The family of hierarchical classes models:

A state-of-the-art overview

Iven Van Mechelen, Eva Ceulemans, Iwin Leenen,

and Jan Schepers

University of Leuven

Iven.VanMechelen@psy.kuleuven.be

Overview of the talk

1. data
2. models

    2.1 context

    2.2 hierarchical classes models

        2.2.1 basic model

        2.2.2 justification of the Max operator

        2.2.3 the HICLAS family

3. research topics

    3.1 models

    3.2 estimation

    3.3 model selection and model checking

    3.4 quantification of uncertainty

4. references

Overview of the talk

1. data
2. models

   2.1 context

   2.2 hierarchical classes models

      2.2.1 basic model

      2.2.2 justification of the Max operator

      2.2.3 the HICLAS family
3. research topics

   3.1 models

   3.2 estimation

   3.3 model selection and model checking

   3.4 quantification of uncertainty
4. references

1. Data

- $I_1 \times I_2 \times \ldots \times I_N$ *N*-way *N*-mode data array $\underline{\mathbf{X}}$

  with entries $x_{i_1 i_2 \ldots i_N}$

- binary, rating-valued, real-valued

- array-conditionality

Note:

We leave aside:

(1) structurally incomplete data

(2) multiblock data

Overview of the talk

2. Models

2.1 Context

- *N*-way Tucker model:

$$x_{i_1 i_2 \ldots i_N} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \cdots \sum_{p_N=1}^{P_N} \left[\left(\prod_{n=1}^{N} a^n_{i_n p_n} \right) g_{p_1 p_2 \ldots p_N} \right] + e_{i_1 i_2 \ldots i_N}$$

with $\mathbf{A}^n$ denoting the $I_n \times P_n$ n^{th} component matrix

and $\underline{\mathbf{G}}$ denoting the core array

$$x_{i_1 i_2 \ldots i_N} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \cdots \sum_{p_N=1}^{P_N} \left[\left(\prod_{n=1}^{N} a^n_{i_n p_n} \right) g_{p_1 p_2 \ldots p_N} \right] + e_{i_1 i_2 \ldots i_N}$$

- options:

1. $\mathbf{A}^n$ identity matrix for 1, 2, …, N-1 component matrices

   $\rightarrow$ N-way $\begin{cases} \text{Tucker } N\text{-1} \\ \text{Tucker } N\text{-2} \\ \ldots \\ \text{Tucker 1} \end{cases}$ model

2. $\underline{\mathbf{G}}$ superdiagonal or superidentity

   $\rightarrow$ N-way CANDECOMP/PARAFAC

- models studied in research group:

  (1) primary constraint:

  $\mathbf{A}^n$ binary for 1, 2, …, N component matrices

  $\rightarrow$ overlapping clustering for n^{th} mode

  (2) operator: Σ or Max (Min)

$$x_{i_1 i_2 \ldots i_N} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \cdots \sum_{p_N=1}^{P_N} \left[\left(\prod_{n=1}^{N} a^n_{i_n p_n} \right) g_{p_1 p_2 \ldots p_N} \right] + e_{i_1 i_2 \ldots i_N}$$

$$x_{i_1 i_2 \ldots i_N} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \ldots \operatorname*{Max}_{p_N=1}^{P_N} \left[\left(\prod_{n=1}^{N} a^n_{i_n p_n} \right) g_{p_1 p_2 \ldots p_N} \right] + e_{i_1 i_2 \ldots i_N}$$

- unifying model: Van Mechelen & Schepers (submitted)

- focus of present talk:

(1) $\mathbf{A}^n$ binary for all component matrices

$\rightarrow$ overlapping clustering for all modes

(2) operator: Max

(3) $\underline{\mathbf{G}}$ takes values in $\mathbb{R}^+$

Overview of the talk

1. data
2. models

    2.1 context

    2.2 hierarchical classes models

        2.2.1 basic model

        2.2.2 justification of the Max operator

        2.2.3 the HICLAS family

3. research topics

    3.1 models

    3.2 estimation

    3.3 model selection and model checking

    3.4 quantification of uncertainty

4. references

2. Models

2.2 Hierarchical classes models

2.2.1 Basic model

- two-way two-mode

$$x_{i_1 i_2} = \underbrace{\operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \, a^2_{i_2 p_2} \, g_{p_1 p_2} \right)}_{\text{reconstructed data } \hat{x}_{i_1 i_2}} + e_{i_1 i_2}$$

$$\hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

- three cases:

1. **G** binary (0/1) $\Rightarrow$ **G** is identity matrix

$$\hat{x}_{i_1 i_2} = \operatorname*{Max}_{p=1}^{P} \left(a^1_{i_1 p} \; a^2_{i_2 p} \right)$$

$$\hat{x}_{i_1 i_2} = \bigoplus_{p=1}^{P} \left(a^1_{i_1 p} \; a^2_{i_2 p} \right) \quad \text{(Boolean sum)}$$

2. **G** rating-valued (i.e., values in {1, 2, …, V})

3. **G** real-valued (i.e., values in $\mathbb{R}^+$)

2. Models

2.2 Hierarchical classes models

2.2.2 Justification of the Max operator

Three (interrelated reasons):

1. representation by component matrices of quasi-orders that can be naturally defined on each of the modes
2. substantive interpretation of decomposition rule
3. decision rule for overlapping biclusters (triclusters etc.)

2. Models

2.2 Hierarchical classes models

2.2.2 Justification of the Max operator

Three (interrelated reasons):

1. representation by component matrices of quasi-orders that can be naturally defined on each of the modes
2. substantive interpretation of decomposition rule
3. decision rule for overlapping biclusters (triclusters etc.)

- binary case (**G** is 0/1)

illustrative example: 7 $\times$ 6 feature by object reconstructed data

$\hat{\mathbf{X}} =$

	O_1	O_2	O_3	O_4	O_5	O_6
F_1	1	0	0	0	1	0
F_2	0	1	1	0	1	1
F_3	0	0	0	1	0	1
F_4	0	0	0	1	0	1
F_5	1	1	1	0	1	1
F_6	1	1	1	0	1	1
F_7	1	1	1	1	1	1

implication relation $\preccurlyeq$ among features

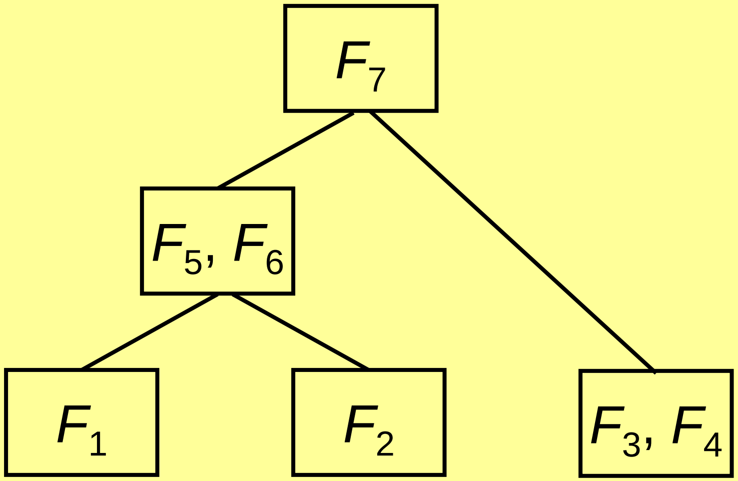
e.g., $F_1 \preccurlyeq F_5$, $F_3 \preccurlyeq F_4$, $F_5 \preccurlyeq F_7$

- similar implication relation among objects

- quasi-order: reflexive, transitive

  → implies partial order on resulting equivalence classes

- graphical representation: (quasi-) Hasse diagram



→ HIERARCHICAL CLASSES !!!

- representation of quasi-orders by component matrices:

quasi-order on $\{F_1, \ldots, F_7\}$ as implied by $\hat{\mathbf{X}}$ $=$ quasi-order on $\{F_1, \ldots, F_7\}$ as implied by $\mathbf{A}^1$

$\hat{\mathbf{X}} =$

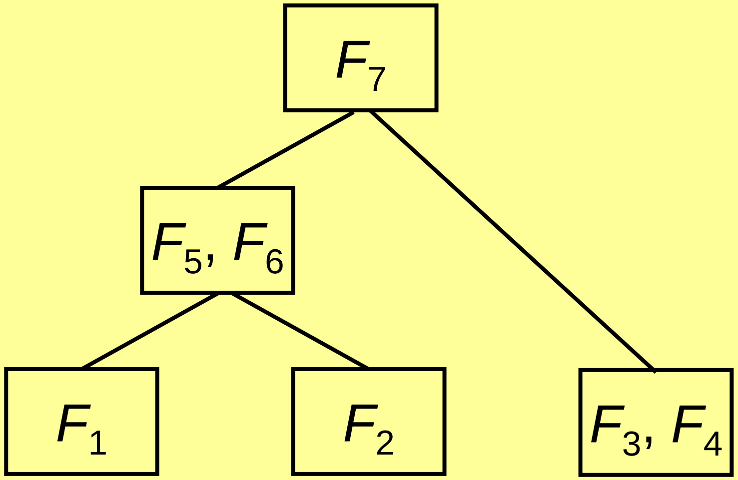
	O_1	O_2	O_3	O_4	O_5	O_6
F_1	1	0	0	0	1	0
F_2	0	1	1	0	1	1
F_3	0	0	0	1	0	1
F_4	0	0	0	1	0	1
F_5	1	1	1	0	1	1
F_6	1	1	1	0	1	1
F_7	1	1	1	1	1	1

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$	$A^1_{\bullet 3}$
F_1	1	0	0
F_2	0	1	0
F_3	0	0	1
F_4	0	0	1
F_5	1	1	0
F_6	1	1	0
F_7	1	1	1

$= \mathbf{A}^1$

- representation of quasi-orders by component matrices:

quasi-order on $\{F_1, \ldots, F_7\}$ as implied by $\hat{\mathbf{X}}$ $=$ quasi order-on $\{F_1, \ldots, F_7\}$ as implied by $\mathbf{A}^1$



	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$	$A^1_{\bullet 3}$
F_1	1	0	0
F_2	0	1	0
F_3	0	0	1
F_4	0	0	1
F_5	1	1	0
F_6	1	1	0
F_7	1	1	1

$= \mathbf{A}^1$

- representation of quasi-orders by component matrices:

quasi-order on $\{O_1, \ldots, O_6\}$ as implied by $\hat{\mathbf{X}}$ $=$ quasi-order on $\{O_1, \ldots, O_6\}$ as implied by $\mathbf{A}^2$

$\hat{\mathbf{X}} =$

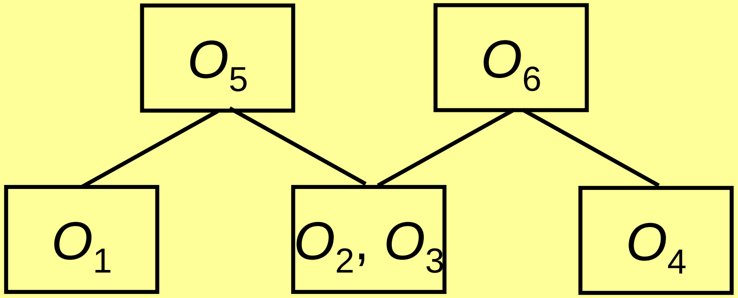
	O_1	O_2	O_3	O_4	O_5	O_6
F_1	1	0	0	0	1	0
F_2	0	1	1	0	1	1
F_3	0	0	0	1	0	1
F_4	0	0	0	1	0	1
F_5	1	1	1	0	1	1
F_6	1	1	1	0	1	1
F_7	1	1	1	1	1	1

	$A^2_{\bullet 1}$	$A^2_{\bullet 2}$	$A^3_{\bullet 3}$
O_1	1	0	0
O_2	0	1	0
O_3	0	1	0
O_4	0	0	1
O_5	1	1	0
O_6	0	1	1

$= \mathbf{A}^2$

- representation of quasi-orders by component matrices:

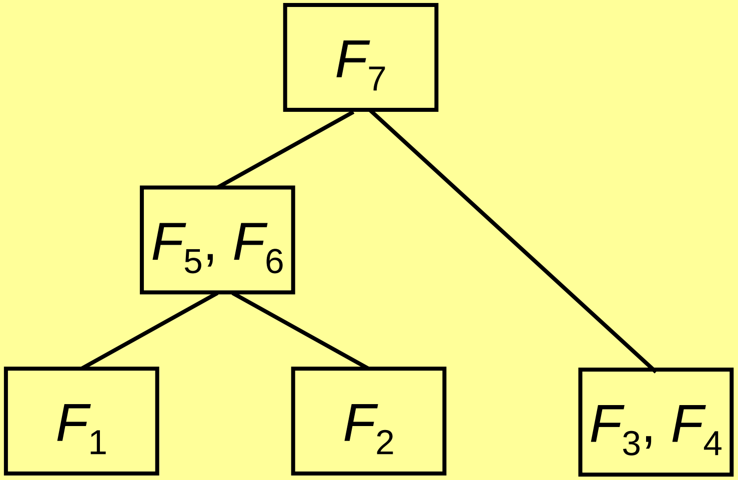
quasi-order on $\{O_1, \ldots, O_6\}$ as implied by $\hat{\mathbf{X}}$ $=$ quasi-order on $\{O_1, \ldots, O_6\}$ as implied by $\mathbf{A}^2$



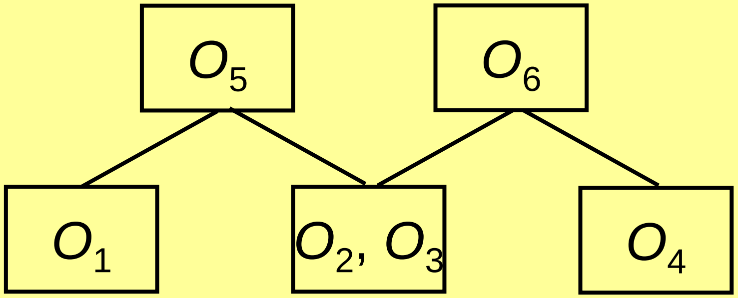
	$A^2_{\bullet 1}$	$A^2_{\bullet 2}$	$A^3_{\bullet 3}$
O_1	1	0	0
O_2	0	1	0
O_3	0	1	0
O_4	0	0	1
O_5	1	1	0
O_6	0	1	1

$= \mathbf{A}^2$

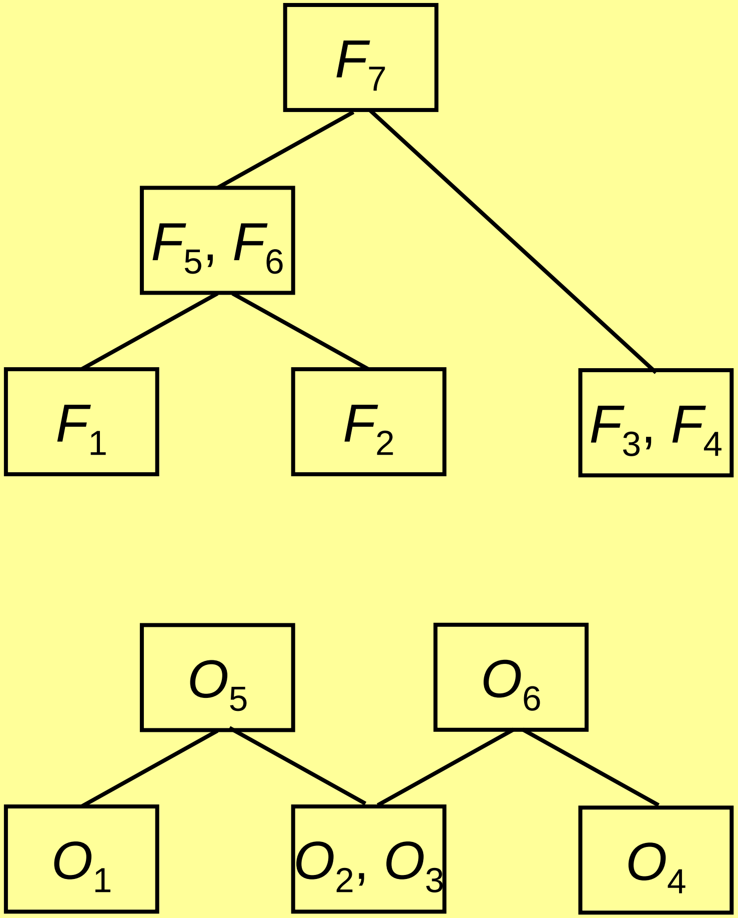
- Note: (quasi) Hasse diagram as implied by $\mathbf{A}^1$



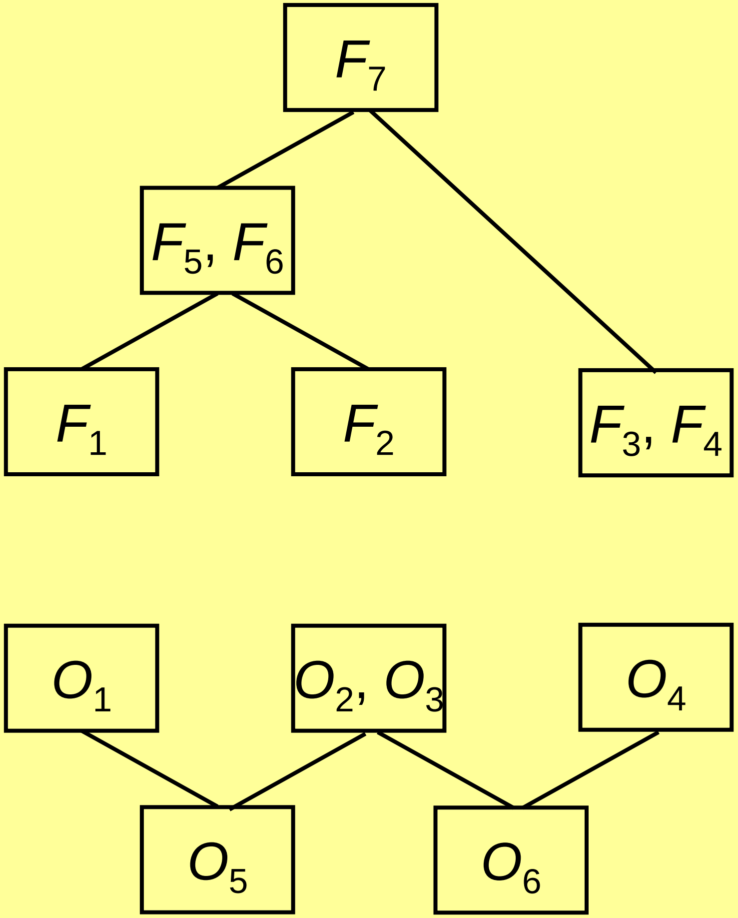
- Note: (quasi) Hasse diagram as implied by $\mathbf{A}^2$



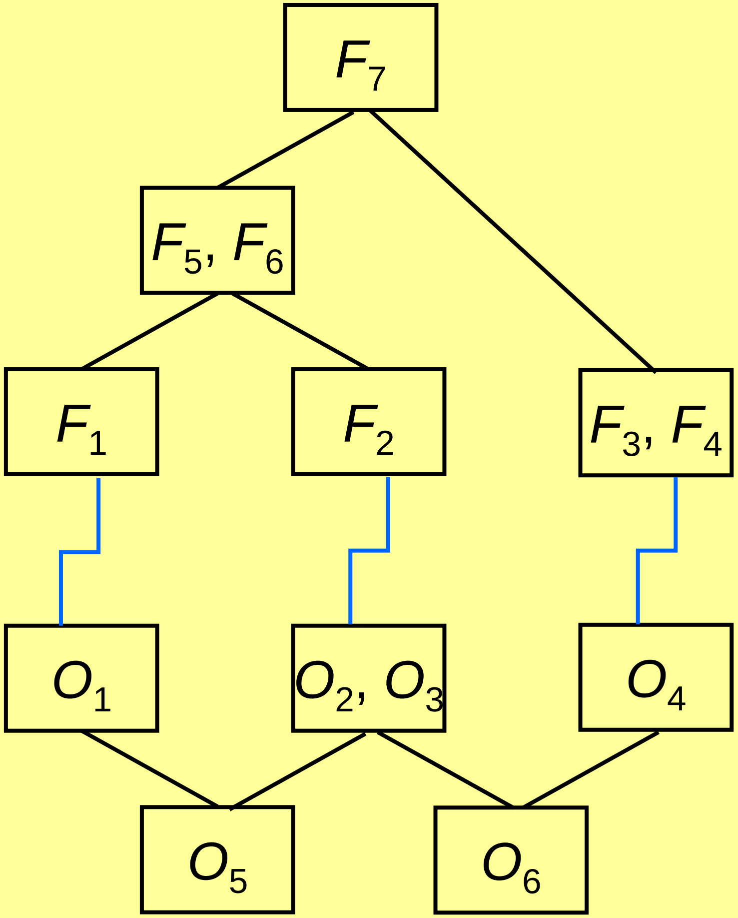
- Note: (quasi) Hasse diagrams as implied by $\mathbf{A}^1$ and $\mathbf{A}^2$



- Note:



- Note: comprehensive graphical representation of HICLAS model !!!



- rating- or positively real-valued case

illustrative example: 7 $\times$ 6 feature by object reconstructed data

$\hat{\mathbf{X}} =$

	O_1	O_2	O_3	O_4	O_5	O_6
F_1	4	1	1	0	4	1
F_2	0	5	5	0	5	5
F_3	0	0	0	2	0	2
F_4	0	0	0	2	0	2
F_5	4	5	5	0	5	5
F_6	4	5	5	0	5	5
F_7	4	5	5	2	5	5

generalized implication relation $\preccurlyeq$ among features

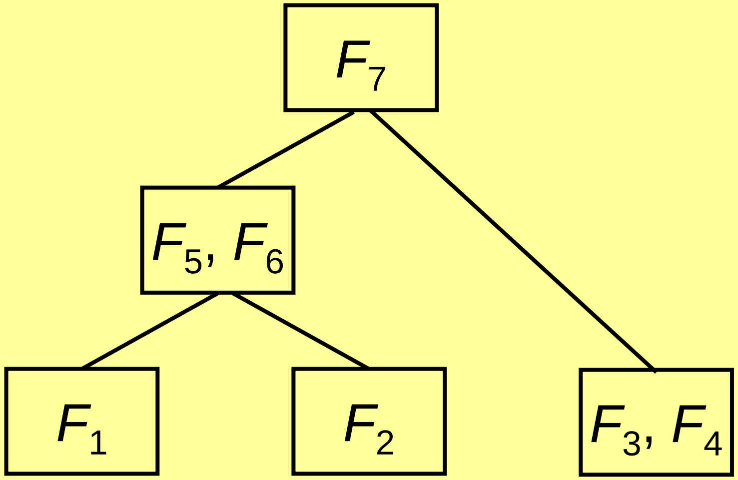
e.g., $F_1 \preccurlyeq F_5$, $F_3 \preccurlyeq F_4$, $F_5 \preccurlyeq F_7$

- generalized implication relation is again a quasi-order

- similar quasi-order on objects

- quasi-orders can be graphically represented by (quasi-) Hasse diagrams

  e.g.,



- quasi-orders are to be represented by component matrices

- representation of quasi-orders by component matrices:

quasi-order on $\{F_1, \ldots, F_7\}$ as implied by $\hat{\mathbf{X}}$ $=$ quasi-order on $\{F_1, \ldots, F_7\}$ as implied by $\mathbf{A}^1$

$\hat{\mathbf{X}} =$

	O_1	O_2	O_3	O_4	O_5	O_6
F_1	4	1	1	0	4	1
F_2	0	5	5	0	5	5
F_3	0	0	0	2	0	2
F_4	0	0	0	2	0	2
F_5	4	5	5	0	5	5
F_6	4	5	5	0	5	5
F_7	4	5	5	2	5	5

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$	$A^1_{\bullet 3}$
F_1	1	0	0
F_2	0	1	0
F_3	0	0	1
F_4	0	0	1
F_5	1	1	0
F_6	1	1	0
F_7	1	1	1

$= \mathbf{A}^1$

- Note:

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$	$A^1_{\bullet 3}$
F_1	1	0	0
F_2	0	1	0
F_3	0	0	1
F_4	0	0	1
F_5	1	1	0
F_6	1	1	0
F_7	1	1	1

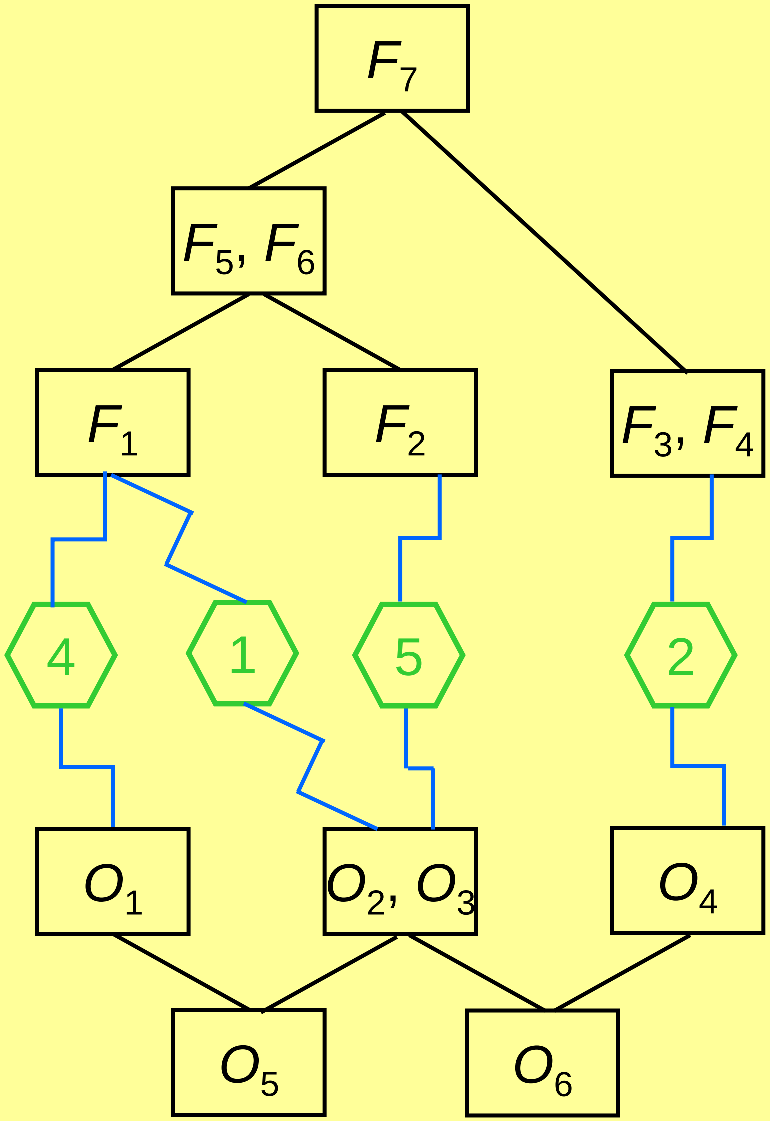
$\mathbf{A}^2$

	$A^2_{\bullet 1}$	$A^2_{\bullet 2}$	$A^3_{\bullet 3}$
O_1	1	0	0
O_2	0	1	0
O_3	0	1	0
O_4	0	0	1
O_5	1	1	0
O_6	0	1	1

G

	$A^2_{\bullet 1}$	$A^2_{\bullet 2}$	$A^3_{\bullet 3}$
$A^1_{\bullet 1}$	4	1	0
$A^1_{\bullet 2}$	0	5	0
$A^1_{\bullet 3}$	0	0	2

- Note: comprehensive graphical representation of HICLAS model



- Note:

  Max-operator can be shown to be only operator that allows representations of quasi-orders !!!

2. Models

2.2 Hierarchical classes models

2.2.2 Justification of the Max operator

Three (interrelated reasons):

1. representation by component matrices of quasi-orders that can be naturally defined on each of the modes
2. substantive interpretation of decomposition rule
3. decision rule for overlapping biclusters (triclusters etc.)

- consider binary (0/1) case

- assume (reconstructed) data pertain to person by problem failure/success:

  $\hat{x}_{i_1 i_2} = 0$: person i_1 fails for problem i_2

  $\hat{x}_{i_1 i_2} = 1$: person i_1 succeeds for problem i_2

- underlying mechanism:

  - P (latent) solution strategies

  - person i_1 may master strategy p or not: $a^1_{i_1 p} = 1$ or 0

  - strategy p may be suitable for solving problem i_2 or not: $a^2_{i_2 p} = 1$ or 0

- $$\hat{x}_{i_1 i_2} = \operatorname*{Max}_{p=1}^{P} \left(a^1_{i_1 p} \; a^2_{i_2 p} \right)$$

  $$\hat{x}_{i_1 i_2} = \bigoplus_{p=1}^{P} \left(a^1_{i_1 p} \; a^2_{i_2 p} \right)$$

- $\hat{x}_{i_1 i_2} = 1$ iff $\left(\exists p : \; a^1_{i_1 p} = 1 \text{ and } a^2_{i_2 p} = 1 \right)$

  person i_1 succeeds for problem i_2 iff there is at least one strategy *p* that person i_1 masters and that is suitable for solving problem i_2

Note:

existential quantifier $\rightarrow$ disjunctive model

2. Models

2.2 Hierarchical classes models

2.2.2 Justification of the Max operator

Three (interrelated reasons):

1. representation by component matrices of quasi-orders that can be naturally defined on each of the modes
2. substantive interpretation of decomposition rule
3. decision rule for overlapping biclusters (triclusters etc.)

- consider positively real-valued (reconstructed) object by variable data

- consider the following two models:

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \, g_{p_1 p_2} \right)$$

$$\hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \, g_{p_1 p_2} \right)$$

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$$\hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

$\mathbf{A}^1$

	$A^2_{\bullet 1}$	$A^2_{\bullet 2}$
V_1	0	0
V_2	1	0
V_3	1	0
V_4	1	1
V_5	0	1
V_6	0	1
V_7	0	0

$\mathbf{A}^2$

	$A^2_{\bullet 1}$	$A^2_{\bullet 2}$
$A^1_{\bullet 1}$	2	0
$A^1_{\bullet 3}$	0	3

$\mathbf{G}$

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right) \qquad \hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

A^1

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

A^2

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right) \qquad \hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

A^1

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
							O_1
							O_2
							O_3
							O_4
							O_5
							O_6

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

A^2

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \, g_{p_1 p_2} \right) \qquad \hat{x}_{i_1 i_2} = \underset{p_1=1}{\overset{P_1}{\mathrm{Max}}} \; \underset{p_2=1}{\overset{P_2}{\mathrm{Max}}} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \, g_{p_1 p_2} \right)$$

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
							O_1
							O_2
							O_3
							O_4
							O_5
							O_6

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

$\mathbf{A}^2$

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right) \qquad \hat{x}_{i_1 i_2} = \mathop{\mathrm{Max}}_{p_1=1}^{P_1} \mathop{\mathrm{Max}}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
							O_1
							O_2
							O_3
							O_4
							O_5
							O_6

$\mathbf{G}$

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

$\mathbf{A}^2$

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right) \qquad \hat{x}_{i_1 i_2} = \underset{p_1=1}{\overset{P_1}{\mathrm{Max}}} \; \underset{p_2=1}{\overset{P_2}{\mathrm{Max}}} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

A^1

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
0	0	0	0	0	0	0	O_1
0	2	2	2	0	0	0	O_2
0	2	2	2	0	0	0	O_3
0	2	2	2	0	0	0	O_4
0	0	0	0	0	0	0	O_5
0	0	0	0	0	0	0	O_6

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

A^2

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right) \qquad \hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
0	0	0	0	0	0	0	O_1
0	0	0	0	0	0	0	O_2
0	0	0	0	0	0	0	O_3
0	0	0	3	3	3	0	O_4
0	0	0	3	3	3	0	O_5
0	0	0	0	0	0	0	O_6

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$\mathbf{A}^2$

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$$\hat{x}_{i_1 i_2} = \operatorname*{Max}_{p_1=1}^{P_1} \operatorname*{Max}_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
0	0	0	0	0	0	0	O_1
0	2	2	2	0	0	0	O_2
0	2	2	2	0	0	0	O_3
0	2	2	?	3	3	0	O_4
0	0	0	3	3	3	0	O_5
0	0	0	0	0	0	0	O_6

G

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3

$\mathbf{A}^2$

	V_1	V_2	V_3	V_4	V_5	V_6	V_7
$A^2_{\bullet 1}$	0	1	1	1	0	0	0
$A^2_{\bullet 2}$	0	0	0	1	1	1	0

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$$\hat{x}_{i_1 i_2} = \underset{p_1=1}{\overset{P_1}{\mathrm{Max}}} \; \underset{p_2=1}{\overset{P_2}{\mathrm{Max}}} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$	V_1	V_2	V_3	V_4	V_5	V_6	V_7	
O_1	0	0	0	0	0	0	0	0	0	O_1
O_2	1	0	0	2	2	2	0	0	0	O_2
O_3	1	0	0	2	2	2	0	0	0	O_3
O_4	1	1	0	2	2	5	3	3	0	O_4
O_5	0	1	0	0	0	3	3	3	0	O_5
O_6	0	0	0	0	0	0	0	0	0	O_6

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

$\mathbf{A}^2$

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

$$\hat{x}_{i_1 i_2} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$$\hat{x}_{i_1 i_2} = \underset{p_1=1}{\overset{P_1}{\mathrm{Max}}} \; \underset{p_2=1}{\overset{P_2}{\mathrm{Max}}} \left(a^1_{i_1 p_1} \; a^2_{i_2 p_2} \; g_{p_1 p_2} \right)$$

$\mathbf{A}^1$

	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$
O_1	0	0
O_2	1	0
O_3	1	0
O_4	1	1
O_5	0	1
O_6	0	0

V_1	V_2	V_3	V_4	V_5	V_6	V_7	
0	0	0	0	0	0	0	O_1
0	2	2	2	0	0	0	O_2
0	2	2	2	0	0	0	O_3
0	2	2	3	3	3	0	O_4
0	0	0	3	3	3	0	O_5
0	0	0	0	0	0	0	O_6

G

$A^2_{\bullet 1}$	2	0
$A^2_{\bullet 2}$	0	3
	$A^1_{\bullet 1}$	$A^1_{\bullet 2}$

$\mathbf{A}^2$

0	1	1	1	0	0	0	$A^2_{\bullet 1}$
0	0	0	1	1	1	0	$A^2_{\bullet 2}$
V_1	V_2	V_3	V_4	V_5	V_6	V_7	

Note:

taxonomic overview of biclustering: see Van Mechelen, Bock, & De Boeck (2004)

2. Models

2.2 Hierarchical classes models

2.2.3 The HICLAS family

- key features:

1. operator = Max

$$x_{i_1 i_2 \ldots i_N} = \underset{p_1=1}{\overset{P_1}{\text{Max}}}\ \underset{p_2=1}{\overset{P_2}{\text{Max}}} \ldots \underset{p_N=1}{\overset{P_N}{\text{Max}}} \left[\left(\prod_{n=1}^{N} a^{n}_{i_n p_n} \right) g_{p_1 p_2 \ldots p_N} \right] + e_{i_1 i_2 \ldots i_N}$$

2. all $\mathbf{A}^n$ 0/1
3. entries of $\underline{\mathbf{G}} \in \mathbb{R}^+$
4. representation of quasi-orders by $\mathbf{A}^n$

- unconstrained models:

$\underline{\mathbf{G}}$ 0/1

- $N = 2 \Rightarrow \mathbf{G} = \mathbf{I}$

(disjunctive) hierarchical classes model (De Boeck and Rosenberg, 1988)

Notes:

* dual conjunctive model with Min-operator: Van Mechelen, De Boeck & Rosenberg, 1995)

* stochastic extension: Bayesian HICLAS (Leenen, Van Mechelen, Gelman & De Knop, submitted)

- *N=3*

  * **G** superidentity (CANDECOMP/PARAFAC case):

  INDCLAS (Leenen, Van Mechelen, De Boeck & Rosenberg, 1999)

  * **G** general:

  - Tucker3 HICLAS (Ceulemans, Van Mechelen & Leenen, 2003)

  - Tucker2 HICLAS (Ceulemans & Van Mechelen, 2004)

G rating-valued (*N* =2)

- HICLAS-R (Van Mechelen, Lombardi & Ceulemans, conditionally accepted)

  Note:

  option to put limit on number of distinct values in **G** (‘optimal coarsening ’) entries $\underline{\mathbf{G}} \in \mathbb{R}^{+}$

- real-valued HICLAS (Schepers & Van Mechelen, submitted)

- constrained models

taxonomy of constraints (Ceulemans, Van Mechelen & Kuppens, 2004), based on:

1. locus: component matrices – core
2. nature: value (e.g., values from previous study) vs. structure (e.g., Guttman scale, consecutive ones, decomposability, lower bound on number of ones, etc.)
3. extent: full vs. partial
4. use of external information: no vs. yes

Note:

varying amount of required algorithmic adaptations

Overview of the talk

1. data
2. models

    2.1 context

    2.2 hierarchical classes models

        2.2.1 basic model

        2.2.2 justification of the Max operator

        2.2.3 the HICLAS family

3. research topics

    3.1 models

    3.2 estimation

    3.3 model selection and model checking

    3.4 quantification of uncertainty

4. references

3. Research topics

3.1 Models

- mathematical properties (e.g., rank)

  relevant structures: Boolean algebra and tropical semiring $\left(\mathbb{R}^{+}, \text{Max}, \times\right)$

Notes:

1. differences between these structures and common algebras over $\mathbb{R}$

   e.g., row, column, and decomposition (Schein) rank may differ

   e.g., only invertible Boolean matrices are permutation matrices

2. even with Σ-operator rank issues considerably change if one or more of the $\mathbf{A}^{n}$ are constrained to be 0/1 ($\neq$ -1/1)

- identifiability / uniqueness:
  - only invertible Boolean matrices are permutation matrices

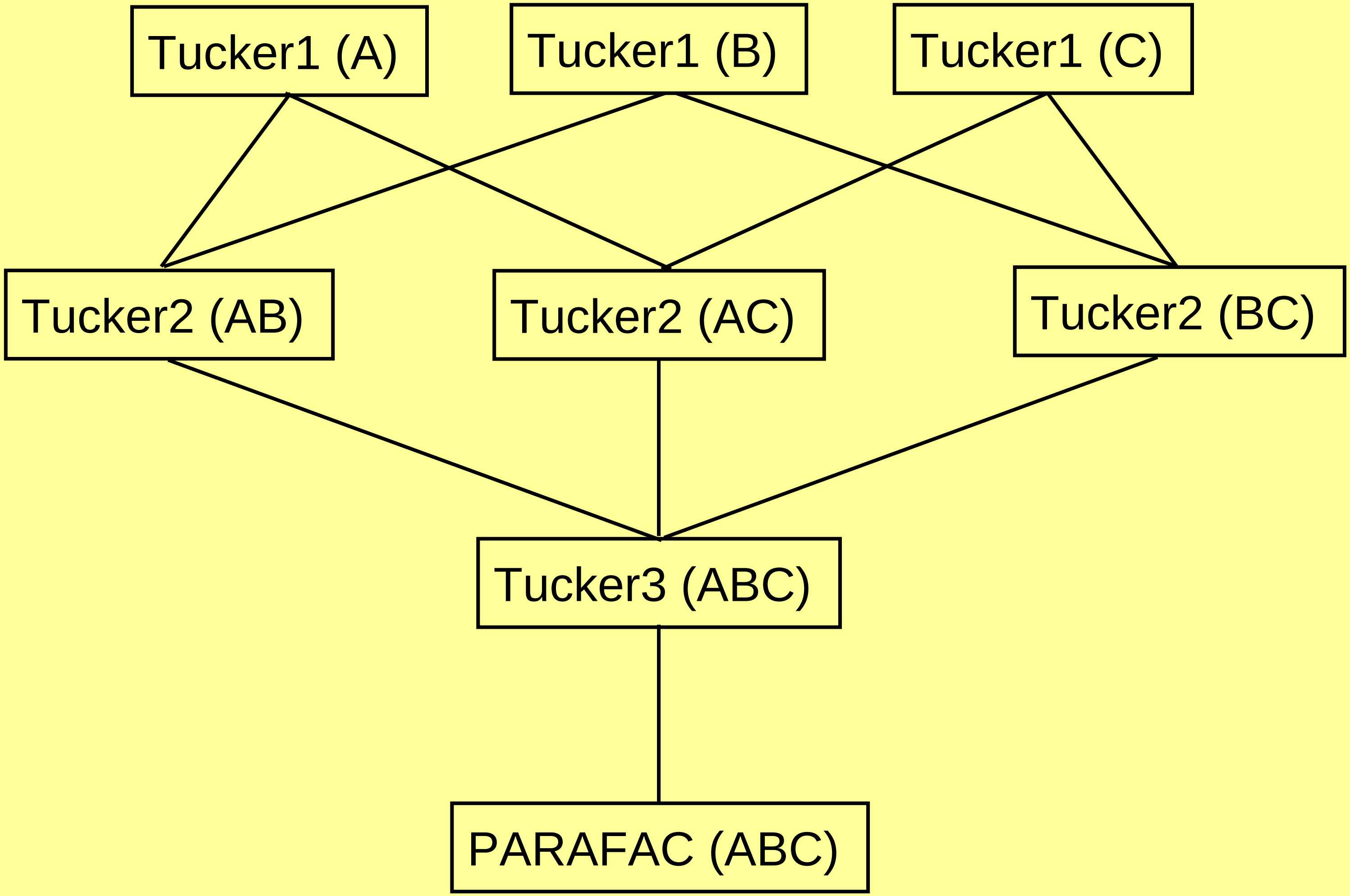
    $\rightarrow$ less identifiability problems than for common *N*-way Tucker models

    $\rightarrow$ in general only permutational freedom rather than full rotational freedom
  - remaining amount of nonuniqueness pertains to particular decompositions
  - part of remaining nonuniqueness removed by requirement of representation of quasi-order
  - at this moment only sufficient condition for uniqueness in 0/1 case (Van Mechelen, De Boeck & Rosenberg, 1995; Ceulemans & Van Mechelen, 2003)

    HICLAS/INDCLAS: presence of all component patterns that contain single 1

    Tucker3 HICLAS: idem + no core slice is subset of union of other core slices

- model interrelations (Ceulemans & Van Mechelen, 2005):



Overview of the talk

1. data
2. models

   2.1 context

   2.2 hierarchical classes models

      2.2.1 basic model

      2.2.2 justification of the Max operator

      2.2.3 the HICLAS family
3. research topics

   3.1 models

   3.2 estimation

   3.3 model selection and model checking

   3.4 quantification of uncertainty
4. references

3. Research topics

3.2 Estimation

- $$L_1 = \sum_{i_1, i_2, \ldots, i_N} \left| \hat{x}_{i_1 i_2 \ldots i_N} - x_{i_1 i_2 \ldots i_N} \right|$$

  $$L_2 = \sum_{i_1, i_2, \ldots, i_N} \left(\hat{x}_{i_1 i_2 \ldots i_N} - x_{i_1 i_2 \ldots i_N} \right)^2$$

  Note:

  in pure 0/1 case $L_1 = L_2$

- evaluation criteria for algorithms:
  - primary: goodness of fit; ?global optimum
  - secondary: goodness of recovery

- types of algorithms under study:
  - alternating least L_1 / L_2
  - simulated annealing (Ceulemans, Van Mechelen & Leenen, submitted)
  - MCMC (Metropolis) (Leenen, Van Mechelen, Gelman, & De Knop, submitted)

- algorithmic issues

  - parametrization of solution space (and partitioning of parameter vector, if applicable) (Schepers, Van Mechelen & Ceulemans, in press)

  - starts

    * number

    * nature: rational from data / rational from data + noise / random from data / purely random
    (Ceulemans, Van Mechelen & Leenen, submitted)

  - iterative process

    - choice metaparameters in simulated annealing and MCMC

    - greedy vs. branch and bound in conditional updating of alternating least L_p algorithms (Leenen & Van Mechelen, 2001)

Overview of the talk

3. Research topics

3.2 Model selection and model checking

- issues
  - model type
  - rank
  - correct representation quasi-orders
  - assumptions about error process

- types of methods under study
  - scree-test based approaches (including extended convex hull methods (Leenen & Van Mechelen, 2001; Ceulemans & Van Mechelen, 2005; see also Ceulemans & Kiers, in press)
  - pseudo AIC approach (Ceulemans & Van Mechelen, 2005)
  - Bayesian approaches with posterior predictive checks (Leenen, Van Mechelen, Gelman & De Knop, submitted)

Overview of the talk

1. data
2. models

    2.1 context

    2.2 hierarchical classes models

        2.2.1 basic model

        2.2.2 justification of the Max operator

        2.2.3 the HICLAS family

3. research topics

    3.1 models

    3.2 estimation

    3.3 model selection and model checking

    3.4 quantification of uncertainty

4. references

3. Research topics

3.4 Quantification of uncertainty

- issues:
  - telling apart strong and weak parts in obtained representations; ‘confidence intervals’
  - not only single parameters; also aspects derived from multiple parameters (e.g., classes, hierarchical relations)
  - confidence intervals in 0/1 case are tricky!

- types of methods under study
  - nonparametric Bootstrap approach
  - Bayesian approach based on simulated posterior (Leenen, Van Mechelen, Gelman & De Knop, submitted)

Overview of the talk

1. data
2. models

    2.1 context

    2.2 hierarchical classes models

        2.2.1 basic model

        2.2.2 justification of the Max operator

        2.2.3 the HICLAS family

3. research topics

    3.1 models

    3.2 estimation

    3.3 model selection and model checking

    3.4 quantification of uncertainty

4. references

4. References

Ceulemans, E., & Kiers, H.A.L. (in press). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *British Journal of Mathematical and Statistical Psychology*.

Ceulemans, E., & Van Mechelen, I. (2003). Uniqueness of *N*-way *N*-mode hierarchical classes models. *Journal of Mathematical Psychology*, 47, 259-264.

Ceulemans, E., Van Mechelen, I., & Leenen, I. (2003). Tucker3 hierarchical classes analysis. *Psychometrika*, *68*, 413-433.

Ceulemans, E., & Van Mechelen, I. (2004). Tucker2 hierarchical classes analysis. *Psychometrika*, *69*, 375-399.

Ceulemans, E., & Van Mechelen, I. (2005). Hierarchical classes models for three-way three-mode binary data: Interrelations and model selection. *Psychometrika*, *70,* 461-480.

Ceulemans, E., Van Mechelen, I., & Leenen, I. (submitted). The local minima problem in hierarchical classes analysis: An evaluation of a simulated annealing algorithm and various multistart procedures.

De Boeck, P., & Rosenberg, S. (1988). Hierarchical classes: Model and data analysis. *Psychometrika, 53*, 361-381.

Leenen, I., Van Mechelen, I., De Boeck, P., & Rosenberg, S. (1999). INDCLAS: Individual differences hierarchical classes analysis. *Psychometrika*, *64*, 9–24.

Leenen, I., Van Mechelen, I., & De Boeck, P. (1999). A generic disjunctive/conjunctive decomposition model for *n*-ary relations. *Journal of Mathematical Psychology*, 43, 102-122.

Leenen, I., & Van Mechelen, I. (2001). An evaluation of two algorithms for hierarchical classes analysis. *Journal of Classification*, 18, 57-80.

Leenen, I., Van Mechelen, I., Gelman, A., & De Knop, S. (submitted). Bayesian hierarchical classes analysis.

Schepers, J., & Van Mechelen, I. (submitted). The real-valued model of hierarchical classes.

Schepers, J., Van Mechelen, I., & Ceulemans, E. (in press). Three-mode partitioning. *Computational Statistics and Data Analysis*.

Van Mechelen, I., De Boeck, P., & Rosenberg, S. (1995). The conjunctive model of hierarchical classes. *Psychometrika*, *60*, 505-521.

Van Mechelen, I., Bock, H.-H., & De Boeck, P. (2004). Two-mode clustering methods: A structured overview. *Statistical Methods in Medical Research*, *13*, 363-394.

Van Mechelen, I., Lombardi, L., & Ceulemans, E. (conditionally accepted). Hierarchical classes modeling of rating data. *Psychometrika*.

Van Mechelen, I., & Schepers, J. (submitted). A unifying model involving a categorical and/or dimensional reduction for multimode data.