Dioids in Data Mining

Pauli Miettinen
10 March 2014
What is a dioid?

• Dioid is not a diode

• Dioid is an **idempotent semiring**

  \[ S = (A, \oplus, \otimes, 0, 1) \]

  • 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?

• Because I can

• MFs model the whole data using sums of rank-1 components

• Dioids change how these components interact

Siegfried said they’re a hot topic
Some examples (1)

• The **Boolean algebra** $B = (\{0,1\}, \lor, \land, 0, 1)$

• The **subset lattice** $L = (2^U, \cup, \cap, \emptyset, U)$ is isomorphic to $B^n$

• The **Boolean matrix factorization** expresses matrix $A$ as $A \approx B \otimes_B C$ where all matrices are Boolean
BMF example

\[
\begin{pmatrix}
1 & 1 & 0 \\
1 & 1 & 1 \\
0 & 1 & 1
\end{pmatrix}
= \begin{pmatrix}
1 & 1 \\
1 & 1 \\
0 & 1
\end{pmatrix}
\otimes_B \begin{pmatrix}
1 & 1 & 0 \\
1 & 1 & 1 \\
0 & 1 & 1
\end{pmatrix}
\]
Some examples (2)

- **Fuzzy logic** $F = ([0, 1], \text{max}, \text{min}, 0, 1)$
- 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 & 1 & 1 \\
0 & 1 & 1 & 1
\end{pmatrix}
\approx
\begin{pmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{pmatrix}
\otimes_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**
  \[ M = (\mathbb{R}_{\geq 0}, \max, \times, 0, 1) \]

• Isomorphic to the **tropical algebra**
  \[ T = (\mathbb{R} \cup \{ -\infty \}, \max, +, -\infty, 0) \]

• \( T = \log(M) \) and \( M = \exp(T) \)
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

\[
\begin{pmatrix}
1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 1 & 1
\end{pmatrix}
\approx
\begin{pmatrix}
1 & 0 \\
1 & 1 \\
0 & 2/3 \\
0 & 1
\end{pmatrix} \otimes_M
\begin{pmatrix}
1 & 1 & 0 & 1 \\
0 & 1 & 2/3 & 1
\end{pmatrix}
\]

\[
= \begin{pmatrix}
1 & 1 & 0 & 0 \\
1 & 1 & 2/3 & 1 \\
0 & 2/3 & 4/9 & 2/3 \\
0 & 1 & 2/3 & 1
\end{pmatrix}
\]
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 $\|X - \tilde{X}\| \leq \alpha$ in max-plus then
  $\|X' - \tilde{X}'\| \leq M^2 \alpha$ in max-times, where

  $M = \exp(\max_{i,j} \{X_{ij}, \tilde{X}_{ij}\})$
More max-plus

• Max-plus linear functions: \( f(x) = f^T \otimes x \)
  \[
  = \max \{ f_i + x_i \}
  \]
  
  • \( f(\alpha \otimes x \oplus \beta \otimes y) = \alpha \otimes f(x) \oplus \beta \otimes f(y) \)

• Max-plus eigenvectors and values:
  \[
  X \otimes v = \lambda \otimes v \quad (\max_j \{ x_{ij} + v_j \} = \lambda + v_i \text{ for all } i)
  \]

• Max-plus linear systems: \( A \otimes x = b \)
  
  • Solving in pseudo-P for integer \( A \) and \( b \)
Computational complexity

• If exact $k$-factorization over semiring $K$ implies exact $k$-factorization over $B$, then finding the $K$-rank of a matrix is NP-hard (even to approximate)

• Includes fuzzy, max-times, and tropical

• N.B. feasibility results in $T$ 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