Why

How can we construct distributions which factor according to a directed graph?[1] Related: how can we compactly represent complex distributions over high-dimensional spaces?

Definition

Let $\bar{G} = (G, A)$ be a typed graph (see Typed Graphs) with directed and acyclic G. For source vertices i, let $g_i : A_i \to [0, 1]$ be a distribution and otherwise let $g_i : A \times A_{\mathrm{pa}_i} \to [0, 1]$ denote a function satisfying $g_i(\cdot, x)$ is a distribution for every $x \in X_{\mathrm{pa}_i}$.

We call the ordered pair $(\bar{G}, g)$ a *distribution graph* .

The *distribution* of $(\bar{G}, g)$ is the function $p : \prod_i A_i \to [0, 1]$ defined by

$$p(a) = \prod_{\mathrm{pa}_i = \varnothing} g_i(a_i) \prod_{\mathrm{pa}_i \neq \varnothing} g_i(a_i, a_{\mathrm{pa}_i}).$$

It is, of course, a distribution. And it factors according to the directed and acyclic graph G. Also, the g_i, $i = 1, \ldots, n$, are the conditionals.[2]

In other words, a distribution graph represents a probability distributions via products of smaller, "local", conditional probability distributions.

Other terminology

Other terminology includes *distribution network, conditional distribution network, conditional distribution graph, bayesian network,*[3] *bayes net, directed probabilistic graphical model, directed graphical model.*

---

[1]Future editions might flip the order of this sheet with that of directed graph distributions since, in the genetic approach, in may be more natural to think of constructing such distributions before analyzing them. This is partially motivated by the acyclic constraint here, which restricts the graphs according to which a distribution can factor.

[2]Future editions will elaborate and give a proof.

[3]Indeed, this term is near universal in certain literatures. We avoid it in these sheets as a result of the Bourbaki project's policy on naming.

Necessity of acyclicity

If G above is not taken to be acyclic, then the "distribution" of the distribution graph need not be a proper probability distribution (the condition which will fail is normalization).