Distances, Norms and Error Propagation in Idempotent Semirings

Roland Glück
roland.glueck@dlr.de
Deutsches Zentrum für Luft- und Raumfahrt
Groningen, 28nd October 2018

Applications of Idempotent Semirings

- Graph Theory
  - Berghammer, Stucke and Winter (RAMiCS 2015)
  - Brunet, Pous and Stucke (ITP 2016)
  - Glück (RAMiCS 2017)
  - Guttmann (RAMiCS 2017)
- Energy Optimization (Ésik, Fahrenberg, Leagy and Quaas at ATVA 2013)
- Language Problems (Backhouse, JLAP 2006)
- Fuzzy Logic (Kawahara, Furusawa and Winter at Various Occasions)
- Program Verification and Analysis
  - Armstrong, Struth and Weber (ITP 2013)
  - Glück and Krebs (RAMiCS 2015)
  - Michels, Joosten, Joosten, van der Woude (RAMiCS 2011)
  - Oliveira (JLAP 2014)
- Database Theory (Litak, Mikulás and Hiddders (RAMiCS 2014))
- Tropical Optimization (Krivulin, Various Occasions)

Alas,

- used data my be defective by
  - errors in measurement
  - estimated data
  - human shortcomings

Classical mathematics/physics have

- error propagation
- numerical/statistical methods

However, no such tools are known for idempotent semirings!

Definition Idempotent Semiring

Definition

An *idempotent semiring* is a structure $S_S = (M_S, +_S, 0_S, \cdot_S, 1_S)$ with internal operations $+_S$ and $\cdot_S$ and constants 0_S and 1_S in M_S such that

- $+_S$ is commutative, associative and idempotent with neutral element 0_S,
- $\cdot_S$ is associative with neutral element 1_S and annihilator 0_S, and
- $\cdot_S$ distributes from both sides over $+_S$.

$\sqsubseteq_S$ with $x \sqsubseteq_S y \ \Leftrightarrow_{df} \ x +_S y = y$ is the *natural order* of ${}_SS$.

Examples:

- tropical semiring $\mathbb{R}_{\geq 0}^{\mathsf{min},+} =_{def} (\mathbb{R}_{\geq 0} \cup \{\infty\}, \mathbf{min}, \infty, +, 0)$
- max-min semiring $\mathbb{R}_{\geq 0}^{\mathsf{max},\mathsf{min}} =_{def} (\mathbb{R}_{\geq 0} \cup \{\infty\}, \mathbf{max}, 0, \mathbf{min}, \infty)$
- semiring of finite languages, $LAN_{\Sigma}^{\mathbf{fin}} =_{def} (\Sigma^{\mathbf{fin}}, \cup, \emptyset, \cdot, \{\varepsilon\})$

Definition Measure System

Definition

A *measure sysem* is a structure $S_m = (M_m, +_m, 0_m, \cdot_m, \sqsubseteq_m)$ with internal operations $+_m$ and $\cdot_m$ and a constant 0_m such that

- $+_m$ is commutative and associative with neutral element 0_m,
- $\cdot_m$ is associative and distributes from both sides over $+_m$,
- $\sqsubseteq_m$ is an order on M_m with least element 0_m, and
- $+_m$ and $\sqsubseteq_m$ are isotone in both arguments wrt. $\sqsubseteq_m$.

No annihilation by 0_m, no neutral element of $\cdot_m$!

Examples:

- $\mathbf{m}\mathbb{R}_{\geq 0}^{\mathsf{max},+} =_{def} (\mathbb{R}_{\geq 0}, \mathsf{max}, 0, +, \leq)$
- $\mathbf{m}\mathbb{R}_{\geq 0}^{\mathsf{max},\mathsf{max}} =_{def} (\mathbb{R}_{\geq 0} \cup \{\infty\}, \mathsf{max}, 0, \mathsf{max}, \leq)$
- $\mathbf{m}\mathbb{N}_0^{+,\cdot} =_{def} (\mathbb{N}_0, +, 0, \cdot, \leq)$
- $\mathbf{m}LAN_\Sigma^{\mathbf{fin}} =_{def} (\Sigma^{\mathbf{fin}}, \cup, \emptyset, \cdot, \subseteq)$

DLR

Definition Distance

Definition

Given an idempotent semiring $S_s = (M_s, +_s, 0_s, \cdot_s, 1_s)$ and a measure system $S_m = (M_m, +_m, 0_m, \cdot_m, \sqsubseteq_m)$ we call a mapping $\boldsymbol{d} : M_s \times M_s \to M_m$ an *S_m-distance on S_s*. It is called

- *additive*, if $\boldsymbol{d}(x_1 +_s x_2, y_1 +_s y_2) \sqsubseteq_m \boldsymbol{d}(x_1, y_1) +_m \boldsymbol{d}(x_2, y_2)$,
- *multiplicative*, if $\boldsymbol{d}(x_1 \cdot_s x_2, y_1 \cdot_s y_2) \sqsubseteq_m \boldsymbol{d}(x_1, y_1) \cdot_m \boldsymbol{d}(x_2, y_2)$,
- *order preserving*, if $x \sqsubseteq_s y \wedge y \sqsubseteq_s z \Rightarrow \boldsymbol{d}(x, y) \sqsubseteq_m \boldsymbol{d}(x, z)$, and
- *strict*, if $\boldsymbol{d}(x, y) = 0_m \Leftrightarrow x = y$ holds.

A *complete distance* is an additive, multiplicative, order preserving and strict distance.

Examples:

- $\boldsymbol{d}_s(x, y) =_{def} |x - y|$ is a complete $\mathbf{m}\mathbb{R}_{\geq 0}^{\max,+}$-distance on $\mathbb{R}_{\geq 0}^{\min,+}$.
- $\boldsymbol{d}_m(x, y) =_{def} |x - y|$ is a complete $\mathbf{m}\mathbb{R}_{\geq 0}^{\max,\max}$-distance on $\mathbb{R}_{\geq 0}^{\max,\min}$.
- $\boldsymbol{d}_w(L_1, L_2) =_{def} L_1 \triangle L_2$ is an additive, order preserving and strict (but no multiplicative) $\mathbf{m}LAN_{\Sigma}^{\mathbf{fin}}$-distance on $LAN_{\Sigma}^{\mathbf{fin}}$.

Distance Properties

Lemma

Let $\boldsymbol{d}$ *be a complete* S_m*-distance on* S_s*. Then the following properties hold for all* $x, y, z \in M_m$*,* $l, n \in \mathbb{N}_0$ *and mappings* $f, g, h : M_s \to M_s$*:*

- $\boldsymbol{d}(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} y_i) \sqsubseteq_m \sum_{i=1}^{n} \boldsymbol{d}(x_i, y_i)$
- $\boldsymbol{d}(x, y +_s z) \sqsubseteq_m \boldsymbol{d}(x, y) +_m \boldsymbol{d}(x, z)$
- $\boldsymbol{d}(x, y) \sqsubseteq_m \boldsymbol{d}(x, z) \Rightarrow \boldsymbol{d}(x^n, y^n) \sqsubseteq_m \boldsymbol{d}(x, z)^n$
- $\boldsymbol{d}(x, y) \sqsubseteq_m \boldsymbol{d}(x, z) \Rightarrow \boldsymbol{d}(\sum_{i=l}^{n} x^i, \sum_{i=l}^{n} y^i) \sqsubseteq_m \sum_{i=l}^{n} \boldsymbol{d}(x, z)^i$
- *If* f *and* g *are isotone with respect to* $\sqsubseteq_s$ *and* $f(x) \sqsubseteq_s g(x) \sqsubseteq_s h(x)$ *holds for all* $x \in M_s$ *then* $\boldsymbol{d}(f^n(x), g^n(x)) \sqsubseteq_m \boldsymbol{d}(f^n(x), h^n(x))$ *holds.*

On the Way to Matrices

- applications of matrices over idempotent semirings:
  - shortest path: matrices over $(\mathbb{R}_{\geq 0} \cup \{\infty\}, \mathbf{min}, \infty, +, 0)$
  - maximum capacity path: matrices over $(\mathbb{R}_{\geq 0} \cup \{\infty\}, \mathbf{max}, 0, \mathbf{min}, \infty)$
  - relations: matrices over $(\mathtt{BOOL}, \vee, \mathtt{FALSE}, \wedge, \mathtt{TRUE})$
  - automata theory: matrices over $(\boldsymbol{\Sigma}^{\mathbf{fin}}, \cup, \emptyset, \cdot, \{\varepsilon\})$
- therefore extension of distances to matrices
- writing $S^{l \times m}$ for $l \times m$-matrices over S
- matrix operations defined routinely (as in traditional linear algebra)
- order defined entrywise (provided S is ordered)
- folklore: if S is an idempotent semiring then so is $S^{n \times n}$
- holds analogously for measure spaces

Definition Matrix Distance

Definition

Given an idempotent semiring $S_s = (M_s, +_s, 0_s, \cdot_s, 1_s)$, a measure system $S_m = (M_m, +_m, 0_m, \cdot_m, \sqsubseteq_m)$, an S_m-distance $\boldsymbol{d}$ on S_s of any kind and two matrices $A, B \in M_s^{l\times n}$ we define the distance $\boldsymbol{d}^{l\times n}(A, B)$ entrywise as a matrix of the type $M_m^{l\times n}$, i.e., $(\boldsymbol{d}^{l\times n}(A, B))_{ij} =_{def} \boldsymbol{d}(A_{ij}, B_{ij})$.

How do properties of $\boldsymbol{d}$ carry over to $\boldsymbol{d}^{l\times n}$?

Matrix Distance Properties

Theorem

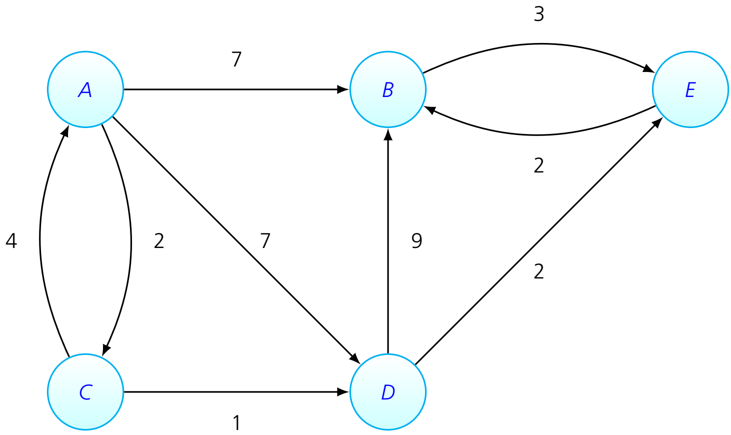
Let $\mathbf{d}$ *be an* S_m*-distance on* S_s *and consider the idempotent semiring* $S_s^{n\times n}$ *of* $n \times n$*-matrices over* S_s *and let* $S_m^{n\times n}$ *be the set of* $n \times n$*-matrices over* S_m*. Then the following properties hold:*

- *If* $\mathbf{d}$ *is additive then* $\mathbf{d}^{n\times n}$ *is an additive* $S_m^{n\times n}$*-distance on* $S_s^{n\times n}$*.*
- *If* $\mathbf{d}$ *is order preserving then* $\mathbf{d}^{n\times n}$ *is an order preserving* $S_m^{n\times n}$*-distance on* $S_s^{n\times n}$*.*
- *If* $\mathbf{d}$ *is strict then* $\mathbf{d}^{n\times n}$ *is a strict* $S_m^{n\times n}$*-distance on* $S_s^{n\times n}$*.*
- *If* $\mathbf{d}$ *is both additive and multiplicative then* $\mathbf{d}^{n\times n}$ *is a multiplicative* $S_m^{n\times n}$*-distance on* $S_s^{n\times n}$*.*

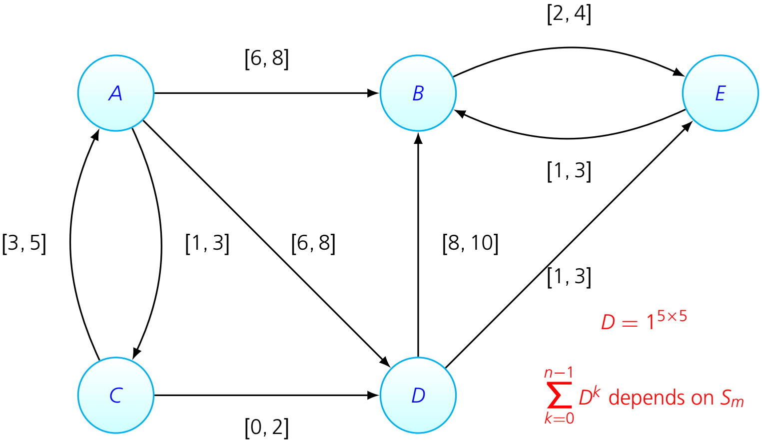
Application of Matrix Distance

- Consider an S_m-distance $\boldsymbol{d}$ on S_s.
- Then $\boldsymbol{d}^{n\times n}$ is a complete $S_m^{n\times n}$-distance on $S_s^{n\times n}$.
- Assume three matrices A, $\hat{A}$ and D from $S_s^{n\times n}$ with $\boldsymbol{d}_s^{n\times n}(A,\hat{A}) \leq D$.
- Then we have $\boldsymbol{d}_s^{n\times n}(\sum_{k=0}^{n-1} A^k, \sum_{k=0}^{n-1} \hat{A}^k) \leq \sum_{k=0}^{n-1} D^k$.
- Applies to maximum capacity paths, shortest paths, maximum reliability paths, ...

Original Network



Perturbated Network



$D = 1^{5\times 5}$

$\sum_{k=0}^{n-1} D^k$ depends on S_m

Definition Norm

Definition

Given an idempotent semiring $S_s = (M_s, +_s, 0_s, \cdot_s, 1_s)$ and a measure system $S_m = (M_m, +_m, 0_m, \cdot_m, \sqsubseteq_m)$ we call a mapping $||\cdot|| : M_s \to M_m$ an *S_m-norm on S_s*. It is called

- *additive* if $||x +_s y|| \sqsubseteq_m ||x|| +_m ||y||$,
- *multiplicative* if $||x \cdot_s y|| \sqsubseteq_m ||x|| \cdot_m ||y||$,
- *order preserving* if $x \sqsubseteq_s y \Rightarrow ||x|| \sqsubseteq_m ||y||$, and
- *strict* if $||x|| = 0_m \Leftrightarrow x = 0_s$ holds.

A *complete norm* is an additive, multiplicative, order preserving and strict norm.

Examples:

- For every idempotent semiring $S_s = (M_s, +_s, 0_s, \cdot_s, 1_s)$ the identity is a complete S_m-norm with $S_m = (M_s, +_s, 0_s, \cdot_s, \sqsubseteq_s)$.
- $||L|| =_{def} |L|$ is a complete $\mathbf{m}\mathbb{N}_0^{+,\cdot}$-norm on $LAN_\Sigma^{\mathbf{fin}}$.

Induced Norm

Theorem

Let $S_s = (M_s, +_s, 0_s, \cdot_s, 1_s)$ be an idempotent semiring and let $\boldsymbol{d}$ by an (additive, multiplicative, order preserving, strict) S_m-distance on S_s for some measure system $S_m = (M_m, +_m, 0_m, \cdot_m, \sqsubseteq_m)$. Then the mapping $||\cdot||_{\boldsymbol{d}} : M_s \to M_m$, defined by $||x||_{\boldsymbol{d}} =_{def} \boldsymbol{d}(0_s, x)$, is an (additive, multiplicative, order preserving, strict) S_m-norm on S_s. It is called the norm induced by $\boldsymbol{d}$.

Example:

- $||\cdot||$ on $\mathbb{R}^n$ is the norm induced by $\boldsymbol{d}(x, y) =_{df} ||x - y||$.
- The identity norm on $LAN_{\Sigma}^{\mathbf{fin}}$ is the norm induced by the distance $\boldsymbol{d}_w =_{df} L_1 \triangle L_2$.

Norm-Distributivity

Definition

An S_m-distance $\boldsymbol{d}$ on S_s of any kind is called *norm-distributive* if its induced norm fulfills the property $\boldsymbol{d}(x \cdot_s y, x \cdot_s z) \sqsubseteq_m ||x||_{\boldsymbol{d}} \cdot_m \boldsymbol{d}(y, z)$ for all x, y and z.

- compare $|cx - cy| = |c| \cdot \boldsymbol{d}(x, y)$ or
- $|f(x) - f(x + h)| \approx |f'(x)| \cdot |h|$
- $\boldsymbol{d}_w$ is a norm-distributive distance on $LAN^{\mathbf{fin}}_{\Sigma}$.

Composing Norms and Distances

Theorem

Consider the distance $\mathbf{d}_{\|\ \|} =_{def} \|\ \| \circ \mathbf{d}$. Then the following holds:

- *If $\mathbf{d}$ is additive and $\|\ \|$ is additive and order preserving then $\mathbf{d}_{\|\ \|}$ is additive.*
- *If $\mathbf{d}$ is multiplicative and $\|\ \|$ is multiplicative and order preserving then $\mathbf{d}_{\|\ \|}$ is multiplicative.*
- *If both $\mathbf{d}$ and $\|\ \|$ are order preserving then $\mathbf{d}_{\|\ \|}$ is order preserving.*
- *If $\mathbf{d}$ and $\|\ \|$ are strict then $\mathbf{d}_{\|\ \|}$ is strict.*
- *If $\mathbf{d}$ is norm-distributive and $\|\ \|$ is multiplicative and order preserving then $\mathbf{d}_{\|\ \|}$ is norm-distributive, too.*

Also well-behaved properties with respect to matrices.

Final Example

- Consider again finite languages.
- Automata theory considers matrices of languages.
- Consider three such $n \times n$-matrices L_1, L_2 and L_3.
- Then we have $\boldsymbol{d}_W(L_1 \cdot L_2, L_1 \cdot L_3) \subseteq L_1 \cdot \boldsymbol{d}_W(L_2, L_3)$ and
- $\boldsymbol{d}_N(L_1 \cdot L_2, L_1 \cdot L_3) \leq |L_1| \cdot \boldsymbol{d}_N(L_2, L_3)$.

Further Work

- Kleene star
- eigenvalues/bideterminants
- randomized/stochastic errors
- topological considerations
- ...

Review

... some of his results may be "straightforward", some estimations "rough" and some proofs not particularly deep and rather tedious. Doing science is not quite an endless sequence of fireworks, ecstasy and champaign; by nature, there must be more moments of toil and frustration than insight and epiphany. We are so often pushed to present a distorted image or our work and focus on sell talk ...