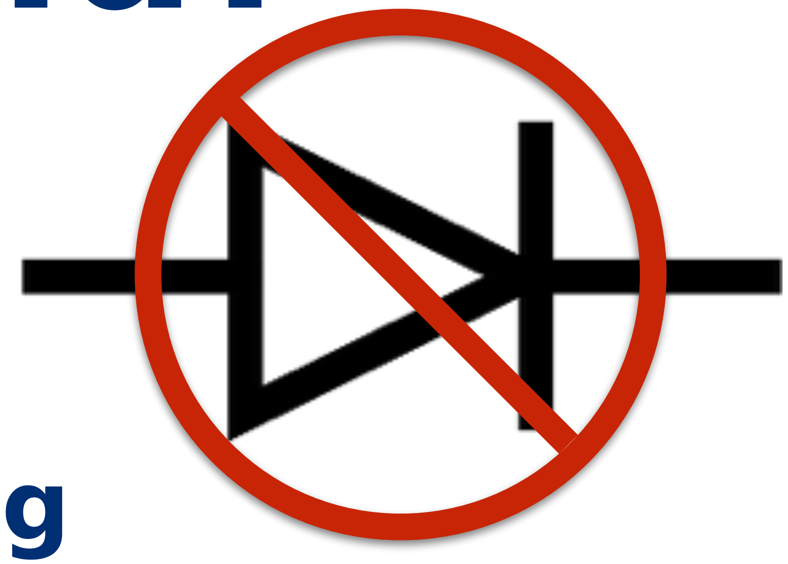
Dioids in Data Mining

Pauli Miettinen

10 March 2014

max planck institut
informatik

What is a dioid?



- Diode is not a diode

- Dioid is an **idempotent semiring**

  $\boldsymbol{S} = (A, \oplus, \otimes, ⓪, ①)$

  - Addition $\oplus$ is idempotent

    - $a + a = a$ for all $a \in A$

  - Addition is not invertible

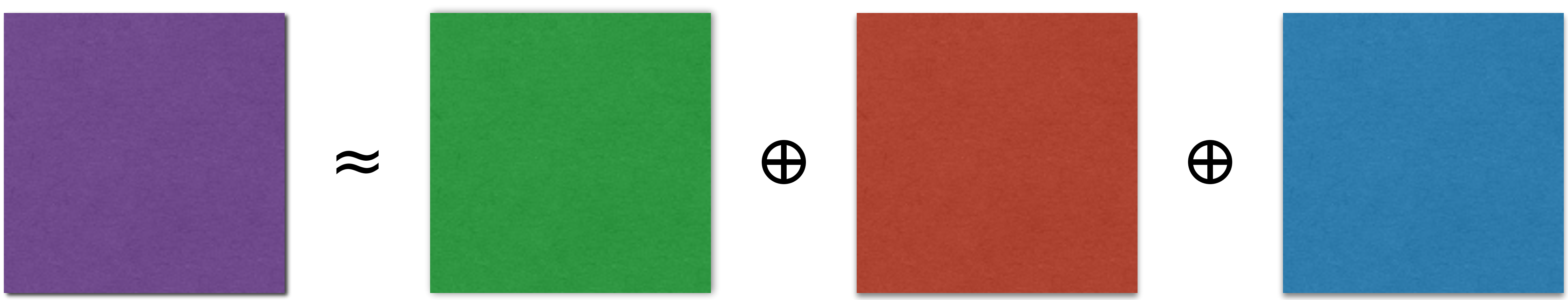
Why dioids in DM?

- **What happens if we replace normal algebra with some dioid?**
  - Non-linear structure
  - Computationally harder problems
- Matrix-factorization type problems

Why matrix factorizations?

Siegfried said they're a hot topic

- Because I can
- MFs model the whole data using sums of rank-1 components
- Dioids change how these components interact



Some examples (1)

- The **Boolean algebra** $\boldsymbol{B} = (\{0,1\}, \vee, \wedge, 0, 1)$
  - The **subset lattice** $\boldsymbol{L} = (2^U, \cup, \cap, \varnothing, U)$ is isomorphic to $\boldsymbol{B}^n$
- The **Boolean matrix factorization** expresses matrix **A** as $\mathbf{A} \approx \mathbf{B} \otimes_{\boldsymbol{B}} \mathbf{C}$ where all matrices are Boolean

BMF example

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \otimes_{\boldsymbol{B}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Some examples (2)

- **Fuzzy logic** $\boldsymbol{F}$ = ([0, 1], max, min, 0, 1)
- Generalizes (relaxes) Boolean algebra
  - Exact k-decomposition under fuzzy logic implies exact k-decomposition under Boolean algebra

Fuzzy example

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \approx \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \otimes_{\boldsymbol{F}} \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 2/3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 2/3 & 1 \\ 0 & 1 & 2/3 & 1 \\ 0 & 1 & 2/3 & 1 \end{pmatrix}$$

Some examples (3)

- The **max-times algebra**

  $\boldsymbol{M} = (\mathbb{R}_{\geq 0}, \max, \times, 0, 1)$

  - Isomorphic to the **tropical algebra**

    $\boldsymbol{T} = (\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$

    - $\boldsymbol{T} = \log(\boldsymbol{M})$ and $\boldsymbol{M} = \exp(\boldsymbol{T})$

Why max-times?

- One interpretation: *Only strongest reason matters*
  - Normal algebra: rating is a linear combination of movie’s features
  - Max-times: rating is determined by the most-liked feature

Max-times example

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \approx \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2/3 \\ 0 & 1 \end{pmatrix} \otimes_{\boldsymbol{M}} \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 2/3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 2/3 & 1 \\ 0 & 2/3 & 4/9 & 2/3 \\ 0 & 1 & 2/3 & 1 \end{pmatrix}$$

On max-times algebra

- Max-times algebra relaxes Boolean algebra (but not fuzzy logic)
- Rank-1 components are “normal”
  - Easy to interpret?
- Not much studied

On tropical algebras

- A.k.a. max-plus, extremal, maximal algebra
- Much more studied than max-times
  - Can be used to solve max-times problems, but needs care with the errors
    - If $\|\boldsymbol{X} - \tilde{\boldsymbol{X}}\| \leq \alpha$ in max-plus then $\|\boldsymbol{X}' - \widetilde{\boldsymbol{X}'}\| \leq M^2\alpha$ in max-times, where $M = \exp(\max_{i,j}\{\boldsymbol{X}_{ij}, \tilde{\boldsymbol{X}}_{ij}\})$

More max-plus

- Max-plus linear functions: $f(\boldsymbol{x}) = \boldsymbol{f}^T \otimes \boldsymbol{x}$

  $= \max\{f_i + x_i\}$

  - $f(\alpha \otimes \boldsymbol{x} \oplus \beta \otimes \boldsymbol{y}) = \alpha \otimes f(\boldsymbol{x}) \oplus \beta \otimes f(\boldsymbol{y})$

- Max-plus eigenvectors and values:

  $\boldsymbol{X} \otimes \boldsymbol{v} = \lambda \otimes \boldsymbol{v}$ ($\max_j\{x_{ij} + v_j\} = \lambda + v_i$ for all i)

- Max-plus linear systems: $\boldsymbol{A} \otimes \boldsymbol{x} = \boldsymbol{b}$

  - Solving in pseudo-P for integer $\boldsymbol{A}$ and $\boldsymbol{b}$

Computational complexity

- If exact k-factorization over semiring $\boldsymbol{K}$ implies exact k-factorization over $\boldsymbol{B}$, then finding the $\boldsymbol{K}$-rank of a matrix is NP-hard (even to approximate)
  - Includes fuzzy, max-times, and tropical
    - N.B. feasibility results in $\boldsymbol{T}$ often require finite matrices

Anti-negativity and sparsity

- A semiring is **anti-negative** if no non-zero element has additive inverse
  - Some dioids are anti-negative, others not
- Anti-negative semirings yield sparse factorizations of sparse data

Conclusions

- Idempotent semirings capture non-linear structure
- Some are already used in DM
- More abstract view should help finding connections
- Max-plus algebras can provide tools for other problems

Sept. 15-19

ECML PKDD 2014

www.ecmlpkdd2014.org

Nancy France

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases

Abstract DL	12 April
Paper DL	16 April