

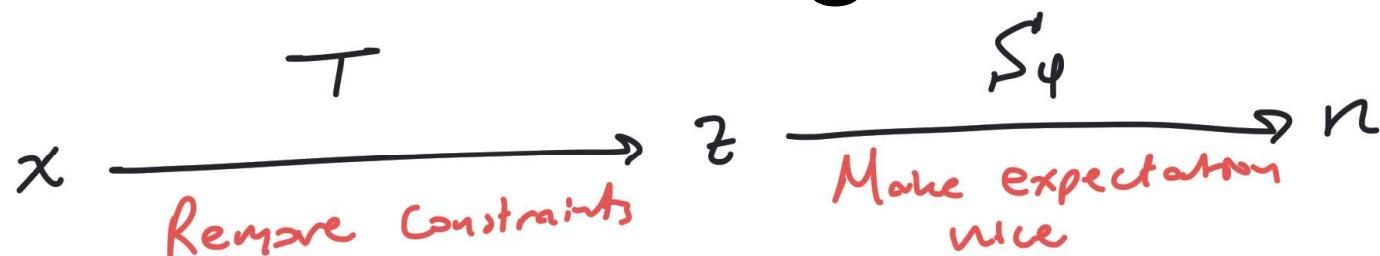
Lecture 28:

Variational Inference

Professor Ilias Bilonis

**Solving the optimization problem
with stochastic gradient descent**

Solving the optimization problem with stochastic gradient descent



$$ELBO(\varphi) = H[q(z; \varphi)] + \mathbb{E}_{n \sim N(0, I)} \left[\ln p(T^{-1}(S_\varphi^{-1}(n))) + \ln |\det J_{T^{-1}(S_\varphi^{-1}(n))}| \right]$$

- φ_0 initialize

- For steps $t=1, 2, \dots$
 - sample M n 's from $N(0, I)$, n_1, n_2, \dots, n_M

$$\varphi_{t+1} = \varphi_t + \alpha_t \cdot \frac{1}{M} \sum_{m=1}^M \nabla_{\varphi} (\dots) \big|_{n=n_m}$$

learning rate

$$\sum_{t=1}^{\infty} \alpha_t = +\infty, \quad \sum_{t=1}^{\infty} \alpha_t^2 < \infty$$