Chapter 7

Lossy Data Compression

Po-Ning Chen

Department of Communications Engineering

National Chiao-Tung University

Hsin Chu, Taiwan 30050

General lossy source compression for block codes

Definition 7.1 (lossy compression block code) Given a finite source alphabet $\mathcal{Z}$ and a finite reproduction alphabet $\hat{\mathcal{Z}}$, a block code for data compression of blocklength n and size M is a mapping $f_n(\cdot) : \mathcal{Z}^n \to \hat{\mathcal{Z}}^n$ that results in $\|f_n\| = M$ codewords of length n, where each codeword is a sequence of n reproducing letters.

Definition 7.2 (distortion measure) A distortion measure $\rho_n(\cdot, \cdot)$ is a mapping

$$\rho_n : \mathcal{Z}^n \times \hat{\mathcal{Z}}^n \to \Re^+ \stackrel{\triangle}{=} [0, \infty).$$

We can view the distortion measure as the cost of representing a source n-tuple z^n by a reproduction n-tuple $f_n(z^n)$.

Definition 7.3 (distortion inf-spectrum) Let $(\boldsymbol{Z}, \hat{\boldsymbol{Z}})$ and $\{\rho_n(\cdot, \cdot)\}_{n \geq 1}$ be given. The *distortion inf-spectrum* $\underline{\lambda}_{\boldsymbol{Z}, \hat{\boldsymbol{Z}}}(\theta)$ is defined by

$$\underline{\lambda}_{\boldsymbol{Z}, \hat{\boldsymbol{Z}}}(\theta) \stackrel{\triangle}{=} \liminf_{n \to \infty} Pr\left\{ \frac{1}{n} \rho_n(Z^n, \hat{Z}^n) \leq \theta \right\}.$$

Definition 7.4 (distortion inf-spectrum for lossy compression code $\boldsymbol{f}$) Let $\boldsymbol{Z}$ and $\{\rho_n(\cdot, \cdot)\}_{n \geq 1}$ be given. Let $\boldsymbol{f}(\cdot) \stackrel{\triangle}{=} \{f_n(\cdot)\}_{n=1}^{\infty}$ denote a sequence of (lossy) data compression codes. The *distortion inf-spectrum* $\underline{\lambda}_{\boldsymbol{Z}, \boldsymbol{f}(\boldsymbol{Z})}(\theta)$ for $\boldsymbol{f}(\cdot)$

is defined by

$$\underline{\lambda}_{\boldsymbol{Z},\boldsymbol{f}(\boldsymbol{Z})}(\theta) \stackrel{\triangle}{=} \liminf_{n\to\infty} Pr\left\{\frac{1}{n}\rho_n(Z^n, f_n(Z^n)) \leq \theta\right\}.$$

General lossy source compression for block codes

Definition 7.5 Fix $D > 0$ and $1 > \varepsilon > 0$. R is a *ε-achievable data compression rate at distortion D* for a source $\boldsymbol{Z}$ if there exists a sequence of data compression codes $f_n(\cdot)$ with

$$\limsup_{n\to\infty} \frac{1}{n} \log \|f_n\| \le R,$$

and

$$\sup\left[\theta \ : \ \underline{\lambda}_{\boldsymbol{Z},\boldsymbol{f}(\boldsymbol{Z})}(\theta) \le \varepsilon\right] \le D. \tag{7.1.1}$$

- Note that (7.1.1) is equivalent to stating that the limsup of the probability of excessive distortion (i.e., distortion larger than D) is smaller than $1 - \varepsilon$.
- The *infimum ε-achievable data compression rate at distortion D* for $\boldsymbol{Z}$ is denoted by $T_\varepsilon(D, \boldsymbol{Z})$.

Theorem 7.6 (general data compression theorem) Fix $D > 0$ and $1 > \varepsilon > 0$. Let $\boldsymbol{Z}$ and $\{\rho_n(\cdot,\cdot)\}_{n\geq 1}$ be given.

$$R_\varepsilon(D) \leq T_\varepsilon(D, \boldsymbol{Z}) \leq R_\varepsilon(D-\gamma),$$

for any $\gamma > 0$, where

$$R_\varepsilon(D) \triangleq \inf_{\left\{P_{\hat{\boldsymbol{Z}}|\boldsymbol{Z}} \; : \; \sup\left[\theta \; : \; \underline{\lambda}_{\boldsymbol{Z},\hat{\boldsymbol{Z}}}(\theta) \leq \varepsilon\right] \leq D\right\}} \bar{I}(\boldsymbol{Z}; \hat{\boldsymbol{Z}}),$$

where the infimum is taken over all conditional distributions $P_{\boldsymbol{Z}|\hat{\boldsymbol{Z}}}$ for which the joint distribution $P_{\boldsymbol{Z},\hat{\boldsymbol{Z}}} = P_{\boldsymbol{Z}} P_{\boldsymbol{Z}|\hat{\boldsymbol{Z}}}$ satisfies the distortion constraint.

Example

- Probability-of-error distortion measure $\rho_n : \mathcal{Z}^n \to \mathcal{Z}^n$:

$$\rho_n(z^n, \tilde{z}^n) = \begin{cases} n, & \text{if } z^n \neq \tilde{z}^n; \\ 0, & \text{otherwise.} \end{cases}$$

- We define a data compression code $f_n : \mathcal{Z}^n \to \mathcal{Z}^n$ based on a chosen (asymptotic) lossless fixed-length data compression code book $\mathcal{C}_n \subset \mathcal{Z}^n$:

$$f_n(z^n) = \begin{cases} z^n, & \text{if } z^n \in \mathcal{C}_n; \\ \underline{0}, & \text{if } z^n \notin \mathcal{C}_n, \end{cases}$$

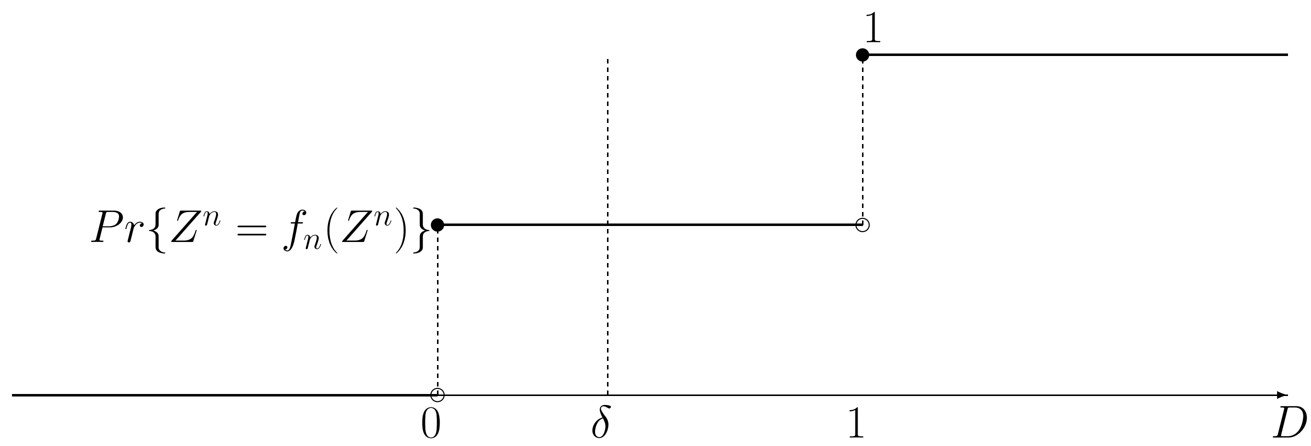
where $\underline{0}$ is some default element in $\mathcal{Z}^n$.

- Then $(1/n)\rho_n(z^n, f_n(z^n))$ is either 1 or 0 which results in a cumulative distribution function as shown in Figure 7.2.

- Consequently, for any $\delta \in [0, 1)$,

$$Pr\left\{\frac{1}{n}\rho_n(Z^n, f_n(Z^n)) \leq \delta\right\} = Pr\left\{Z^n = f_n(Z^n)\right\}.$$

Example



The CDF of $(1/n)\rho_n(Z^n, f_n(Z^n))$ for the probability-of-error distortion measure.

Example

- By comparing the (asymptotic) lossless and lossy fixed-length compression theorems under the probability-of-error distortion measure, we observe that

$$
\begin{aligned}
R_\varepsilon(\delta) &= \inf_{\{P_{\hat{\boldsymbol{Z}}|\boldsymbol{Z}} : \sup[\theta \ : \ \underline{\lambda}_{\boldsymbol{Z},\hat{\boldsymbol{Z}}}(\theta) \le \varepsilon] \le \delta\}} \bar{I}(\boldsymbol{Z};\hat{\boldsymbol{Z}}) \\
&= \begin{cases} 0, & \delta \ge 1; \\ \displaystyle\inf_{\{P_{\hat{\boldsymbol{Z}}|\boldsymbol{Z}} : \liminf_{n\to\infty} Pr\{Z^n = \hat{Z}^n\} > \varepsilon\}} \bar{I}(\boldsymbol{Z};\hat{\boldsymbol{Z}}), & \delta < 1, \end{cases} \\
&= \begin{cases} 0, & \delta \ge 1; \\ \displaystyle\inf_{\{P_{\hat{\boldsymbol{Z}}|\boldsymbol{Z}} : \limsup_{n\to\infty} Pr\{Z^n \ne \hat{Z}^n\} \le 1-\varepsilon\}} \bar{I}(\boldsymbol{Z};\hat{\boldsymbol{Z}}), & \delta < 1, \end{cases}
\end{aligned}
$$

  where

$$
\underline{\lambda}_{\boldsymbol{Z},\hat{\boldsymbol{Z}}}(\theta) = \begin{cases} \liminf_{n\to\infty} Pr\left\{Z^n = \hat{Z}^n\right\}, & 0 \le \theta < 1; \\ 1, & \theta \ge 1. \end{cases}
$$

- In particular, in the extreme case where ε goes to one,

$$
\bar{H}(\boldsymbol{Z}) = \inf_{\{P_{\hat{\boldsymbol{Z}}|\boldsymbol{Z}} : \limsup_{n\to\infty} Pr\{Z^n \ne \hat{Z}^n\} = 0\}} \bar{I}(\boldsymbol{Z};\hat{\boldsymbol{Z}}).
$$

Example

- Therefore, in this case, the data compression theorem reduces (as expected) to the asymptotic lossless fixed-length data compression theorem.