IERG 6154: Network Information Theory
Homework 2

Due: Jan 24, 2019

1. Show that $\log(2^x + 2^y)$ is convex in (x, y). This is used to prove conditional EPI.
2. Let X and Z be independent, zero-mean, continuous random variables with $\mathrm{E}(X^2) = P, \mathrm{E}(Z^2) = Q$.
   (a) Let $E(X|X+Z)$ be the conditional expectation of X given $X+Z$. (If you are not familiar with conditional expectation, please take $E(X|X+Z)$ to be a random variable $\hat{Y}$ that is the function of $X+Z$ that minimizes $E\left((X-\hat{Y})^2\right)$ among all such functions. Show that
   $$\mathrm{E}\left((X - E(X|X+Z))^2\right) \leq \frac{PQ}{P+Q}.$$
   (Hint: Use the best linear estimator of X given $X+Z$ and upper bound the variance using this estimator. In other words let $Y_1 = a(X+Z)$ and optimize over the choice of a)
   (b) Show that $h(X|X+Z) \leq \frac{1}{2}\log\left(2\pi e \frac{PQ}{P+Q}\right)$. Further show that equality holds *iff* X and Z are Gaussians. (For this part, EPI may be useful).
   (c) Let $\tilde{X}, \tilde{Z} \sim \mathrm{N}(0, P), \mathrm{N}(0, Q)$ repectively and assume that they are independent of each other and of X, Z. Show that $I(\tilde{X}; \tilde{X} + Z) \geq I(\tilde{X}; \tilde{X} + \tilde{Z})$ (Hint: Use EPI)
3. . Show that the following two inequalities are equivalent:
   - Let X and Y be independent and $\lambda \in [0, 1]$. then
   $$J(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \leq \lambda J(X) + (1-\lambda)J(Y). \tag{1}$$
   - Let X and Y be independent, then
   $$\frac{1}{J(X+Y)} \geq \frac{1}{J(X)} + \frac{1}{J(Y)}. \qquad \text{(Stam's inequality)} \tag{2}$$
4. **Costa's concavity of entropy power**: Let X be an arbitrary continuous random variable with density $g(x)$ such that $|\log g(x)| \leq C(1 + x^2)$. Let Z be an independent Gaussian with zero mean and unit variance. Let the bivariate function $f(x, t)$ be the pdf of the random variable $X + \sqrt{t}Z$ and h be the differential entropy of f. (It is easy to show that for every

positive t, the density function $f(x,t)$ and all of its derivatives viewed as a function of x decays to 0 as $|x| \to \infty$; however, for this homework, you may assume this is so without needing to establish it).

Subscripts denote partial derivatives ($\frac{d^2}{dx^2} f = f_{xx}$). We use natural logarithms and the bounds for the Integrals $\pm\infty$.

(a) Show that f satisfies the partial differential (heat) equation:

$$f_t = \frac{1}{2} f_{xx}$$

(b) Show that

$$h_t = \frac{1}{2} \int f \left(\frac{f_x}{f}\right)^2 dx \tag{3}$$

Hint: $\int_{-\infty}^{\infty} UV' dx = -\int_{-\infty}^{\infty} U'V dx$ when $UV \to 0$ as $|x| \to \infty$.

(c) Show that (Hint: use the previous hint)

$$\int f \left(\frac{f_x}{f}\right)^2 \frac{f_{xx}}{f} dx = \frac{2}{3} \int f \left(\frac{f_x}{f}\right)^4 dx \tag{4}$$

(d) Show that

$$h_{tt} = -\frac{1}{2} \int f \left(\frac{f_{xx}}{f} - \left(\frac{f_x}{f}\right)^2\right)^2 dx.$$

Thus $h(X + \sqrt{t}Z)$ is concave in t. However, something stronger is true as we shall see below.

(e) Show that $e^{2h(X+\sqrt{t}Z)}$ is concave in t. (Hint: Use $\int f(\frac{f_{xx}}{f} - \frac{f_x^2}{f^2} + \beta)^2 dx \geq 0$ and choose an appropriate β to infer the result.) This result was originally established by Costa and the proof outlined in this homework is due to C. Villani.

5. (Rioul's proof of Entropy Power Inequality) One ingredient that one needs to be aware of while doing this work is the following map, called the Knothe map (the construction was mentioned in class).

We consider a generic random variable X with a differentiable density function, $f(x)$ satisfying

$$\int_{\mathbb{R}^n} f(x) \log(1 + f(x)) dx < \infty.$$

Then one can construct a map, called the Knothe map, $\Phi : \mathbb{R}^n \mapsto \mathbb{R}^n$ such that the distribution of $\Phi(Z)$ is same as the distribution of X and further the Jacobian matrix is upper triangular with positive diagonal entries. In the above, Z is the standard Gaussian, $\mathcal{N}(0, I)$.

If $g(x)$ is the Gaussian density, then we have the identity that

$$g(z) = f(\Phi(z)) \det(J_\Phi(z))$$

where $\det(J_\Phi(z))$ is the determinant of the Jacobian of the transformation.

(a) Show that $h(X) = h(Z) + \mathrm{E}(\log(\det(J_\Phi(Z))))$

(b) Let X and Y be independent random variables (with the properties on densities as stated above) such that Φ and Ψ denote their Knothe maps, respectively. Let Z, W be two i.i.d. standard Gaussians.

Show that (using Lieb's equivalent form of EPI), EPI is equivalent to showing

$$h\left(\sqrt{\lambda}\Phi(Z) + \sqrt{1-\lambda}\Psi(W)\right) - h(Z) \geq \lambda \mathrm{E}(\log(\det(J_\Phi(Z)))) + (1-\lambda)\mathrm{E}(\log(\det(J_\Psi(W)))).$$

(c) Define $\hat{Z} = \sqrt{\lambda}Z + \sqrt{1-\lambda}W$ and $\hat{W} = -\sqrt{1-\lambda}Z + \sqrt{\lambda}W$. Then justify the steps marked $(a), (b), (c), (d)$ below

$$\begin{aligned}
&h\left(\sqrt{\lambda}\Phi(Z) + \sqrt{1-\lambda}\Psi(W)\right) \\
&\quad = h\left(\sqrt{\lambda}\Phi(\sqrt{\lambda}\hat{Z} - \sqrt{1-\lambda}\hat{W}) + \sqrt{1-\lambda}\Psi(\sqrt{1-\lambda}\hat{Z} + \sqrt{\lambda}\hat{W})\right) \\
&\quad \stackrel{(a)}{\geq} h\left(\sqrt{\lambda}\Phi(\sqrt{\lambda}\hat{Z} - \sqrt{1-\lambda}\hat{W}) + \sqrt{1-\lambda}\Psi(\sqrt{1-\lambda}\hat{Z} + \sqrt{\lambda}\hat{W}) | \hat{W}\right) \\
&\quad \stackrel{(b)}{=} h(Z) + \mathrm{E}\left(\mathrm{E}\left(\log\left(\lambda \det(J_\Phi(\sqrt{\lambda}\hat{Z} - \sqrt{1-\lambda}\hat{W})) + (1-\lambda)\det(J_\Psi(\sqrt{1-\lambda}\hat{Z} + \sqrt{\lambda}\hat{W}))\right) \Big|_{\hat{W}}\right)\right) \\
&\quad \stackrel{(c)}{=} h(Z) + \mathrm{E}\left(\log\left(\lambda \det(J_\Phi(\sqrt{\lambda}\hat{Z} - \sqrt{1-\lambda}\hat{W})) + (1-\lambda)\det(J_\Psi(\sqrt{1-\lambda}\hat{Z} + \sqrt{\lambda}\hat{W}))\right)\right) \\
&\quad \stackrel{(d)}{\geq} h(Z) + \lambda \mathrm{E}\left(\log\left(\lambda \det(J_\Phi(\sqrt{\lambda}\hat{Z} - \sqrt{1-\lambda}\hat{W}))\right)\right) + (1-\lambda)\mathrm{E}\left(\log\left(\det(J_\Psi(\sqrt{1-\lambda}\hat{Z} + \sqrt{\lambda}\hat{W}))\right)\right) \\
&\quad = h(Z) + \lambda \mathrm{E}\left(\log\left(\lambda \det(J_\Phi(Z))\right)\right) + (1-\lambda)\mathrm{E}\left(\log\left(\det(J_\Psi(W))\right)\right).
\end{aligned}$$