**Variational Inference**

\[ \log Z \geq \mathbb{E}_{x \sim q}[\log p(x) - \log q(x)] \]

Any choice of \( q \) gives a lower bound

**Choice of \( q \)**

- **Analytic optimization:**
  - mean field / structured mean field
- **Stochastic optimization:**
  - neural networks
  - more expressive, but requires sampling

Continuous Settings

- draw a sample
- follow gradients

Stochastic optimization: Reparameterization trick: low variance

Discrete Settings

- draw a sample
- follow gradients

No info “around” the sample

Large variance in high dimensions

**Sum Product Networks**

Avoids sampling

Expressive distribution

**Sum Node**

Selectivity / Determinism:

\[ -\mathbb{E}_{x \sim q_i}[\log g_i(x)] = - \sum_{j \in \text{children}} \alpha_{ij} \log q_j(x) + \alpha_{ij} \mathbb{E}_{x \sim q_i}[\log g_j(x)] \]

entropy of sum

sum of (constant terms + entropy)

Computes the expectation analytically:

\[ \log Z \geq \mathbb{E}_{x \sim q}[\log p(x) - \log q(x)] \]

**Product Node**

Children are independent

**Continuous Settings**

Sampling works well

Reparameterization trick: low variance

**Discrete Settings**

**Experiments**

- Tree Reweighted BP
- Loopy Belief Propagation
- Structured Mean Field
- Mean Field

Diff of Log of Partition Fn

Interaction Strength

*closest to 0 is best

**Computation cost:**

- \( O(t m) \) for discrete settings (structured mean field)
- \([Lowd 2010]\) \( O(m^2) \) for continuous settings

Expressed in log-polynomial form

**Paper / Code**