

Unimodal Category and Topological Statistics

Yuliy Baryshnikov[†] and Robert Ghrist[‡]

[†]Mathematical and Algorithmic Sciences
Bell Laboratories, Murray Hill NJ, USA

[‡]Departments of Mathematics & Electrical/Systems Engineering
University of Pennsylvania, Philadelphia PA, USA
Email: ymb@research.bell-labs.com, ghrist@math.upenn.edu

Abstract—We consider the problem of decomposing a compactly supported distribution $f : \mathbb{R}^n \rightarrow [0, \infty)$ into a minimal number of unimodal components by means of some convex operation (*e.g.*, sum or sup). The resulting “unimodal category” of f is a topological invariant of the distribution which shares a number of properties with the Lusternik-Schnirelmann category of a topological space. This work introduces the concept of a unimodal category, provides fundamental examples (ℓ^p unimodal category), and computes various unimodal categories for distributions on \mathbb{R}^n with n small.

1. Introduction

This notes initiates a novel topological approach to the problem of decomposing distributions $f : \mathbb{R}^n \rightarrow [0, \infty)$ into a convex combination of basis distributions. Instead of employing a decomposition into analytically defined (*e.g.*, normal) distributions, we propose a decomposition into topologically defined factors. Specifically, we consider the decomposition of a distribution into a sum of **unimodal** distributions: those with a single maximum value and no other extrema. Such a decomposition is not uniquely defined; however, the minimal number of unimodal summands is. This **unimodal category**, $\text{ucat}(f)$, is a coarse measure of complexity for a distribution.

1.1. Motivation

There are several contexts within which the question about the number of summands in a unimodal decomposition arises. The first such context is statistics: unimodal distributions are the primal building blocks of statistical models. Essentially all classical probability distributions, including normal, Poisson, Gamma, Beta, Bernoulli, and more, are all unimodal. The methodology of statistical modeling essentially forces one to assume that the presence of several modes in a distribution is a consequence of its being a mixture of several distributions, and the relationship between the number of modes (essentially, the number of local maxima of the density, in the multivariate case) and the number of “components” of the mixture, *i.e.*, the number of summands in the convex decomposition, has been studied by many authors.

To justify the radical difference of our setup from the traditional statistical one — our lack of any assumptions about the structure of the summands beyond unimodality — we can invoke two considerations: (1) The variety of unimodal building blocks in standard statistical models suggests that one should try to abstract away any specific distribution, retaining only the minimal topological properties; and (2) Any specific analytic form of a density binds the distribution to some fixed coordinate system, the absence of which will inexorably force a topological approach. We foresee applications of this topological decomposition to a variety of contexts within statistics, as well as visualization (where decomposition can play a role in efficient encoding of the images or of multidimensional data).

1.2. Statement of Results

We define the unimodal category, give a complete characterization and method of computation in the univariate case, and provide some key steps to understanding decompositions of higher-dimensional distributions. The specific contributions include (1) that the unimodal category is invariant under the right-action of a homeomorphism on the domain; (2) a simple algorithm for its computation in the univariate case; and (3) a bivariate characterization as a function of the Reeb graph of the distribution, labeled by critical values. We close with a conjecture concerning the monotonicity of the unimodal category with respect to pointwise norms used.

1.3. Related work

Morse structures associated to mixtures of multivariate normal distributions are discussed in [8]. This question already seems to be of interest to statisticians — even the simple univariate case is discussed in detail across several papers [4, 1, 9], while the mixture of non-Gaussian unimodal densities is considered in [5, 6]. In particular, it is known that in mixtures of normal univariate distributions, the number of modes cannot exceed the number of components, a result which does not hold in higher dimensions (compare [2]).

2. Unimodal category

Definition 1 For X a topological space, the **Lusternik-Schnirelmann category** of X , $\mathbf{LScat}(X)$, is the minimum number of open sets contractible in X which cover X . The **geometric category** of X , $\mathbf{gcat}(X)$, is the minimum number of open contractible sets which cover X .

Some authors (including those of [3]) use a reduced category, which measures the minimal number of open sets minus one. We do not follow this convention. Geometric category \mathbf{gcat} is a homeomorphism invariant of a space, and L-S category \mathbf{LScat} is a homotopy invariant. There are numerous deep connections between category and critical point theory (the classical motivation for the subject), dynamical systems, homotopy theory, and symplectic topology. We introduce a new variant of category for distributions based on decomposition into unimodal factors. For X a topological space, let $\mathfrak{D} = \mathfrak{D}(X)$ denote the set of all compactly supported continuous distributions $f : X \rightarrow [0, \infty)$.

Definition 2 A distribution $u \in \mathfrak{D}$ is said to be **unimodal** if the upper excursion sets $u^c = u^{-1}([c, \infty))$ have the homotopy type of a point for all $0 < c \leq M$ and are empty for all $c > M$. Such a u has M as its maximal value.

We will refer to the nonempty upper excursion sets $u^c \subset X$ as being **contractible**, though it must be clarified that such sets are contractible *in themselves* as opposed to contractible in X . The latter would be more in line with the definitions used in Lusternik-Schnirelmann theory, but would render the theory useless for most applications (where $X = \mathbb{R}^n$).

Definition 3 Fix a norm $\nu = \|\cdot\|$ on \mathbb{R}^N . The **unimodal ν -category** of a distribution $f \in \mathfrak{D}(X)$ is defined as the minimal number \mathbf{ucat}^ν of unimodal distributions $u_\alpha, \alpha = 1, \dots, \mathbf{ucat}^\nu$ on X such that f is pointwise the ν -norm of the collection (u_α) . Specifically, $f(x) = \|(u_\alpha(x))\|$ for all $x \in X$.

The most natural and fundamental example is the minimal number of unimodal distributions required to represent a given distribution as a sum of unimodals. Summation of the components corresponds to the 1-norm on vectors, leading to the following generalization.

Definition 4 The **unimodal p -category** of a distribution $f \in \mathfrak{D}$ is the minimal number of unimodal distributions $u_\alpha, \alpha = 1, \dots, \mathbf{ucat}^p$ such that f is pointwise an ℓ^p combination. Specifically,

$$0 < p < \infty : f(x) = \left(\sum_{\alpha} (u_{\alpha}(x))^p \right)^{\frac{1}{p}} \quad (1)$$

$$p = \infty : f(x) = \max_{\alpha} \{u_{\alpha}(x)\} \quad (2)$$

The category \mathbf{ucat}^2 measures something akin to *energy* of a distribution, which \mathbf{ucat}^∞ is a natural measure for problems in which mode interference is negligible. All of these unimodal categories can be viewed as a deformation from the geometric category of the support, $\mathbf{gcat}(\mathbf{supp}(f))$, which we will identify later as $\mathbf{ucat}^0(f) = \lim_{p \rightarrow 0^+} \mathbf{ucat}^p(f)$.

The following result, while trivial, is the topological invariance of unimodal category that makes it applicable to problems in which distribution data is without a fixed coordinate system.

Lemma 5 Any unimodal ν -category is invariant under the right-action of the homeomorphism group.

Proof: Let $u \in \mathfrak{D}$ be unimodal and let $h \in \mathbf{Homeo}(X)$. Then $(u \circ h)^c = h(u^c)$, which, being the homeomorphic image of a contractible set, is contractible. \square

3. Unimodal 1-category

The most natural variant of unimodal category appears to be \mathbf{ucat}^1 , for which the modal decomposition problem is additive.

Lemma 6 Assume that f is a Morse function on a manifold M . Then $\mathbf{ucat}^1(f) \leq \#\max(f)$, the number of local maxima of f .

Proof: The unstable manifolds $W^u(p_j)$ of the local maxima p_j of f are disjoint discs whose closures cover M . Let u_j be the characteristic function of $W^u(p_j)$ convolved with a bump function so as to smooth it to zero near the boundary. These are clearly unimodal. Standard bump function methods yield that $\sum_j u_j = f$. \square

For non-Morse functions, one can replace the upper bound with the geometric category of the set of maxima. In either context, the upper bound is often not sharp. An exact answer can be computed in some cases.

Definition 7 Assume that a unimodal decomposition of $f = \sum_{\alpha} u_{\alpha}$ is given. A set $U \subset \mathbf{supp}(f)$ is called **max-free** if U does not contain any of the critical points of any f_{α} . For any max-free U we denote by $\mathbf{dep}(U)$ the **depth** of U , the number of functions of the unimodal decomposition not vanishing on U :

$$\mathbf{dep}(U) = \#\{\alpha : U \cap \mathbf{supp}(u_{\alpha}) \neq \emptyset\}.$$

The following trivial lemma is useful:

Lemma 8 The depth of an open max-free set U is bounded from below by

$$\mathbf{dep}(f) \leq \left\lceil \frac{\max_U f}{\max_{\partial U} f} \right\rceil.$$

Proof: On a max-free U , all of the functions u_α attain their maxima on the boundary. \square

Computing unimodal 1-category solves other p -category problems:

Lemma 9 For all $f \in \mathfrak{D}$, $\mathbf{ucat}^p(f) = \mathbf{ucat}^1(f^{1/p})$.

Proof: Unimodality is preserved under powers, and

$$f = \left(\sum_{\alpha} u_{\alpha}^p \right)^{\frac{1}{p}} \Leftrightarrow f^p = \sum_{\alpha} u_{\alpha}^p.$$

\square

We commence with a computation of the unimodal 1-category on $\mathfrak{D}(\mathbb{R}^1)$. We assume that this function has isolated critical points. Up to a homeomorphism, this distribution is completely characterized by the *up-down* sequence of critical values corresponding to local minima and maxima,

$$0 = n_0 < m_0 > n_1 < \dots < m_k > n_k = 0;$$

where $n_i = f(x_{2i})$; $m_i = f(x_{2i+1})$; $x_0 < x_1 < \dots < x_{2k}$; and $i = 0, \dots, k$, counting the initial and final point of the support as local minima.

Proposition 10 If an open interval (x_{2i}, x_{2j}) bounded by local minima is max-free (for some unimodal decomposition of f), then

$$n_i - m_i + n_{i+1} - \dots - m_{j-1} + n_j \geq 0. \quad (3)$$

Proof: The result follows immediately from the combination of the following facts: (1) The total variation is subadditive; and (2) The total variation of a function monotone on an interval is at most its maximal value on the interval (attained at one of the boundary points). Computing the total variations for f concludes the proof. \square

Consider the open intervals with endpoints at the local minima $\{x_{2i}\}_1^k$. Call such an interval **forced-max** if the inequality (3) is violated there. Obviously, forced max intervals form an ideal: any interval containing a forced-max interval is itself forced-max.

Theorem 11 Let $f \in \mathfrak{D}(\mathbb{R})$ have maximal values $(m_i)_1^k$ and minimal values $(n_i)_0^k$ ordered according to the critical point order in the domain. Then $\mathbf{ucat}^1(f)$ is equal to the maximal number of non-intersecting forced-max intervals:

$$\mathbf{ucat}^1(f) = \max \{N : (x_{2i_0}, x_{2i_1}), \dots, (x_{2i_{N-1}}, x_{2i_N}) \text{ forced max}\}. \quad (4)$$

Proof: It is clear that for any collection of N nonintersecting forced-max intervals, the number of summands in the unimodal decomposition cannot be less than N : each forced-max interval, trivially, contains a critical point of at least one of the functions of the decomposition. The following algorithm yields an explicit unimodal decomposition *and*, simultaneously, a collection of nonintersecting forced-max intervals, one for each summand.

We construct the functions u_α iteratively, left to right, according to Algorithm **Sweep**. This entails sweeping f from the left and pulling out unimodal factors which, on their descent, compensate for the remaining factors as much as possible by descending according to the (positive) slope df .

Algorithm 1 $\{u_\alpha\} = \mathbf{Sweep}(f)$

Require: $f \in \mathfrak{D}(\mathbb{R})$ with minima $n_i = f(x_{2i})_0^k$ and maxima $m_i = f(x_{2i-1})_1^k$

- 1: $u_0 \leftarrow 0$; $\alpha \leftarrow 1$; $g_\alpha \leftarrow f$
- 2: **while** $g_\alpha \neq 0$ **do**
- 3: $y_\alpha \leftarrow$ first maximum of g_α from left
- 4: $u_\alpha|_{(-\infty, y_\alpha]} \leftarrow g_\alpha$
- 5: $du_\alpha|_{(y_\alpha, \infty)} \leftarrow \min(df, 0)$
- 6: $u_\alpha \leftarrow \max(u_\alpha, 0)$
- 7: increment α
- 8: $g_\alpha \leftarrow f - \sum_{\beta < \alpha} u_\beta$
- 9: **end while**
- 10: **return** $\{u_\beta\}_{\beta < \alpha}$

The proof that this construction is minimal follows from the observation that the local minima spanning the curves in the graphical construction form a max-forced partition of the support of f . \square

We have observed that, in the univariate case, the unimodal 1-category is a function of the critical values and the order in which they appear; however, there is not a strict dependence on this ordering. For example, in accordance with Lemma 5, \mathbf{ucat}^1 must be invariant under reversing the order of the critical points ($x \mapsto -x$). This in itself is an interesting observation: running Algorithm **Sweep** on $f(-x)$ reveals the limits to uniqueness of modal decompositions. There is in fact a weaker dependence on the labeled Reeb graph of the distribution.

Definition 12 The **Reeb graph** Γ_f of a $f \in \mathfrak{D}(X)$ with image $[0, L]$ is the rooted metric graph given by the quotient space $\Gamma_f = X \times [0, L] / \sim$, where $(x, c) \sim (x', c)$ iff $c = f(x) = f(x')$ and x, x' lie in the same connected component of $f^{-1}(c)$. The metric on Γ_f is induced by that on $[0, L]$ and the root of Γ_f is the point at height zero. To each edge of the Reeb graph we may associate the diffeomorphism type of the corresponding connected component of $f^{-1}(c)$. We will refer to this system of marks as the **Reeb data**.

In particular, the vertices of the graph correspond to the connected singular level sets of f and the edges correspond to non-singular level sets. For $f \in \mathfrak{D}(\mathbb{R})$ of reasonable non-degeneracy, Γ_f is a finite tree. The construction of this tree from an excursion is well-known in combinatorics and probability theory (cf. [7]). It follows from the proof of Theorem 11 that the dependence of \mathbf{ucat}^1 on the critical values is independent of their ordering; in other words, one can swap the subexcursions bordering at a local maxima without affecting the unimodal number.

Corollary 13 On $\mathfrak{D}(\mathbb{R})$, \mathbf{ucat}^1 is a function of the isomorphism type of the Reeb graph (as a rooted, metric graph).

4. Toward multivariate decompositions

Multivariate distributions introduce a number of complexities. We sketch a few results in this section, reserving a more complete treatment for a later work. An indication of the troubles arising in the multivariate case, we notice that an analogue of Proposition 10 is not valid there. For example, a convex combination of two unimodal bivariate functions ($\exp(-(y/100)^2 - (x - \sin(y)/5)^2)$ and $\exp(-(y/100)^2 - (x + 0.3)^2)$ with weights $\frac{1}{2}$) reveals a distribution with higher-appearing complexity. The dependence of the unimodal category on the Reeb graph noted in Corollary 13 can be extended to planar distributions:

Proposition 14 The unimodal 1-category of $f \in \mathfrak{D}(\mathbb{R}^n)$ for $n = 1, 2$ is a function of the combinatorial type of the Reeb graph of f labeled by critical values.

Proof: For dimension one, this is Corollary 13. In dimension two, the result follows from the fact that a Morse function on the plane can be reconstructed, up to a diffeomorphism of the domain, from the isotopy class of the figures eight formed by the connected components of the level sets containing critical points of f , and the critical values at these points. \square

In higher dimensions, the Reeb graph alone does not carry enough information: for example, a critical point of signature $(2, 1)$ is a vertex of valence 2 on the Reeb graph, with no information about the change of topology under the surgery defined by the point. In dimension two, the change of topology is unambiguous; in higher dimensions it is necessary to specify the Reeb data.

Proposition 15 Fix a (non-unimodal) C^1 distribution $f \in \mathfrak{D}(\mathbb{R}^n)$. For any $0 < p < \infty$, the unimodal p -category of the shifted distribution obtained by adding $C\mathbf{1}_{\text{supp}f}$ to f for constant $C \geq 0$ is a non-increasing function of C which stabilizes to

$$\lim_{C \rightarrow \infty} \mathbf{ucat}^p(f + C\mathbf{1}_{\text{supp}f}) = 1 + \mathbf{gcat}(\text{supp}f). \quad (5)$$

Proof: By Lemma 9 it suffices to prove the result for $p = 1$. Monotonicity in C is clear: a constant added to a unimodal function is still unimodal. Assume by compactness and smoothness that the differential satisfies $\|df\| \leq M$. Let $x_0 \in \text{supp}f$ be the maximum of f and let κ denote the conical function $\kappa(x) = M\|x - x_0\|$ (smoothed at the cone point if desired). Then

$$f + C\mathbf{1}_U = (f + (C - \kappa)\mathbf{1}_U) + (\kappa\mathbf{1}_U),$$

and these summands are unimodal for C sufficiently large. \square

If one normalizes the distributions $f + C\mathbf{1}_{\text{supp}f}$ to have unit mass, then one sees clearly the effect of increasing C is to reduce the total variation.

5. Unimodal ∞ -category

The norm used in a unimodal category can be considered as a model of interference between modes, with \mathbf{ucat}^1 measuring additive mode interference. One natural limit is perfect non-interference, as measured by the ∞ -norm: only the dominant mode matters. Computing the unimodal ∞ -category in the univariate setting is trivial.

Lemma 16 For any $f \in \mathfrak{D}(\mathbb{R}^1)$, $\mathbf{ucat}^\infty(f)$ equals the number of local maxima of f .

Proof: The upper bound is obtained via Lemma 6; the lower bound is trivial. \square

Lemma 17 \mathbf{ucat}^∞ is invariant under the left action of $\mathbf{Homeo}[0, \infty)$.

Proof: For $h \in \mathbf{Homeo}[0, \infty)$, $\max_\alpha \{u_\alpha\} = \max_\alpha \{h \circ u_\alpha\}$. \square

In cases where exact computation of unimodal category is difficult, one turns to bounds based on analytic, geometric, or topological features of the distribution. We believe that \mathbf{ucat}^∞ will prove easier to compute on multivariate distributions, and that the following will prove instrumental in obtaining clear bounds.

Conjecture 18 For any $0 < p < q \leq \infty$ and any $f \in \mathfrak{D}(X)$, $\mathbf{ucat}^p(f) \leq \mathbf{ucat}^q(f)$.

Acknowledgments

The work reported here was supported by DARPA HR0011-07-1-0002 and by ONR N000140810668.

References

- [1] J. Behboodan, On the modes of a mixture of two normal distributions. *Technometrics*, 12, 1970, 131–139.
- [2] M. Carreira-Perpinan and C. Williams, “On the number of modes of a Gaussian mixture,” in *Scale-Space Methods in Computer Vision, Lecture Notes in Comput. Sci.*, vol. 2695, 2003, 625–640.
- [3] O. Cornea, G. Lupton, J. Oprea, and D. Tanré, *Lusternik-Schnirelmann Category*, Amer. Math. Soc., 2003.
- [4] I. Eisenberger, “Genesis of bimodal distributions,” *Technometrics*, 6, 1964, 357–363.
- [5] I. Kakiuchi, “Unimodality conditions of the distribution of a mixture of two distributions,” *Math. Sem. Notes Kobe Univ.*, 9, 1981, 315–325.
- [6] J. Kemperman, “Mixtures with a limited number of modal intervals,” *Ann. Statist.* 19, 1991, 2120–2144.
- [7] J. Pitman, *Combinatorial Stochastic Processes*. Springer Lecture Notes in Mathematics, 1875, Springer-Verlag, Berlin, 2006.
- [8] S. Ray and B. Lindsay, “The Topography Of Multivariate Normal Mixtures”, *Annals of Statistics*, Vol. 33, No. 5, 2005, 2042–2065.
- [9] C. Robertson and J. Fryer “Some descriptive properties of normal mixtures,” *Skand. Aktuarietidskr*, 1969, 137–146.