EE/Ma 126b Information Theory - Homework 4.1

Ling Li, ling@cs.caltech.edu

February 13, 2001

Here are 3 versions of the solution to homework 4.1. The first one is what I handed in. I made a conceptual error there since after deciding the boundary be $\bar{x}$, the two points $\bar{x} \pm \delta$ may no longer be optimal. Malik pointed out my error and gave me a hint. That resulted in the second version. However, I still can not work it out — I was stuck when I try to get the derivative to be 0, or something similar. Again Malik helped me. The third version is almost what he said.

1. *One bit quantization of a single Gaussian random variable.* Let the two reproduction points for 1 bit quantization of $X \sim \mathcal{N}(0, \sigma^2)$ are $\bar{x} - \delta$ and $\bar{x} + \delta$, and $\delta \geq 0$. To minimize the distortion, $\bar{x} - \delta$ is used to represent $X \leq \bar{x}$ and $\bar{x} + \delta$ is used to represent $X > \bar{x}$, since squared error measure is used. Thus the distortion is

$$\begin{aligned}
D(\bar{x}, \delta) &= \int_{-\infty}^{\bar{x}} f(x)(x - \bar{x} + \delta)^2 dx + \int_{\bar{x}}^{\infty} f(x)(x - \bar{x} - \delta)^2 dx \\
&= \sigma^2 + \bar{x}^2 + \delta^2 + 2\delta \left(\bar{x} + 2 \int_{-\infty}^{\bar{x}} f(x) x dx - 2\bar{x} \int_{-\infty}^{\bar{x}} f(x) dx \right) \\
&= \sigma^2 + \bar{x}^2 + \delta^2 + 2\delta \left(\bar{x} + 2 \int_{-\infty}^{0} x f(x + \bar{x}) dx \right) \qquad (1) \\
&= \sigma^2 + \left(\delta + \bar{x} + 2 \int_{-\infty}^{0} x f(x + \bar{x}) dx \right)^2 + \bar{x}^2 - \left(\bar{x} + 2 \int_{-\infty}^{0} x f(x + \bar{x}) dx \right)^2 \\
&\geq \sigma^2 - 4 \left(\bar{x} + \int_{-\infty}^{0} x f(x + \bar{x}) dx \right) \int_{-\infty}^{0} x f(x + \bar{x}) dx. \qquad (2)
\end{aligned}$$

To minimize D, let $\delta = -\bar{x} - 2 \int_{-\infty}^{0} x f(x + \bar{x}) dx$, making the equality of (5) hold. Notice that

$$\begin{aligned}
-\bar{x} + 2 \int_{-\infty}^{0} x f(x - \bar{x}) dx &= 2 \int_{-\infty}^{\infty} (x + \bar{x}) f(x + \bar{x}) dx - \bar{x} + 2 \int_{-\infty}^{0} x f(\bar{x} - x) dx \\
&= \bar{x} + 2 \int_{-\infty}^{\infty} x f(x + \bar{x}) dx - 2 \int_{0}^{\infty} x f(\bar{x} + x) dx \\
&= \bar{x} + 2 \int_{-\infty}^{0} x f(x + \bar{x}) dx.
\end{aligned}$$

Thus from (1) we have $D(-\bar{x}, \delta) = D(\bar{x}, \delta)$. This implies that we can study $D(\bar{x}, \delta)$ for just $\bar{x} \leq 0$.

However, I can not find a way to prove that $\bar{x} = 0$ gives the minimum distortion, though this seems very natural. In fact, I have spent more than 12 hours on this problem and really got mad. So, assume that $\bar{x} = 0$. We get $\delta = -2\int_{-\infty}^{0} xf(x)dx = \boxed{\sqrt{\frac{2}{\pi}}\sigma}$ and from (5), the minimum distortion is $\sigma^2 - 4\left(\int_{-\infty}^{0} xf(x)dx\right)^2 = \sigma^2 - \delta^2 = \boxed{\frac{\pi-2}{\pi}\sigma^2}$.

2. *One bit quantization of a single Gaussian random variable.** Let the boundary be x. We will use one symbol to represent all $t \le x$ and another symbol for $t > x$. Since squared error measure is used, we should use the conditional expectations as the two reproduction points. That is, $x_0 = \int_{-\infty}^{x} \frac{f(t)}{A(x)} tdt$ for $t \le x$ and $x_1 = \int_{x}^{\infty} \frac{f(t)}{1-A(x)} tdt$ for $t > x$, where $f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{t^2/2\sigma^2}$ is the probability density function and $A(x) = \int_{-\infty}^{x} f(t)dt$. Thus the distortion is

$$D = \int_{-\infty}^{x} f(t)(t-x_0)^2 dt + \int_{x}^{\infty} f(t)(t-x_1)^2 dt.$$

If $x \ne (x_0+x_1)/2$, then we know we can decrease the distortion by moving the boundary to $x' = (x_0+x_1)/2$, since for those t between x and x', the distortion decreases by $|2(x_0-x_1)(x-x')|$, and for other t, the distortion remains unchanged. For this new boundary x', we can also find two conditional expectations $x_0' = \int_{-\infty}^{x'} \frac{f(t)}{A(x')} tdt$ and $x_1' = \int_{x'}^{\infty} \frac{f(t)}{1-A(x')} tdt$. We can further decrease the distortion D by setting x_0' and x_1' as two new reproduction points. We can repeat this procedure to decrease D until for some boundary $\bar{x}$ and conditional expectations $\bar{x}_0$ and $\bar{x}_1$, $\bar{x} = (\bar{x}_0 + \bar{x}_1)/2$. Thus we know the minimum of D requires (Simple calculations give $\int_{-\infty}^{x} f(t)tdt = -\int_{x}^{\infty} f(t)tdt = -\sigma^2 f(x)$.)

$$2x = \int_{-\infty}^{x} \frac{f(t)}{A(x)} tdt + \int_{x}^{\infty} \frac{f(t)}{1-A(x)} tdt = \sigma^2 f(x)\left(\frac{1}{1-A(x)} - \frac{1}{A(x)}\right) \tag{3}$$

for boundary x.

Another way. D is the weighted sum of two conditional variances,

$$\begin{aligned}
D(x) &= A(x)\left[\int_{-\infty}^{x} \frac{f(t)}{A(x)} t^2 dt - \left(\int_{-\infty}^{x} \frac{f(t)}{A(x)} tdt\right)^2\right] \\
&\quad + (1-A(x))\left[\int_{x}^{\infty} \frac{f(t)}{1-A(x)} t^2 dt - \left(\int_{x}^{\infty} \frac{f(t)}{1-A(x)} tdt\right)^2\right] \\
&= \int_{-\infty}^{\infty} f(x)t^2 dt - \frac{\sigma^4 f^2(x)}{A(x)} - \frac{\sigma^4 f^2(x)}{1-A(x)} \\
&= \sigma^2 - \frac{\sigma^4 f^2(x)}{A(x)(1-A(x))} = \sigma^2 - \frac{\sigma^2 e^{-x^2/\sigma^2}}{2\pi A(x)(1-A(x))}.
\end{aligned} \tag{4}$$

From $f'(x) = -\frac{x}{\sigma^2} f(x)$ and $A'(x) = f(x)$, we have

$$\begin{aligned}
\left[\frac{f^2(x)}{A(x)(1-A(x))}\right]' &= \frac{-2xf^2(x)}{\sigma^2 A(x)(1-A(x))} - \frac{f^2(x)f(x)(1-2A(x))}{[A(x)(1-A(x))]^2} \\
&= -\frac{f^2(x)}{A(x)(1-A(x))}\left[\frac{2x}{\sigma^2} + f(x)\left(\frac{1}{A(x)} - \frac{1}{1-A(x)}\right)\right].
\end{aligned}$$

*Prof. Malik Magdon-Ismail gave me the idea of first setting the boundary, instead of setting the two reproduction points. The last way wasted me more than 12 hours.

Now we returned back to (3). However, I don't know whether $x = 0$ is the only solution to (3). So, assume it is. We get $x_{0,1} = \pm 2\int_{-\infty}^{0} f(t)tdt = \boxed{\pm\sqrt{\frac{2}{\pi}}\sigma}$ and from (5), the minimum distortion is $\boxed{\frac{\pi-2}{\pi}\sigma^2}$.

3. *One bit quantization of a single Gaussian random variable.*[†] Let the boundary be x. We use one symbol to represent all $t \le x$ and another symbol for $t > x$. Since squared error measure is used, the two conditional expectations should be the reproduction points. That is, $x_0 = \int_{-\infty}^{x} \frac{f(t)}{1/2+A(x)} tdt$ for $t \le x$ and $x_1 = \int_{x}^{\infty} \frac{f(t)}{1/2-A(x)} tdt$ for $t > x$, where $f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{t^2/2\sigma^2}$ is the probability density function and $A(x) = \int_0^x f(t)dt$. Thus the distortion D is the weighted sum of two conditional variances:

$$\begin{aligned} D(x) &= (1/2 + A(x))\left[\int_{-\infty}^{x} \frac{f(t)}{1/2+A(x)} t^2 dt - \left(\int_{-\infty}^{x} \frac{f(t)}{1/2+A(x)} tdt\right)^2\right] \\ &\quad + (1/2 - A(x))\left[\int_{x}^{\infty} \frac{f(t)}{1/2-A(x)} t^2 dt - \left(\int_{x}^{\infty} \frac{f(t)}{1/2-A(x)} tdt\right)^2\right] \\ &= \int_{-\infty}^{\infty} f(x)t^2 dt - \frac{\sigma^4 f^2(x)}{1/2+A(x)} - \frac{\sigma^4 f^2(x)}{1/2-A(x)} \\ &= \sigma^2 - \frac{4\sigma^4 f^2(x)}{1-4A^2(x)} = \sigma^2 - \frac{2\sigma^2 e^{-x^2/\sigma^2}}{\pi(1-4A^2(x))}. \end{aligned} \tag{5}$$

Now we want to prove $e^{-x^2/\sigma^2} \le 1 - 4A^2(x)$. Without loss of generality, assume $x \ge 0$.

$$\begin{aligned} 4A^2(x) &= \left(\int_{-x}^{x} f(t)dt\right)^2 = \int_{-x}^{x}\int_{-x}^{x} f(u)f(v)dudv \\ &\le \frac{1}{2\pi\sigma^2}\int_{-\sqrt{2}x}^{\sqrt{2}x}\left(\int_{-\sqrt{2x^2-u^2}}^{\sqrt{2x^2-u^2}} e^{-\frac{u^2+v^2}{2\sigma^2}} dv\right) du \qquad (6) \\ &= \frac{1}{2\pi\sigma^2}\int_0^{2\pi} d\theta \int_0^{\sqrt{2}x} e^{-\frac{r^2}{2\sigma^2}} rdr = 1 - e^{-x^2/\sigma^2}. \end{aligned}$$

The reason for (6) is that $f > 0$ and the square of $[-x, x] \times [-x, x]$ is inside the circle with center $(0, 0)$ and radius $\sqrt{2}x$. So $e^{-x^2/\sigma^2} \le 1 - 4A^2(x)$ with equality iff $x = 0$. Then from (5), $D(x) \ge \frac{\pi-2}{\pi}\sigma^2$ with equality iff $x = 0$. So, the minimum distortion is $\boxed{\frac{\pi-2}{\pi}\sigma^2}$, with $x = 0$ and $x_{0,1} = \pm 2\int_{-\infty}^{0} f(t)tdt = \boxed{\pm\sqrt{\frac{2}{\pi}}\sigma}$.

[†] Prof. Malik Magdon-Ismail gave me the idea of first setting the boundary, instead of setting the two reproduction points. The last way wasted me more than 12 hours.