Data source X: real-value or discrete random variable $X: \Omega \to \mathbb{R}^M$

Def 1. For any distortion d, divergence D, and a random variable X, the (information) rate-distortion-perception function (RDPF) is defined as

$$R(\theta_d, \theta_p) = inf_{P_{\hat{X}|x}} I[X, \hat{X}]\ s.t. E[d(X, \hat{X})] \leq \theta_d\ and\ D[P_x, P_{\hat{x}}] \leq \theta_D$$

It can be generalized for an arbitrary set of constraints on $P_{X,\hat{X}}$.

Def 2. For a source $X \sim P_X$ and a set of real valued functions D_i of joint distributions $P_{X,\hat{X}}$, the rate functions $(IRF) is\ defined\ as$

$$R(\theta) = inf_{P_{\hat{X}|X}} I[X, \hat{X}]\ s.t. \forall i: D_i[P_{X,\hat{X}}] \leq \theta_i$$

The RDPF is a special case of IRF where we choose

$$D_1[P_{X,\hat{X}}] = E[d(X, \hat{X})]\ and\ D_2[P_{X,\hat{X}}] = D[P_X, P_{\hat{X}}]$$

Given a polish metric space $\mathcal{X}$, a Borel measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is induced by the metric

X: a random variable with distribution $p_X \in P(\mathcal{X})$

Δ: $\mathcal{X} \times \mathcal{X} \to [0, \infty)$ a measurable distortion function $\Delta(x, \hat{x}) = 0\ iff\ x = \hat{x}$

d: $P(\mathcal{X}) \times P(\mathcal{X}) \to [0, \infty)$ a divergence with $d(p_x, p_{\hat{x}}) = 0\ iff\ p_x = p_{\hat{x}}$

$$R(D, P) \coloneqq \inf_{P_{\hat{X}|X}} I(X: \hat{X})\ subject\ to$$

$$E[\Delta(X, \hat{X})] \leq D, d(P_X, P_{\hat{X}}) \leq P$$

Let S be uniformly distributed over the unit circle $S \coloneqq \{S \in \mathbb{R}^2: \|s\|_2 = 1\}$. $\|\cdot\|_p$ is the p-norm

The question?

Minimize the expected distortion $E[\|s - \hat{s}\|_2^2]$ if S is encoded using 1 bit where $\hat{S}$ is required to meet the perfect perceptual quality ($\hat{S}$ is also uniform distribution) e.g.

$$K := \begin{cases} 0, & \frac{\theta(S)}{\pi} + W \in [0,1) \cup [2,3), \\ 1, & \frac{\theta(S)}{\pi} + W \in [1,2) \end{cases}$$

$$\hat{S} := \left(\cos\left((K+W)\pi\right), \sin\left((K+W)\pi\right)\right)$$

Where W is uniform distribution over $[0,1)$ and independent to S.

$$E[\|s - \hat{s}\|^2] = 2 - \frac{8}{\pi^2}$$

This is the minimum achievable with private randomness only.

If not private randomness (a.k.a. common randomness), then

$$K := \begin{cases} 0, & \frac{\theta(S)}{\pi} + W \in [0,1) \cup [2,3), \\ 1, & \frac{\theta(S)}{\pi} + W \in [1,2) \end{cases}$$

$$\hat{S} := \left(\cos\left((K-W)\pi\right), \sin\left((K-W)\pi\right)\right)$$

Then

$$E[\|s - \hat{s}\|^2] = 2 - \frac{4}{\pi}$$

Classical Rate-Distortion Theory

Trade off between compression rate and quality loss.

Rate distortion function $R(D)$ minimum rate to encode data with an expected distortion D.

D: distortion constraint

P: perception constraint

R is achievable with common randomness if $(big\ n)$ there exist shared on a polish space Q,

with seed distribution P_Q, encoding distribution $P_{Z|X^nQ}$, decoding distribution $P_{\hat{X}^n|ZQ}$ with $\hat{X} = X$,

such that the joint distribution

$$P_{X^nQZ\hat{X}^n} := P_X^n P_Q P_{Z|X^n} P_{\hat{X}^n|ZQ}$$

Satisfies

$$\frac{1}{n} H(Z \mid Q) \leq R$$

$$\frac{1}{n}\sum_{t=1}^{n} E\left[\Delta\left(X_t, \hat{X}_t\right)\right] \leq D$$

$$d\left(P_X, P_{\hat{X}_t}\right) \leq p, t \in [1:n]$$

The infimum of R is denoted by $R_{cr}(D, P)$.

When the encoder and decoder does not share the (random) seed, $R_{pr}(D, P)$.

If the encoder and decoder is deterministic, $R_{nr}(D, P)$

※Perceptual Quality (Realism)

Measured in terms of **divergence** between source and target.

Distortion is measured using semimetrics.

Remark 1. $\frac{1}{n}H(Z, Q) \leq R \Leftrightarrow \frac{1}{n}\log|Z| \leq R$ (Z is alphabet size)

Z can be represented by a variable length code of $H(Z \mid Q = q) + 1$

Normalize by n and take the expectation with Q gives us

$\frac{1}{n}H(Z \mid Q) + \frac{1}{n}$ with $n \to \infty$ $\frac{1}{n}H(Z \mid Q)$

Remark 2. By definition,

$$R_{cr}(D, P) \leq R_{pr}(D, P) \leq R_{nr}(D, P)$$

Assumption 1. $d(\cdot,\cdot)$ is convex in its second argument. (taking this assumption from now on)

Them 1. $R_{cr}(D, P) = R(D, P)$ for $D \geq 0, P \geq 0$.

Remark 3. If

$$\frac{1}{n}\sum_{t=1}^{n} d(P_X, P_{\hat{X}}) \leq P \; or \; d\left(P_X, \frac{1}{n}\sum_{t=1}^{n} P_{\hat{X}}\right) \leq P$$

Then Them 1. Holds.

Assumption 2. For any $D > 0$ and $P > 0$ we have $R(D,P) < \infty$.

There exists a discrete random variable $\tilde{\mathcal{X}}$ satisfying $\tilde{\mathcal{X}} \subseteq \mathcal{X}$ and $|\tilde{\mathcal{X}}| < \infty$ such that

$I(X{:}\tilde{X}) \leq R(D,P) + \epsilon$

$E[\Delta(X,\tilde{X})] \leq D + \epsilon$

$E\left[\max_{\tilde{x}\in\tilde{\mathcal{X}}} \Delta(x,\tilde{x})\right] < \infty$

$d(P_X, P_{\tilde{X}}) \leq P + \epsilon$

$d(P_X, \gamma) < \infty$ for all distribution of possible $\tilde{\mathcal{X}}$.

With assumptions 1,2

Them 2. $R_{nr}(D,P) = R(D,P)$

With asymptotic setting it is possible to leverage the aggregated randomness to simultaneously shape the marginal distributions of all output symbols into the desired form via proper deterministic encoding and decoding

Them 3. For $P \geq 0$,

$$R_{cr}(0,P) = R_{pr}(0,P) = R_{nr}(0,P) = R(0,P)$$

Where

$$R(0,P) = \begin{cases} H(X), & P_X \text{ is a discrete distribution} \\ & \infty \end{cases}$$

Assumption 3. For any $D > 0$ and $\epsilon > 0$ there exists a discrete random variable $\tilde{\mathcal{X}}$ and an arbitrary random variable $\hat{X} \sim \mathcal{X}$ such that $X \leftrightarrow \tilde{X} \leftrightarrow \hat{X}$ form a Markov chain.

The support of $\tilde{X}$ satisfies $|\tilde{\mathcal{X}}| < \infty$ and

$I(X;\tilde{X}) \leq R(D,0) + \epsilon$

$E[\Delta(X,\tilde{X}) \leq D$

$P_{\hat{X}} = P_X$

In short, assumption 3 basically postulates the existence of a discrete random variable $\tilde{X}$ sitting between X and $\tilde{X}$ with $I(X;\tilde{X}) \approx I(X,\hat{X})$

Them 4. With assumption 1,3 $R_{pr}(D,0) = R(D,0)$ for $D > 0$.