



Tree-Structured Density Approximation

1 Why

2 Problem

Tree-structured density approximation is a mathematical optimization problem in which we find a tree-structured density which minimizes its differential entropy relative to a given density over the set of all densities which factor according to trees.

2.1 Notation

Let $g : \mathbf{R}^n \rightarrow \mathbf{R}$ be a density. Let d denote the differential relative entropy.

We want to find a density $f : \mathbf{R}^n \rightarrow \mathbf{R}$ and tree T on $\{1, \dots, n\}$ to

$$\begin{array}{ll} \text{minimize} & d(q, p) \\ \text{subject to} & p \text{ factors according to the tree } T \end{array}$$

3 Solution

Proposition 1. *Let $g : \mathbf{R}^n \rightarrow \mathbf{R}$ be a density. Let T be a tree on $\{1, \dots, n\}$. Let p_j be the parent of vertex j for the T rooted at vertex i , $j = 1, \dots, n$ and $j \neq i$. Then the density $f : \mathbf{R}^d \rightarrow \mathbf{R}$ defined by*

$$f = g_i \prod_{j \neq i} g_{j|p_j}$$

achieves minimum differential entropy relative to g among all densities which factor according to T .

Proposition 2. *Let $g : \mathbf{R}^n \rightarrow \mathbf{R}$ be a density. A tree T is a solution to the problem above if and only if it is a minimum spanning tree of the differential mutual information graph of g .*