# Bethe Free Energy

accts for pairwise corr

- Exact on trees

Params $b_i(x_i) \cong P(x_i)$

$b_{ij}(x_i, x_j) \cong P(x_i, x_j)$

Abuse notation + $b = \{b_i, b_{ij}\}$ wrt $G$

Ideally, wanna search over globally consistent marg wrt $G$

**Def**

$\text{Marg}(G) \triangleq \{ b : \{b_i\}_{i \in V}, \{b_{ij}\}_{(i,j) \in E}$ s.t. $\exists P(x)$ w.

$b_i(x_i) = \sum_{x_{-i}} P(x) \quad \forall i \in V$

$b_{ij}[x_i, x_j] = \sum_{x_{-i,j}} P(x) \quad \forall (i,j) \in E$

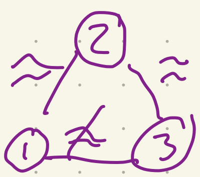
but checking if $b \in \text{Marg}$ is NP-hard ☹

Instead,

**Def**

$\text{Loc}(G) \triangleq \{ b : \sum_{x_i} b_i(x_i) = 1 \quad \forall i \in V$

$\sum_{x_j} b_{ij}(x_i, x_j) = b_i(x_i) \quad \forall (i,j) \in E \}$




Local $\not\Rightarrow$ Global

$X = \{0, 1\}$ $\quad b_i = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$ $\quad b_{12} = b_{23} \begin{bmatrix} .49 & .01 \\ .01 & .49 \end{bmatrix}$ $\quad b_{13} = \begin{bmatrix} .01 & .49 \\ .49 & .01 \end{bmatrix}$

For general G


When G is a tree, $\forall$ locally consistent b $\exists$ P(x) which is globally consistent for b

$$\tilde{P}(x) = \prod_{i \in V} b_i(x_i) \prod_{(i,j) \in E} \frac{b_{ij}(x_i, x_j)}{b_i(x_i)\, b_j(x_j)}$$

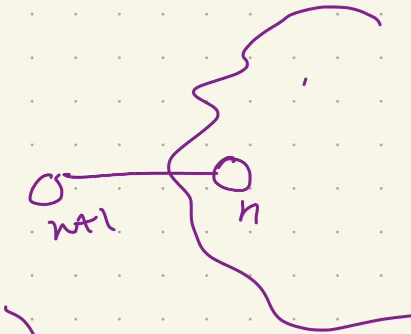
Claim: $\tilde{P}$ is consistent if G is a tree

Pf By induction on n.

$n=1$ $\quad \tilde{P}(x_1) = b_1(x_1)$ ✓

suppose true $\forall$ trees of size $\leq n$

Fix a tree of size $n+1$. It has a leaf.




$$\tilde{P}(x_1^n, x_{n+1}) = \tilde{P}(x_1^n) \cdot b_{n+1}(x_{n+1}) \cdot \frac{b_{n,n+1}(x_n, x_{n+1})}{b_n \cdot b_{n+1}}$$

$$= \tilde{P}(x_1^n) \frac{b_{n,n+1}}{b_n}$$

$$\tilde{P}(x_n, x_{n+1}) = \sum_{x \in V \setminus n, n+1} \tilde{P}(x_1^n, x_{n+1}) = \tilde{P}(x_n) \frac{b_{n,n+1}(x_n, x_{n+1})}{b_n \cdot b_{n+1}} b_{n+1}$$

$$\stackrel{IH}{=} \cancel{b_n(x_n)} \frac{b_{n,n+1}}{\cancel{b_n} \cdot \cancel{b_{n+1}}} \cdot \cancel{b_{n+1}}$$ ✓

**Def** Bethe Free Energy on a tree

$$\mathbb{F}(b) \triangleq G_{tree}(b) = -\text{Energy} + \text{Entropy}$$

$$= \sum_{(i,j)\in E} \sum_{x_i,x_j} b_{ij}(x_i,x_j) \log(f_{ij}(x_i,x_j)) - \log[b_{ij}(x_i,x_j)] +$$

$$\sum_{i\in V} (\deg(i)-1) \sum_{x_i} b_i(x_i) (\log b_i(x_i))$$

Claim: if $G$ is a tree, then

$$\sup_{b\in LOC(G)} \mathbb{F}(b) = \sup_{b\in Marg(G)} G(b) = \Phi$$

This is the Bethe variational prob

(opt on a non-tree is called Bethe appx)

## Bethe free energy & BP (not just trees!)

Claim: (1) fixed pts of BP are one-to-one w. Stationary pts of Beth.

(2) BP messages $\{m_{i\to j}(x_i)\}$ are the exponentials of lagrangians $\{\lambda^*_{ij}(x_i)\}$

Gibbs Free Energy for gen $b$ is

$$G_{Ftotal}(b) = -\underbrace{\mathbb{E}_b[-\log(-f\,total(b))]}_{\text{energy}} + \underbrace{\mathbb{E}_b[-\log b(x)]}_{\text{entropy}}$$

Evaluate on

$$b(x) = \prod_{i\in V} b_i(x_i) \prod_{ij\in E} \frac{b_{ij}(x_i,x_j)}{b_i(x_i)\, b_j(x_j)}$$

$$\text{Energy} = -\sum_{(i,j)\in E} \sum_{x_i,x_j} b_{ij}(x_i,x_j] \log f_i(x_i,x_j)$$

$$\text{Entropy} = \mathbb{E}_b\left[-\log\left[\prod b_i(x_i) \prod_{i,j\in E} \frac{b_{ij}(x_i,x_j)}{b_i(x_i)b_j(x_j)}\right]\right]$$

$$= \sum_i \sum_{x_i} -b_i \log b_i(x_i)$$

$$- \sum_{\substack{i,j\in E\\ x_i,x_j}} b_{ij} (\log b_{ij} - \log b_i - \log b_j)$$

$$\sum_{\substack{(i,j)\in E\\ x_i,x_j}} b_{ij} \log b_{ij} - \sum_{i\in V} \sum_{x_i} \deg(i)\, b_i \log b_i$$

Pf Lagrangian mults: $\lambda_i$ for $\sum_{x_i} b_i(x_i) = 1$

$\lambda_{i\to j}(x_i)$ for $\sum_{x_j} b_{ij}(x_i\, x_j) = b(x_i)$

$$Lag(b,\lambda) = \mathbb{H}_b - \lambda_i\left[\sum_{x_i} b_i(x_i) - 1\right] - \sum_{i,j}\sum_{x_i} \lambda_{i\to j}(x_i)\left[\sum_{x_j} b_{ij}(x_i\, x_j) - b_i(x_i)\right]$$

$$\nabla_{b_{ij}(x_i x_j)} Lag(b,\lambda) = -1 - \log b_{ij}(x_i,x_j) + \log f_{ij}(x_i,x_j) - \lambda_{i\to j}(x_i) - \lambda_{j\to i}(x_j)$$

$$\nabla_{b_i(x_i)} Lag(b,\lambda) = -(1-\deg(i))\log(b_i(x_i)\cdot e) - \lambda_i + \sum_{j\in\partial i}\lambda_{i\to j}(x_i)$$

setting $=0$ & solving $\Rightarrow$

$$b^*_{ij}(x_i,x_j) = f_{ij}(x_i,x_j)\exp\{-1 - \lambda_{i\to j}(x_i) - \lambda_{j\to i}(x_j)\}$$

$$b^*_i(x_i) \propto \exp\left\{-\frac{1}{\deg(i)-1}\sum_{j\in\partial i}\lambda_{i\to j}(x_i)\right\}$$

$$\sum_{x_j} b^*_{ij}(x_i,x_j) = b^*_i(x_i)$$

Change variables $m_{i\to j}(x_i) \propto e^{-\lambda_{i\to j}(x_i)}$

1. $b^*_{ij} \propto m_{ij}(x_i)\, f_{ij}(x_i\, x_j)\, m_{j\to i}(x_j)$
2. $b^*_i \propto \prod_{j\in\partial i}(m_{i\to j}(x_j))^{\deg(i)-1}$
3. $\sum_{x_i} b^*_{ij}(x_i,x_j) = b^*_i(x_i)$

To show:

$$\prod_{k \in \partial i \setminus j} \left\{ \sum_{x_k} b^*_{ik}(x_i\, x_k)\right\}^2 = \prod_{k \in \partial i \setminus j} b^*_i(x_i) = b^*_i(x_i)^{\deg(i)-1}$$

Local consistency

$\Updownarrow$ stationarity of $F_b$ ②

$\Updownarrow$ Stationarity ①

$$\prod_{k \in \partial i \setminus j} \cancel{m_{i\to k}(x_i)} \sum_{x_k} m_{k\to i}(x_k)\, f_{ik}(x_i, x_k) \quad \propto \quad \prod_{k \in \partial i} \cancel{m_{i\to k}(x_i)}$$

cancelling for $k \neq j$

$$\prod_{k \in \partial i \setminus j} \sum_{x_k} m_{k\to i}(x_k)\, f_{ki}(x_k, x_i) \quad \propto \quad m_{i\to j}(x_i)$$

The BP update!

Naive MF $\max_{b \in (\Delta_{|X|-1})^n} G(b)$ $\leq$ Bethe FE on tree $\max_{b \in Loc(T)} G(b)$ $\leq$ right on tree Gibbs FE $\max_{b(\cdot) \in \Delta_{|X|^n - 1}} G b$ $\leq$ Tree-BP Weighted $\sum_{T \in \mathcal{T}} \max_{b \in Loc(T)} G(b)$

$\approx$ Region-based FE

$$\max_{b(\cdot) \in Loc(GR)} G(b)$$

$\boxed{\text{Generalized BP}}$

$\approx$ (Not UB or LB) Beth FE on non-tree

$$\max_{b(\cdot) \in Loc(G)} G(b)$$

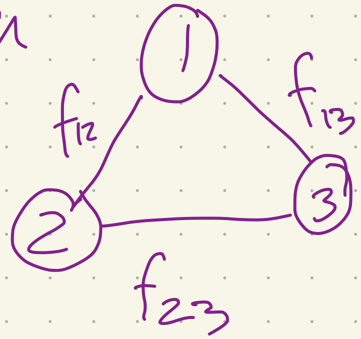
$\boxed{\text{BP}}$

# Tree weighted BP

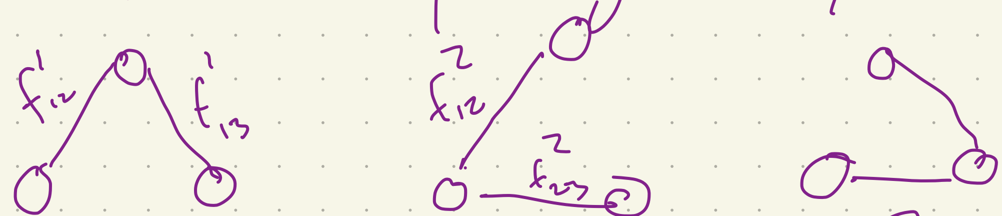
(slides have more deets)

"A new class of upper bds on the log partition fn" Wainwright, Jaakkola, Willsky

example) GM




Consider all spanning trees / wts c for trees




$T_1$ $\qquad$ $T_2$ $\qquad$ $T_3$

$c_1 = \quad c_2 = \quad = c_3 \quad = 1/3$

$$\boxed{\begin{aligned} f_{12} &= \left(f_{12}^{(1)}\right)^{c_1} \cdot \left(f_{12}^{(2)}\right)^{c_2} \\ &= \left(f_{12}^{3/2}\right)^{1/3} \left(f_{12}^{3/2}\right)^{1/3} \end{aligned}}$$

Consider all S.T.s. $T_k \in T(G)$

& wts. $c_k$ s.t.

$\sum c_k = 1$ &

$\forall ij \quad \log f_{ij}(x_i, x_j) = \sum_{k=1}^{|T(G)|} c_k \log f_{ij}^{(k)}(x_i, x_j)$

Then, energy of G Free energy decomposes:

$$E_b\left[-\sum_{ij \in E} \log f_{ij}(x_i, x_j)\right] = E_b\left[-\sum_E \sum_k c_k \log f_{ij}^k(x_i, x_j)\right]$$

$$= \sum_k c_k E_b\left[-\sum_{E_k} \log f_{ij}^k(x_i, x_j)\right]$$

energy of one tree model

Claim: $\log Z \le \sum_k c_k \log Z_k$

Pf

$$\log Z = \max_{b \in \Delta_{|\mathcal{X}|^n - 1}} \mathbb{E}_b\Big[\sum_{(i,j)\in E} \log f_{ij}(x_i, x_j)\Big] + \text{Entropy}(b)$$

$$= \max \sum_k c_k \mathbb{E}_b\Big[\sum_{(i,j)\in E_k} \log f^{k}_{ij}(x_i, x_j)\Big] + \text{Entropy}(b)$$

$$\le \sum_k c_k \max_b \Big\{ \mathbb{E}_b\Big[\sum \qquad\qquad \Big] + \text{Entropy}(b) \Big\}$$

$$= \sum c_k \max_{b \in Loc(T_k)} \Big\{ \qquad '' \qquad\qquad '' \qquad \Big\}$$

GM on a tree $T_k$!

Can be solved exactly w. BP.

Details that must be addressed (but we don't here):

- $c_k$, $f_{ij}^{(k)}$ must be chosen to min RHS
- # STs explodes