Information Theory with Applications,Math6397 Lecture Notes from October 14, 2014

taken by Sabrine Assi

4 Rate Distortion Theory

4.1 Distortion measures

4.1.1 Problem. Given a DMS $\{X_j\}_{j=1}^{\infty}$ with $H(X_1)$ is greater than the channel capacity C, can we control the transmission quality?

Strategy: Insert a lossy compression map$x_j \to \hat{x}_j$ so that $H(\hat{x}_j) < C$ so that the transmission is perfect with overwhelming probability ,so errors are intentional to reduce the amount of data

4.1.2 Definition. A distortion measure of a map (a cost function)
α:$\mathbb{A}$ x $\hat{\mathbb{A}}$ $\to$ R ,where $\mathbb{A}$ is the alphabet of the sequence of the source,$\hat{A}$ is the reproduction alphabet. Often $\hat{\mathbb{A}}$$\subset$ $\mathbb{A}$ as part a data reduction strategy .

4.1.3 Example. Assuming measure Q is uniform on $\mathbb{A}$=$\{1, 2, 3, 4\}$, we want to find $\hat{\mathbb{A}} \subset \mathbb{A}$, $\left|\hat{A}\right|$ =2 and

d(i,j)=$\begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i, j \in \{1, 2\}\, or, i, j \in \{3, 4\} \\ 2 & \text{if } else \end{cases}$
what map x $\to$ $\hat{x}$ minimizes $\mathbb{E}\left[d\left(x, \hat{x}\right)\right]$?

If we choose $\hat{\mathbb{A}}$ =$\{1, 3\}$ and let
$\hat{x}$=$\begin{cases} 1 & \text{if } x \in \{1, 2\} \\ 3 & \text{if } x \in \{3, 4\} \end{cases}$
H(X)=$\ln 4$
$H\left(\hat{X}\right)$=$\ln 2$ and
$\mathbb{E}\left[d\left(x, \hat{x}\right)\right]$=1/2
This grouping strategy gives minimum expected distortion because any other partition would give at least one case where $d\left(x, \hat{x}\right)$=2 and thus
$\mathbb{P} = \left[d\left(x, \hat{x}\right) = 2\right] \geq 1/4$ giving $\mathbb{E}\left[d\left(x, \hat{x}\right)\right] > 1/2$.

Several distortion measures are common.

4.1.4 Definition. The hamming distortion is given by $\mathbb{A} = \hat{\mathbb{A}}$ and

$$d(x, \hat{x}) = \begin{cases} 0 & \text{if } x = \hat{x} \\ 1 & \text{if } else \end{cases}$$

So
$\mathbb{E}\left[d\left(X, \hat{X}\right)\right] = \mathbb{P}\left(X \neq \hat{X}\right)$ .If $\mathbb{A}$, $\hat{\mathbb{A}} \subset \mathbb{R}$,then the square error distortion is
$d(x, \hat{x}) = |x - \hat{x}|^2$
then the expected distortion is the mean squared error.

4.1.5 Definition. Given sequences $x = \{x_1, x_2......., x_n\}$ with values in $\mathbb{A}$ and $\hat{x} = \left\{\hat{x_1}, \hat{x_2}....., \hat{x_n}\right\}$ with values in $\hat{\mathbb{A}}$,then the additive distortion is
$d_n(x, \hat{x}) = \sum_{j=1}^{n} d(x_j, \hat{x_j})$
and the maximum distortion is
$d_n(x, \hat{x}) = \max_j d(x_j, \hat{x_j})$

4.1.6 Question. what is the best possible compression rate given allowable (expected) distortion?

4.1.7 Definition. $\mathbb{A}_n$ (n,m,D) fixed length lossy compression code for a source $\{x_j\}_{j=1}^{\infty}$ with alphabet $\mathbb{A}$ and a distortion measure d_n is given by a map

$$\phi : \mathbb{A}^n \to \hat{\mathbb{A}}$$

$|\phi(\mathbb{A}^n)| = m$ and $\mathbb{E}\left[\frac{1}{n} d_n(x, \hat{x})\right] \leq D$
Remark:The size of $\phi(\mathbb{A}^n)$ is m, so $H(\phi(x_1, x_2, ...x_n)) \leq \ln m$ and the coding rate needed is at most $\frac{1}{n} \ln m$

4.1.8 Definition. For a sequence of distortion measures $\{d_n\}_{n=1}^{\infty}$ a source $\{x_j\}_{j=1}^{\infty}$ with alphabet $\mathbb{A}$ .The rate-distortion pair(R,D)is achievable if there exists a sequence fixed-length (n,m_n,D) code with $\limsup \frac{1}{n} \ln m_n \leq R$ .The rate -distortion function is R(D)=$\inf\{R' \in \mathbb{R} : (R', D)\, achievable\}$
Remark:Motivated by the proofs of the following theorems, we also consider the random maps $\phi : \mathbb{A}^n * \Omega \to \hat{\mathbb{A}}^n$ especially the case of ϕ being a discrete memoryless channel.

4.1.9 Definition. Let $d : \mathbb{A} \times \hat{\mathbb{A}} \to \mathbb{R}$ be integrable with respect to measure induced by X_j and $\hat{X}_j = \phi(x_j)$, where ϕ is a discrete memoryless channel and let d_n be additive distortion , then we define the distortion typical set

$$\mathcal{D}_\delta^n = \{(a, \hat{a}) \in \mathbb{A} \times \hat{\mathbb{A}}^n :$$

$$\left|\frac{1}{n} \ln \mathbb{P}_X(a) + H(X_1)\right| \leq \delta,$$

$$\left|\frac{1}{n} \ln \mathbb{P}_{\hat{X}(\hat{a})} + H(\hat{X}_1)\right| < \delta,$$

$$\left|\frac{1}{n}\ln \mathbb{P}_(X,\hat{X})(a,\hat{a}) + H(X,\hat{X}_1)\right| < \delta,$$

$$\left|\frac{1}{n}d_n\left(a,\hat{a}\right) - \mathbb{E}[d(X_1,\hat{X}_1)]\right| < \delta\right\}.$$