparisons: Integration rule of separate stimulus evaluations is known (e.g., additive)

1. Combining Absolute and Relative Judgments

- Relative Judgment: $y_{ijk} = \mu_{ij} - \mu_{ik} + \epsilon_{ijk}$
- Absolute Judgment: $y_{ij} = \eta_{ij} + \epsilon_{ij}$
- Hypothesis: $\mu_{ij} = \eta_{ij}$

$$\begin{pmatrix} y_{ijk} \\ y_{ijl} \\ y_{ijm} \\ y_{ikl} \\ y_{ikm} \\ y_{ilm} \\ y_{ij} \\ y_{ik} \\ y_{il} \\ y_{im} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{ij} \\ \mu_{ik} \\ \mu_{il} \\ \mu_{im} \end{pmatrix} + \begin{pmatrix} \epsilon_{ijk} \\ \epsilon_{ijl} \\ \epsilon_{ijm} \\ \epsilon_{ikl} \\ \epsilon_{ikm} \\ \epsilon_{ilm} \\ \epsilon_{ij} \\ \epsilon_{ik} \\ \epsilon_{il} \\ \epsilon_{im} \end{pmatrix}$$

- Only one binary judgment is needed

- Nested models for testing equality of absolute and comparative judgments

$$\mathbf{A}^{(R)}\boldsymbol{\mu}_i = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_{ij} \\ \mu_{ik} \\ \mu_{il} \\ \mu_{im} \end{pmatrix}$$

$$\mathbf{A}_i^{(F)}\boldsymbol{\mu}_i^{(F)} = \begin{pmatrix} 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \eta_{ij} \\ \eta_{ik} \\ \mu_{ij} - \mu_{im} \\ \mu_{ik} - \mu_{im} \\ \mu_{il} - \mu_{im} \end{pmatrix}$$

- Tests for equality of mean and covariance structures
- Minimum of two absolute judgments to conduct tests

2.Specifying the scale origin a priori

- Choice between "doing nothing" or following a specific course of action. Status-quo option is a natural zero point
  - Two lotteries: (j, p_s) and (k, p_u) [win j with prob. p_s and "nothing" with prob. $(1 - p_s)$]
  - "winning nothing" defines the origin of the individual item utility scales (μ_{i0})
- Potential fields of applications include choices among (medical) treatment plans, insurance policies, and options with future outcomes.

Inferring the scale origin from judgments of stimulus bundles

- Multiple items presented as a set or bundle, e.g.,
  - Consumer goods (keyboard and mouse bundle)
  - Meal package
- Scale origin can be determined if the overall evaluation of an item bundle can be predicted by an additive combination of the separate stimuli that constitute the bundle

- $y_{il(jk)} = \mu_{il} - \mu_{i(jk)} + \epsilon_{il(jk)}$

$$\begin{pmatrix} y_{ijk} \\ y_{ijl} \\ y_{ikl} \\ y_{ij(kl)} \\ y_{ik(jl)} \\ y_{il(jk)} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{ij} \\ \mu_{ik} \\ \mu_{il} \\ \mu_{i(kl)} \\ \mu_{i(jl)} \\ \mu_{i(jk)} \end{pmatrix} + \begin{pmatrix} \epsilon_{ijk} \\ \epsilon_{ijl} \\ \epsilon_{ikl} \\ \epsilon_{ij(kl} \\ \epsilon_{ik(jl} \\ \epsilon_{il(jk} \end{pmatrix}$$

- $\mu_{i(jk)} = \mu_{ij} + \mu_{ik}$

$$\begin{pmatrix} y_{ijk} \\ y_{ijl} \\ y_{ikl} \\ y_{ij(kl)} \\ y_{ik(jl)} \\ y_{il(jk)} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \mu_{ij} \\ \mu_{ik} \\ \mu_{il} \end{pmatrix} + \begin{pmatrix} \epsilon_{ijk} \\ \epsilon_{ijl} \\ \epsilon_{ikl} \\ \epsilon_{ij(kl)} \\ \epsilon_{ik(jl)} \\ \epsilon_{il(jk)} \end{pmatrix}$$

- Nested models for testing stimulus combination rule

Some Remarks on the Identifying–of–Scale–Origin Approaches

- Approaches can be combined in various ways.
- But – underlying assumptions may not be satisfied (e.g., absolute and relative judgments may focus on different stimulus features; stimulus features may be evaluated differently depending on whether the choice mode is risky or non-risky; overall evaluation of a stimulus bundle may not be an additive function of the separate stimulus evaluations.)
- Statistical tests of assumptions are straightforward.

Discussion

- Between– and within–judge effects are easy to analyze with current software packages but interpretation is difficult.
- Two Solutions:
    1. Distance interpretation and multidimensional–scaling display
    2. Identification of scale origin by augmenting comparative judgment data.
- For studying individual preference differences, it is beneficial to combine absolute and comparative judgments and/or to enhance the comparative judgment task.