

# 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  $g: \mathbb{R}^n \to \mathbb{R}$  be a density. Let d denote the differential relative entropy. We want to find a density  $f: \mathbb{R}^n \to \mathbb{R}$  and tree T on  $\{1, \ldots, n\}$  to

minimize d(q, p)subject to p factors according to the tree T.

## 3 Solution

**Proposition 1.** Let  $g : \mathbb{R}^n \to \mathbb{R}$  be a density. Let T be a tree on  $\{1, \ldots, n\}$ . Let  $p_j$  be the parent of vertex j for the T rooted

at vertex i, j = 1, ..., n and  $j \neq i$ . Then the density  $f : \mathbb{R}^d \to \mathbb{R}$  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  $g: \mathbb{R}^n \to \mathbb{R}$  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.