Formal Adventures in Convex and Conical Spaces

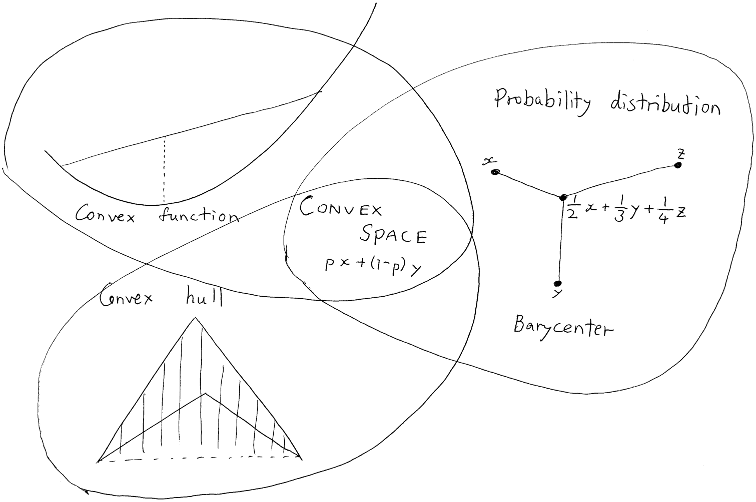
Reynald Affeldt[1] Jacques Garrigue[2] Takafumi Saikawa[2]

[1] National Institute of Advanced Industrial Science and Technology, Japan

[2] Nagoya University, Japan

July 28, 2020

Background


Convex function
Convex hull
Convex Space
$px + (1-p)y$
Probability distribution
x
z
y
$\frac{1}{2}x + \frac{1}{3}y + \frac{1}{4}z$
Barycenter

Overview

This work builds a hub to connect many interesting domains of formalization: probability, analysis, program semantics, etc.
Applications:

- Convex functions
- Barycenter
- Convex hulls
- Semantics of probabilistic programs

Contribution: we formalized different interfaces and lemmas connecting them, and defined derived constructions for applications.

Convex combination and convex space

Convexity notions such as

- Convex functions:
  $f(px + (1-p)y) \leq pf(x) + (1-p)f(y)$
- Convex sets: $x, y \in X \Rightarrow px + (1-p)y \in X$

are based on convex combinations in vector spaces:

$$x \triangleleft_p \triangleright y = px + (1-p)y$$

Many authors axiomatized convex combinations, capturing essentially the same concept:

- Barycentric calculus [Stone, 1949]
- Semiconvex algebra [Flood, 1981]
- Abstract probabilistic domain [Jones and Plotkin, 1989]
- Convex Spaces [Fritz, 2015]

Convex space

- Carrier set X
- Convex combination operations $(\_ \triangleleft_p \triangleright \_) : X \times X \to X$ indexed by $p \in [0, 1]$
- Some appropriate laws

Laws come from a geometric intuition: convex combination defines a barycenter of points.

Barycenter = n-ary combination

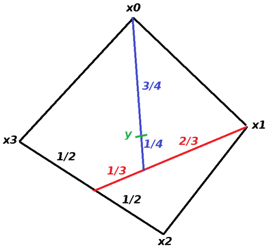
Let

- d : a finite distribution ($\sum_i d_i = 1$)
- x : sequence of points in a convex space X

The barycenter of n points x_i with weights d_i is recursively defined:

$$\underset{i<n}{\diamondsuit} d_i x_i = \begin{cases} x_0 & \text{if } d_0 = 1 \\ x_0 \triangleleft d_0 \triangleright \left(\underset{i<n-1}{\diamondsuit} d'_i x_{i+1} \right) & \text{otherwise} \end{cases}$$

where d' is a new distribution: $d'_i = d_{i+1}/(1-d_0)$



$$y = \frac{1}{4}x_0 + \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_3$$

Hull = set of barycenters

Let A be a convex space. For a subset X of A, its hull is

$$\left\{ \underset{i<n-1}{\diamond} d_i x_i \;\middle|\; \begin{array}{l} \quad (n \in \mathbb{N}) \\ \wedge\, (d : n\text{-point distribution}) \\ \wedge\, (x : \text{sequence of } n \text{ points in } X) \end{array} \right\}$$

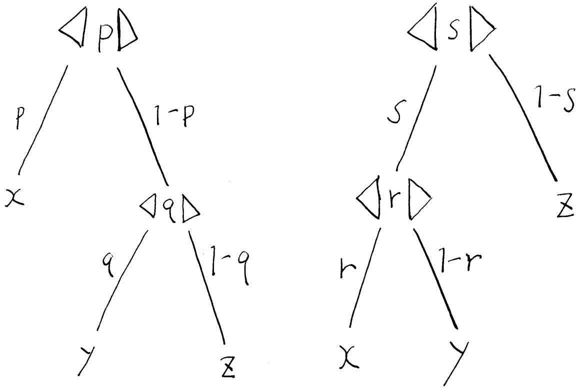
Convex space

- Carrier set X
- Convex combination operations $(\_ \triangleleft_p \triangleright \_) : X \times X \to X$ indexed by $p \in [0,1]$
- Unit law: $x \triangleleft_1 \triangleright y = x$
- Idempotence law: $x \triangleleft_p \triangleright x = x$
- Skewed commutativity law: $x \triangleleft_{1-p} \triangleright y = y \triangleleft_p \triangleright x$
- Quasi-associativity law:
  $x \triangleleft_p \triangleright (y \triangleleft_q \triangleright z) = (x \triangleleft_r \triangleright y) \triangleleft_s \triangleright z$, where
  $s = 1 - (1-p)(1-q)$ and $r = \begin{cases} p/s & \text{if } s \neq 0 \\ 0 & \text{otherwise} \end{cases}$
  Note: r is irrelevant to the value of $(x \triangleleft_r \triangleright y) \triangleleft_s \triangleright z$ if $s = 0$.

Convex space

- Carrier set X
- Convex combination operations $(\_ \triangleleft_p \triangleright \_) : X \times X \to X$ indexed by $p \in [0,1]$
- Unit law: $x \triangleleft_1 \triangleright y = x$
- Idempotence law: $x \triangleleft_p \triangleright x = x$
- Skewed commutativity law: $x \triangleleft_{1-p} \triangleright y = y \triangleleft_p \triangleright x$
- Quasi-associativity law:
  $x \triangleleft_p \triangleright (y \triangleleft_q \triangleright z) = (x \triangleleft_r \triangleright y) \triangleleft_s \triangleright z$, where
  $s = 1 - (1-p)(1-q)$ and $r = \begin{cases} p/s & \text{if } s \neq 0 \\ 0 & \text{otherwise} \end{cases}$
  Note: r is irrelevant to the value of $(x \triangleleft_r \triangleright y) \triangleleft_s \triangleright z$ if $s = 0$.

A hell of a quasi-associativity



- $s = 1 - (1-p)(1-q)$ and $r = \frac{p}{s}$
- Zero introduces a special case due to the division.
- Recursive barycenter is even more difficult to formally reason about.

The solution: conical spaces [Flood, 1981]

A conical space is a semimodule over the semiring $\mathbb{R}_{\geq 0}$:

- Carrier set X
- Zero $0 : X$.
- Addition operation $\_ + \_ : X \times X \to X$
- Scaling operations $c\_ : X \to X$ indexed by $c \in \mathbb{R}_{\geq 0}$
- Associativity law for addition: $x + (y + z) = (x + y) + z$
- Commutativity law for addition: $x + y = y + x$
- Associativity law for scaling: $c(dx) = (cd)x$
- Left-distributivity law: $(c + d)x = cx + dx$
- Right-distributivity law: $c(x + y) = cx + cy$
- Zero law for addition: $0 + x = x$
- Zero law for scaling: $0x = 0$
- One law for scaling: $1x = x$

Embedding convex spaces into cones [Flood, 1981]

A conical space is canonically made from convex space X:

$$S_X := (\mathbb{R}_{>0} \times X) \cup \{0\}$$

with an accompanying embedding:

$$\begin{array}{rccc} \iota : & X & \rightarrowtail & S_X \\ & x & \mapsto & (1, x) \end{array}$$

Addition is defined for $a, b \in S_X$,

$$a + b := \begin{cases} (r + q, x \triangleleft_{r/(r+q)} \triangleright y) & \text{if } a = (r, x) \text{ and } b = (q, y) \\ a & \text{if } b = 0 \\ b & \text{if } a = 0 \end{cases}$$

Scalar multiplication is for $a \in S_X$ and $p \in \mathbb{R}_{\geq 0}$,

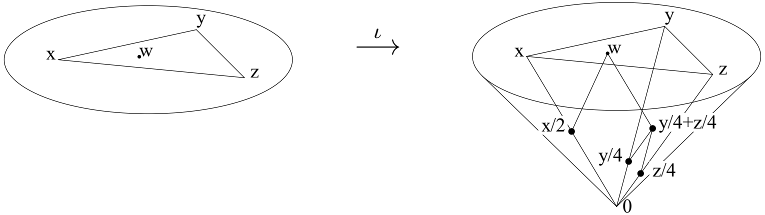
$$pa := \begin{cases} (pq, x) & \text{if } p > 0 \text{ and } a = (p, x) \\ 0 & \text{otherwise} \end{cases}$$

Key lemma for barycenter computation

Lemma (S1_convn)

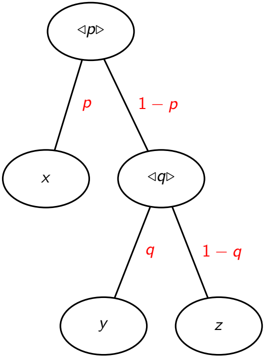
$$\iota(\underset{i<n}{\diamondsuit} d_i x_i) = \sum_{i<n} d_i \iota(x_i)$$

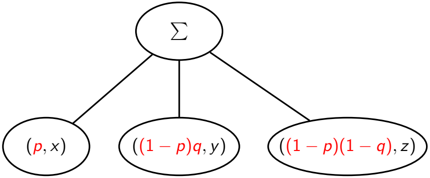
Example ($\iota(w) = \frac{1}{2}\iota(x) + \frac{1}{4}\iota(y) + \frac{1}{4}\iota(z)$)



Welcome to ~~quasi~~-associativity paradise!

The embedding moves weights from edges to leaves:





Proof of a typical property: entropic identity

```
Lemma convACA (a b c d : T) p q :
  (a <|q|> b) <|p|> (c <|q|> d) = (a <|p|> c) <|q|> (b <|p|> d).
apply S1_inj; rewrite ![in LHS]S1_conv !convptE.
rewrite !scalept_addpt !scalept_comp //.
rewrite !(mulRC p) !(mulRC p.~) addptA addptC
rewrite (addptC (scalept (q*p) _)) !addptA -addptA
rewrite !(addptC (scalept (_.~ * _.~) _)) {1}addptC.
by rewrite !S1_conv !convptE !scalept_addpt !scalept_comp.
```

Injectivity of ι:

$\iota((a \triangleleft_q \triangleright b) \triangleleft_p \triangleright (c \triangleleft_q \triangleright d)) = \iota((a \triangleleft_p \triangleright c) \triangleleft_q \triangleright (b \triangleleft_p \triangleright d))$

Key lemma:

$\mathrm{LHS} = ((pq)\iota(a) + (p\overline{q})\iota(b)) + ((\overline{p}q)\iota(c) + (\overline{p}\overline{q})\iota(d))$

$\mathrm{RHS} = ((qp)\iota(a) + (q\overline{p})\iota(c)) + ((\overline{q}p)\iota(b) + (\overline{q}\overline{p})\iota(d))$

Associativity and Commutativity:

$\mathrm{LHS} = \mathrm{RHS}$

Multiary axiomatizations

Multiary axiomatizations take the barycenter operation ($\lhd\!\rhd_{i<n} d_i x_i$) as primitive.
Multiary is preferred in many work (e.g. [Bonchi et al., 2017]).
However,

- Different authors use different axiomatizations. How do they compare?
- Their equivalence to the binary axiomatization is often assumed without an explicit proof, citing a lemma by Stone [Stone, 1949, Lemma 4].

We want to prove that these axiomatizations are really equivalent.
This also ensures that our whole formalization is correctly done.

Two multiary axiomatizations

Standard:

- Projection law: if $d_j = 1$, $\underset{i<n}{\diamond} d_i x_i = x_j$.
- Barycenter law: $\underset{i<n}{\diamond} d_i \left(\underset{j<m}{\diamond} e_{i,j} x_j \right) = \underset{j<m}{\diamond} \left(\sum_{i<n} d_i e_{i,j} \right) x_j$.

Beaulieu:

- Partition law: $\underset{i \in I}{\diamond} \lambda_i x_i = \underset{j \in J}{\diamond} \rho_j \left(\underset{k \in K_j}{\diamond} \frac{\lambda_k}{\rho_j} x_k \right)$ where $\{K_j \mid j \in J\}$ is a partition of I, and $\rho_j = \sum_{k \in K_j} \lambda_k \neq 0$.
- Idempotence law: $\underset{i \in I}{\diamond} \lambda_i A_i = A$ if $A_i = A$ for all $\lambda_i > 0$.

Standard ⇔ Beaulieu

Sketch of the proof

Equivalence between Standard and Beaulieu:

- Barycenter ⊢ Partition
- Projection, Barycenter ⊢ Idempotence
- Idempotence ⊢ Projection
- Idempotence, Partition ⊢ Barycenter (difficult)

Binary ⇔ Standard

Bi-interpretability between Binary (convex space) and Standard:

[Binary ⊨ Standard]

- $\diamond_{i<n} d_i x_i := \begin{cases} x_0 & \text{if } d_0 = 1 \\ x_0 \triangleleft d_0 \triangleright (\diamond_{i<n-1} d'_i x_{i+1}) & \text{otherwise} \end{cases}$
  where $d'_i = d_{i+1}/(1 - d_0)$
- Prove Standard axioms for this definition

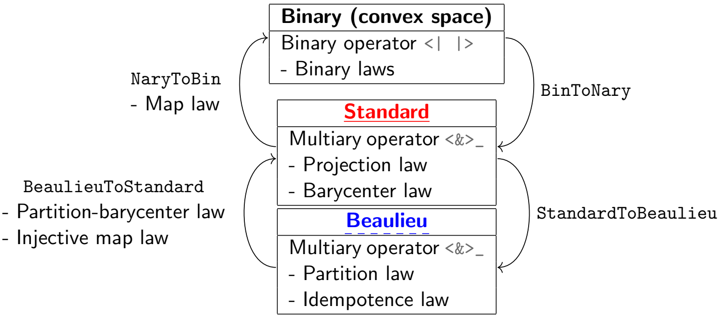
[Standard ⊨ Binary]

- $x_0 \triangleleft p \triangleright x_1 := \diamond_{i<2} d_i x_i$ where $d_0 = p$ and $d_1 = 1 - p$
- Prove Binary axioms for this definition

[Interpretations are inverse to each other]

- $\left\{ \begin{matrix} \text{Binary } \triangleleft_p \triangleright \\ \text{Standard } \diamond_i \end{matrix} \right\} \xrightarrow{\text{def}} \left\{ \begin{matrix} \diamond_i \\ \triangleleft_p \triangleright \end{matrix} \right\} \xrightarrow{\text{def}} \left\{ \begin{matrix} \triangleleft_p \triangleright' \\ \diamond'_i \end{matrix} \right\}$
- Prove $\triangleleft_p \triangleright = \triangleleft_p \triangleright'$ and $\diamond_i = \diamond'_i$

Overview of the equivalences



Conclusion

We formalized:

- Convex spaces with several interfaces
- Equivalence lemmas connecting them
- Further useful constructions such as convex sets and hulls

On-going work includes:

- Convexity of information theoretic functions [Infotheo, 2020]
- Category of convex spaces and probabilistic programming [Affeldt et al., 2019]

Future work:

- What is the true difference between conical and convex spaces?
- Ordered convex spaces with monotonicity
- Topological convex spaces
- Affine Lie algebras

Affeldt, R., Garrigue, J., Nowak, D., Saikawa, T., Sauvage, C., and Tanaka, K. (2019).
Monadic equational reasoning in Coq.
https://github.com/affeldt-aist/monae/.
Coq scripts.

Bonchi, F., Silva, A., and Sokolova, A. (2017).
The Power of Convex Algebras.
In Meyer, R. and Nestmann, U., editors, 28th International Conference on Concurrency Theory (CONCUR 2017), volume 85 of Leibniz International Proceedings in Informatics (LIPIcs), pages 23:1–23:18. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Flood, J. (1981).
Semiconvex geometry.
Journal of the Australian Mathematical Society, 30(4):496–510.

Fritz, T. (2015).
Convex spaces I: Definition and examples.
Available at https://arxiv.org/abs/0903.5522. First version: 2009.

Infotheo (2020).
A Coq formalization of information theory and linear error-correcting codes.
https://github.com/affeldt-aist/infotheo/.
Coq scripts.

Jones, C. and Plotkin, G. D. (1989).
A probabilistic powerdomain of evaluations.
In [1989] Proceedings. Fourth Annual Symposium on Logic in Computer Science, pages 186–195.

Stone, M. H. (1949).
Postulates for the barycentric calculus.
Ann. Mat. Pura Appl., 29(1):25–30.