TTIC 31230 Fundamentals of Deep Learning

Problems for Rate Distortion Autoencoders.

Problem 1 The mutual information between two random variables x and y is defined by

$$I(x,y) = E_{x,y}\ \ln \frac{p(x,y)}{p(x)p(y)} = KL(p(x,y), p(x)p(y))$$

Mutual information has an interpretation as a channel capacity.

(a) Suppose that we draw a random bit $y \in \{0,1\}$ with $P(0) = P(1) = 1/2$ and send it across a noisy channel to a receiver who gets $y' = y \oplus \epsilon$ where ϵ is an independent "noise variable" with $\epsilon \in \{0,1\}$, where $\oplus$ is exclusive or (y gets flipped when $\epsilon = 1$), and where the "noise" ϵ has a probability P of being 1.

(a) Solve for the channel capacity $I(y, y')$ as a function of P in units of bits. When measured in bits, this channel capacity has units of bits received per message sent.

(b) Explain why your answer to part (a) makes sense in terms of what the receiver knows for $P = 1/2$ and when $P = 1$.

Problem 2. Consider a rate-distortion autoencoder.

$$\Phi^* = \operatorname*{argmin}_{\Phi}\ I_\Phi(y,z) + \lambda E_{y\sim \mathrm{Pop},\ z\sim p_\Phi(z|y)}\ \mathrm{Dist}(y, y_\Phi(z)).$$

Here $I_\Phi(y,z)$ is defined by the distribution where we draw y from Pop and z from $P_\Phi(z|y)$. The distribution $p_\Phi(z|y)$ is typically defined by $z = z_\Phi(y) + \epsilon$ for some form of random noise ϵ.

(a) Starting from the definition of $I_\Phi(y,z)$ given in problem 1, show

$$I_\Phi(y,z) = E_{y\sim \mathrm{Pop}} KL(p_\Phi(z|y), p_\Phi(z))$$

where $p_\Phi(z) = \sum_y \mathrm{Pop}(y) P_\Phi(z|y)$.

(b) Show the variational equation

$$I(y,z) = \inf_q\ E_{y\sim \mathrm{Pop}} KL(p_\Phi(z|y), q(z)).$$

Hint: It suffices to show

$$I(y,z) \leq\ E_{y\sim \mathrm{Pop}} KL(p_\Phi(z|y), q(z))$$

and that there exists a q achieving equality.

Problem 3. Consider a rate-distortion autoencoder

$$\Phi^*, \Psi^* = \operatorname*{argmin}_{\Phi,\Psi}\ E_{y\sim \mathrm{Pop}} KL(p_\Phi(z|y), p_\Psi(z)) + \lambda E_{y\sim \mathrm{Pop},\ z\sim p(z|y)}\ \mathrm{Dist}(y, y_\Phi(z)).$$

Define $p_\Phi(z|y)$ by $z = z_\Phi(y) + \epsilon$ with $z_\Phi[y] \in \mathbb{R}^d$ and ϵ drawn uniformly from $[0,1]^d$. In other words, we add noise drawn uniformly from $[0,1]$ to each component of $z_\Phi(y)$.

Define $p_\Psi(z)$ to be log-uniform in each dimension. More specifically $p_\Psi(z)$ is defined by drawing $s[i]$ uniformly from the interval $[1, s_{\max}]$ and then setting $z[i] = e^s$ so that $\ln z[i]$ is uniformly distributed over the interval $[0, s_{\max}]$. This gives

$$\begin{aligned} dz &= e^s ds &= z ds \\ dp &= \frac{1}{s_{\max}}\, ds \\ p_\Psi(z[i]) &= \frac{dp}{dz} &= \frac{1}{s_{\max} z[i]} \end{aligned}$$

Assume That we have that $z_\Phi(y) \in [0, e^{s_{\max}-1}]^d$ so that with probability 1 over the draw of ϵ $P_\Psi(z_\Phi(y)+\epsilon) > 0$.

(a) For $z \in [z_\Phi(y), z_\Phi(y)+1]$ what is $p_\Phi(z|y)$?

(b) Solve for $KL(p_\Phi(z|y), p_\Psi(z))$ in terms of $z_\Phi(y)$ under the above specifications.

(b) Explain how these specifications model rounding down each number in $z_\Phi(y)$ to the nearest integer.