Why

We can approximate a density with a tree density similar to how we can approximate a distribution with a tree distribution.

Definition

We use the differential relative entropy as a criterion of approximation. An *optimal tree approximator* of density for a tree is a density which factors according to a tree and minimizes its differential relative entropy with the given density.

Notation

Let $g : \mathbf{R}^n \to \mathbf{R}$ be a density and T be a tree on $\{1, \ldots, n\}$. An optimal tree approximator of g for T is a density f that factors according to T and minimizes $d(g, f)$. In other words, given g and T we want to find f to

$$\begin{array}{ll} \text{minimize} & d(g, f) \\ \text{subject to} & f \text{ factors according to } T. \end{array}$$

Result

Prop. 1. *Let $g : \mathbf{R}^n \to \mathbf{R}$ be a density and T be a tree on $\{1, \ldots, n\}$. The density $f_T^* : \mathbf{R}^d \to \mathbf{R}$ defined by*

$$f_T^* = g_1 \prod_{i \neq 1} g_{i|\mathbf{pa}\, i}$$

minimizes the differential relative entropy with g among all densities on $\mathbf{R}^n$ which factor according to T ($\mathbf{pa}\, i$ is the parent

of i in T rooted at vertex 1, $i = 2, \ldots, n$).

Proof. Let $f : \mathbf{R}^d \to \mathbf{R}$ be a density factoring according to T. First, express

$$f = f_1 \prod_{i=1} f_{i|\mathbf{pa}\,i}.$$

Second, recall that $d(g, f) = h(g, f) - h(g)$. Since $h(g)$ does not depend on f, f is a minimizer of $d(g, f)$ if and only if f is a minimizer of $h(g, f)$.

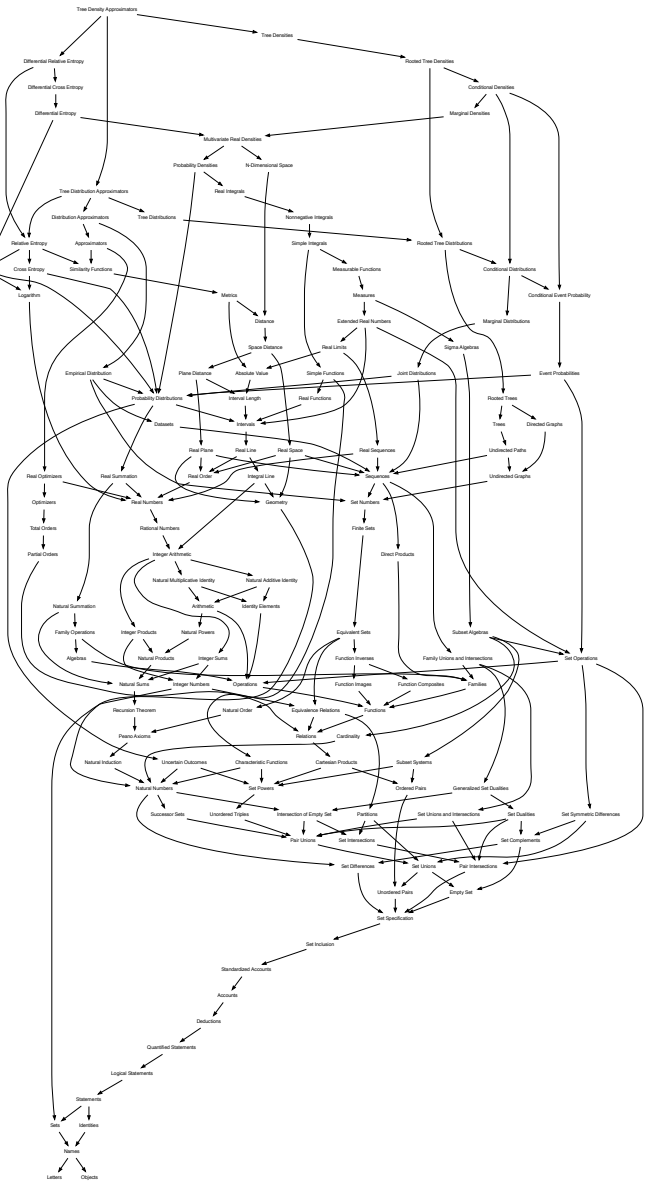
Third, express

$$\begin{aligned} h(g, f) &= -\int_{\mathbf{R}^d} g \log f \\ &= -\int_{\mathbf{R}^d} g(x) \left(\log f_i(x_i) + \sum_{i \neq 1} \log f_{i|\mathbf{pa}\,i}(x_i, x_{\mathbf{pa}\,i}) \right) dx \\ &= h(g_1, f_1) + \sum_{i \neq 1} \left(\int_{\mathbf{R}} g_{\mathbf{pa}\,i}(\xi) h\big(g_{i|\mathbf{pa}\,i}(\cdot, \xi), f_{i|\mathbf{pa}\,i}(\cdot, \xi)\big) d\xi \right) \end{aligned}$$

which separates across f_1 and $f_{i|\mathbf{pa}\,i}(\cdot, \xi)$ for $i = 1, \ldots, n$ and $\xi \in \mathbf{R}$. In particular, since $g_{pai} \geq 0$, we can minimize the integrand pointwise.

Fourth, recall $h(\phi, \psi) \geq 0$ for densities ϕ, ψ of any dimension, and zero if $\phi = \psi$. So $f_1 = g_1$ and $f_{i|\mathbf{pa}\,i} = g_{i|\mathbf{pa}\,i}$ are solutions.

□


Tree Density Approximators
Tree Densities
Differential Relative Entropy
Rooted Tree Densities
Differential Cross Entropy
Conditional Densities
Differential Entropy
Marginal Densities
Multivariate Real Densities
Probability Densities
N-Dimensional Space
Real Integrals
Tree Distribution Approximators
Distribution Approximators
Tree Distributions
Nonnegative Integrals
Relative Entropy
Approximators
Rooted Tree Distributions
Simple Integrals
Entropy
Cross Entropy
Similarity Functions
Measurable Functions
Conditional Distributions
Logarithm
Metrics
Measures
Conditional Event Probability
Distance
Extended Real Numbers
Marginal Distributions
Space Distance
Real Limits
Sigma Algebras
Empirical Distribution
Plane Distance
Absolute Value
Simple Functions
Joint Distributions
Event Probabilities
Probability Distributions
Interval Length
Real Functions
Rooted Trees
Trees
Directed Graphs
Datasets
Intervals
Real Plane
Real Line
Real Space
Real Sequences
Undirected Paths
Real Optimizers
Real Summation
Real Order
Integral Line
Sequences
Undirected Graphs
Optimizers
Real Numbers
Geometry
Set Numbers
Total Orders
Rational Numbers
Finite Sets
Partial Orders
Integer Arithmetic
Direct Products
Natural Multiplicative Identity
Natural Additive Identity
Natural Summation
Arithmetic
Identity Elements
Equivalent Sets
Subset Algebras
Family Operations
Integer Products
Natural Powers
Algebras
Natural Products
Integer Sums
Function Inverses
Family Unions and Intersections
Set Operations
Natural Sums
Integer Numbers
Operations
Function Images
Function Composites
Families
Recursion Theorem
Natural Order
Equivalence Relations
Functions
Peano Axioms
Relations
Cardinality
Natural Induction
Uncertain Outcomes
Characteristic Functions
Cartesian Products
Subset Systems
Natural Numbers
Set Powers
Ordered Pairs
Generalized Set Dualities
Successor Sets
Unordered Triples
Intersection of Empty Set
Partitions
Set Unions and Intersections
Set Dualities
Set Symmetric Differences
Pair Unions
Set Intersections
Set Complements
Set Differences
Set Unions
Pair Intersections
Empty Set
Unordered Pairs
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects