Probabilistic Circuits for Variational Inference in Discrete Graphical Models

NeurIPS 2020

Andy Shih

Stefano Ermon
Variational Inference

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

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

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

Any choice of $q$ gives a lower bound
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
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
Variational Inference

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

Choice of \( q \)

- Analytic optimization:
  - mean field
  - structured mean field

- Stochastic optimization:
  - neural networks
Variational Inference

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

Choice of $q$

Analytic optimization:
- mean field
- structured mean field

Stochastic optimization:
- neural networks

Continuous: ✔️
Discrete: ❌

[Zhang 2017]
Sampling — Continuous Settings

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

Proposal distribution \( q \)

Target distribution \( p \)

Stanford University
Sampling — Continuous Settings

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

Proposal distribution \( q \)
Target distribution \( p \)

Stanford University
Sampling — Continuous Settings

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

Proposal distribution \( q \)

Target distribution \( p \)
Sampling — Continuous Settings

Proposal distribution \( q \)

Target distribution \( p \)

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

draw a sample
Sampling — Continuous Settings

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

Proposal distribution \( q \)

Target distribution \( p \)
Sampling — Continuous Settings

Proposal distribution $q$

Target distribution $p$

$log Z \geq E_{x \sim q}[\log p(x) - \log q(x)]$

draw a sample
Sampling — Continuous Settings

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

Proposal distribution \( q \)

Target distribution \( p \)
Sampling — Discrete Settings

Each circle is a point in discrete space
Larger circle = high probability mass

Proposal distribution $q$
Target distribution $p$
Each circle is a point in discrete space
Larger circle = high probability mass

Proposal distribution $q$

Target distribution $p$
Sampling — Discrete Settings

Each circle is a point in discrete space
Larger circle = high probability mass

Proposal distribution \( q \)
Target distribution \( p \)
Sampling — Discrete Settings

Each circle is a point in discrete space
Larger circle = high probability mass

- No information “around” the samples
- Very high variance

Cannot easily optimize!

missing out on the big mode
Discrete Settings

Choice of $q$

Analytic optimization:
- mean field
- structured mean field

Stochastic optimization:
- neural networks
Discrete Settings

**Avoid sampling**

**Choice of q**

Analytic optimization:
- mean field
- structured mean field

Stochastic optimization:
- neural networks
Discrete Settings

Avoid sampling

Expressive distribution

Choice of $q$

Analytic optimization:
- mean field
- structured mean field

Stochastic optimization:
- neural networks
Discrete Settings

Avoid sampling

Expressive distribution

**Choice of \( q \)**

Analytic optimization:
- mean field
- structured mean field
- sum product networks

Stochastic optimization:
- neural networks
Sum Product Networks

Proposal distribution $q$

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

Proposal distribution $q$

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

compute analytically!

exact gradients
Sum Product Networks

Proposal distribution $q$

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

compute analytically!
Sum Product Networks

Proposal distribution $q$

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

compute analytically!
Experiments

16x16 Ising Model

*closest to 0 is best

SPN (our method)
Experiments

16x16 Ising Model

- Tree Reweighted Belief Propagation
- Loopy Belief Propagation
- SPN (our method)
- Structured Mean Field
- Mean Field

*closest to 0 is best
Summary

Discrete settings: sampling
Summary

Discrete settings: sampling

Probabilistic Circuits (e.g. Sum Product Networks)
• Expressive family of distributions!
• Can compute gradients analytically — no sampling!

Thanks!
andyshih@cs.stanford.edu