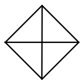


Why

Which is the optimal tree to use for tree density approximation?

Definition

We want to choose the tree whose corresponding approximator for the given density achieves minimum relative entropy with the given density among all tree density approximators. We call such a density an *optimal tree approximator* of the given density. We call a tree according to which an optimal tree approximator factors and optimal approximator tree.

Result

Prop. 1. *Let $g : \mathsf{R}^n \to [0, 1]$ be a density. A tree T on $\{1, \dots, n\}$ is an optimal approximator tree if and only if it is a maximal spanning tree of the differential mutual information graph of g.*

Proof. First, denote the optimal approximator of g for tree T by f_T^*. Recall

$$f_T^* = f_1 \prod_{i \neq 1} f_{i|\textsf{pa}\, i}$$

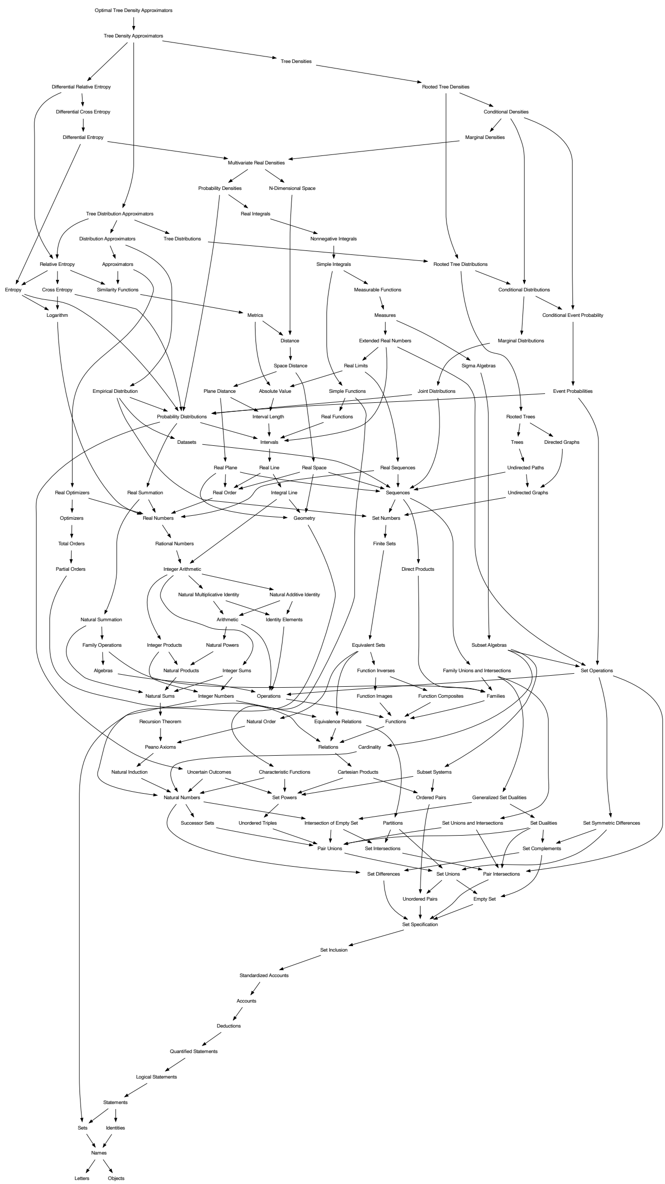
Second, recall $d(g, f) = H(g, f) - H(g)$. Since $H(g)$ does not depend on f, f is a minimizer of $d(g, f)$ if and only if it is a minimizer of $H(g, f)$.

Third, express the cross entropy of f_T^* relative to g as

$$
\begin{aligned}
H(q, p_T^*) &= h(q_1) - \sum_{j \neq i} \left(\int_{\mathbf{R}^d} g(x) \log g_{i|pai}(x_i, x_{\mathsf{pa}\, i}) dx \right) \\
&= H(q_1) - \sum_{i \neq 1} \sum_{a \in A} q(a)(\log q_{i,\mathsf{pa}\, i}(a_i, a_{\mathsf{pa}\, i}) - \log q_{\mathsf{pa}\, i}(a_{\mathsf{pa}\, i})) \\
&= H(q_1) - \sum_{i \neq 1} \sum_{a \in A} q(a)(\log q_{i,\mathsf{pa}\, i}(a_i, a_{\mathsf{pa}\, i}) - \log q_{\mathsf{pa}\, i}(a_{\mathsf{pa}\, i}) - \log q \\
&= \sum_{i=1}^{n} H(q_i) - \sum_{i \neq 1} I(q_i, q_{\mathsf{pa}\, i}) \\
&= \sum_{i=1}^{n} H(q_i) - \sum_{\{i,j\} \in T} I(q_i, q_j)
\end{aligned}
$$

where $\mathsf{pa}\, i$ denotes the parent of vertex i in T $(i = 2, \ldots, n)$. $H(g_i)$ does not depend on the choice of tree. Choosing a tree to minimize the second term in the final expression above is equivalent to choosing a maximal spanning tree from the weighted graph with differential mutual information edge weights; namely, the mutual information graph of g.

$\square$


Optimal Tree Density Approximators
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
Tree Distribution Approximators
Real Integrals
Distribution Approximators
Tree Distributions
Nonnegative Integrals
Relative Entropy
Approximators
Simple Integrals
Rooted Tree Distributions
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
Datasets
Intervals
Trees
Directed Graphs
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
Family Operations
Integer Products
Natural Powers
Equivalent Sets
Subset Algebras
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
Unordered Pairs
Empty Set
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects