R. Belohlavek, M. Krmelova: Beyond Binary Matrix Decompositions: Work in Progress

Department of Computer Science, Palacky University, Olomouc (17. listopadu 12, CZ–77146 Olomouc, Czech Republic)

General Matrix Decomposition

- input: $n \times m$ object-attribute matrix I with entries I_{ij} expressing grades to which object i has attribute j
- output: an $n \times k$ object-factor matrix A and a $k \times m$ factor-attribute matrix B
- grades are taken from a bounded scale L
- use the calculus of matrices over residuated lattices
- goal: find A and B with k (# factors) as small as possible

$$I = A \circ B$$

- motivation: factor analysis

Preliminaries

- equivalent pairs For $a_1, a_2, b_1, b_2 \in L$, we put

$$\langle a_1, b_1\rangle \equiv \langle a_2, b_2\rangle \text{ iff } a_1 \otimes b_1 = a_2 \otimes b_2,$$

in which case we say that $\langle a_1, b_1\rangle$ and $\langle a_2, b_2\rangle$ are equivalent
- $\equiv$ is an equivalence on $L \times L$
- the classes of $\equiv$ are just the sets

$$fac_{\otimes}(a) = \{\langle b, c\rangle \in L \times L \mid a = b \otimes c\}$$

- more generally, for $C_1, C_2 \in L^X$, $D_1, D_2 \in L^Y$:

$$\langle C_1, D_1\rangle \equiv \langle C_2, D_2\rangle \text{ iff } C_1 \otimes D_1 = C_2 \otimes D_2,$$

- $\langle C_1, D_1\rangle$ and $\langle C_2, D_2\rangle$ are equivalent, or determined the same rectangle
- for $J \in L^X \times L^Y$:

$$fac_{\otimes}(J) = \{\langle C, D\rangle \in L^X \times L^Y \mid J = C \otimes D\}$$

$$fac^M_{\otimes}(J) = \{\langle C, D\rangle \in L^X \times L^Y \mid J = C \otimes D$$
$$\langle C, D\rangle \text{ is maximal with } J = C \otimes D\}$$

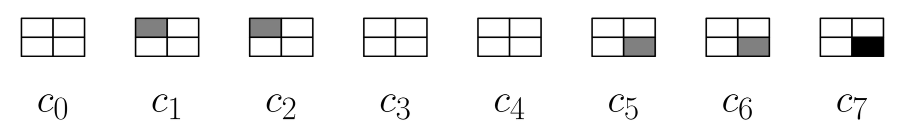
- for $C \subseteq L^X$ and $D \subseteq L^Y$ $\gamma(C) = \langle\{C\}^{\uparrow\downarrow}, \{C\}^{\uparrow}\rangle$ and $\mu(D) = \langle\{D\}^{\downarrow}, \{D\}^{\downarrow\uparrow}\rangle$
- object concepts: $\mathcal{O}(X, Y, I) = \{\gamma\{^a/_x\} | x \in X, a \in L\}$
- attribute concepts: $\mathcal{A}(X, Y, I) = \{\mu\{^a/_y\} | y \in Y, a \in L\}$
- formal concepts which are both object and attribute covers such I_{xy}, which could be covered by only one maximal rectangle

Small Example

- residuated lattice: three-element Łukasiewicz chain $L = \{0, 0.5, 1\}$
- input matrix: $I = \begin{pmatrix} 0.5 & 0 \\ 0 & 1 \end{pmatrix}$
- set of all concepts:

c_i	Extent	Intent
c_0	$\{x_1, x2\}$	$\{\}$
c_1	$\{x_1, {}^{0.5}/_{x_2}\}$	$\{{}^{0.5}/_{y_1}\}$
c_2	$\{{}^{0.5}/_{x_1}\}$	$\{y_1, {}^{0.5}/_{y_2}\}$
c_3	$\{\}$	$\{y_1, y_2\}$
c_4	$\{{}^{0.5}/_{x_1}, {}^{0.5}/_{x_2}\}$	$\{{}^{0.5}/_{y_1}, {}^{0.5}/_{y_2}\}$
c_5	$\{{}^{0.5}/_{x_2}\}$	$\{{}^{0.5}/_{y_1}, y_2\}$
c_6	$\{{}^{0.5}/_{x_1}, x_2\}$	$\{{}^{0.5}/_{y_2}\}$
c_7	$\{x_2\}$	$\{y_2\}$

- set of all object concepts and set of all attribute concept: $\mathcal{O}(X, Y, I) = c_1, c_2, c_5, c_7$ and $\mathcal{A}(X, Y, I) = c_1, c_2, c_6, c_7$
- concepts as rectangular pattern:



- concepts c_1, c_2 and c_7 are both object and attribute concepts
- concepts c_1 and c_2 have same maximal rectangle
- two different optimal decomposition $\mathcal{F} = \{c_1, c_7\}$ and $\mathcal{F}' = \{c_2, c_7\}$ and in both of them is concept c_7

Essential Concepts

- interval: for $C \subseteq L^X$ and $D \subseteq L^Y$

$$\mathcal{I}_{C,D} = [\gamma(C), \mu(D)]$$

- $[\gamma(C), \mu(D)] = \{\langle E, F\rangle \in \mathcal{B}(X, Y, I) \mid \gamma(C) \leq \langle E, F\rangle \leq \mu(D)\}$
- in the Boolean case holds:

$I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$ iff for every $\langle i, j\rangle$ there exists $\langle C, D\rangle \in \mathcal{F}$ such that $I_{ij} = C(i) \otimes D(j)$ (1)

The following lemma shows that with linearly ordered residuated lattices, factorization behaves as in the Boolean case:

Lemma 1 Let $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$. Let $\mathbf{L}$ be linearly ordered.
(a) If $\mathcal{F}$ is finite then (1) holds.
(b) If $\mathbf{L}$ is finite then (1) holds.

Lemma 2 (a) For $J \in L^X \times L^Y$ we have $fac(J) \neq \emptyset$ iff J is rectangular.
(b) If J is a maximal rectangle contained in I then $fac^M(J) = \{\langle C, D\rangle \in \mathcal{B}(X, Y, I) \mid J = C \otimes D\}$.
(c) $\mathcal{B}(X, Y, I) = \bigcup\{fac(J) \mid$
J is a maximal rectangle contained in $I\} \cup \{\langle\emptyset^{\uparrow\downarrow}, \emptyset^{\uparrow}\rangle, \langle\emptyset^{\downarrow}, \emptyset^{\downarrow\uparrow}\rangle\}$.

Lemma 3 If $\langle E, F\rangle \in \mathcal{I}_{C,D}$ then $C \otimes D \subseteq E \otimes F$. In particular, if $\langle E, F\rangle \in \mathcal{I}_{\{^a/i\},\{^b/j\}}$ then $a \otimes b \leq E(i) \otimes F(j)$.

Remark The converse claim to previous Lemma does not hold. Note that the converse claim holds in the Boolean case [2].

Lemma 4 Let $\langle E, F\rangle \in \mathcal{B}(X, Y, I)$.
(a) $\mathcal{I}_{E,F} = \{\langle E, F\rangle\}$.
(b) For every i, j, $\langle E, F\rangle \in \mathcal{I}_{\{^{E(i)}/i\},\{^{F(j)}/j\}}$.

The following shows two kinds of converse claim to the second part of Lemma 3

Lemma 5 Let $\langle E, F\rangle \in \mathcal{B}(X, Y, I)$. If $a \otimes b \leq E(i) \otimes F(j)$ then for some c, d with $a \otimes b \leq c \otimes d$ we have $\langle E, F\rangle \in \mathcal{I}_{\{^c/i\},\{^d/j\}}$.

The second part of Lemma 3 may also be: if $\langle E, F\rangle \in \mathcal{I}_{\{^c/i\},\{^d/j\}}$ for some $c \otimes d \geq a \otimes b$ then $a \otimes b \leq E(i) \otimes F(j)$. This is closer in form to Lemma 5.

Definition We say that $\langle E, F\rangle \in L^X \times L^Y$ covers $\langle i, j\rangle$ in $I \in L^{X \times Y}$ if $I_{ij} \leq E(i) \otimes F(j)$.
Consider a given $I \in L^{X \times Y}$. Let

$$\mathbf{I}_{ij} = \{\mathcal{I}_{\{^a/i\},\{^b/j\}} \mid a \otimes b = I_{ij}\}$$

and put

$$\mathcal{I}_{ij} = \bigcup \mathbf{I}_{ij}.$$

Then we have:

Lemma 6 If $\langle E, F\rangle \in \mathcal{B}(X, Y, I)$ covers $\langle i, j\rangle$ then $\langle E, F\rangle \in \mathcal{I}_{ij}$.

As a corollary we obtain:

Theorem $\langle E, F\rangle \in \mathcal{B}(X, Y, I)$ covers $\langle i, j\rangle$ iff $\langle E, F\rangle \in \mathcal{I}_{ij}$.

Lemma 7 If $\otimes$ is $\wedge$, then $\mathcal{I}_{ij}$ is an interval in $\mathcal{B}(I)$, namely, $\mathcal{I}_{ij} = \mathcal{I}_{\{^{I_{ij}}/i\},\{^{I_{ij}}/j\}}$.

Essential $\langle i, j\rangle$ is such $\mathcal{I}_{ij}$, which is non-empty and minimal w.r.t. $\subseteq$.

Definition $J \leq I$ is called essential (or an essential part of I) if for every $\mathcal{F} \subseteq \mathcal{B}(I)$ we have: if $J \leq A_{\mathcal{F}} \circ B_{\mathcal{F}}$ then $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$.
Denote by $\mathcal{E}(I)$ the matrix defined by

$$(\mathcal{E}(I))_{ij} = \begin{cases} I_{ij} \text{if } \mathcal{I}_{ij} \text{ is non-empty and minimal w.r.t. } \subseteq. \\ 0 \text{ otherwise.} \end{cases}$$

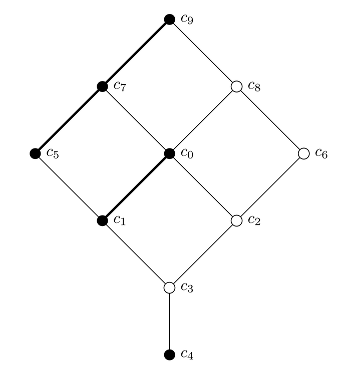
Theorem $\mathcal{E}(I)$ is an essential part of I.

Illustrative Example

- set of objects $X = \{x_1, x_2, x_3\}$, set of attributes $Y = \{y_1, y_2, y_3\}$
- residuated lattice: five-element Łukasiewicz chain $L = \{0, 0.25, 0.5, 0.75, 1\}$
- input matrix: $I = \begin{pmatrix} 1.00 & 0.75 & 1.00 \\ 1.00 & 0.50 & 0.50 \\ 0.50 & 0.50 & 1.00 \end{pmatrix}$
- the classes of $\equiv$ are:

$$fac_{\otimes}(0.25) = \{\langle 1, 0.25\rangle, \langle 0.75, 0.5\rangle, \langle 0.5, 0.75\rangle, \langle 0.25, 1\rangle\}$$
$$fac_{\otimes}(0.5) = \{\langle 1, 0.5\rangle, \langle 0.75, 0.75\rangle, \langle 0.5, 1\rangle\}$$
$$fac_{\otimes}(0.75) = \{\langle 1, 0.75\rangle, \langle 0.75, 1\rangle\}$$
$$fac_{\otimes}(1) = \{\langle 1, 1\rangle\}$$

- $\mathcal{I}_{x_2,y_2} = \mathcal{I}_{\{^{0.5}/x_2\},\{^{0.5}/y_2\}}$:



- essential matrix: $\mathcal{E}(I) = \begin{pmatrix} 0.00 & 0.75 & 0.00 \\ 1.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$
- I_{x_1,y_3} is covered by concepts:
$c_0 = \langle\{x_1, {}^{0.75}/x_2, {}^{0.75}/x_3\}, \{{}^{0.75}/y_1, {}^{0.75}/y_2, {}^{0.75}/y_3\}\rangle$,
$c_1 = \langle\{x_1, {}^{0.75}/x_2, {}^{0.5}/x_3\}, \{y_1, {}^{0.75}/y_2, {}^{0.75}/y_3\}\rangle$,
$c_2 = \langle\{x_1, {}^{0.5}/x_2, {}^{0.75}/x_3\}, \{{}^{0.75}/y_1, {}^{0.75}/y_2, y_3\}\rangle$,
$c_3 = \langle\{x_1, {}^{0.5}/x_2, {}^{0.5}/x_3\}, \{y_1, {}^{0.75}/y_2, y_3\}\rangle$,
$c_4 = \langle\{{}^{0.75}/x_1, {}^{0.5}/x_2, {}^{0.5}/x_3\}, \{y_1, y_2, y_3\}\rangle$
- I_{x_2,y_1} is covered by concept:
$c_5 = \langle\{x_1, x_2, {}^{0.5}/x_3\}, \{y_1, {}^{0.5}/y_2, {}^{0.5}/y_3\}\rangle$
- I_{x_3,y_3} is covered by concept:
$c_6 = \langle\{x_1, {}^{0.5}/x_2, x_3\}, \{{}^{0.5}/y_1, {}^{0.5}/y_2, y_3\}\rangle$
- decomposition of matrix contains c_5, c_6 and one of c_0, c_1, c_2, c_3, c_4
- output matrices:

$$A = \begin{pmatrix} 1.00 & 1.00 & 1.00 \\ 1.00 & 0.50 & 0.75 \\ 0.50 & 1.00 & 0.75 \end{pmatrix}, B = \begin{pmatrix} 1.00 & 0.50 & 0.50 \\ 0.50 & 0.50 & 1.00 \\ 0.75 & 0.75 & 0.75 \end{pmatrix}$$

Our Previous Work

[1] Belohlavek R., Vychodil V.: Discovery of optimal factors in binary data via a novel method of matrix decomposition. Journal of Computer and System Sciences 76(1)(2010), 3–20.

[2] Belohlavek, Trnecka: From-below approximations in boolean matrix factorization: geometry and new algorithm (submitted).

[3] Belohlavek R.: Optimal decompositions of matrices with entries from residuated lattices. Journal of Logic and Computation, doi: 10.1093/logcom/exr023, online: September 7, 2011.

[4] Belohlavek R., Vychodil V.: Factor analysis of incidence data via novel decomposition of matrices. LNAI 5548(2009), 83-97.

[5] Belohlavek R., Krmelova M.: Factor analysis of ordinal data via decomposition of matrices with grades (submitted).

International Center for Information and Uncertainty
SSIU 2013: International Summer School "Information and Uncertainty"
http://mcin.upol.cz/SSWIU-2013/