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 $h(g, f)$.

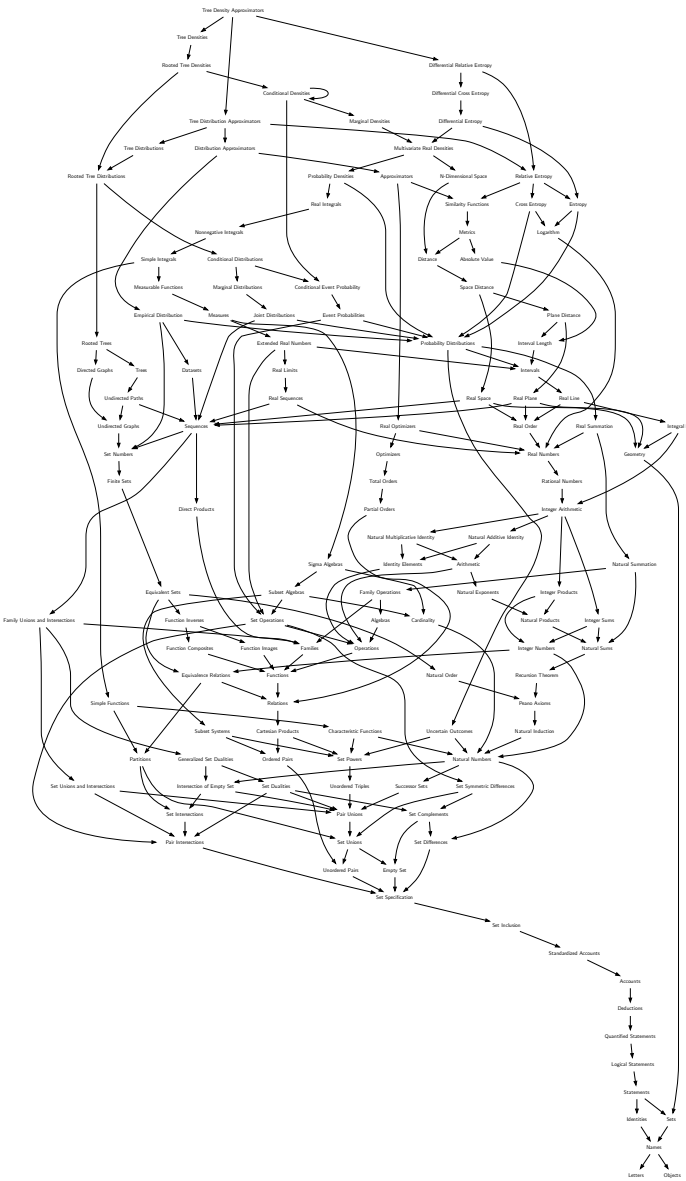
Third, express

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

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