Statistical Techniques in Robotics (16-831, F10) Lecture#7 (Tuesday September 14)

Gibbs Fields & Markov Random Fields

Lecturer: Drew Bagnell *Scribe: Chris Dellin & Bruno Hexsel* [1]

1 Gibbs Fields

Like a Bayes' Net, a Gibbs Field is a representation of a set of random variables and their relationships. An example of a Gibbs Field is given in Figure 1; edges are undirected, and connote some correlation between the connected nodes. As with a Bayes' Net, *fewer* connections means *more* structure. Gibbs Fields are powerful because they imply a way to write the joint probability of the random variables as functions over cliques in the graph.

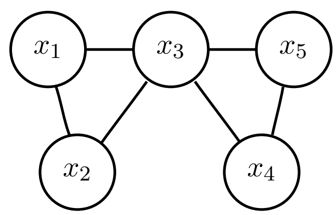


Figure 1: A Gibbs Field with nodes x_1, x_2, x_3, x_4, x_5.

1.1 Cliques and Joint Probability

A clique is any fully connected subset of the the graph (e.g. $\{x_4\}$, $\{x_1, x_2, x_3\}$, or $\{x_3, x_5\}$). We denote the set of all cliques in a graph as C, with a clique $c_i \in C$ comprising its nodes (e.g. $c_3 = \{x_3, x_5\}$). The joint probability for any set of random variables $\mathbf{x} = \{x_1, \dots x_n\}$ represented by a Gibbs Field can be written as the product of clique potentials ϕ_i:

$$P(\mathbf{x}) = \frac{1}{Z} \prod_{c_i \in C} \phi_i(c_i), \tag{1}$$

with $\phi_i(c_i)$ the ith clique potential, a function only of the values of the clique members in c_i. Each potential function ϕ_i must be positive, but unlike probability distribution functions, they need not sum to 1. A normalization constant Z is required in to create a valid probability distribution $Z = \sum_x \prod_{c \in C} \phi_i(c_i)$.

For any Gibbs Field, there is a subset $\hat{C}$ of C consisting of only *maximal cliques* which are not proper subsets of any other clique. For example, the Gibbs Field in Figure 1 has two maximal cliques: $\hat{c}_1 = \{x_0, x_1, x_2\}$ and $\hat{c}_2 = \{x_2, x_3, x_4\}$. We can write a clique potential $\hat{\phi}$ for each maximal clique that is the product of all the potentials of its sub-cliques. In this way, we can write the joint probability as only a product over these maximal clique potentials:

$$P(\mathbf{x}) = \frac{1}{Z} \prod_{c_i \in \hat{C}} \hat{\phi}_i(c_i). \tag{2}$$

[1]Some content adapted from previous scribes: Byron Boots

We usually take these potentials to be only functions over the maximal cliques, as in (2).

1.2 Clique Potentials as Energy Functions

Often, clique potentials of the form $\phi_i(c_i) = \exp(-f(c_i))$ are used, with $f_i(c_i)$ an energy function over values of c_i. The energy assigned by the function $f_i(c_i)$ is an indicator of the likelihood of the corresponding relationships within the clique, with a higher energy configuration having lower probability and vice-versa. If this is the case, (1) can be written as

$$P(\mathbf{x}) = \frac{1}{Z} \exp\left[- \sum_{c_i \in C} f_i(c_i) \right]. \tag{3}$$

For example, we can write energy functions over the cliques in the example graph from Figure 1. Let $f_1(\{x_0, x_1, x_2\}) = x_0^2 + (x_1 - 5x_2 - 3)^2$, and $f_2(\{x_2, x_3, x_4\}) = (x_2 - x_3)^2 + (x_2 + x_3 + x_4)^2$. Then the joint probability can be written as

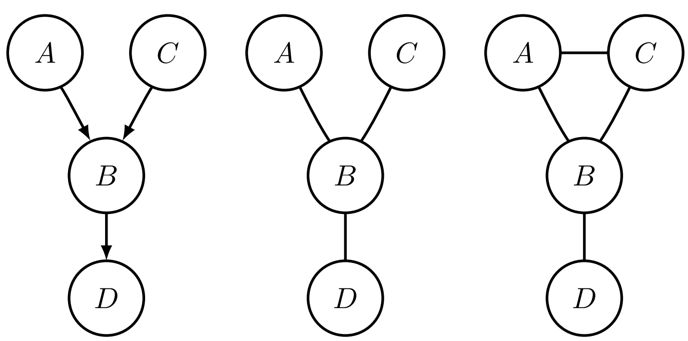
$$P(\mathbf{x}) = \frac{1}{Z} \exp\left[-(x_0^2 + (x_1 - 5x_2 - 3)^2) - ((x_2 - x_3)^2 + (x_2 + x_3 + x_4)^2) \right].$$

In this form (f_i quadratic in the xs), the Gibbs Field is known as a Gaussian Gibbs Field.

1.3 Moralizing: Converting a Bayes' Net to a Gibbs Field

Consider the Bayes' Net in Figure 2(a). Simply removing the arrows (Figure 2(b)) to create a Gibbs Field is not sufficient! In particular, in the resulting Gibbs Field, observing node B causes nodes A and C to become independent. This is the opposite of what the original Bayes' Net represented!

Instead, we need to *moralize* the graph. Whenever there are two parents that are not connected (married), we connect them. Thus, Figure 2(c) shows the correct representation of the original Bayes net. Note that during this conversion we actually lose information, namely that A and C are marginally independent.

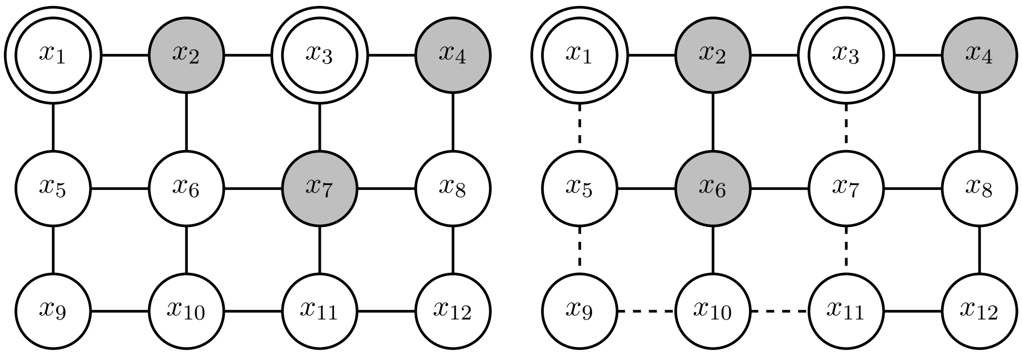


(a) Bayes' Net (b) Incorrect! (c) Correct, with moralizing

Figure 2: Moralizing process.

2 Markov Random Fields (MRFs)

A Markov Random Field (MRF) is an undirected graphical model that explicitly expresses the conditional independence relationships between nodes. Two nodes are conditionally independent if all paths between them are blocked by given nodes. See Figures 3(a) and 3(b) for examples. Note that this rule is much simpler than for Bayes' Nets. Due to the way that Markov Random Fields express the relationship between nodes, they make a lot of sense as a representation of physical space.



(a) Node x_1 is conditionally independent to node x_3 given x_2, x_4 and x_7. There is no unblocked path between x_1 and x_3.

(b) There is at least one path from x_1 to x_3 given x_2, x_4 and x_6; x_1 and x_3 are not (necessarily) conditionally independent.

Figure 3: Markov Random Field example.

2.1 Markov Random Field Constraints

The question is, given a Markov Random Field (and its associated conditional independence relationships), what is the form of the joint probability distribution? Indeed, can we even show that such a distribution exists? For example, consider the Field from Figure 3(a). We could write a series of conditional independence relationships that are asserted from the Field. For example:

$$
\begin{gathered}
x_1 \perp x_3 \mid x_2, x_5 \\
x_1 \perp x_7 \mid x_3, x_6, x_{10} \\
x_9 \perp x_{12} \mid x_2, x_7, x_{10} \\
\vdots
\end{gathered}
$$

What can we say about the form of the joint probability function $P(\mathbf{x})$ in this case?

There is a trivial example of a suitable joint probability distribution – when all the nodes are independent:

$$P(\mathbf{x}) = \prod_i P(x_i).$$

This would be a possible solution even for a fully connected Markov Random Field – the weakest possible Field, in which no conditional independencies are specified. But any given field, such as the one from Figure 3(a), has fewer edges (and therefore more structure). We should hope that there exists a stronger form of $P(\mathbf{x})$ that follows from a Markov Random Field.

The Hammersley-Clifford theorem proves that a Markov Random Field and Gibbs Field are equivalent with regard to the same graph.[2] In other words:

- Given any Markov Random Field, all joint probability distributions that satisfy the conditional independencies can be written as clique potentials over the maximal cliques of the corresponding Gibbs Field.
- Given any Gibbs Field, all of its joint probability distributions satisfy the conditional independence relationships specified by the corresponding Markov Random Field.

[2] Actually, this is true only as long as $P(\mathbf{x}) \geq 0 \; \forall \mathbf{x}$; that is, as long as all configurations of values are possible.