Categorical Probability and Stochastic Dominance in Metric Spaces

Submitted PhD Thesis

Paolo Perrone*

Max Planck Institute for Mathematics in the Sciences
Leipzig, Germany

5th July 2018

Abstract

In this work we introduce some category-theoretical concepts and techniques to study probability distributions on metric spaces and ordered metric spaces. The leading themes in this work are *Kantorovich duality* [Vil09, Chapter 5], *Choquet theory* [Win85, Chapter 1], and the categorical theory of *monads and their algebras* [Mac00, Chapter VI].

Categorical Probability. In Chapter 1 we give an overview of the concept of a *probability monad*, first defined by Giry [Gir82].

Probability monads can be interpreted as a categorical tool to talk about *random elements* of a space. Given a space X, we can assign to it a space PX, which extends X by allowing extra elements, *random elements*. We can consider these random elements as *formal convex combinations, or mixtures, of elements of X*. For example, the law of a fair coin flip is 1/2 "heads" + 1/2 "tails". Of course, in general, such mixtures are given

*Correspondence: perrone [at] mis.mpg.de

by integrals rather than just sums. Probability monads allow to iterate the construction, and talk about the space PPX of *random elements with random law.* Given such an element of PPX, one can always *integrate* it to obtain a simple probability measure in PX. In other words, integration always defines a map $E : PPX \to PX$.

Spaces where the convex combinations can be *actually evaluated*, so that they are well-defined operations, are called *algebras of the probability monad.* These are the spaces, for example $\mathbb{R}$, where one can take *expectation values* of random variables. The set {"heads", "tails"} is *not* an algebra of the monad: there is no element, or *deterministic* state which correspond to "halfway between heads and tails".

As it is known, to every monad corresponds an adjunction. For probability monads, this adjunction can be interpreted in terms of Choquet theory [Win85, Chapter 1]: given any object X and any algebra A, there is a natural bijection between maps $X \to A$ and *affine* maps $PX \to A$.

The Kantorovich Monad. In Chapter 2 we define a probability monad on the category of complete metric spaces and 1-Lipschitz maps called the *Kantorovich monad*, extending a previous construction due to van Breugel [vB05]. This monad assigns to each complete metric space X its *Wasserstein space* PX, which is itself a complete metric space [Vil09].

It is well-known [Vil09, Chapter 6] that finitely supported probability measures with rational coefficients, or *empirical distributions of finite sequences*, are dense in the Wasserstein space. This density property can be translated into categorical language as a *universal property* of the Wasserstein space PX, namely, as a *colimit* of a diagram involving certain powers of X. The monad structure of P, and in particular the integration map E, is uniquely determined by this universal property, without the need to define it in terms of integrals or measure theory. In some sense, the universal property makes the integration map *inevitable*, it arises directly from the characterization of P in terms of finite powers.

We prove that the algebras of the Kantorovich monad are exactly the *closed convex subsets of Banach spaces*. In the spirit of categorical probability, these can be interpreted as the *complete metric spaces with a well-defined notion of convex combinations.* The "Choquet adjunction" that we obtain is then the following: *given a complete metric space X and a Banach space A, there is a natural bijection between short maps $X \to A$ and short affine maps $X \to A$.*

In the end of the chapter we show that both the integration map $E : PPX \to PX$ and the marginal map $\Delta : P(X \times Y) \to PX \times PY$ are *proper maps.* This means in particular that the set of probability measures *over* a Wasserstein space PX which integrate to a given measure $p \in PX$ is always compact, and analogously, that the set of couplings of any two probability measures p and q is compact as well. As a consequence, on every complete metric space, every Kantorovich duality problem admits an optimal solution.

Stochastic Orders. In Chapter 3 we extend the Kantorovich monad of Chapter 2 to metric spaces equipped with a partial order. The order is inherited by the Wasserstein space, and is called the *stochastic order*. Differently from most approaches in the literature, we define a compatibility condition of the order with the *metric itself*, rather then with the topology it induces. We call the spaces with this property *L-ordered spaces*.

On L-ordered spaces, the stochastic order induced on the Wasserstein spaces satisfies itself a form of Kantorovich duality: given two measures p, q, we can say that $p \leq q$ if and only if they admit a coupling r such that for all the points (x, y) in the support of r we have $x \leq y$. An interpretation is that *there exists a transport plan that moves the mass only upwards in the order, not downwards.* Alternatively, we can say that $p \leq q$ if and only if for all *monotone* 1-Lipschitz functions $\int_X f\, dp \leq \int_X f\, dq$.

This Kantorovich duality property implies that the stochastic order on L-ordered spaces is always a partial order, i.e. it is antisymmetric.

The Kantorovich monad of Chapter 2 can be extended naturally to the category of L-ordered metric spaces. We prove that its algebras are the closed convex subsets of *ordered* Banach spaces, i.e. Banach spaces equipped with a partial order induced by a closed cone. The integration map on ordered Banach spaces is always monotone, and we prove that it is even *strictly* monotone: if $p \leq q$ for the stochastic order and p and q have the same expectation value, then $p = q$. This generalizes a result which is long known for the real line.

We can consider the category of L-ordered metric spaces as *locally posetal 2-categories*, with the 2-cells given by the pointwise order of the functions. This gives an order-theoretical version of the "Choquet adjunction": given an L-ordered complete metric space X and an ordered Banach space A, there is a natural *isomorphism of partial orders* between short *monotone* maps $X \to A$ and short affine *monotone* maps $X \to A$.

Moreover, in this 2-categorical setting, we can describe *concave and convex maps* categorically, exactly as the *lax and oplax morphisms of algebras.*

Convex Orders. In Chapter 4 we study a different order between probability measures, which can be interpreted as *pointing in the direction of increasing randomness.*

We have seen that probability monads can be interpreted in terms of formal convex combinations, and that their algebras can be interpreted as spaces where such convex combinations can be evaluated. Here we develop a *new* categorical formalism to describe *operations evaluated partially*. For example, "5+4" is a partial evaluation of the sum "2+3+4". We prove that partial evaluations for the Kantorovich monad, or *partial expectations*, define a closed partial order on the Wasserstein space PA over every algebra A, and that the resulting ordered space is itself an algebra.

We prove that, for the Kantorovich monad, these partial expectations correspond to conditional expectations *in distribution*. This implies that the partial evaluation order is equivalent to the order known in the literature as the *convex* or *Choquet order* [Win85].

A useful consequence of this equivalence and of the fact that the integration map E is proper is that *bounded monotone nets in the partial evaluation order always converge*. This fact can be interpreted as a result of *convergence in distribution for martingales and inverse martingales over general Banach spaces*.

Given an algebra A, we can compare the partial evaluation order and the stochastic order on PA. We show that the two orders are *transverse*, in the sense that every two probability distributions comparable for both orders are necessarily equal. We can also combine the two orders to form a new order, which we call the *lax partial evaluation order*. The space PA with this order also forms an algebra.

Finally, we study the relation between these partial evaluation orders and convex functions. As is well-known [Win85], the Choquet order is dual to convex functions. We know from Chapter 3 that convex functions are the oplax morphisms of algebras. This is not a coincidence: as we show, the partial evaluation order and convex functions are related by the "ordered Choquet adjunction" of Chapter 3. This permits to characterize the partial evaluation order in terms of a universal property, as an *oplax codescent object* [Lac02]. From this universal property we can derive a general duality result valid on *all ordered Banach spaces*, which says that over every ordered Banach space A, the lax partial evaluation order is dual to *monotone convex functions*. In other words, for every two probability measures p and q over A, $\int f dp \leq \int f dq$ for all convex monotone functions f *if and only if* $p \preceq_l q$ for the lax partial evaluation order. As far as we know, this result in its full generality is new.

Sources. Part of this work is contained in the papers [FP17] and [FP18]. The rest will appear in two papers which are currently in preparation.

References

[FP17] Tobias Fritz and Paolo Perrone. A Probability Monad as the Colimit of Finite Powers, 2017. Submitted. arXiv:1712.05363.

[FP18] Tobias Fritz and Paolo Perrone. Bimonoidal Structure of Probability Monads. *Proceedings of MFPS 34, ENTCS*, 2018. arXiv:1804.03527.

[Gir82] Michèle Giry. A Categorical Approach to Probability Theory. In *Categorical aspects of topology and analysis*, volume 915 of *Lecture Notes in Mathematics.* 1982.

[Lac02] Stephen Lack. Codescent objects and coherence. *Journal of Pure and Applied Algebra*, 175(1-3), 2002.

[Mac00] Saunders Mac Lane. *Categories for the Working Mathematician.* Springer, 2000.

[vB05] Franck van Breugel. The Metric Monad for Probabilistic Nondeterminism. Available at http://www.cse.yorku.ca, 2005.

[Vil09] Cédric Villani. *Optimal transport: old and new*, volume 338 of *Grundlehren der mathematischen Wissenschaften.* Springer, 2009.

[Win85] Gerhard Winkler. *Choquet order and simplices with applications in probabilistic models.* Lecture Notes in Mathematics. Springer, 1985.