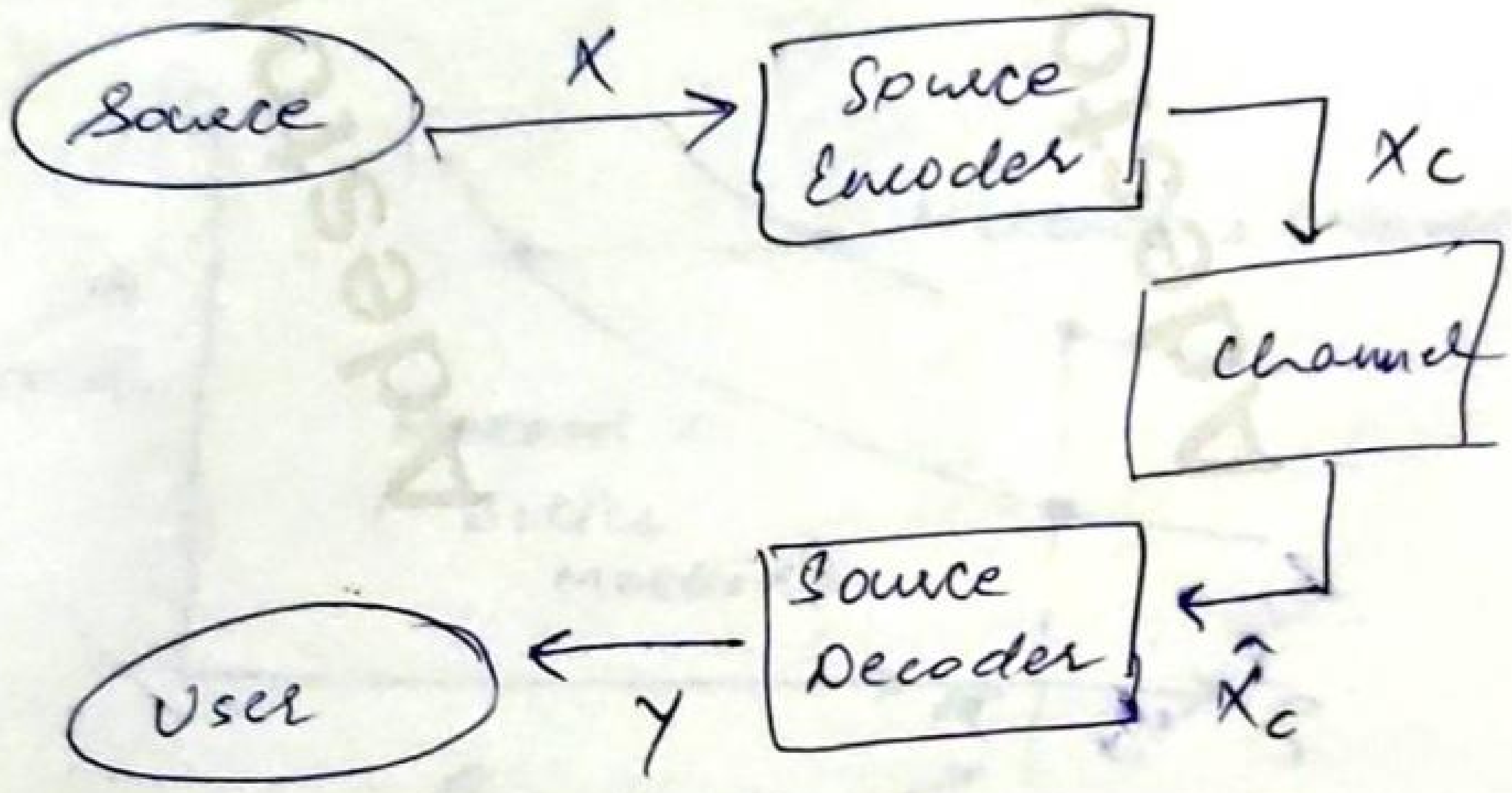
UNIT-4. LOSSY COMPRESSION TECHNIQUE (Irreversible Compression)

→ Distortion Criteria
→ Models

} Mathematical Preliminaries of Lossy Coding.

- Some popular measures of fidelity or closeness between the informance sequence $\{x_n\}$ and the reconstructed sequence $\{y_n\}$

Generic Compression Scheme



→ The fidelity of reconstructed sequence can be checked by the difference between original & reconstructed (uncompressed) value that will tell us the Distortion Ratio.

→ 2 measures of distortion are — 1) Squared Error Measure
— 2) Absolute Difference Measure.

1) Squared Error measure :

$$d(x_n, y_n) = (x_n - y_n)^2$$

2) Absolute Error measure :

$$d(x_n, y_n) = |x_n - y_n|$$

It is very difficult to calculate on term basis. So need an average.

Mean squared Error

3) For the sequence of length N, MSE will be: (mse).

$$D = E\left\{\frac{1}{N}\sum_{n=1}^{N}(X_n - Y_n)^2\right\} = \frac{1}{N}\sum_{n=1}^{N} E\left[(X_n - Y_n)^2\right]$$

E, expectation is w.r.t the joint distribution of X_n & Y_n.
Assuming the sequences are i.i.d (independent & identically distributed)
then,

$$D = E\left[(X-Y)^2\right] = \sigma^2_{err}$$

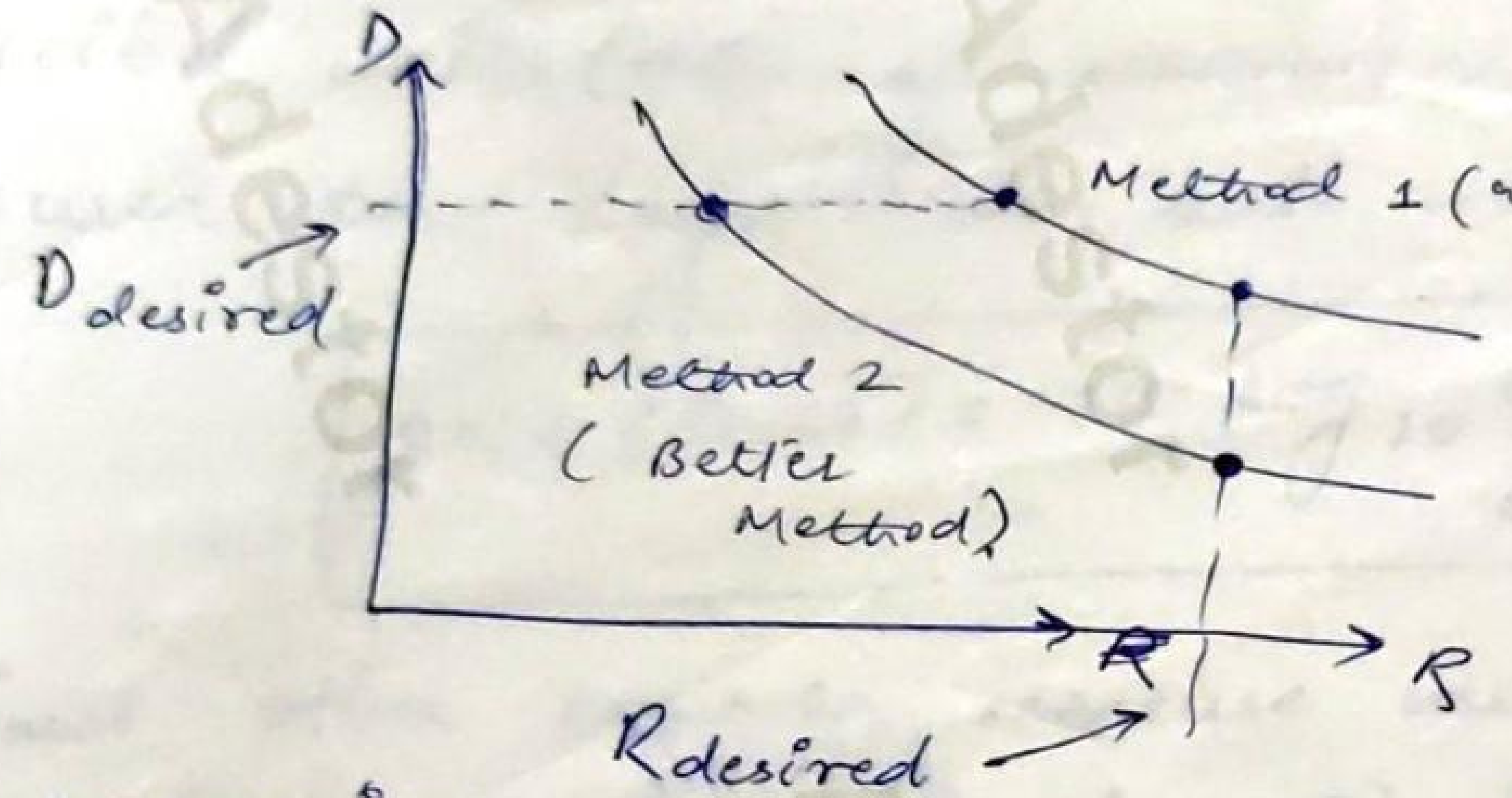
Goals of Lossy Compression:

1) Reduce distortion for a fixed rate
2) Reduce rate for a fixed distortion.

} works for constraint optimization problem

Rate is nothing but measure of how many bits are used to represent the compressed signal.

- bits / sample
- bits / second
- bits / fixed duration
-



2 kinds of compression problems:

* Distortion - Rate Problem
↓
→ Given a constraint on transmitted data rate or storage capacity but the problem is to compress the source file at or below this rate but at the highest fidelity possible.

→ Ex: Areas of voicemail, digital cellular Mobile Radio & video-conferencing

* Rate - Distortion Problem
↓
→ Given the requirement to achieve a certain pre-specified fidelity, the problem is to meet the requirement with as few bits / second

→ Ex: compression in areas of CD-Quality Audio & Motion picture quality video.

→ Many existing Audio; Speech; Image; & Video compression techniques have transform, quantization, & bit-rate allocation procedure that capitalizes on general shape of Rate-Distortion Functions.

* RATE-DISTORTION THEORY: It gives analytical expression for how much compression can be achieved using lossy compression method.

→ Rate is usually - no. of bits / data sample - to be stored / sent.

→ Distortion - expected value of the square of diff. b/w I/p & O/p signal (MSE).

→ If one is interested in the size of the error relative to the signal, we can measure the signal-to-noise ratio (SNR) by taking the ratio of average square of the original data sequence & the mean square error

$$\boxed{SNR = \frac{\sigma_x^2}{\sigma_d^2}}$$

→ In decibel units (dB) of measurement; the SNR is measured on logarithmic scale - with base 10.

$$\boxed{SNR\ (\text{in dB}) = 10 \log_{10} \frac{\sigma_x^2}{\sigma_d^2}}$$

→ At times, other common measure used for distortion is Peak-signal-to-Noise Ratio (PSNR). Because at times one is interested in the size of Error Relative to the peak value of the signal x_{peak} than in the size of the error relative to the average squared value of the signal.

$$\boxed{PSNR\ (dB) = 10 \log_{10} \frac{x_{peak}^2}{\sigma_d^2}}$$

→ Other measure used after MSE, for distortion evaluation of Image compression Algorithms is Absolute Difference.

$$\boxed{d_1 = \frac{1}{N} \sum_{n=1}^{N} |x_n - y_n|}$$

→ At times, distortion is not perceptible as long as it is below some threshold. This requires the maximum value of the error magnitude.

$$\boxed{d_\infty = \max_n |x_n - y_n|}$$

* INFORMATION THEORY

→ We need to discuss about information relationship b/w the random variables that take on value from 2 different alphabets. As might the variables'll be distinct.

ⓐ CONDITIONAL ENTROPY: We already've understood entropy earlier

$$\boxed{H(X) = -\sum_{i=0}^{N-1} P(x_i)\,(\log_2 P(x_i))}$$

where X, random variable taken source alphabet $\mathcal{X} = \{x_0, x_1 \ldots x_{N-1}\}$.

$$\boxed{H(Y) = -\sum_{j=0}^{M-1} P(y_j) \log_2 P(y_j)}$$

where Y, random variable taken from reconstruction alphabet

$\mathcal{Y} = \{y_0, y_1, \ldots y_{M-1}\}$

→ Thus, Conditional Entropy is the average value of the conditional self-information b/w 2 random variables.

self Information: $\boxed{i(A) = \log \frac{1}{P(A)} = -\log P(A)}$

Conditional self-Information: $\boxed{i(A|B) = \log \frac{1}{P(A|B)} = -\log P(A|B)}$

Ex: Event A: "Frank is thirsty"

Event B: "Frank has not drunk anything in 2 days".

then $P(A|B) = 1$. but $\underbrace{i(A|B) = 0}$

This information is of no surprise that as he hasn't drank for 2 days so he will be thirsty surely

→ One is generally interested in average value of the conditional self-information & that average value is called as conditional entropy.

Thus, the conditional Entropies of source & reconstruction alphabets are given as:

Amount of uncertainty in random variable X

$$H(X|Y) = -\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} P(x_i|y_j)\, P(y_j) \log_2 P(x_i|y_j)$$

$$H(Y|X) = -\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} P(y_j|x_i)\, P(x_i) \log_2 P(y_j|x_i)$$

UNIT-4 : MODELS in LOSSY COMPRESSION TECHNIQUE

→ Probability Models
→ Linear System Model
→ Physical Models

* PROBABILITY MODELS: Here, in lossy compression schemes the focus is more towards general approach rather than exact correspondance, which was for lossless ones.
The reasons are more pragmatic than theoretical some pdf (probability distribution function) are more analytically tractable than others & our focus is towards relating the source distribution to the following 4 probability models. They are as:

→ Uniform → Gaussian → Laplacian → Gamma

$\underbrace{\hspace{6cm}}_{\text{Distributions}}$

Uniform

→ When we don't know anything about the distribution of source i/p except the range of values, we prefer this model

→ The pdf is as:

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$

→ distributed b/w random variables a and b.

Gaussian

→ It is tractable & by Central Limit Theorem;

→ The pdf of GD is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-\mu)^2}{2\sigma^2}$$

→ Mean, μ

→ variance, σ^2

→ Due to above 2 reasons Gaussian distribution is prefered.

Laplacian

→ At times, we've sharp peak at zero. Ex: Speech consists of silence. So, sample of speech will be zero or close to zero with high probability.

→ In these case Laplacian dist. works well than Gaussian.

→ The pdf is as:

$$f_X(x) = \frac{1}{\sqrt{2\sigma^2}} \exp \frac{-\sqrt{2}|x|}{\sigma}$$

$\mu = 0$, mean
σ^2, variance

Gamma

→ More peaked distribution.

→ less tractable

→ The pdf is as:

$$f_X(x) = \frac{\sqrt[4]{3}}{\sqrt{8\pi\sigma|x|}} \exp \frac{-\sqrt{3}|x|}{2\sigma}$$

Linear System Models :

→ Earlier we assumed that information sequence $\{x_n\}$ can be modeled by a sequence of iid random variables.

→ But, in general most sequences derived from real sources such as speech will contain dependencies. (correlation b/w samples is easy way).

→ To view the information sequence as o/p of linear system governed by a difference equation with an iid input. ← method

→ This structure of linear system as reflected in the parameters of the difference eq. introduces the correlation

→ Information sequence :
$$x_n = \sum_{i=1}^{N} a_i x_{n-i} + \sum_{j=1}^{M} b_j \epsilon_{n-j} + \epsilon_n$$

x_n, samples of the process.

ϵ_n, white nose sequence.

→ Physical Models : They're just depended on the source and how the physics is built for the specific approach. No as such special variety is being defined in this but the physical appearance of the model plays an important role in the complete evaluation.
