1 Reading (optional)

1. Read [1, Chapter 10]

2 Exercises

NOTE: Each exercise is 10 points. Only 3 exercises per assignment will be graded. If you submit more than 3 solved exercises please indicate which ones you want to be graded.

1 Consider a standard Gaussian vector $S^n$. Answer the following question when $n$ is large.

1. Let $S_{max} = \max_{1 \leq i \leq n} S_i$. Show that $\mathbb{E}[\left(S_{max} - \sqrt{2\log n}\right)^2] \to 0$ when $n \to \infty$.

2. Suppose you are given a budget of $\log n$ bits. Consider the following scheme: Let $i^*$ denote the index of the largest coordinate. The compressor stores the index $i^*$ which costs $\hat{\log} n$ bits and the decompressor outputs $S^n$ where $\hat{S}_i = \sqrt{2\log n}$ for $i = i^*$ and $S_i = 0$ otherwise. Show that distortion in terms of mean-square error satisfies $\mathbb{E}[\|\hat{S}^n - S^n\|_2^2] = n - 2\log n + o(1)$ when $n \to \infty$.

3. Using the rate-distortion function, show that the above scheme is asymptotically optimal.

2 Non-asymptotic $R(D)$. Our goal is to show that convergence to $R(D)$ happens much faster than convergence to capacity in channel coding. Consider binary uniform $X = \text{Bern}(1/2)$ with Hamming distortion and show:

1. Show that there exists a lossy code $X^n \rightarrow W \rightarrow \hat{X}^n$ with $M$ codewords and

$$\mathbb{P}[d(X^n, \hat{X}^n) > D] \leq (1 - p(nD))^M,$$

where

$$p(s) = 2^{-n} \sum_{j=0}^{s} \binom{n}{j}.$$  

2. Show that there exists a lossy code with $M$ codewords and

$$\mathbb{E}[d(X^n, \hat{X}^n)] \leq \frac{1}{n} \sum_{s=0}^{n-1} (1 - p(s))^M$$  (1)

Hint: for a non-negative integer valued random variable $A$ we have

$$\mathbb{E}[A] = \sum_{a=0}^{\infty} \mathbb{P}[A > a].$$
3. Show that there exists a lossy code with $M$ codewords and

$$E[d(X^n, \hat{X}^n)] \leq \frac{1}{n} \sum_{s=0}^{n-1} e^{-Mp(s)}$$

(Note: For $M \approx 2^{nR}$, numerical evaluation of (1) for large $n$ is challenging. At the same time (2) is only slightly slacker.)

4. For $n = 10, 50, 100$ and $200$ compute the upper bound on $\log M^*(n, 0.11)$ via (2). Compare with the lower bound

$$\log M^*(n, D) \geq nR(D).$$

3 As in the previous problem let $X = \text{Bern}(1/2)$. Using Stirling formula and (2)-(3) show

$$\log M^*(n, D) = nR(D) + O(\log n).$$

This result holds for many other memoryless sources as well.

4 Let $X$ takes values on a finite alphabet $\mathcal{A}$ and $P_X[a] > 0$ for all $a \in \mathcal{A}$. Suppose the distortion metric satisfies $d(x, y) = D_0 \implies x = y$. Show that

$$R(D_0) = \log |\mathcal{A}|,$$

while

$$R(D_0+) = H(X).$$

5 Consider Bernoulli(1/2) source $S \in \{0, 1\}$, reproduction alphabet $\hat{\mathcal{A}} = \{0, e, 1\}$ and distortion metric

$$d(a, \hat{a}) = \begin{cases} 
0, & a = \hat{a}, \\
1, & \hat{a} = e, \\
\infty, & a \neq \hat{a}, \hat{a} \neq e.
\end{cases}$$

Find rate-distortion function $R(D)$. (Note: since $D_p = D_{max} = \infty$ you cannot blindly use achievability results from lectures).

6 Consider transmitting a stationary memoryless Gaussian source $S^{k+i.d.} \sim \mathcal{N}(0, 1)$ over $n$ uses of the stationary memoryless AWGN channel with additive noise $Z^{n+i.d.} \sim \mathcal{N}(0, \sigma^2)$ and average transmission power $P$. Consider the asymptotic regime of $n \to \infty$ and rate $R = \frac{k}{n}$.

1. Fix $R > 0$. What is the smallest achievable distortion for reconstructing the source in terms of the mean-square error?

2. Now consider the special case of $R = 1$. Find an explicit scheme to achieve the optimal distortion. What is the blocklength of your scheme? (Hint: linear processing).

References

6.441 Information Theory
Spring 2016

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.