-> Presentation slot: please chuck
→ Dec 7th (Final project report-5pages) see Piazza note
see l'iazza note
-> Dec 1st HW4.
Generative models/Sampling
→ LDA, GMM Last time → ML estimates · EM algorithm
> ML estimates EM algorithm
> EM algorithm: application to GMMs
-> Variational Inference (VAEs implement approximate inference)
Indirect sampling: MCMC Markov Chain Monte Carlo (Not cover: adaptive MCMC) hoal: Score Generative model / Differsion model
(Not coxer: adaptive MCMC)
noal: Score Jenerative model / Differsion model
Distinction blw inference methods, Sampling algorithms & Generative modeling
In Jerence Via sampling: Target II () Latent voriable
Input: $\pi(z) = 9(z/x)$
Generative $P_{\theta}(x,z) = P_{\theta}(z) P_{\theta}(x/z)$ assumption likelihood
VI: argmin $D_{KL}(2 \ P_{\theta}(z x))$ $2 \in Q$ $\pi(z)$
-Mean-field assumption
→ VAE (Lecture 18)
VI : approximate Bayosias inference
Sampling: (Latest variable models/ beyond inference) Want: Samples from TT
Want. Samples from 1
Wart: Samples from TT Input: Unnormalized TT is available
Ward: Samples from TT Input: Unnormalized TT is available $TT(x) = \frac{e^{-V(x)}}{e^{-V(x)}}$
Input: Unnormalized π is available $T(x) = \frac{-V(x)}{e}$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(P_{\Theta}(z x) = \frac{P_{\Theta}(x,z)}{P_{\Theta}(x)} \right)$ $-V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(P_{\Theta}(z x) = \frac{P_{\Theta}(x,z)}{P_{\Theta}(x)} \right)$ $-V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(P_{\theta}(z x) = \frac{P_{\theta}(x,z)}{P_{\theta}(x)}\right)$ Input: $V(x)$ Score _{T} $(x) = -V(x)$ $V(x) = -V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(P_{\theta}(z x) = \frac{P_{\theta}(x,z)}{P_{\theta}(x)}\right)$ Input: $V(x)$ $Core_{T}(x) = -V(x)$ $V(x)$ $V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(f_{\theta}(z x) = \frac{P_{\theta}(x,z)}{ f_{\theta}(x) }\right)$ Input: $V(x)$ $Score_{\pi}(x) = -V(x)$ $V(x)$
Input: Unnormalized π is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(P_{\Theta}(z x) = \frac{P_{\Theta}(x,z)}{ P_{\Theta}(x) }\right)$ Input: $V(x)$ Score, $Y(x)$ Probability $Z \in \mathbb{R}^d$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(f_{\theta}(z x) = \frac{P_{\theta}(x,z)}{ f_{\theta}(x) } \right)$ Input: $V(x)$ Score_{\pi}(x) = $V(x)$ Probability Aevity $X \in \mathbb{R}^d$ $X $
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\left(f_{\theta}(z x) = \frac{P_{\theta}(x,z)}{F_{\theta}(x)}\right)$ Input: $V(x)$ $Score_{\pi}(z) = -V(x)$ $V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\left(\int_{\mathcal{B}} (z x) = \frac{P_{\mathcal{B}}(x,z)}{P_{\mathcal{B}}(x)}\right)$ Input: $V(x)$ $Score_{T}(x) = -V(x)$ $V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\left(f_{\theta}(z x) = \frac{P_{\theta}(x,z)}{F_{\theta}(x)}\right)$ Input: $V(x)$ $Score_{\pi}(z) = -V(x)$ $V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\begin{pmatrix} f_{\theta}(z x) = P_{\theta}(x,z) \\ \hline{P_{\theta}(x)} \end{pmatrix}$ Input: $V(x)$ Score, $V(x)$ Probability $V(x)$ Probability $V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\begin{pmatrix} f_{\theta}(z x) = P_{\theta}(x,z) \\ \hline{P_{\theta}(x)} \end{pmatrix}$ Input: $V(x)$ Score, $V(x)$ Probability $V(x)$ Probability $V(x)$
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\left(\int_{0}^{\infty} (z x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) } \right)$ Input: $V(x)$ Score $T(x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Probability $T(x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Probability $T(x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Probability $T(x,z) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ $T(x,z) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Transport - based sampling adjointhms: Triangular transport Optimal transport Normalizing flows Particle-based sampling Optimal transport Largevin dynamics & variants
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\begin{pmatrix} f_{\theta}(z x) = P_{\theta}(x,z) \\ \hline{P_{\theta}(x)} \end{pmatrix}$ Input: $V(x)$ Score, $V(x)$ Probability $V(x)$ Probability $V(x)$
Input: Unnormalized π is available $\pi(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(f_{\theta}(z x) = \frac{P_{\theta}(x,z)}{ f_{\theta}(x) }\right)$ Input: $V(x)$ $Score_{\pi}(x) = -V(x)$ $Probability \qquad x \in \mathbb{R}^d$ $Probability \qquad x \in \mathbb{R}^d$ $V: \mathcal{X} \to \mathbb{R}^d$ $V: \mathcal{X} \to \mathbb{R}^d$ $V: \mathcal{X} \to \mathbb{R}^d$ Transport - based sampling adjointhms: • Triangular transport • Optimal transport • Normalizing flows • Particle-based sampling • Optimal transport • Largevin dynamics & variants • MCMC variants • MCMC variants
Input: Unnormalized T is available $T(x) = \frac{e^{-V(x)}}{Z}$ Normalization constate $\left(\int_{0}^{\infty} (z x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) } \right)$ Input: $V(x)$ Score $T(x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Probability $T(x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Probability $T(x) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Probability $T(x,z) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ $T(x,z) = \int_{0}^{\infty} (x,z) \frac{z}{ f_{0}(x) }$ Transport - based sampling adjointhms: Triangular transport Optimal transport Normalizing flows Particle-based sampling Optimal transport Largevin dynamics & variants
Input: Unnormalized π is available $\pi(x) = \frac{e^{-V(x)}}{Z}$ Normalization constant $\left(f_{0}(z x) = \frac{P_{0}(x,z)}{ F_{0}(x) }\right)$ Input: $V(x)$ $S_{0}(x) = \frac{P_{0}(x,z)}{ F_{0}(x) }$ Input: $V(x)$ $S_{0}(x) = \frac{P_{0}(x,z)}{P_{0}(x)}$ Probability $S_{0}(x) = \frac{P_{0}(x,z)}{P_{0}(x)}$ Probability $S_{0}(x) = \frac{P_{0}(x,z)}{P_{0}(x)}$ $V: \mathcal{X} \Rightarrow R $ $V: \mathcal{X} \Rightarrow R $ I variant Transport - based sampling adjoints. Total palar trasport Optimal transport Normalizing flaus Particle-based sampling Optimal transport Largevin dynamics & Variants MCMC variants Nemerative model Want: pamples from π
Input: Unnormalized π is available $\pi(x) = \frac{e^{-V(x)}}{Z}$ Normalization constat $\left(f_{\theta}(z x) = \frac{P_{\theta}(x,z)}{ f_{\theta}(x) }\right)$ Input: $V(x)$ $Score_{\pi}(x) = -V(x)$ $Probability \qquad x \in \mathbb{R}^d$ $Probability \qquad x \in \mathbb{R}^d$ $V: \mathcal{X} \to \mathbb{R}^d$ $V: \mathcal{X} \to \mathbb{R}^d$ $V: \mathcal{X} \to \mathbb{R}^d$ Transport - based sampling adjointhms: • Triangular transport • Optimal transport • Normalizing flows • Particle-based sampling • Optimal transport • Largevin dynamics & variants • MCMC variants • MCMC variants

Expectation maximization convergence

Iterative algorithm:

 $2_{t+l,i}^{(z)} = P_{\theta_t}(z/x_i)$ E-step:

M-step:

 $\theta_{t+1} = \underset{0}{\operatorname{argmax}} \underbrace{\sum_{i=1}^{m} ELBO(2_{t+j})}_{\theta_{i}}$ ELBO(2, 0, x) = - E log2 + E logPo(x, 2)

EM algorithm increases the likelihood at enry step

 $\log \rho_{\ell(x)} = l(\theta_{t+1}, x) > l(\theta_{t}, x)$

 $\cdot l(\theta_{t+1,2}) = ELBO(\theta_{t+2},\theta_{t+1},z)$ > ELBO(2+1, O+1, 2)

(M-step) > ELBO(2+1, Ot, 2) $= l(\theta_t, x)$

Po (·/×) - argmax ELBO(9, 0, x) (E-step) Granssian Mixture model

Generative assumption / Probabitions
model

 $\Theta = \left(\begin{array}{c} \mu_1, \dots, \mu_k, \frac{\mathcal{Z}}{\sigma_1}, \dots, \frac{\mathcal{Z}}{\kappa}, \frac{\mathcal{T}}{\kappa} \\ \mathbb{R}^d & \mathbb{R}^{d \times d} & \dots & \mathbb{T}_k \end{array} \right)$

 $P_{\theta}(X/Z=j) = \frac{-(x-\mu_{j})^{T_{0}}}{(2\pi)^{\frac{4}{2}}} = \frac{1}{|Z|^{\frac{1}{2}}}$

2: Latent variables

Z: clustering assignment

 $P_{Q}(X, Z_{j}) = \sum_{j=1}^{k} T_{j} e^{-(x-\mu_{j})} \Sigma_{j}^{-1}$ 0 observed Latent variable 0 parameters $X \in \mathbb{R}^{d}$ $Z \in [k]$ k : no of components

arg max ELBO (
$$2_{tH}$$
, θ)

 θ
 $(\pi_{1},...,\pi_{k}, \mu_{1},...,\mu_{k})$
 $(-\frac{E}{2}\log 2 + \frac{E}{2}\log \beta_{0}(x_{2}^{2}))$
 $ELBO(2_{tH}, \theta) = \sum_{i=1}^{m} \sum_{j=1}^{k} 2_{tH}(ij|x_{i})\log \beta_{0}(x_{2}, x_{2}^{2})$
 $\log \pi_{i} - \frac{1}{2}(x_{i} - \mu_{i}) \sum_{j=1}^{m} (x_{i} - \mu_{i})$
 $(\sum_{j=1}^{k} \pi_{j} = 1)$
 $ELBO(2_{tH}, 2\pi_{j}, \mu_{i}3_{j=1}^{k})$
 $+\lambda (\sum_{j=1}^{k} \pi_{j} - 1)$

$$\begin{aligned}
& \mathcal{L}_{t+1} \left(\frac{z}{x} \right) = \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \cdot \mathcal{L}_{t+1} \left(\frac{z}{x} \right) \\
&= \mathcal{L}_{t+1} \left(\frac{z$$

ML estimation voing EM

$$\frac{P_{\theta_{t}}(z=J) \cdot P_{t}(x=Z-J)}{P_{t}(x-P_{t})} = \frac{-1(x-P_{t})}{Z} \frac{-1(x-P_{t})}{(2\pi)^{d/2}} = \frac{-1(x-P_{t})}{(2\pi)^{d/2}} = \frac{-1(x-P_{t})}{(2\pi)^{d/2}}$$

$$H_{j} = H_{j,t} \quad T_{j} = T_{j,t}$$

$$\Rightarrow M \quad step$$

$$arg ma \times ELBO(2_{tH}, \theta)$$

$$\theta$$

$$(- F_{j} \log 2 + E \log$$

ELBO(
$$Q_{t+1}$$
, Θ) = $\sum_{k=1}^{\infty} \sum_{k=1}^{\infty} Q_{t+1}^{k} (i|x_{i}) \log P_{\theta}(x_{i})$

$$|\log \pi_{i} - \frac{1}{2}(x_{i} - \mu_{i})| \leq \frac{1}{2} |\log \pi_{i}(x_{i})|$$

$$|\log \pi_{i} - \frac{1}{2}(x_{i} - \mu_{i})| \leq \frac{1}{2} (x_{i} - \mu_{i})|$$

$$\frac{2}{2t+1}$$

$$\frac{$$

Space-shuttle O-ring model
Dalal et al 1989, A. Grandy lecture
notes

-) Observations of failure from tests at various temperatures

at various temperatures $P_{\theta}(x/z=1) = \frac{e^{\alpha} e^{x\beta}}{2\pi B}$

 $1 + e^{\alpha} e^{\alpha \beta}$ x : temperature $\theta : (\alpha, \beta)$

Poior on θ $P(\alpha, \beta) = \frac{1}{b} e^{\alpha} e^{-\frac{c^{\alpha}}{b}}$ Uniform on β .

 $P_{\theta}(z|x) = \frac{P_{\theta}(z,x)}{P_{\theta}(x)}$ $P_{\theta}(z) = \int P_{\theta}(x|z) P_{\theta}(z) dz$

$$\frac{\sqrt{I}}{}$$

$$\frac{1}{2(\mu,\pi)} = \frac{k}{\pi} \mathcal{N}(\mu_{k}; m_{k}, s_{k}^{2}) \frac{k}{\pi} \mathcal{T}_{j}$$

ELBO
$$(2, \theta, x)$$

$$= \sum_{j=1}^{k} \mathbb{E} \log \mathcal{N}(\mu_{k}; m_{k}, S_{k}^{2})$$

Markov Chain Monte Carlo: Construct a Markov Chain whose invariant) stationary measure is the toyet TT. $\rightarrow \pi = e^{-V}/Z$ Have: $e^{-V(x)}$ -> Markov Chain: X1,..., Xn ... $P(X_{t+1} | X_t, ..., X_l) = P(X_{t+1} | X_t)$ $\rightarrow k(x, A) = P(X_{t+1} \in A | X_t = x)$ -> (Xn) is Harris recurrent if I a measure IT s.t. & A with $T(A) > 0 = X_n$ visits A infinitely often $\pi: \mathbb{R}^d \to \mathbb{R}^+ \quad \pi(A) = \mathbb{P}(X \in A)$ $\pi(A) = \int_{Y} K(x, A) \pi(dx)$ $\underbrace{\sum_{X_t} \pi(X_t) \cdot P(X_{t+1} = \mathcal{Y}|X_t)}_{X_t} = \pi(\mathcal{Y})$ Ergodie theorem: f ∈ L¹(T) then $\frac{1}{n} \sum_{i=1}^{n} f(X_i) \rightarrow \int f(x) \pi(dx)$ $= \underbrace{F} f(X)$ $X \sim \pi$. Let $g \in L^{1}(\pi)$ with π $Vor(g(X)) < \infty$ $\hat{\mathcal{J}}_{N} = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{S}(X_{n})$ $\sigma^2 := \lim_{N \to \infty} Var \left(\sqrt{N} \hat{g}_N \right)$ $N \to \infty$ $= Vax \left(g(X)\right) + 2 \leq Cov(g(X_1), n=1)$ $g(X_{1+N})$ $\left(\widehat{g}_{N}^{2} - \underset{X \sim \Pi}{E} g(X)\right) \sqrt{N} \xrightarrow{d} \mathcal{N}(0, \sigma^{2})$ Convergence to stationary distribution Harris recurrent, aperiodic Markov chain with invariant measure T,

lim || SK (20,0) m(dx) - TH = 0

11700 || HI- M2 || N = Sup | M(A) - M2 (A) |

A