Mathematics of Big Data & Machine Learning

Associative Arrays – Theory & Applications

Hayden Jananthan & Jeremy Kepner

January 10, 2024

Matrices

Two-dimensional arrays of values from a field (e.g., $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \ldots$) or ring (e.g., $\mathbb{Z}, \mathbb{R}[x], \ldots$):

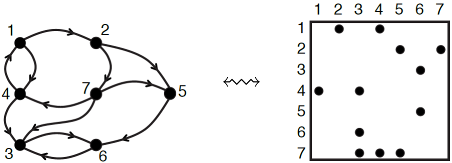
$$A := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

Important uses:

- Representations of linear maps $\mathbb{R}^n \to \mathbb{R}^m$.

$$\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + z \begin{bmatrix} \\ 1 \end{bmatrix} = \begin{bmatrix} x + y \\ y + z \end{bmatrix}$$

- Adjacency & incidence matrices of finite graphs.



Defining Matrices Formally

How to formally define notion of matrix? E.g.,

$$\mathsf{A} := \begin{bmatrix} 1 & 1 & \\ & 1 & 1 \end{bmatrix} \qquad \mathsf{B} := \begin{bmatrix} 1 & & & 1 \\ & -2 & & 3 \\ 7 & 7 & 7 & \end{bmatrix}$$

Definition (matrix)

An *$m \times n$-matrix over a ring* R is a map $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to R$.

Formal description of B above:

Domain. $\{1, 2, 3\} \times \{1, 2, 3, 4\}$

Codomain. $\mathbb{R}$ (or $\mathbb{Z}, \mathbb{Q}, \mathbb{C}$, etc.)

Graph.
$$\begin{array}{llll} \{((1,1),1), & ((1,2),\ \ 0), & ((1,3),0), & ((1,4),1), \\ \ ((2,1),0), & ((2,2),-2), & ((2,3),0), & ((2,4),3), \\ \ ((3,1),7), & ((3,2),\ \ 7), & ((3,3),7), & ((3,4),0)\ \} \end{array}$$

Matrix Operations

Suppose A and B are $m \times n$ matrices over a ring $\mathrm{R} = (R, +, \cdot, -, 0, 1)$.

Matrix Addition.

$$\mathrm{A} + \mathrm{B} \in R^{m \times n} \quad \Big| \quad (i, j) \mapsto \mathrm{A}(i, j) + \mathrm{B}(i, j),$$

Matrix Element-wise Multiplication.

$$\mathrm{A} \cdot \mathrm{B} \in R^{m \times n} \quad \Big| \quad (i, j) \mapsto \mathrm{A}(i, j) \cdot \mathrm{B}(i, j),$$

Now suppose A is $m \times n$ and B is $n \times p$.

Matrix Multiplication.

$$\mathrm{AB} \in R^{m \times p} \quad \Big| \quad (i, j) \mapsto \sum_{k=1}^{n} \mathrm{A}(i, k)\mathrm{B}(k, j).$$

Matrix Operations – Examples

Matrix Addition.

$$\begin{bmatrix} 1 & -1 & \\ 2 & & 7 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1+1 & -1+1 & 0+1 \\ 2+2 & 0+1 & 7+2 \end{bmatrix} = \begin{bmatrix} 2 & & 1 \\ 4 & 1 & 9 \end{bmatrix}$$

Matrix Element-wise Multiplication.

$$\begin{bmatrix} 1 & -1 & \\ 2 & & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1\cdot 1 & -1\cdot 1 & 0\cdot 1 \\ 2\cdot 2 & 0\cdot 1 & 7\cdot 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & \\ 4 & & 14 \end{bmatrix}$$

Matrix Multiplication.

$$\begin{bmatrix} 1 & -1 & \\ 2 & & 7 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1\cdot 1 - 1\cdot 1 + 0\cdot 1 & 1\cdot 2 - 1\cdot 1 + 0\cdot 2 \\ 2\cdot 1 + 0\cdot 1 + 7\cdot 1 & 2\cdot 2 + 0\cdot 1 + 7\cdot 2 \end{bmatrix} = \begin{bmatrix} & 1 \\ 9 & 18 \end{bmatrix}$$

From Matrices to Associative Arrays

Two shortcomings of matrices over fields/rings:

1. Useful operations between matrices not examples of previous matrix operations, but similar:
    - Element-wise maximum and minimum.
    - Path optimizations in graphs using min-max/max-min or min-plus/max-plus.

   $\Longrightarrow$ allow more general value structures.
2. Need to keep track of indices.
    - Adding rows/columns to make incompatible matrices compatible.
    - Extracting submatrices.
    - Partitioning matrix to make distributed.
    - Broadcasting ambiguity.

   $\Longrightarrow$ allow for more general indices, treat as sparse by default.

1. Background & Motivation
2. Associative Arrays – Basics
3. Associative Arrays in Practice
4. Associative Arrays – Theory

Semirings

Definition (semiring)

A *semiring* is a quintuple $\mathbb{V} = (V, \oplus, \otimes, 0, 1)$ where:

- $\oplus$ (*addition*) is associative, commutative, and has identity 0.

$$u \oplus (v \oplus w) = (u \oplus v) \oplus w, \quad u \oplus v = v \oplus u, \quad u \oplus 0 = 0 \oplus u = u$$

- $\otimes$ (*multiplication*) is associative, has identity 1, and has annihilator 0.

$$u \otimes (v \otimes w) = (u \otimes v) \otimes w, \quad u \otimes 1 = 1 \otimes u = u, \quad u \otimes 0 = 0 \otimes u = 0$$

- $\otimes$ distributes over $\oplus$.

$$u \otimes (v \oplus w) = (u \otimes v) \oplus (u \otimes w), \quad (v \oplus w) \otimes u = (v \otimes u) \oplus (w \otimes u)$$

Sometimes called "rigs" as they are "ri**n**gs" without **n**egatives.

Semirings – Examples

Example (reduct of a ring)

All rings become semirings after ‘forgetting’ negation. E.g.,

- Plus-times algebra $(\mathbb{R}, +, \cdot, 0, 1)$.

Example (bounded, distributive lattice)

$(L, \vee, \wedge, \inf(L), \sup(L))$ whenever $(L, \leq)$ is a bounded, distributive lattice. E.g.,

- Power set algebras $(\mathscr{P}(S), \cup, \cap, \emptyset, S)$.
- Bounded totally ordered sets $(L, \max, \min, \min(L), \max(L))$.

Example (tropical semiring)

- Max-plus algebra. $(\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$.
- Min-plus algebra. $(\mathbb{R} \cup \{\infty\}, \min, +, \infty, 0)$.

Semirings – Examples (continued)

Example (semirings of nonnegatives)

$(\mathbb{N}, +, \cdot, 0, 1)$, $(\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$, etc.

Example (string algebra)

$(\Sigma^\star \sqcup \{\infty\}, \frown, \min, \langle\rangle, \infty)$, where Σ is a totally ordered alphabet.

Nonexamples:

Example

$(\mathscr{P}(S), \cup, \cup, \emptyset, \emptyset)$.

Example

$(\Sigma^\star, \frown, \max, \langle\rangle, \langle\rangle)$, where Σ is a totally ordered alphabet.

Associative Arrays

Fix a semiring $\mathbb{V} = (V, \oplus, \otimes, 0, 1)$.

Definition (associative array)

An *associative array* is a function of the form $\mathbb{A} : I \times J \to V$ such that $\mathbb{A}$ has finite support — i.e., $\mathrm{card}(\{(i,j) \in I \times J \mid \mathbb{A}(i,j) \neq 0\}) < \infty$.

Benefits of finite support condition:

- Needed for associative array algebra well-definedness.
- May treat associative arrays as *extremely* sparse — given $\mathbb{A} : I \times J \to V$, we implicitly set

$$\mathbb{A}(i,j) := 0 \quad \text{whenever} \quad (i,j) \notin I \times J.$$

In effect, all associative arrays over $\mathbb{V}$ can be treated as having the same index sets.

Associative Arrays as Linear Maps

Given I, set $V^{\boxplus I} := \{\mathrm{v} \in V^I \mid \mathrm{v}(i) \neq 0$ for all but finitely many $i \in I\}$.

Definition (free V-semimodule)

The *free* V*-semimodule indexed by* I, $\mathrm{V}^{\boxplus I} = (V^{\boxplus I}, \oplus, \otimes, 0)$, is defined by setting, for all $u \in V$, $\mathrm{v}, \mathrm{w} \in V^{\boxplus I}$, and $i \in I$,

$$(\mathrm{v} \oplus \mathrm{w})(i) := \mathrm{v}(i) \oplus \mathrm{w}(i), \quad (u \otimes \mathrm{v})(i) := u \otimes \mathrm{v}(i), \quad 0(i) := 0.$$

Proposition

There is a one-to-one correspondence between linear maps $\varphi : \mathrm{V}^{\boxplus J} \to \mathrm{V}^{\boxplus I}$ *(i.e.,* $\varphi(\mathrm{v} \oplus \mathrm{w}) = \varphi(\mathrm{v}) \oplus \varphi(\mathrm{w})$ *and* $\varphi(u \otimes \mathrm{v}) = u \otimes \varphi(\mathrm{v})$ *for all* $u \in V$ *and* $\mathrm{v}, \mathrm{w} \in V^{\boxplus J}$*) and associative arrays of the form* $I \times J \to V$.

Associative Array Operations

Suppose $\mathrm{A} : I \times J \to V$ and $\mathrm{B} : I' \times J' \to V$ are associative arrays.

Array Addition.

$$\mathrm{A} \oplus \mathrm{B} : (I \cup I') \times (J \cup J') \to V \quad \Big| \quad (i,j) \mapsto \mathrm{A}(i,j) \oplus \mathrm{B}(i,j)$$

Array Element-wise Multiplication.

$$\mathrm{A} \otimes \mathrm{B} : (I \cap I') \times (J \cap J') \to V \quad \Big| \quad (i,j) \mapsto \mathrm{A}(i,j) \otimes \mathrm{B}(i,j)$$

Array Multiplication.

$$\mathrm{A} \oplus.\!\otimes \mathrm{B} : I \times J' \to V \quad \Big| \quad (i,j) \mapsto \bigoplus_{k \in J \cap I'} \mathrm{A}(i,k) \otimes \mathrm{B}(k,j)$$

Associative Array Operations – Examples

Work in $(\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$.

Array Addition.

$$\begin{array}{cccc} & a & b & d \\ a & 0 & -1 & 1 \\ b & 2 & & \\ c & & 1 & 0 \end{array} \max \begin{array}{cccc} & a & b & e \\ a & 1 & & 0 \\ c & 2 & 0 & \\ d & -1 & 2 & 1 \end{array}$$

$$= \begin{array}{ccccc} & a & b & d & e \\ a & \max(0,1) & \max(-1,-\infty) & \max(1,-\infty) & \max(-\infty,0) \\ b & \max(2,-\infty) & \max(-\infty,-\infty) & \max(-\infty,-\infty) & \max(-\infty,-\infty) \\ c & \max(-\infty,-1) & \max(1,2) & \max(0,-\infty) & \max(-\infty,1) \\ d & \max(-\infty,-1) & \max(-\infty,2) & \max(-\infty,-\infty) & \max(-\infty,1) \end{array}$$

$$= \begin{array}{ccccc} & a & b & d & e \\ a & 1 & -1 & 1 & 0 \\ b & 2 & & & \\ c & -1 & 2 & 0 & 1 \\ d & -1 & 2 & & 1 \end{array}$$

Associative Array Operations – Examples

Work in $(\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$.

Array Element-wise Multiplication.

$$\begin{array}{c|ccc} & a & b & d \\ \hline a & 0 & -1 & 1 \\ b & 2 & & \\ c & & 1 & 0 \end{array} + \begin{array}{c|ccc} & a & b & e \\ \hline a & 1 & & 0 \\ c & 2 & 0 & \\ d & -1 & 2 & 1 \end{array}$$

$$= \begin{array}{c|cc} & a & b \\ \hline a & 0+1 & -1+-\infty \\ c & -\infty+1 & 1+0 \end{array}$$

$$= \begin{array}{c|cc} & a & b \\ \hline a & 1 & \\ c & & 1 \end{array}$$

Associative Array Operations – Examples

Work in $(\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$.

Array Multiplication.

$$\begin{array}{cc} & \begin{array}{ccc} a & b & d \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \left[\begin{array}{ccc} 0 & -1 & 1 \\ 2 & & \\ & 1 & 0 \end{array}\right] \end{array} \max.+ \begin{array}{cc} & \begin{array}{ccc} a & b & e \end{array} \\ \begin{array}{c} a \\ c \\ d \end{array} & \left[\begin{array}{ccc} 1 & & 0 \\ 2 & 0 & \\ -1 & 2 & 1 \end{array}\right] \end{array}$$

$$= \begin{array}{cc} & \begin{array}{ccc} a & b & e \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \left[\begin{array}{ccc} \max(0+1, 1+-1) & \max(0+-\infty, 1+2) & \max(0+0, 1+1) \\ \max(2+1, -\infty-1) & \max(2+-\infty, -\infty+2) & \max(2+0, -\infty+1) \\ \max(-\infty+1, 0-1) & \max(-\infty+-\infty, 0+2) & \max(-\infty+0, 0+1) \end{array}\right] \end{array}$$

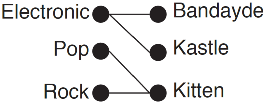
$$= \begin{array}{cc} & \begin{array}{ccc} a & b & e \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \left[\begin{array}{ccc} 1 & 3 & 2 \\ 3 & & 2 \\ -1 & 2 & 1 \end{array}\right] \end{array}$$

Associative Arrays & Big Data

Unified treatment of spreadsheets, databases, matrices, and graphs:

	Artist	Date	Duration	Genre
053013ktnA1	Bandayde	2013-05-30	5:14	Electronic
053013ktnA2	Kastle	2013-05-30	3:07	Electronic
063012ktnA1	Kitten	2010-06-30	4:38	Rock
082812ktnA1	Kitten	2012-08-28	3:25	Pop

	Bandayde	Kastle	Kitten
Electronic	1	1	
Pop			1
Rock			1

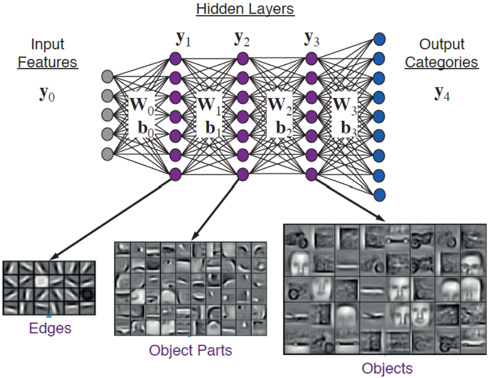


$$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & 0.5 & 1.1 \\ 2 & 1.1 & 0.5 \end{array}$$

Additional benefits:

- Supports distributed storing, no need to record partitioning details.
- Sparsity built into the very concept.
- Meaningful, nonarbitrary indices may be used, aiding comprehension.

Associative Arrays & Machine Learning



Inference Step/Forward Propagation:

$$\mathrm{y}_{k+1} = h(\mathrm{W}_k \mathrm{y}_k + \mathrm{b}_k) = \max(\mathrm{W}_k \mathrm{y}_k + \mathrm{b}_k, 0) = \mathrm{W}_k \mathrm{y}_k \otimes \mathrm{b}_k \oplus 0,$$

where $\oplus$, $\otimes$ correspond to max-plus semiring.

Nullspaces

Definition (nullspace)

The *nullspace* of an $m \times n$-matrix A is the set $\{\mathrm{v} \in V^n \mid \mathrm{Av} = 0\}$.

Definition (zero-sum free)

A semiring V is *zero-sum free* if $x \oplus y = 0$ if and only if $x = y = 0$.

Theorem

If A *is an* $m \times n$*-matrix over a zero-sum-free commutative semiring* V*, then the nullspace of* A *is given by* $N_1 \times N_2 \times \cdots \times N_n$*, where for each* $j \in \{1, 2, \ldots, n\}$ *the subset* $N_j \subseteq V$ *is defined by*

$$N_j := \left\{ v \in V \;\middle|\; v \otimes \bigoplus_{i=1}^{n} \mathrm{A}(i,j) = 0 \right\}.$$

Linear Systems

Definition (join-blank semiring)

A semiring V is a *join-blank semiring* if:

- $\oplus$ is the join operation $\vee$ for a complete lattice order $\leq$ on V.
- $\otimes$ satisfies $v \otimes \bigvee U = \bigvee \{v \otimes u \mid u \in U\}$ for any $v \in V$ and $U \subseteq V$.

Let V be a join-blank semiring and $\leq$ its complete lattice ordering. Given $\mathrm{A} \in V^{m \times n}$ and $\mathrm{w} \in V^m$, write $X(\mathrm{A}, \mathrm{w}) := \{\mathrm{v} \in V^n \mid \mathrm{Av} = \mathrm{w}\}$.

Theorem (Join-Blank Structure Theorem)

Suppose $\mathrm{A} \in V^{m \times n}$ *and* $\mathrm{w} \in V^m$*. If* $X(\mathrm{A}, \mathrm{w}) \neq \emptyset$*, then* $\mathrm{x} := \max(X(\mathrm{A}, \mathrm{w}))$ *exists and there is a subset* $U \subseteq V^m$ *such that*

$$X(\mathrm{A}, \mathrm{w}) = \bigcup_{\mathrm{u} \in U} [\mathrm{u}, \mathrm{x}] = \bigcup_{\mathrm{u} \in U} \{\mathrm{v} \in V^n \mid \mathrm{u} \leq \mathrm{v} \leq \mathrm{x}\}.$$

Proving the Join-Blank Structure Theorem

Lemma

$X(\mathrm{A}, \mathrm{w})$ *is closed under taking suprema of nonempty subsets.*

Proof.

Suppose $\emptyset \neq U \subseteq X(\mathrm{A}, \mathrm{w})$. Infinite distributivity implies $u \otimes - : V \to V$ is order-preserving for every $u \in V$. Then

$$\begin{aligned}\left(\mathrm{A} \bigvee U\right)(i) &= \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes (\bigvee U)(j) = \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes \bigvee_{\mathrm{v} \in U} \mathrm{v}(j) \\ &= \bigvee_{j=1}^{m} \bigvee_{\mathrm{v} \in U} \mathrm{A}(i,j) \otimes \mathrm{v}(j) = \bigvee_{\mathrm{v} \in U} \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes \mathrm{v}(j) \\ &= \bigvee_{\mathrm{v} \in U} \mathrm{w}(i) = \mathrm{w}(i).\end{aligned}$$

Thus, $\bigvee U \in X(\mathrm{A}, \mathrm{w})$. □

Proving the Join-Blank Structure Theorem

Lemma

$X(\mathrm{A},\mathrm{w})$ *is convex, i.e., if* $\mathrm{v}_1 \leq \mathrm{v}_2 \leq \mathrm{v}_3$ *and* $\mathrm{v}_1, \mathrm{v}_3 \in X(\mathrm{A},\mathrm{w})$, *then* $\mathrm{v}_2 \in X(\mathrm{A},\mathrm{w})$.

Proof.

Infinite distributivity implies $u \otimes - : V \to V$ is order-preserving for every $u \in V$. Implies for all i, j:

$$\mathrm{A}(i,j) \otimes \mathrm{v}_1(j) \leq \mathrm{A}(i,j) \otimes \mathrm{v}_2(j) \leq \mathrm{A}(i,j) \otimes \mathrm{v}_3(j).$$

Taking suprema over j gives

$$\mathrm{w}(i) = \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes \mathrm{v}_1(j) \leq \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes \mathrm{v}_2(j) \leq \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes \mathrm{v}_3(j) = \mathrm{w}(i),$$

so $\mathrm{w}(i) = \bigvee_{j=1}^{m} \mathrm{A}(i,j) \otimes \mathrm{v}_2(j)$ for every i. Thus, $\mathrm{v}_2 \in X(\mathrm{A},\mathrm{w})$. □

Proving the Join-Blank Structure Theorem

Proof of the Join-Blank Structure Theorem.

If $X(\mathrm{A},\mathrm{w}) \neq \emptyset$, then Lemma 1 implies

$$\mathrm{x} := \bigvee X(\mathrm{A},\mathrm{w}) \in X(\mathrm{A},\mathrm{w}).$$

Thus, $\mathrm{x} = \max(X(\mathrm{A},\mathrm{w}))$.
Taking $U := X(\mathrm{A},\mathrm{w})$ and any $\mathrm{v} \in U$, Lemma 2 implies $[\mathrm{v},\mathrm{x}] \subseteq X(\mathrm{A},\mathrm{w})$.
Hence,

$$\bigcup_{\mathrm{v}\in U} [\mathrm{v},\mathrm{x}] \subseteq X(\mathrm{A},\mathrm{w}).$$

Reverse inclusion follows since $\mathrm{v} \in [\mathrm{v},\mathrm{x}]$, so

$$X(\mathrm{A},\mathrm{w}) = \bigcup_{\mathrm{v}\in U} [\mathrm{v},\mathrm{x}].$$

☐

Linear Systems – Max-Blank Structure Theorem

Definition (max-blank semiring)

A *max-blank semiring* is a join-blank semiring whose complete lattice ordering is total (i.e., any two elements are comparable).

Let V be a max-blank semiring and $\leq$ its complete total ordering.

Theorem (Max-Blank Structure Theorem)

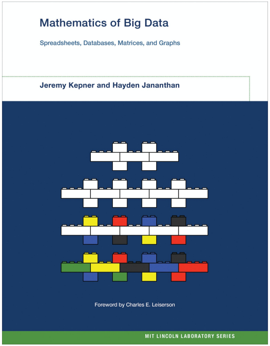
Suppose $\mathrm{A} \in V^{m\times n}$ *and* $\mathrm{w} \in V^m$. *Assume that sets of the form*

$$\{v \in V \mid \mathrm{A}(i,j) \otimes v = \mathrm{w}(i)\}$$

are closed intervals whenever nonempty. If $X(\mathrm{A},\mathrm{w}) \neq \emptyset$, *then* $X(\mathrm{A},\mathrm{w})$ *is a finite union of closed intervals with common terminal point.*

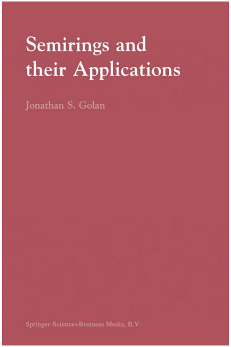
Further Reading

Mathematics of Big Data
Kepner & Jananthan
(MIT Press)



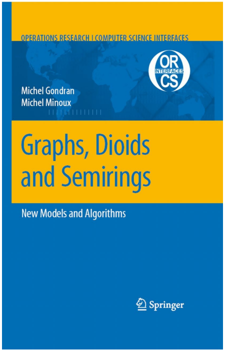
Mathematics of Big Data
Spreadsheets, Databases, Matrices, and Graphs
Jeremy Kepner and Hayden Jananthan
Foreword by Charles E. Leiserson
MIT LINCOLN LABORATORY SERIES

Semirings and their Applications
Golan
(Springer)



Semirings and their Applications
Jonathan S. Golan
Springer-Science+Business Media, B.V.

Graphs, Dioids and Semirings
Gondran & Minoux
(Springer)



OPERATIONS RESEARCH | COMPUTER SCIENCE INTERFACES
Michel Gondran
Michel Minoux
Graphs, Dioids and Semirings
New Models and Algorithms
Springer

Math of Big Data Summer UROP

Email Jeremy (`kepner@ll.mit.edu`) or Hayden (`haydenj@mit.edu`) if you're interested in exploring the mathematical foundations and applications of associative arrays and semirings!

Some recent previous student work:

- *Testing RadiX-Nets: Advances in Viable Sparse Topologies* (Kevin Kwak, Zack West)
- *Fuzzy Relational Databases via Associative Arrays* (Kevin Min)
- *From Bit to Insights: Exploring Network Traffic, Traffic Matrices, and Heavy-Tailed Data* (Christopher Howard)
- *Algebraic Conditions on One-Step Breadth-First Search* (Emma Fu)