IT for Statistics and Learning
2023

Assignment 3
Assigned: Thursday, Nov 23, 2023
Due: Friday, Dec 1, 2023 M. Skoglund

Problem 3.1: Prove that

$$I(X;Z|Y) = \min_{Q_{XYZ}} D(P_{XYZ} \| Q_{XYZ})$$

where the min is over all Q_{XYZ} s.t. $X \to Y \to Z$ and where $I(X;Z|Y)$ is computed based on P_{XYZ}

Problem 3.2: Let $I(P_X, P_{Y|X})$ denote $I(X;Y)$ when $(X,Y) \sim P_{Y|X} \times P_X$. Assume there is a P_X^* such that $\sup_{P_X} I(P_X, P_{Y|X}) = I(P_X^*, P_{Y|X})$ and let $P_Y^* = P_{Y|X} \circ P_X^*$. Then, prove that for all P_X and Q_Y we have

$$D(P_{Y|X} \| P_Y^* | P_X) \leq D(P_{Y|X} \| P_Y^* | P_X^*) \leq D(P_{Y|X} \| Q_Y | P_X^*)$$

Problem 3.3: Use the result in Problem 3.2 to prove that

$$\max_{P_X} I(X;Y) = \max_{P_X} \min_{Q_Y} D(P_{Y|X} \| Q_Y | P_X) = \min_{Q_Y} \max_{P_X} D(P_{Y|X} \| Q_Y | P_X)$$

Problem 3.4: Assume $\mathcal{X} = \{X_i\}_{i=1}^n$ are i.i.d. Gaussian with $E[X_i] = m$ and $E[(X_i - m)^2] = \sigma^2$, where both m and σ are fixed and deterministic. Show that

$$T(\mathcal{X}) = \frac{1}{n} \sum_{i=1}^{n} X_i$$

is a sufficient statistic of $\mathcal{X}$ for m.

Problem 3.5: Consider problem 3.4, but assume now that $X_i = M + Z_i$ where M is Gaussian with $E[M] = 0$ and $E[M^2] = 1$, and $\{Z_i\}$ are zero-mean i.i.d. Gaussian independent of M and with variance σ^2. Prove that

$$T(\mathcal{X}) = \frac{1}{n} \sum_{i=1}^{n} X_i$$

is a sufficient statistic of $\mathcal{X}$ for M by demonstrating that $I(M;\mathcal{X}) = I(M;T)$.

Problem 3.6: Assume that X and Y are jointly Gaussian with $E[X] = E[Y] = 0$, $E[X^2] = E[Y^2] = 1$ and $E[XY] = \rho$. Consider the information bottleneck problem

$$\inf_{P_{T|X}} \left(I(X;T) - \lambda I(Y;T) \right)$$

for $Y \to X \to T$. Characterize analytically the optimal kernel $P_{T|X}^*$ for this problem, as well as the corresponding points $(I(X;T), I(Y;T))$ on the information curve. You can use without proof the fact that the optimal kernel results in (Y, X, T) being jointly Gaussian.