Methodology & Statistics (Psychology)

Tom F. Wilderjans

Universiteit Leiden
The Netherlands

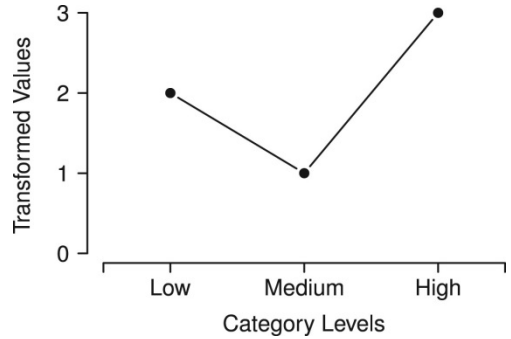
Discover the world at Leiden University

FSW – Methodology & Statistics

Develops data analysis techniques for behavioural, social, and health sciences

- Specific Challenges
  - Assumptions are often not tenable
  - Data are often categorical
  - Samples are heterogeneous and the heterogeneity is of specific interest

- Solutions
  - Resampling techniques for stability and model selection
  - Optimal scaling
  - Develop specific methods for individual differences



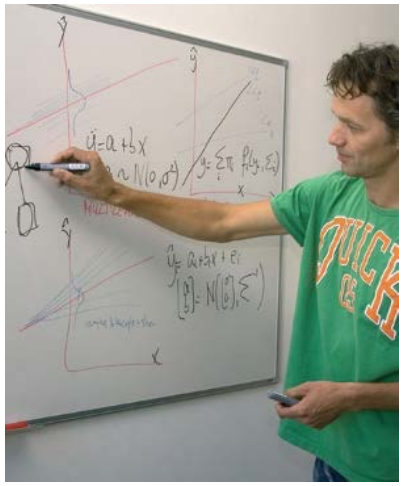
FSW – Methodology & Statistics

New data poses new challenges

- `New' data formats:
  - fMRI
  - Eye tracking
  - Wearables
  - Omics data, genetic data
  - And combinations of these with traditional data

- Promises
  - Personalized treatment
  - Personalized health
  - Personalized learning

FSW – Methodology & Statistics



How

- Think and Derive
  - Mathematical derivations
  - Developing Algorithms

- Simulate and Analyze
  - Generate data with known properties
  - Analyze how the algorithms behave

- Apply and Compare
  - Application of new techniqes to empirical data sets
  - Comparison of new techniques with standard

Data: $\mathbf{G}, p, \lambda, \alpha, \epsilon$
Result: $\boldsymbol{\mu}, \mathbf{U}, \mathbf{D}, \mathbf{V}$
$t = 0$;
Compute $\mathbf{W}$ from missing values in $\mathbf{G}$;
Initialize $\boldsymbol{\mu}$: $\boldsymbol{\mu}^{(0)} = n^{-1}\mathbf{J}_c\mathbf{G}'\mathbf{1}$;
Compute the SVD of $\mathbf{JGJ}_c$: $\mathbf{JGJ}_c = \mathbf{P\Phi Q}'$;
Initialize $\mathbf{U}$: $\mathbf{U}^{(0)} = \mathbf{P}$;
Initialize $\mathbf{V}$: $\mathbf{V}^{(0)} = \mathbf{Q}$;
Initialize $\mathbf{D}$: $d_{ss}^{(0)} = \max(0, \phi_{ss} - \lambda)$;
Compute $\mathbf{\Pi}$ by (4) ;
Compute $L^{(0)} = L(\boldsymbol{\mu}^{(0)}, \mathbf{U}^{(0)}, \mathbf{D}^{(0)}, \mathbf{V}^{(0)})$;
while $t = 0$ *or* $(L^{(t)} - L^{(t-1)})/L^{(t)} \geq \epsilon$ **do**
  $t = t + 1$;
  $\mathbf{Z} = [(\mathbf{1}\boldsymbol{\mu}^{(t-1)'} + \mathbf{U}^{(t-1)}\mathbf{D}^{(t-1)}\mathbf{V}^{(t-1)'}) + 2(\mathbf{G} - \mathbf{W} \odot \mathbf{\Pi})]\mathbf{J}_c$;
  Compute update $\boldsymbol{\mu}$: $\boldsymbol{\mu}^{(t)} = n^{-1}\mathbf{Z}'\mathbf{1}$;
  Compute the SVD of $\mathbf{JZ}$: $\mathbf{JZ} = \mathbf{P\Phi Q}'$;
  Update $\mathbf{U}$: $\mathbf{U}^{(t)} = \mathbf{P}$;
  Update $\mathbf{V}$: $\mathbf{V}^{(t)} = \mathbf{Q}$;
  Update $\mathbf{D}$: $d_{ss} = (1 + \max(0, \phi_{ss} - \lambda)$;
  Compute $\mathbf{\Pi}$ by (4) ;
  Compute $L^{(t)} = L(\boldsymbol{\mu}^{(t)}, \mathbf{U}^{(t)}, \mathbf{D}^{(t)}, \mathbf{V}^{(t)})$;
end

Algorithm 1: The majorizing algorithm for MMCA. ϵ is here a small positive value, for example, $\epsilon = 10^{-8}$.