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

$$\text{minimize} \quad d(g, f)$$
$$\text{subject to} \quad f \text{ factors according to } T.$$

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 (**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.

$\square$

Tree Density Approximators

Tree Densities

Rooted Tree Densities

Conditional Densities

Marginal Densities

Differential Relative Entropy

Differential Cross Entropy

Differential Entropy

Tree Distribution Approximators

Tree Distributions

Multivariate Real Densities

Distribution Approximators

Rooted Tree Distributions

Probability Densities

N-Dimensional Space

Approximators

Relative Entropy

Conditional Distributions

Real Integrals

Similarity Functions

Entropy

Cross Entropy

Conditional Event Probability

Nonnegative Integral

Metrics

Logarithm

Marginal Distributions

Simple Integral

Distance

Absolute Value

Measurable Functions

Space Distance

Empirical Distribution

Event Probabilities

Measures

Joint Distribution

Plane Distance

Probability Distributions

Rooted Trees

Extended Real Numbers

Interval Length

Directed Graphs

Trees

Real Limits

Intervals

Datasets

Undirected Paths

Real Sequences

Real Space

Real Plane

Real Line

Undirected Graphs

Sequences

Real Order

Integral Line

Real Summation

Real Optimizers

Set Numbers

Geometry

Real Numbers

Optimizers

Sigma Algebras

Direct Products

Finite Sets

Rational Numbers

Total Orders

Integer Arithmetic

Natural Addition Identity

Natural Multiplicative Identity

Partial Orders

Natural Summation

Identity Elements

Arithmetic

Subset Algebras

Equivalent Sets

Family Operations

Natural Exponents

Integer Products

Set Operations

Family Unions and Intersections

Function Inverses

Algebras

Natural Products

Integer Sums

Families

Function Composition

Function Images

Operations

Natural Sums

Integer Numbers

Function

Equivalence Relations

Natural Order

Recursion Theorem

Simple Functions

Cardinality

Relations

Peano Axioms

Subset Systems

Cartesian Products

Characteristic Functions

Uncertain Outcomes

Natural Induction

Generalized Set Dualities

Ordered Pairs

Set Powers

Partitions

Natural Numbers

Set Symmetric Differences

Set Unions and Intersections

Set Dualities

Unordered Triples

Intersection of Empty Set

Successor Sets

Set Complements

Pair Unions

Set Intersections

Pair Intersections

Set Unions

Set Differences

Empty Set

Unordered Pairs

Set Specification

Set Inclusion

Standardized Accounts

Accounts

Deductions

Quantified Statements

Logical Statements

Statements

Identities

Sets

Names

Letters

Objects