Tropical Algebras, Idempotent Analysis and System Theory
A Story of Continuous Mathematics in The Limit

Amir Daneshgar

Department of Mathematical Sciences
Sharif University of Technology

http://www.sina.sharif.ir/~daneshgar
daneshgar@sharif.ir

Bahman 1393 (February, 2015)
Prologue: Good Mathematics Over Semirings!
1. Prologue: Good Mathematics Over Semirings!
2. Semirings: Rings and Dioids
1. Prologue: Good Mathematics Over Semirings!
2. Semirings: Rings and Dioids
3. A Couple of Examples
1 Prologue: Good Mathematics Over Semirings!
2 Semirings: Rings and Dioids
3 A Couple of Examples
4 Tropical Algebra/Geometry
1. Prologue: Good Mathematics Over Semirings!
2. Semirings: Rings and Dioids
3. A Couple of Examples
4. Tropical Algebra/Geometry
5. Idempotent Analysis
1 Prologue: Good Mathematics Over Semirings!

2 Semirings: Rings and Dioids

3 A Couple of Examples

4 Tropical Algebra/Geometry

5 Idempotent Analysis

6 A Nonlinear System Theory
[D. Hilbert, 1900]

“Mathematical science is in my opinion an indivisible whole, an organism whose vitality is conditioned upon the connection of its parts.”
Prologue:
Good Mathematics
Over Semirings!
Good Mathematics Over Semirings!


“As a mathematical discipline travels far from its empirical source, or still more, if it is a second and third generation only indirectly inspired from ideas coming from ‘reality’, it is beset with very grave dangers. ... But there is a grave danger that the subject will develop along the line of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities.”
[T. Tao, What is Good Mathematics 2007]
“However, there is the remarkable phenomenon (ref. Wigner unreasonable effectiveness of mathematics) that good mathematics in one of the above senses tends to beget more good mathematics in many of the other senses as well, leading to the tentative conjecture that perhaps there is, after all, a universal notion of good quality mathematics, and all the specific metrics listed above represent different routes to uncover new mathematics, or different stages or aspects of the evolution of a mathematical story.”
Good Mathematics Over Semirings!

General References for the Origins:

- Vandiver H. S.; *Note on a simple type of algebra in which cancellation law of addition does not hold*, Bull. Amer. Math. Soc. 40 (1934), 914-920.
<table>
<thead>
<tr>
<th>Category</th>
<th>Author(s)</th>
<th>Year(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy set theory</td>
<td>Lotfi AskarZadeh</td>
<td>1965,</td>
</tr>
<tr>
<td></td>
<td>Joseph Goguen</td>
<td>1968/9</td>
</tr>
<tr>
<td></td>
<td>Wolfgang Wechler</td>
<td>1983,</td>
</tr>
<tr>
<td></td>
<td>Jonathan S, Golan</td>
<td>1999</td>
</tr>
<tr>
<td>Theory of Computation</td>
<td>Samuel Eilenberg</td>
<td>1974,</td>
</tr>
<tr>
<td></td>
<td>Arto Salomaa</td>
<td>1978,</td>
</tr>
<tr>
<td></td>
<td>Jan A. Bergstra, Jan W. Klop</td>
<td>1984,</td>
</tr>
<tr>
<td>Discrete Optimization</td>
<td>Raymond A. Cuninghame-Green</td>
<td>1979,</td>
</tr>
<tr>
<td></td>
<td>Michel Gondran, Michel Minoux</td>
<td>1984,</td>
</tr>
<tr>
<td></td>
<td>Alexander I. Barvinok</td>
<td>1993</td>
</tr>
<tr>
<td>Idempotent Analysis</td>
<td>Victor P. Maslov, Grigory L. Litvinov</td>
<td>1986,</td>
</tr>
<tr>
<td>Tropical Algebra/Geometry</td>
<td>Imre Simon</td>
<td>1988,</td>
</tr>
</tbody>
</table>
The mysterious duality!

Fuzzy mathematics \( \overset{\text{op}}{\leadsto} \text{Lim} \) \( \overset{\text{Lim}}{\leadsto} \) Conventional Mathematics
The mysterious duality!

Fuzzy mathematics \( \overset{\text{op}}{\Longleftrightarrow} \text{Lim} \) Conventional Mathematics

- \( \infty \)

- 0
The mysterious duality!

\[
\begin{align*}
\text{Fuzzy mathematics} & \quad \text{op} \quad \text{Conventional Mathematics} \\
\{-\infty\} & \quad \sim \quad \text{Lim} \quad \{1\}
\end{align*}
\]
The mysterious duality!

Fuzzy mathematics \( \text{op} \) Conventional Mathematics

\[
\begin{align*}
\{ & -\infty \} & \begin{array}{c}
0 \\
\geq
\end{array} & \begin{array}{c}
0 \\
1
\end{array} \\
\{ & \}
\end{align*}
\]
The mysterious duality!
The mysterious duality!

Fuzzy mathematics

\{ -\infty, 0, \geq \}

toll set

\text{op}

Conventional Mathematics

\{ 0, 1 \}

function

\Leftrightarrow \text{Lim}
The mysterious duality!

Fuzzy mathematics

-∞
+
≥
dioid
toll set

op

Conventional Mathematics

0
1
×
+ 
ring
function
Good Mathematics Over Semirings!

- Is there any difference between mathematics over fields and mathematics over semirings?

Let’s talk about this!
Semirings,

Rings and Dioids
Definition of a semiring

A **semiring** \((S, \oplus, \otimes, 1, 0)\) is an algebraic structure which satisfies the following properties:

1. \(\oplus\) and \(\otimes\) are associative binary operations on \(S\),
2. \((S, \oplus, 0)\) is a commutative monoid with identity \(0\),
3. \((S, \otimes, 1)\) is a monoid with identity \(1\).
4. For all \(a, b, c \in S\) we have
   \[a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)\] and \[(b \oplus c) \otimes a = (b \otimes a) \oplus (c \otimes a)\].
5. For all \(a \in S\) we have \(a \otimes 0 = 0 \otimes a = 0\).
Subclasses of semirings

The class of semirings can be naturally subdivided into two disjoint sub-classes depending on whether the operation $\oplus$ satisfies one of the following two properties:

1. The operation $\oplus$ endows the set $D$ with a group structure;
2. The operation $\oplus$ endows the set $D$ with a canonically ordered monoid structure.

**note**

(1) and (2) cannot be satisfied simultaneously. In case (1), we are led to the well-known **Ring** structure, and in case (2) we are led to the **Dioid** structure.
A **Dioid** is a quintuple \((D, \lor, \otimes, 1, 0)\) that satisfies the following conditions:

1. \((D, \lor)\) is a commutative monoid with neutral element \(0\).
2. \((D, \otimes)\) is a monoid with neutral element \(1\).
3. for all \(d \in D\) we have \(d \otimes 0 = 0 \otimes d = 0\).
4. \(\otimes\) is (right and left) distributive with respect to \(\lor\).
5. \(\lor\) induces an order defined as:
   \[ a \leq b \iff \exists c, \ b = a \lor c. \]

**note**

A **dioid** is a semiring \((D, \lor, \otimes, 1, 0)\) in which the operation \(\lor\) induces a natural order structure.
Idempotent dioids

**note**

In general in a dioid \((D, \vee, \otimes, 1, 0)\) the canonical order is **compatible** with the operations \(\vee\) and \(\otimes\).

A dioid is **idempotent** iff \(\forall a \in D, \ a \vee a = a\).

Note that this implies the following fact,

\[
a \leq b \iff \exists c, \ b = a \vee c \iff a = a \vee b.
\]

**note**

Sometimes this condition is added to the **definition of a dioid** as an axiom.
A Couple of
Motivating Examples
Examples of dioids

Ideals of a commutative ring

- \((\mathbb{N}, +, \times)\)
- \((\mathbb{N} \cup \{+\infty\}, \min, +)\)
- \((\mathbb{Z}, +, \times)\)
- \((\mathbb{N}, \text{lcm}, \gcd)\)

\[\mathbb{R}_{\min} \overset{\text{def}}{=} (\mathbb{R} \cup \{+\infty\}, \min, +)\]
\[\mathbb{R}_{\max} \overset{\text{def}}{=} (\mathbb{R} \cup \{-\infty\}, \max, +)\]
\[\overline{\mathbb{R}} \overset{\text{def}}{=} (\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}, \max, \min)\]

- \((\text{Reg}(\Sigma), \cup, \text{concat})\)
- \((2^{\mathbb{R}^2}, \cup, +)\)

\[A + B \overset{\text{def}}{=} \{y + z : y \in A, \& z \in B\}.\]

R. Dedekind 1894
H. S. Vandiver 1934
A dioid which is not idempotent
Tropical semiring
A ring but not a dioid.
A doubly idempotent dioid.
Min-plus (optimization) algebra
Max-plus algebra.
An idempotent dioid.
A dioid (used in Computer Science).
A dioid (used in Integral Geometry)

Has mathematics gone wild for a century?!
Polynomials in $\mathbb{R}_{\min} [x, y] \overset{\text{def}}{=} (\mathbb{R} \cup \{+\infty\}, \oplus \overset{\text{def}}{=} \min, \otimes \overset{\text{def}}{=} +) [x, y]$

A monomial

$ax^i y^j \overset{\text{def}}{=} a \otimes x \ldots \otimes x \otimes y \ldots \otimes y = a + ix + jy$.

A polynomial

$ax^2 + bxy + e = \min(a + 2x, b + x + y, e)$.

Neutrals

$\infty + x = 0x = x$.

Fact:

The tropical polynomials in $n$ variables $x_1, \ldots, x_n$ are precisely the piecewise-linear concave functions on $\mathbb{R}^n$ with integer coefficients.
Polynomials in $\mathbb{R}_{\min}[x, y] \overset{\text{def}}{=} (\mathbb{R} \cup \{+\infty\}, \oplus \overset{\text{def}}{=} \min, \otimes \overset{\text{def}}{=} +)[x, y]$

Freshman’s dream!

$$3 \min(x, y) = (x \oplus y)^3 = x^3 \oplus y^3 = \min(3x, 3y).$$

More surprises!

$$(x + 1)^2 = x^2 + 17x + 2 = x^2 + 1x + 2.$$  

Fundamental theorem of algebra

- Every tropical polynomial function can be written uniquely as a tropical product of tropical linear functions.
- Unique factorization of tropical polynomials holds in one variable, but it no longer holds in two or more variables.
Polynomials in \( \mathbb{R}_{\min}[x, y] \) def \((\mathbb{R} \cup \{+\infty\}, \oplus \text{ def } \min, \otimes \text{ def } +)[x, y] \)

Cubic polynomials in one variable

Let \( b - a \leq c - b \leq d - c \), then

\[
ax^3 + bx^2 + cx + d = a(x + (b - a))(x + (c - b))(x + (d - c)).
\]
Example

For a matrix

\[
A = \begin{pmatrix}
3 & 7 \\
2 & 4
\end{pmatrix}
\]

in \( \text{Mat}_{n \times n}(\mathbb{R}_{\text{max}}) \) we have

\[
A^2 = \begin{pmatrix}
\max(3 + 3, 7 + 2) & \max(3 + 7, 7 + 4) \\
\max(2 + 3, 4 + 2) & \max(2 + 7, 4 + 4)
\end{pmatrix}
= \begin{pmatrix}
9 & 11 \\
6 & 9
\end{pmatrix}.
\]

In linear algebra we have

\[
A^2 = \begin{pmatrix}
+(3 \times 3, 7 \times 2) & +(3 \times 7, 7 \times 4) \\
+(2 \times 3, 4 \times 2) & +(2 \times 7, 4 \times 4)
\end{pmatrix}
= \begin{pmatrix}
23 & 49 \\
14 & 30
\end{pmatrix}.
\]
The shortest path problem

Let $A \in Mat_{n \times n}(S)$ be the adjacency matrix of the weighted directed graph $G$ on the vertex set $\{v_1, v_2, \ldots, v_n\}$. Then the shortest path problem is equal to find the matrix $D = [d_{ij}]$ such that $d_{ij}$ is the minimum of the weights of all paths starting from $v_i$ and ending at $v_j$. Note that in this setting we have,

$$D = A^* = I + A + A^2 + \cdots + A^k + \cdots = A^{n-1}. $$

Metrics

$A$ is said to be a metric if $A^2 = A$. (Note that this is the triangle inequality!)
The median filter on $\mathbb{R}^Z_3$

Let

$$\mathbb{R}^Z_3 \ni A = \{(0, x_1), (1, x_2), (2, x_3)\},$$

and let $M$ be the median filter acting as

$$M(A) = \{(0, m), (1, m), (2, m)\},$$

where $m = \text{Median}(x_1, x_2, x_3)$.

Note that $m = \max(\min(x_1, x_2), \min(x_2, x_3), \min(x_3, x_1))$, which is a fuzzification of the Boolean median expression $x_1x_2 + x_2x_3 + x_3x_1$. 
Tropical Algebra/Geometry
Tropical Geometry

General references for further reading:

- Speyer David; *Tropical Geometry*, UC Berkeley, 2005.
Recall the definition of a tropical polynomial

\[ \sum_{i_1, \ldots, i_m} a_i x_1^{i_1} \ldots x_m^{i_m} \overset{\text{def}}{=} \sum_i a_i \mathbf{x}^i. \]

Example

- \[ x^3 + 2x^2 + 6x + 11 = \min(3x, 2 + 2x, 6 + x, 11). \]
- \[ ax + by + c = \min(a + x, b + y, c). \]

Convention!

Violet means Tropical!
Prologue
Semirings
Examples
Tropical Algebra/Geometry
Idempotent Analysis
A Nonlinear System Theory

Roots of a tropical polynomial

Definition
A root of a tropical polynomial $F(x) = \sum_i a_i x^i$ is a point $x_0$ such that the minimum is attained in at least two monomials.
The set of roots is denoted by $V(F)$.

Example
$V(x^3 + 2x^2 + 6x + 11) = V(\min(3x, 2 + 2x, 6 + x, 11)) = \{2, 4, 5\}$. 
Assume that $b - a \leq c - b \leq d - c$, then

$$ax^3 + bx^2 + cx + d = a(x + (b - a))(x + (c - b))(x + (d - c)).$$
Roots of a tropical polynomial

Roots of an ordinary polynomial

What are the roots of \(x^2 - x + t = t^0 x^2 - t^0 x + t^1\)?

They are, \(x_1 = \sum_{k=1}^{\infty} \frac{1}{k+1} \binom{2k}{k} t^k\), and \(x_2 = 1 - x_1\).

Roots of the tropicalization

What are the roots of \(x^2 + x + 1 = \min(0 + 2x, 0 + x, 1)\)?

They are \(x_1 = 1\), and \(x_2 = 0\).
Field of Puissuex series, $P[t]$

Let $\mathbb{K}$ be an algebraic closed field of characteristic 0 (e.g. $\mathbb{C}$).

1. **Elements:**

   $$p(t) = \sum_{\tau \in \mathbb{Q}_{\geq 0}, c_\tau \in \mathbb{K}} c_\tau t^\tau,$$

   where summation is over well-ordered subsets of $\mathbb{Q}$.

2. **The set of such series, $P[t]$, with natural operations is also an algebraic closed field.**

3. **What is the dominant term in $p(t)$ when $t \to 0$?**
The order valuation $\nu$ on $P[t]$ is defined as

$$\nu(p) \overset{\text{def}}{=} \min\{\tau \in \mathbb{Q}_{\geq 0} : c_\tau \neq 0\}.$$  

Note that,

1. $\nu(pq) = \nu(p) + \nu(q)$.
2. $\nu(p + q) \geq \min(\nu(p), \nu(q))$.
3. Note that when $t \to 0$ the dominant term in $p(t)$ is $c_{\nu(p(t))} t^{\nu(p(t))}$. 
Given an ordinary polynomial $F(x) = \sum_i a_i x^i$ with $a_i \in P[t]$ and its tropicalization $\tilde{F}(x) = \sum_i \nu(a_i) x^i$, then $\nu(V(F)) = V(\tilde{F})$.

$\nu$ as a limit process!

Note that $\nu(a) = \lim_{t \to 0^+} \frac{\log |a(t)|}{\log t} = \lim_{t \to 0^+} \log_t |a(t)|$. 
Example

Valuation of the roots

What are the roots of \( x^2 - x + t = t^0 x^2 - t^0 x + t^1 \)?

They are, \( \nu(x_1) = \nu\left( \sum_{k=1}^{\infty} \frac{1}{k+1} \binom{2k}{k} t^k \right) = 1 \), and

\( \nu(x_2) = \nu(1 - x_1) = 0 \).

Roots of the tropicalization

What are the roots of \( x^2 + x + 1 = \min(0 + 2x, 0 + x, 1) \)?

They are \( x_1 = 1 \), and \( x_2 = 0 \).
The Log map

\[ \log_t : (\mathbb{C}^*)^2 \to \mathbb{R}^2, \quad -\log_t (x, y) = (-\log_t |x|, -\log_t |y|). \]

The amoebas \(-\log_t(V(x + y + 1))\).

a) \(\text{Log}(\mathcal{L})\)  
b) \(\text{Log}_{t_1}(\mathcal{L})\)  
c) \(\text{Log}_{t_2}(\mathcal{L})\)  
d) \(\lim_{t \to \infty} \text{Log}_t(\mathcal{L})\)

Courtesy of E. Brugallé 2013.
Note
\[ \nu(V(x + y - 1)) = \nu(V(t^0 x + t^0 y + t^0)) = V(x + y + 0). \]

The amoebas \( \log_t(V(x + y + 0)) \).

There is more!
For more on this see Bergman, Maslov, Mikhalkin, Passare and Rullgård, contributions on deformations and dequantization.
Why tropical algebra/geometry?

1. There is a nice geometry (e.g. a nice intersection theory).

2. Mikhalkin’s approach to compute Gromov-Witten invariants through counting tropical curves. Also, Gathmann and Markwig’s tropical/combinatorial proof of Kontsevich recursive formula for this.

3. Multivalued algebra and Viro’s tropical groups.

4. Effective algorithms in tropical settings to compute important applied and theoretical problems (e.g. discriminants and geometry of polytops and varieties).

Tropical algebra/geometry is algebraic geometry in the limit!
This is our first example of this limit process!
We talk more about this!
Idempotent Analysis, Dequantization and Correspondence Principle
Dequantization and Correspondence Principle

**General references for further reading:**

- Litvinov Grigory L.; *Idempotent/tropical analysis, the Hamilton-Jacobi and Bellman equations*, Available at arXiv:1203.0522.
Correspondence principle

Litvinov’s formulation for the limit process. Let’s delve into more details!

Courtesy of Litvinov 2010.
Maslov’s dequantization

Deformation of $\mathbb{R}$

Consider $(\mathbb{R}^+, +, \times)$ and let $h$ be a positive (i.e. Planck’s) constant. Then define $(\mathbb{R}_h, \oplus_h, \otimes_h)$ to be the image of the following map for which the operations $\oplus_h$ and $\otimes_h$ are defined in such a way that make it a natural map, i.e.,

$$\pi_h : \mathbb{R}^+ \to \mathbb{R}_h \quad u \mapsto w = \pi_h(u) \overset{\text{def}}{=} h \ln u,$$

where

$$x \oplus_h y \overset{\text{def}}{=} h \ln(e^{\frac{x}{h}} + e^{\frac{y}{h}}),$$

and

$$x \otimes_h y \overset{\text{def}}{=} h \ln(e^{\frac{x}{h}} \times e^{\frac{y}{h}}) = x + y.$$
Maslov’s dequantization map

\[ w = \hbar \ln u \]

Courtesy of Litvinov 2010.
Maslov’s dequantization

**Identities**

Note that when $h \to 0^+$ then $\mathbb{R}_h \to \mathbb{R}_{\text{max}} \overset{\text{def}}{=} (\mathbb{R} \cup \{-\infty\}, \max, +)$. Clearly in this setting,

$$\pi(1) = 0 \quad \text{and} \quad \pi(0) = -\infty.$$ 

**Aboebas**

Note that a change of variable $h \to \frac{1}{\ln t}$ changes the map $\pi$ to $\pi_t(u) = \nu(u) = \log_t u$ which is the amoebas map.

**Hence**, $\mathbb{R}^+$ can be thought of as a quantization of $\mathbb{R}_{\text{max}}$!
What happens if we dequantize classical dynamics?

Let’s try to find the image in the limit under Maslov’s dequantization.
Consider the heat equation and apply the transform \( w = h \ln u \),

\[
\frac{\partial u}{\partial t} = \frac{h}{2} \frac{\partial^2 u}{\partial x^2} \quad \Rightarrow \quad \frac{\partial w}{\partial t} - \frac{1}{2} \left( \frac{\partial w}{\partial x} \right)^2 - \frac{h}{2} \frac{\partial^2 w}{\partial x^2} = 0,
\]

and pass to the limit \( h \to 0^+ \) to get,

\[
\frac{\partial w}{\partial t} = \frac{1}{2} \left( \frac{\partial w}{\partial x} \right)^2,
\]

which is (a special case) of the Hamilton-Jacobi-Bellman equation, with linear properties for solutions in \( \mathbb{R}_{min} \).
Again consider the discrete time decision (control) process,

\[ x(0) = x_0, \quad x(t + h) = x(t) - u(t), \]

along with the cost function

\[
\min_u \left( \Phi(x(T)) + \sum_{i=0}^{T/h-1} \frac{u(ih)^2}{2h} \right).
\]

This can also be solved through the dynamic programming equation

\[
v(t, x) = \min_u \left( \frac{u^2}{2h} + v(t + h, x - u) \right), \quad v(T, .) = \Phi.
\]
The Brownian decision process II

Use the change of control $u = hw$ in the dynamic programming equation and let $h \mapsto 0$ to get

$$\frac{\partial v}{\partial t} + \min_w \left( -w \frac{\partial v}{\partial x} + \frac{w^2}{2} \right) = 0, \quad v(T, .) = \Phi.$$ 

i.e. equivalent to the Hamilton-Jacobi-Bellman equation,

$$\frac{\partial v}{\partial t} = \frac{1}{2} \left( \frac{\partial v}{\partial x} \right)^2, \quad v(T, .) = \Phi.$$ 

**note**

In this sense the theory of Brownian decision processes is a discretized dequantizations of the theory of heat equations.
Consider a very simple discrete decision (control) process,

\[ x(0) = x_0, \quad x(n+1) = x(n) - u(n), \]

along with the cost function

\[
\text{Cost}(N) = \min_{u(0), u(1), \ldots, u(N-1)} \left( \varphi(x(N)) + \sum_{i=0}^{N-1} c(u(i)) \right).
\]

note

Functions \(c\) and \(\varphi\) satisfy typical conditions as being convex, lower-semicontinuous, and zero at their minimum, etc.
**Definition**

The inf-convolution $f \boxtimes g$ is defined as $(f \boxtimes g)(z) \mapsto \inf_{x+y=z} f(x) + g(y)$.

Example: For any $m \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ define

$$Q_{m,\sigma}(x) \overset{\text{def}}{=} \frac{1}{2} \left( \frac{x - m}{\sigma} \right)^2 \quad \text{for } \sigma \neq 0,$$

$$Q_{m,0}(x) \overset{\text{def}}{=} \delta_m(x) \overset{\text{def}}{=} \begin{cases} 0 & x = m \\ +\infty & \text{otherwise}. \end{cases}$$

Then,

$$(Q_{m_1,\sigma_1} \boxtimes Q_{m_2,\sigma_2})(z) = Q_{m_1 + m_2, \sqrt{\sigma_1 + \sigma_2}}(z).$$
Dynamic programming II

Define

\[
v(n, x) \overset{\text{def}}{=} \min_{u(n), u(1), \ldots, u(N-1)} \left( \varphi(x(N)) + \sum_{i=n}^{N-1} c(u(i)) \mid x(n) = x \right),
\]

and note that it satisfies the dynamic programming equation

\[
v(n, x) = \min_u (c(u) + v(n + 1, x - u)), \quad v(N, x) = \varphi(x).
\]

This equation can be written as follows using inf-convolution \( \boxdot \),

\[
v(n, \cdot) = c \boxdot v(n + 1, \cdot), \quad v(N, \cdot) = \varphi.
\]

note

Hence, the solution is \( v(0, \cdot) = c^N \boxdot \varphi \).
Discrete dynamics in the limit!

Hence dynamics in the discrete case is related to computing iterations (i.e. powers) of an operator.

Let’s concentrate on this in the dequantized world mimicking linear operators!
Why mimicing Linear algebra in $\mathbb{R}_{\text{max}}$

A large number of results of conventional linear algebra can be extended to $\mathbb{R}_{\text{max}}$, and even to matrices over an arbitrary semiring, as

- The Cayley-Hamilton theorem.
- Determinants
- Polynomial functions
- Formal series
Let $S$ be a semiring and $A \in Mat_{n \times n}(S)$ be a matrix with entries in $S$. Then one may naturally identify $A$ with a weighted directed graph structure, on $n$ vertices, where there is a directed edge from the vertex $u$ to the vertex $v$ of weight $\epsilon \neq a_{uv} \in S$ if and only if this value is the entry in the row $u$ and the column $v$ of the matrix $A$.

**note**

Standard notions as paths, cycles and their weights are defined naturally.
The shortest path problem

**definition**

For any matrix $A$ in $(D, \oplus, \otimes)$ we define $A^*$ as follow

$$A^* \overset{\text{def}}{=} I \oplus A \oplus A^2 \oplus \cdots \oplus A^k \oplus \cdots$$

**Questions**

When does this limit $A^*$ *exist*?
How can we *compute* it?
Let $S = \mathbb{R}_{\min} \overset{\text{def}}{=} (\mathbb{R} \cup \{+\infty\}, \land, +)$ and $A \in \text{Mat}_{n \times n}(S)$ be the adjacency matrix of the weighted directed graph $G$ on the vertex set $\{v_1, v_2, \ldots, v_n\}$. Then the shortest path problem is equal to find the matrix $D = [d_{ij}]$ such that $d_{ij}$ is the minimum of the weights of all paths starting from $v_i$ and ending at $v_j$. Note that in this setting we have,

$$D = A^* = I \land A \land A^2 \land \cdots \land A^k \land \cdots$$

**Questions**

When does this limit $A^*$ exist?
How can we compute it?
A simple optimization problem

Let $S = \mathbb{R}_{\max} \overset{\text{def}}{=} (\mathbb{R} \cup \{-\infty\}, \lor, +)$ and $A \in \text{Mat}_{n \times n}(S)$ be the adjacency matrix of the weighted directed graph $G$ on the vertex set \{v_1, v_2, \ldots, v_n\}. Interpret $a_{ij}$ as the profit of going from $v_i$ to $v_j$ and let $f_i \in \mathbb{R}$ be a terminal prize of ending at vertex $v_i$. Therefore, the solution of maximizing the income after $k$ steps is equal to $D_k = A^k f$. Also, an overall maximizing solution is

$$D = A^* f.$$ 

Questions

When does this limit $A^* f$ exists?
How can we compute it?
Simple equations in $\mathbb{R}_{\max}$

Let $\mathbb{R}_{\max} \overset{\text{def}}{=} (\mathbb{R} \cup \{-\infty\}, \lor, +)$ and $A \in \text{Mat}_{n \times n}(\mathbb{R}_{\max})$ be the adjacency matrix of the weighted directed graph $G$. Then, if there are only circuits of nonpositive weight in $G$, there is a solution to $x = Ax \lor b$ which is given by $x = A^*b$. Moreover, if the circuit weights are negative, the solution is unique.

**note**

Hint: $A(A^*b) \lor b = (e \lor AA^*)b = A^*b$. 

Let $A \in \text{Mat}_{n \times n}(\mathbb{R}_{\max})$ and consider the dynamics

$$x(n + 1) = Ax(n), \quad x(0) = x_0,$$

whose solution is $x(n) = A^n x(0)$. Naturally, part of the analysis depend on the existence of eigenvalues and eigenfunctions for $A$, i.e. the existence of $\lambda$ and $f$ such that $Af = \lambda f$. 
If $A \in \text{Mat}_{n \times n}(\mathbb{R}_{\text{max}})$ is irreducible, or equivalently if the corresponding graph $G$ is strongly connected, there exists one and only one eigenvalue (but possibly several eigenvectors). This eigenvalue is equal to the maximum cycle mean of the graph $G$, i.e.

$$\lambda = \max_{c} \frac{|c|_{w}}{|c|_{1}},$$

where $c$ ranges over cycles of the graph $G$. 
Example

Eigenfunctions are not unique.

\[ e = 0, \epsilon = -\infty. \]

\[
\begin{bmatrix}
1 & e \\
\epsilon & 1
\end{bmatrix}
\begin{bmatrix}
e \\
-1
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
e
\end{bmatrix}
= 1 
\begin{bmatrix}
e \\
-1
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & e \\
\epsilon & 1
\end{bmatrix}
\begin{bmatrix}
-1 \\
e
\end{bmatrix}
= 
\begin{bmatrix}
e \\
1
\end{bmatrix}
= 1 
\begin{bmatrix}
-1 \\
e
\end{bmatrix}
\]
Example

Eigenvalues are not unique.

\[
\begin{bmatrix}
1 & \epsilon \\
\epsilon & 2
\end{bmatrix}
\begin{bmatrix}
e \\
\epsilon
\end{bmatrix}
= 1
\begin{bmatrix}
e \\
\epsilon
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & \epsilon \\
\epsilon & 2
\end{bmatrix}
\begin{bmatrix}
\epsilon \\
e
\end{bmatrix}
= 2
\begin{bmatrix}
\epsilon \\
e
\end{bmatrix}
\]

\[e = 0, \epsilon = -\infty.\]
Why idempotent analysis?

1. There is a nice functional analysis and operator theory (giving rise to a nonlinear system theory!).

2. A lot of linear algebra can be extended to this setting.

3. Again provides a nice example for the fact that there is a rich limit theory in which much of the facts reflects!

What are the properties of this new mathematics in the limit?!
A Nonlinear System Theory
References

GENERAL REFERENCES FOR FURTHER READING:

- Litvinov Grigory L.; *Hypergroups and Hypergroup Algebras*, Available at arXiv:1109.6596.
In this part we present an input/output system theory based on semigroups (actually dioids) in which the emphasis is on the role of signals as (fuzzy/toll) sets rather than functions. Hence, this is an example of a system theory based on the set-based approach.

Remark

It is interesting if one can show that there exists a setup in which this is a limit theory of a classical case (e.g. the theory of linear shift-invariant systems).

We will talk more about this!
Basic ideas for an I/O system theory

\[ \Phi \rightarrow F \rightarrow \Phi \]

\[ A \rightarrow F(A) \]

\[ \mathcal{V} \rightarrow \mathcal{V} \]

B
Basic ideas for an I/O system theory

- \( \Phi \) is a function space with some nice algebraic and topologic properties (usually completeness conditions).
Specially $\Phi$ is reconstructable, i.e. can be reconstructed properly by a (relatively small) subspace $B$ of its generic objects.
Usually $\Phi$ is chosen to be a product space $G \times \mathcal{V}_0$. 

- e.g. SPEACH: $G = R$ is the time and $\mathcal{V}_0 = R$ is the space of levels.
- e.g. IMAGE: $G = \mathbb{Z}^2$ is the two dimensional space $\mathcal{V}_0 = \mathbb{Z}$ is the space of gray-scales.
**Basic ideas for an I/O system theory**

$F$ is a natural map (i.e. compatible with the structure), with nice representation properties.
The whole setup should be in coherence with natural phenomena and be able to simulate input-output behaviour of such systems (this is called I/O system theory).
A CLASSICAL EXAMPLE: LSI SYSTEMS (OPERATORS)

- $\Phi$ is a Hilbert space (or some nice Sobolev space).
A classical example: LSI systems (operators)

- $\Phi$ is a Hilbert space (or some nice Sobolev space).
- $F$ is linear and shift invariant, i.e.

$$F(A + \alpha B) = F(A) + \alpha F(B),$$

$$F(T_g(A)) = T_g(F(A)),$$

where $T_g$ is the translation-by-$g$ operator, i.e.

$$T_g(A)(t) = A(t - g).$$
A classical example: LSI systems (operators)

- Φ is a Hilbert space (or some nice Sobolev space).
- F is linear and shift invariant, i.e.
  \[ F(A + αB) = F(A) + αF(B), \]
  \[ F(T_g(A)) = T_g(F(A)), \]
  where \( T_g \) is the translation-by-\( g \) operator, i.e.
  \[ T_g(A)(t) = A(t - g). \]

- We have the following reconstruction:
  \[ F(A)(t) = \text{Conv}(A, δ_F) = \int_{-∞}^{+∞} δ_F(τ)A(t - τ)dτ. \]
Consider the discrete system

\[ T(f)(n) \overset{\text{def}}{=} \frac{1}{3} (f(n) + f(n - 1) + f(n - 2)) \]

as a smoother (low-pass filter) on discrete signals \( f : \mathbb{Z} \rightarrow \mathbb{R} \). The first important observation is that \( T \) is an LSI system and its impulse response \( h \) is the following:

\[
h(n) \overset{\text{def}}{=} \begin{cases} 
\frac{1}{3} & n = 0, 1, 2 \\
0 & n \neq 0, 1, 2.
\end{cases}
\]

Therefore, \( T \) can be expressed as the convolution

\[
T(f)(n) = \sum_{m=0}^{\infty} f(n - m)h(m) = \frac{1}{3} (f(n) + f(n - 1) + f(n - 2)).
\]
Consider

\[ A = \{(t, A(t)) \mid t \in G\}. \]
Then we define the translation operator as follows,

$$A[g, p] = \{(t + g, A(t) \ast p) \mid t \in G\},$$

which means,

$$A[g, p](t) = A(t - g) \ast p.$$
An operator $F$ is translation invariant if

$$F(A[g, p]) = F(A)[g, p].$$

Note: For an LSI operator, translation invariance on the range is equivalent to DC-gain 1.
The set \( A = \{(t, A(t)) \mid t \in G\} \subseteq G \times V_0 \) can be considered as a function or a fuzzy set.
The set \( A = \{(t, A(t)) \mid t \in G\} \subseteq G \times \mathcal{V}_0 \) can be considered as a function or a fuzzy set.

The difference between the two points of view is in the way we look at the range \( \mathcal{V}_0 \).
The set \( A = \{(t, A(t)) \mid t \in G\} \subseteq G \times V_0 \) can be considered as a function or a fuzzy set.

The difference between the two points of view is in the way we look at the range \( V_0 \).

The functional approach is when we consider algebraic properties and the fuzzy set approach is when we consider the comparative structure (order structure) of \( V_0 \).

E.g. The neutral element for summation is 0, while the neutral element for the supremum is \(-\infty\).
Let \((G, +, -, 0)\) be a group and \(V_0 = (\Omega, \leq, *, \div, 0) \cup \{-\infty, +\infty\}\) be a lattice ordered residuated semigroup with the universal bounds \(-\infty\) and \(+\infty\) such that,

\[
\div(-\infty) = +\infty, \quad \div(+\infty) = -\infty,
\]

\[
(-\infty) * (+\infty) = (+\infty) * (-\infty) = (+\infty) * (+\infty) = (+\infty)
\]

\[
(-\infty) * (-\infty) = (-\infty),
\]

\[
\forall p \in \Omega \quad (+\infty) * p = p * (+\infty) = +\infty,
\]

\[
\forall p \in \Omega \quad (-\infty) * p = p * (-\infty) = -\infty.
\]
Minkowski addition and subtraction for L-fuzzy sets are defined as follows

\[
A \oplus B = \sup_{g} A[g, B(g)] \quad , \quad A \ominus B = \inf_{g} A[g, \div B(g)].
\]
Minkowski addition and subtraction for L-fuzzy sets are defined as follows

\[ A \oplus B = \sup_g A[g, B(g)] , \quad A \ominus B = \inf_g A[g, \div B(g)]. \]

e.g. Let

\[ G = \langle a \mid 3a = 0 \rangle , \]
\[ A = \{(0, x_1), (a, x_2), (2a, x_3)\} , \quad B_1 = \{(0, 0), (a, 0), (2a, -\infty)\} . \]

Then,

\[ A \oplus B_1 = \{(0, \sup(x_1, x_3)), (a, \sup(x_2, x_1)), (2a, \sup(x_3, x_2))\} . \]
Minkowski erosion is defined as follows

\[ \text{Er}(A, B) = A \ominus B^s. \]

where,

\[ B^s = \{ (-x, B(x)) \mid x \in G \}. \]
Minkowski erosion is defined as follows

$$Er(A, B) = A \ominus B^s.$$ 

where,

$$B^s = \{(−x, B(x)) \mid x \in G\}.$$ 

Minkowski erosion is non-linear and behaves as a convolution operator.
For a TI operator $F$ the kernel is defined as follows

$$K(F) = \{A \mid 0 \leq F(A)(0)\}.$$
The kernel

- For a TI operator $F$ the kernel is defined as follows

$$K(F) = \{ A \mid 0 \leq F(A)(0) \}.$$ 

- The base of $F$ is the set of all minimal elements of $K(F)$.
For a TI operator $F$ the kernel is defined as follows

$$K(F) = \{A \mid 0 \leq F(A)(0)\}.$$  

The base of $F$ is the set of all minimal elements of $K(F)$.

e.g. Let

$G = \langle a \mid 3a = 0 \rangle$,

$A = \{(0, x_1), (a, x_2), (2a, x_3)\}$, $F(A) = \{(0, m), (a, m), (2a, m)\}$,

where $m = \text{Median}(x_1, x_2, x_3)$.

Then,

$$B(F) = \{\{(0, 0), (a, 0), (2a, -\infty)\}, \{(0, -\infty), (a, 0), (2a, 0)\},$$

$$\{(0, 0), (a, -\infty), (2a, 0)\}\}.$$
A TI operator $F$ is isotone if

$$\forall A, B \ A \leq B \Rightarrow F(A) \leq F(B).$$
A TI operator $F$ is isotone if

$$\forall A, B \quad A \leq B \Rightarrow F(A) \leq F(B).$$

**Strong Reconstruction Theorem** *(D. 1995)*

Let $F$ be an isotone TI operator. Then

$$F(A) = \sup_{D \in K(F)} Er(A, D);$$

and if the base of $F$ exists then

$$F(A) = \sup_{B \in B(F)} Er(A, B).$$
**Example 1: The Mean Filter**

Again consider the discrete system

$$T(f)(n) \overset{\text{def}}{=} \frac{1}{3} (f(n) + f(n - 1) + f(n - 2)).$$

One may note that $T$ is an isotone TI system with the basis consisting of all maps $h_{r,s}, (r, s \in \mathbb{R})$ defined as follows:

$$h_{r,s}(n) \overset{\text{def}}{=} \begin{cases} 
-r - s & n = -2 \\
 0 & n = -1 \\
 s & n = 0 \\
 -\infty & \text{otherwise}
\end{cases}$$

$$T(f)(n) = \sup_{r,s \in \mathbb{R}} Er(f, h_{r,s})(n) = \sup_{r,s \in \mathbb{R}} \inf_{m \in \mathbb{Z}} (f(n + m) - h_{r,s}(m))$$

$$= \sup_{r,s \in \mathbb{R}} \inf_{r,s} (f(n) - r, f(n - 1) - s, f(n - 2) + r + s).$$
Example 2: The Median Filter

Let $G = < a \mid 3a = 0 >$, 
$A = \{(0, x_1), (a, x_2), (2a, x_3)\}$, $F(A) = \{(0, m), (a, m), (2a, m)\}$, where $m = \text{Median}(x_1, x_2, x_3)$. Then,

$$B(F) = \{\{(0, 0), (a, 0), (2a, -\infty)\}, \{(0, -\infty), (a, 0), (2a, 0)\}, \{(0, 0), (a, -\infty), (2a, 0)\}\}$$
**Example 2: The Median Filter**

Let $G = \langle a \mid 3a = 0 \rangle$, 
$A = \{(0, x_1), (a, x_2), (2a, x_3)\}$, $F(A) = \{(0, m), (a, m), (2a, m)\}$, where $m = \text{Median}(x_1, x_2, x_3)$.
Then,

$B(F) = \{(0, 0), (a, 0), (2a, -\infty)\}, \{(0, -\infty), (a, 0), (2a, 0)\},$

$\{(0, 0), (a, -\infty), (2a, 0)\}$

Strong reconstruction theorem implies that
$\text{Median}(x_1, x_2, x_3) = F(A)(0)$
$= \sup(\inf(x_1, x_2), \inf(x_2, x_3), \inf(x_3, x_1))$. 
**Example 2: The median filter**

Let $G = \langle a \mid 3a = 0 \rangle$, 

$A = \{(0, x_1), (a, x_2), (2a, x_3)\}$, $F(A) = \{(0, m), (a, m), (2a, m)\}$, 

where $m = \text{Median}(x_1, x_2, x_3)$.

Then,

$$B(F) = \{(0, 0), (a, 0), (2a, -\infty)\}, \{(0, -\infty), (a, 0), (2a, 0)\},$$

$$\{(0, 0), (a, -\infty), (2a, 0)\}$$

Strong reconstruction theorem implies that

$\text{Median}(x_1, x_2, x_3) = F(A)(0)$

$= \sup(\inf(x_1, x_2), \inf(x_2, x_3), \inf(x_3, x_1))$.

This is a fuzzification of the Boolean expression

$x_1x_2 + x_2x_3 + x_3x_1$. 
Reconstructions are usually as limits of convolutions.
Reconstructions are usually as limits of convolutions.

We know that,
Reconstructions are usually as limits of convolutions.

We know that,

Such limits can be defined in a variety of ways.
Reconstructions are usually as limits of convolutions.

We know that,

- Such limits can be defined in a variety of ways.
- Convolutions are essentially generalized HOM-functors.
Reconstructions are usually as limits of convolutions.

We know that,

- Such limits can be defined in a variety of ways.
- Convolutions are essentially generalized HOM-functors.
- Our general setup will cover both fuzzy and functional approach.
Thank you!

Comments and Criticisms are Welcomed

daneshgar@sharif.ir