WIKIPEDIA

Variational message passing

Variational message passing (**VMP**) is an approximate inference technique for continuous- or discrete-valued Bayesian networks, with conjugate-exponential parents, developed by John Winn. VMP was developed as a means of generalizing the approximate variational methods used by such techniques as Latent Dirichlet allocation and works by updating an approximate distribution at each node through messages in the node's Markov blanket.

Contents

Likelihood Lower Bound

Given some set of hidden variables H and observed variables V, the goal of approximate inference is to lower-bound the probability that a graphical model is in the configuration V. Over some probability distribution Q (to be defined later),

$$\ln P(V) = \sum_H Q(H) \ln \frac{P(H,V)}{P(H|V)} = \sum_H Q(H) \left[\ln \frac{P(H,V)}{Q(H)} - \ln \frac{P(H|V)}{Q(H)} \right].$$

So, if we define our lower bound to be

$$L(Q) = \sum_H Q(H) \ln \frac{P(H,V)}{Q(H)},$$

then the likelihood is simply this bound plus the relative entropy between P and Q. Because the relative entropy is non-negative, the function L defined above is indeed a lower bound of the log likelihood of our observation V. The distribution Q will have a simpler character than that of P because marginalizing over P is intractable for all but the simplest of graphical models. In particular, VMP uses a factorized distribution Q:

where H_i is a disjoint part of the graphical model.

Determining the Update Rule

The likelihood estimate needs to be as large as possible; because it's a lower bound, getting closer $\log P$ improves the approximation of the log likelihood. By substituting in the factorized version of Q, $L(Q)$, parameterized over the hidden nodes H_i as above, is simply the negative relative entropy between Q_j and Q_j^* plus other terms independent of Q_j if Q_j^* is defined as

$$Q_j^*(H_j) = \frac{1}{Z} e^{\mathbb{E}_{-j}\{\ln P(H,V)\}},$$

where $\mathbb{E}_{-j}\{\ln P(H,V)\}$ is the expectation over all distributions Q_i except Q_j. Thus, if we set Q_j to be Q_j^*, the bound L is maximized.

Messages in Variational Message Passing

Parents send their children the expectation of their sufficient statistic while children send their parents their natural parameter, which also requires messages to be sent from the co-parents of the node.

Relationship to Exponential Families

Because all nodes in VMP come from exponential families and all parents of nodes are conjugate to their children nodes, the expectation of the sufficient statistic can be computed from the normalization factor.

VMP Algorithm

The algorithm begins by computing the expected value of the sufficient statistics for that vector. Then, until the likelihood converges to a stable value (this is usually accomplished by setting a small threshold value and running the algorithm until it increases by less than that threshold value), do the following at each node:

1. Get all messages from parents
2. Get all messages from children (this might require the children to get messages from the co-parents)
3. Compute the expected value of the nodes sufficient statistics

Constraints

Because every child must be conjugate to its parent, this limits the types of distributions that can be used in the model. For example, the parents of a Gaussian distribution must be a Gaussian distribution (corresponding to the Mean) and a gamma distribution (corresponding to the precision, or one over σ in more common parameterizations). Discrete variables can have Dirichlet parents, and Poisson and exponential nodes must have gamma parents. However, if the data can be modeled in this manner, VMP offers a generalized framework for providing inference.

References

- Winn, J.M.; Bishop, C. (2005). "Variational Message Passing" (http://www.johnwinn.org/Publications/papers/VMP2004.pdf) (PDF). *Journal of Machine Learning Research*. **6**: 661–694.
- Beal, M.J. (2003). *Variational Algorithms for Approximate Bayesian Inference* (https://web.archive.org/web/20050428173705/http://www.cs.toronto.edu/~beal/thesis/beal03.pdf) (PDF) (PhD). Gatsby Computational Neuroscience Unit, University College London. Archived from the original (http://www.cs.toronto.edu/~beal/thesis/beal03.pdf) (PDF) on 2005-04-28. Retrieved 2007-02-15.

External links

- Infer.NET (http://research.microsoft.com/infernet): an inference framework which includes an implementation of VMP with examples.
- dimple (http://dimple.probprog.org): an open-source inference system supporting VMP.
- An older implementation (http://vibes.sourceforge.net) of VMP with usage examples.