Solution Set 7, Computational Neuroscience

Handed out: 03/23/2004

7.1 Optimal Coding

If \( y = A(x + n) + u \) and \( x, n, \) and \( u \) are gaussians with mean zero and variances \( \sigma^2, \eta^2, \) and \( \nu^2, \) respectively, show that

\[
I(y; x) = \frac{1}{2} \log_2 \frac{A^2(\sigma^2 + \eta^2) + \nu^2}{A^2\eta^2 + \nu^2}
\]  

(1)

The sum of two gaussians is a gaussian, and the product of a gaussian with a scalar is also a gaussian. Hence \( y \) is a gaussian with variance \( A^2(\sigma^2 + \eta^2) + \nu^2 \). This gives us \( b(y) \). Meanwhile, with \( x \) fixed, \( y \) is distributed as a gaussian with mean \( Ax \) and variance \( A^2\eta^2 + \nu^2 \). Following the same steps as in ??, we get the desired answer.

- If \( \nu = 0 \), the \( A \)'s cancel, so \( A \) has no effect on \( I(y; x) \). That is, in the absence of noise, any invertible linear operation is unimportant. If \( \nu \neq 0 \), then for fixed \( \sigma, \eta, \) and \( \nu, I(x; y) \) is maximum when \( A \) is largest; i.e. \( A \to \infty \). This is equivalent to making \( A \) so large that \( \nu \) is negligible— in which case \( I(y; x) = \frac{1}{2} \log_2 \frac{\sigma^2 + \eta^2}{\eta^2} \), just as if \( A \) and \( \nu \) weren’t there at all, because we are back to the \( \nu = 0 \) case.

- Taking \( \frac{dE}{dA} = 0 \), then:

\[
E = C - \frac{k^2}{2}I
\]

\[
\Rightarrow E = A^2(\sigma^2 + \eta^2) + \nu^2 - \frac{k^2}{2} \cdot \frac{1}{2} \log_2 \frac{A^2(\sigma^2 + \eta^2) + \nu^2}{A^2\eta^2 + \nu^2}
\]

So:

\[
\frac{dE}{dA} = \frac{dE}{dA^2} \cdot \frac{dA^2}{dA} = \frac{dE}{dA^2} 2A
\]

\[
\Rightarrow \frac{dE}{dA} = 2A((\sigma^2 + \eta^2) - \frac{k^2}{4} \cdot \frac{\sigma^2\nu^2}{A^2(\sigma^2 + \eta^2) + \nu^2}(A^2\eta^2 + \nu^2))
\]

Because \( \frac{dE}{dA} = 0 \)

\[
\Rightarrow A = 0
\]

or

\[
\Rightarrow (\sigma^2 + \eta^2) - \frac{k^2}{4} \cdot \frac{\sigma^2\nu^2}{(A^2(\sigma^2 + \eta^2) + \nu^2)(A^2\eta^2 + \nu^2)} = 0
\]

\[
\Rightarrow \eta^2(\sigma^2 + \eta^2)A^4 + (\nu^2(\sigma^2 + \eta^2) + \nu^2\eta^2)A^2 + (\nu^4 - \frac{k^2\sigma^2\nu^2}{4(\sigma^2 + \eta^2)}) = 0
\]

So at last:

\[
|A| = \begin{cases} \sqrt{\frac{7}{\eta^2}} \sqrt{\frac{1 + \sqrt{1 + k^2\eta^2/\sigma^2\nu^2}}{2(1 + \eta^2/\sigma^2)}} - 1 & \text{if } \sqrt{\eta} \text{ is not real} \\ 0, & \text{otherwise} \end{cases}
\]

- Plot \( |A| \) as a function of \( \sigma \), for \( \sigma \) from 0.1 to 100, for \( \nu^2 = 1, \eta^2 = 0.001, \) and \( k^2 = 10 \).
7.1.1 Multi-Dimensional Solution

- Let’s assume that \( \nu^2 = 1 \). So now

\[
|A| = \sqrt{\frac{1}{\eta^2}} \sqrt{\frac{1 + \sqrt{1 + k^2 \eta^2 / \sigma^2}}{2(1 + \eta^2 / \sigma^2)}} - 1
\]

The key part is Taylor Expand like this: if \( \epsilon << 1 \), then

\[
(1 + \epsilon)^k = 1 + k\epsilon + o(\epsilon^2)
\]

Now let’s use this on the following part

\[
\sqrt{1 + \frac{k^2 \eta^2}{\sigma^2}} = 1 + \frac{k^2 \eta^2}{2\sigma^2} + o\left(\frac{\eta^4}{\sigma^4}\right)
\]

\[
\frac{1}{1 + \frac{\eta^2}{\sigma^2}} = 1 - \frac{\eta^2}{\sigma^2} + o\left(\frac{\eta^4}{\sigma^4}\right)
\]

\[
\sqrt{\frac{1}{\eta^2}} \sqrt{\frac{1 + \sqrt{1 + \frac{k^2 \eta^2}{\sigma^2}}}{2(1 + \eta^2 / \sigma^2)}} - 1 \approx \sqrt{\frac{1}{\eta^2}} \sqrt{\frac{1 + \frac{k^2 \eta^2}{2\sigma^2} - \frac{\eta^2}{\sigma^2}}{2(1 + \eta^2 / \sigma^2)}}
\]

\[
\Rightarrow \approx \sqrt{\frac{1}{\eta^2}} \sqrt{\frac{(k^2 - 2)\eta^2}{4\sigma^2(1 + \eta^2 / \sigma^2)}}
\]

Since \( k^2 > 2 \) and \( \sigma^2 >> \eta^2 \), so

\[
\Rightarrow \approx \sqrt{\frac{(k^2 - 2)\eta^2}{4\sigma^2}}
\]

\[
\Rightarrow \approx \sqrt{\frac{k^2 - 2}{4\sigma^2}}
\]

So

\[
|A(f, \omega)| \propto \frac{1}{\sqrt{\sigma^2(f, \omega)}}
\]

- In Fig.7 (\( k^2 = 10 \) and \( \eta^2 = 0.02 \)), we can see that in the low noise regime, it is given by whitening filter \( 1 / \sqrt{\sigma^2} \) which achieves spatiotemporal decorrelation; while at high noise regime it asymptotes the low-pass filter which suppresses noise.
Figure 2: How $|A|$ depends on the temporal frequency

- For $k^2 = 32$ and $\eta^2 = 0.032$, the rigorous solution in the previous bullet is close to the approximate solution as we mentioned in the class,

$$|A(w)| \sim \frac{w}{(1 + w^2/w_o^2)^{3/2}}$$

(2)

for $w_o \sim 5Hz$. This approximation is especially good when signal-to-noise ratio is high, i.e., at low temporal frequencies. The two curves are plotted in Fig.8.

Figure 3: Rigorous and approximate solutions in solid and dotted lines, respectively.