STAT 206A: Gibbs Measures | Fall 2006

Lecture 17

Lecture date: Oct 24 | *Scribe: Joel Mefford*

Definition 1 (δ-expander) *A graph $G = (V, E)$ is a δ-expander if,*
$\forall$ subsets $S \subseteq V, |S| \leq \frac{|V|}{2}$ we have $|\delta S| \geq \delta |S|$

Example 2 (Ising model) *Consider the Ising model,*
where $\mathbf{P}_\beta[\sigma] = \frac{1}{Z}\exp(\beta_i \sum_{i,j} \sigma_i \sigma_j)$

Claim 3 *$\forall \delta > 0, \exists \beta_c, \varepsilon > 0$ such that*
$\forall\delta$-expanders, $\forall \beta > \beta_c$ it holds that for a fraction ε of pairs of vertices (i, j) we have $\mathbf{E}_\beta[\sigma_i, \sigma_j] > \epsilon$,

The previous claim implies:

Claim 4 (non-uniqueness) *For any family P_n of Ising models on δ- expander where all vertex degrees $\leq D$,*
if $\beta > \beta_c(\delta)$,
then for P_n we have reconstruction and non-uniqueness uniqueness.

Proof: Let $G = (V, E)$ be a δ-expander graph with $n = |V|$ = number of vertices and

$$S = \frac{1}{n} \sum_{v \in V} \sigma_v$$

Clearly, $\mathbf{E}[S] = 0$.

The claim will be implied if we show that

$$\mathbf{E}[S^2] \geq \varepsilon(\beta)$$

for $\beta > \beta_c$ and $\varepsilon(\beta_c) > 0$, independent of n and β.

In order to prove the last statement it suffices to show that $\exists \xi(\beta_c) > 0$ such that for $\beta > \beta_c$ it holds that $\mathbf{P}[|S| \leq \xi] \leq 2^{-n+5}$.

Consider a σ such that $|S(\sigma)| < \xi$. Then we claim that for apropriate values of ξ and β_c it holds that.

$$\frac{\mathbf{P}_\beta(\sigma)}{\mathbf{P}_\beta(1)} < 4^{-n-5} \tag{1}$$

Since the number of possible σ is at most 2^n, (1) implies the desired result.

We now prove (1) using δ-expansion.

Let $A_\sigma = \{i : \sigma_i = 1\}$. WLOG, let $|A_\sigma| < \frac{n}{2}$.

Then, using expansion,

$$|\partial A_\sigma| \geq \delta |A_\sigma| \geq \frac{\delta(1-\xi)n}{2}$$

This implies that, for β sufficiently large,

$$\frac{\mathbf{P}_\beta(\sigma)}{\mathbf{P}_\beta(1)} \leq \exp\left(-2\beta\delta\left(\frac{(1-\xi)n}{2}\right)\right) \leq 4^{-n-5}$$

□

New topic: Tree factorization and Belief Propagation(BP)

Claim 5 (Factorization 1) *Consider k-factorized distributions over factor graphs: $G_1, \ldots, G_k$.*

Let $v_1 \in V(G_1), \ldots, v_k \in V(G_k)$.

Let P be the distribution obtained by identifying $v_1, \ldots, v_k$ into a single vertex v. Graphically, this is done by connecting the vertices in the vertex-set $v_1, \ldots, v_k$ with hyperedges having potentials ψ.

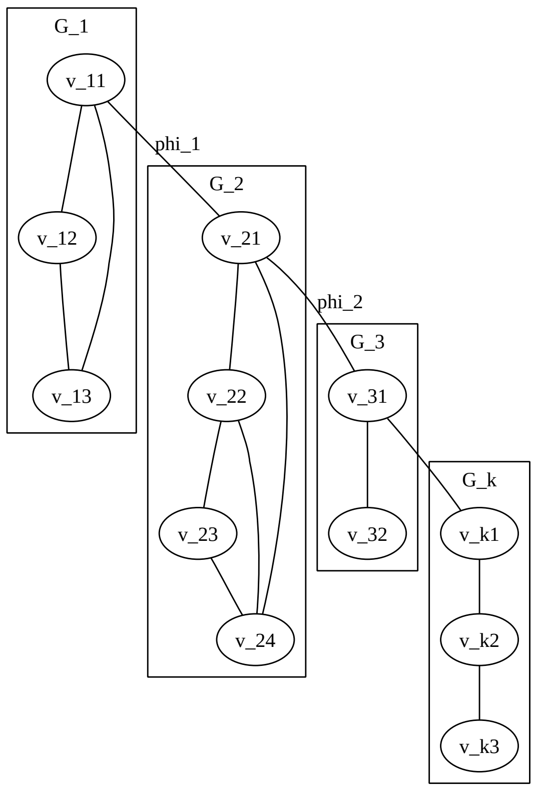

Figure 1:

The following joint distributions are then defined:

$$\begin{aligned}
P(x_1, x_2, \ldots, v) &= \frac{1}{Z} \prod_{j=1}^{k} \prod_{\sigma=1}^{M_n} \psi^j_{\delta\sigma}(x^j, x_v) \\
P_j(x^j, x_{v_j}) &= \frac{1}{Z} \prod_{\sigma=1}^{M_j} \psi_{\delta\sigma}(x^j, x_{v_j})
\end{aligned}$$

Here, Z is a normalization constant, and x^j is a vector of everything except x_{v_j}.

Then,

$$\begin{aligned}
P(x^1, x^2, \ldots, x^k, x_v) &= \frac{1}{Z} \prod_{j=1}^{k} P_j(x^j, x_v) \\
P(x_v) &= \frac{1}{k} \prod_{j=1}^{k} P_j(x_v)
\end{aligned}$$

Claim 6 (Factorization 2) *For a factorized graph $G = G_1, G_2, \ldots, G_k$ and $P_1, P_2, \ldots, P_k$ as above, now consider the vertices $v_1, v_2, \ldots, v_k$ to be connected through a factor node with potential ψ.*

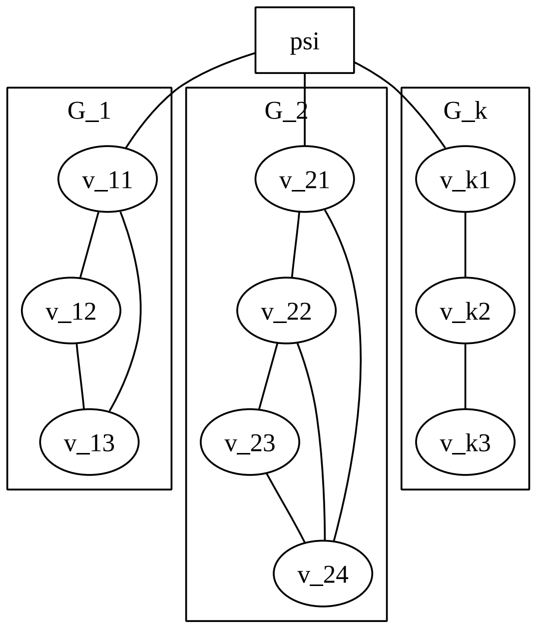

Figure 2:

Then we have the factorization,

$$\begin{aligned} P(x^1, \ldots, x^k, x_{v_1}, x_{v_2}, \ldots, x_{v_k}) &= \frac{1}{Z}\psi(x_{v_1}, \ldots, x_{v_k}) \prod_{j=1}^{k} P_j(x^j, x_{v_j}) \\ &= P(x_{v_1}, \ldots, x_{v_k}) \\ &= \frac{1}{Z}\psi(x_{v_1}, \ldots, x_{v_k}) \prod_{i=1}^{k} P_i(x_{v_i}) \end{aligned}$$

Corollary 7 *For a tree-factor graph one can compute any marginal probabilities in time* $\mathbf{O}(nA_{\max}^{\text{kmax}})$*, where,*

- *n is the size of the tree (number of nodes)*
- *kmax is the maximum degree of a factor node*
- *$A_{\max}$ is the maximum possible number of values of X*

The last result shows that we can calculate **any** marginal probabilities in linear time (in n). In fact

Claim 8 *We can even calculate* **all** *marginal probabilities in linear time.*

Proof: Given a factor node f and a variable node v.

Let $M_{f\to v}$ be the marginal of v at the graph $G_{f\to v}$.

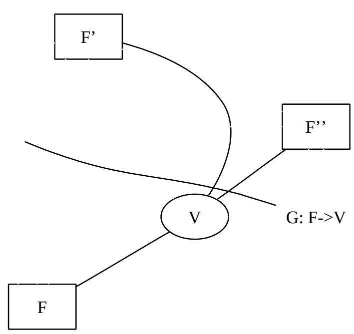

Figure 3:

Let $M_{v\to f}$ be the marginal of V at the graph $G_{V\to F}$.

Then,

$$
\begin{aligned}
M_{v\to f} &= \frac{1}{Z}\prod_{f\neq f'\sim v} M_{f'\to v}(x_v) \\
M_{f\to v} &= \frac{1}{Z}\sum_{x_{w_1},\ldots,x_{w_k}} \psi(x_{w_1},\ldots,x_{w_{k-1}},x_v)\prod_{i=1}^{k-1} M_{w_i\to f}(X_{w_i}) \\
P(X_v) &= \frac{1}{Z}\prod_{f\sim v} M_{f\to v}(X_v)
\end{aligned}
$$

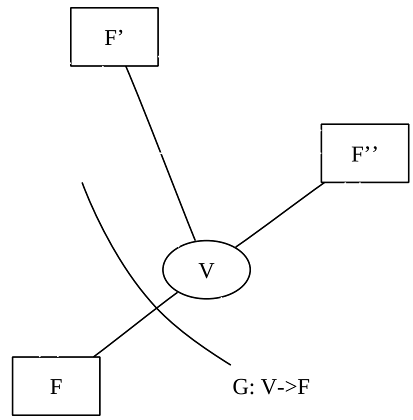

Figure 4:

,

where $f \sim v$ are the factor nodes adjacent to v, $f \neq f' \sim v$ are the factor nodes adjacent to v, excluding v', and $w_1, \ldots, w_k$ are the variable nodes adjacent to f.

If we restrict attention to the neighbors of a given node, and make use of the factorization claim above, we have,

$$\bar{m}_{f \to v}(X_{w_1}, \ldots, X_{w_k}) = \frac{1}{Z} \psi(x_{\sigma_v}, x_{w_1}, \ldots, x_{w_{k-1}}) \prod_{i=1}^{k-1} M_{w_i \to f}(X_{w_i})$$

Marginalization yields the equations above.

If we have all $M_{f \to w}$, we can get all $P(x_v)$ in linear time ($\propto n$). To determine $M_{f \to v}$, we need all values $M_{v \to f}$ leading into it. To determine $M_{v \to f}$, we need all values $M_{f \to v}$ leading into it. If the graph is a tree, we can start the computation at the leaves and work towards the root, finishing in time proportional to the number of vertices on the graph, n.

□

Definition 9 (Loopy Belief Propagation (LBP)) *LBP is a method to calculate marginals on cyclic factor graphs by expanding the graph into a truncated tree and then applying the algorithm above to the expanded graph. More next time.*