符号理論･暗号理論

- No.2 エントロピー -

渡辺　裕

Coding Theory / Cryptography

- No.2 Entropy -

Hiroshi Watanabe

エントロピー

- エントロピーの定義
  - 無記憶情報源 S のN 個の事象 $a_1, a_2, \ldots, a_N$ の出現確率を $p_1, p_2, \ldots, p_N$ としたとき、エントロピー $H(s)$ は次式で与えられる

$$H(S) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

  - 情報源 S の各事象は独立に起こる（無記憶情報源、ex. dice）
  - エントロピーは出現確率に偏りがある場合に小さくなり、偏りが少ない場合には大きくなる（あいまいさの尺度）

Entropy

- Definition of Entropy
  - Let occurrence probabilities of N-events $a_1, a_2, \ldots, a_N$ of memory-less source S be $p_1, p_2, \ldots, p_N$. Entropy $H(s)$ is given as follows.

$$H(S) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

  - Each event of the source S independently occurs. (memory-less source, ex. dice)
  - Entropy becomes small when occurrence probability is biased and vice-versa. (measure of ambiguity)

エントロピーの性質

- エントロピーの値は正（非負）

$$H(s) \geq 0$$
$$since \quad \log_2 p_i \leq 0 \quad (0 \leq p_i \leq 1)$$

- エントロピーの最小値

$$H(S)_{min} = 0$$
$$when \quad p_j = 1 \quad (one\,of\ p_i)$$

Characteristics of Entropy

- Value of Entropy is positive

$$H(s) \geq 0$$
$$since \quad \log_2 p_i \leq 0 \quad (0 \leq p_i \leq 1)$$

- The minimum value of Entropy

$$H(S)_{min} = 0$$
$$when \quad p_j = 1 \quad (one\,of\ p_i)$$

エントロピーの性質(2)

- エントロピーの最大値
  - 平均情報量が最大になるのは、各事象の出現確率が等しい場合（最も不確定な場合）

$$p_1 = p_2 = \cdots = p_N = \frac{1}{N}$$

$$H(S)_{max} = \log_2 N$$

Characteristics of Entropy(2)

- The maximum value of Entropy
  - Mean information content becomes its maximum when the occurrence probability of each event is even(the most uncertain case).

$$p_1 = p_2 = \cdots = p_N = \frac{1}{N}$$

$$H(S)_{max} = \log_2 N$$

エントロピー関数

- 無記憶2元情報源 $S=\{0,1\}$ のエントロピーの算出
- “0” の出現確率を p, “1”の出現確率を q とする
- 事象“0”か“1”が出現するから、$p+q=1$
- エントロピーを p の関数として求める

$$H(p) = -p\log_2 p - q\log_2 q$$
$$= -p\log_2 p - (1-p)\log_2(1-p)$$

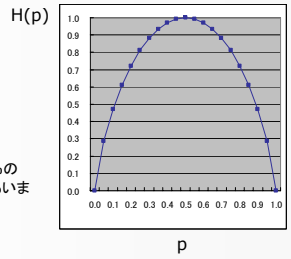
Entropy Function

- Entropy of Memory-less 2-symbol source $S=\{0,1\}$
- Occurrence probabilities of symbol “0” and “1” are p, q
- Total occurrence probability is $p+q=1$
- Obtain entropy as a function of p

$$H(p) = -p\log_2 p - q\log_2 q$$
$$= -p\log_2 p - (1-p)\log_2(1-p)$$

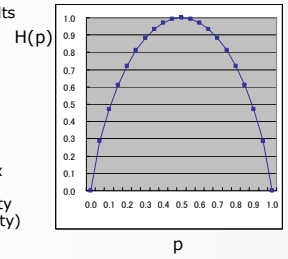
エントロピー関数(2)

- $H(p)$の数値計算結果
  - 最大値
    - $H(0.5)=1$
  - 最小値
    - $H(0)=H(1)=0$
- “0”か“1”の出現確率が50%の場合にエントロピーが最大（あいまいさが最大）



Entropy Function(2)

- Numeric calculation results of $H(p)$
  - Max max value
    - $H(0.5)=1$
  - Min. value
    - $H(0)=H(1)=0$
- Entropy becomes its max value when “0” or “1” occurs with the probability 50% (maximum ambiguity)



計算例

- $P_1=1/4, p_2=1/8, p_3=1/2, p_4=1/8$ のときのエントロピー

$$H(S)=-\frac{1}{4}log_2\frac{1}{4}-\frac{2}{8}log_2\frac{1}{8}-\frac{1}{2}log_2\frac{1}{2}$$
$$=\frac{2}{4}+\frac{3}{4}+\frac{1}{2}$$
$$=1.75$$

Quiz

- Obtain Entropy for $P_1=1/4, p_2=1/8, p_3=1/2, p_4=1/8$

$$H(S)=-\frac{1}{4}log_2\frac{1}{4}-\frac{2}{8}log_2\frac{1}{8}-\frac{1}{2}log_2\frac{1}{2}$$
$$=\frac{2}{4}+\frac{3}{4}+\frac{1}{2}$$
$$=1.75$$

アナログ情報源のエントロピー

- 確率密度 $p(x)$ なる標本値をもつアナログ情報源 S のエントロピー $H(S)$ の定義

$$H(S)\equiv-\int_{-\infty}^{\infty}p(x)log_2p(x)dx$$

Entropy of Analog Source

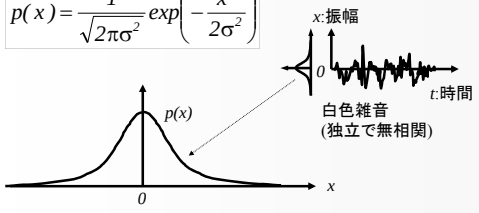
- Definition of entropy $H(S)$ of analog source S when its sample has probability density function $p(x)$.

$$H(S)\equiv-\int_{-\infty}^{\infty}p(x)log_2p(x)dx$$

ガウス分布情報源のエントロピー

- 平均 0, 分散σ^2 の白色ガウス情報源 X の確率分布関数 $p(x)$

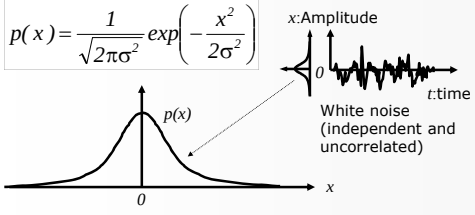
$$p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}exp\left(-\frac{x^2}{2\sigma^2}\right)$$



Entropy of Gaussian Source

- Probability density function $p(x)$ of white gaussian source X with average 0, variance σ^2

$$p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}exp\left(-\frac{x^2}{2\sigma^2}\right)$$



ガウス分布情報源のエントロピー(2)

- エントロピーの算出

$$
\begin{aligned}
H(X) &= -\int_{-\infty}^{\infty} p(x)\log_2 p(x)dx \\
&= -E\left[\log_2 p(x)\right] \\
&= E\left[-\log_2 p(x)\right] \\
&= E\left[-\log_2\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\right] \\
&\ldots \quad (cont.)
\end{aligned}
$$

Entropy of Gaussian Source(2)

- Calculation of entropy

$$
\begin{aligned}
H(X) &= -\int_{-\infty}^{\infty} p(x)\log_2 p(x)dx \\
&= -E\left[\log_2 p(x)\right] \\
&= E\left[-\log_2 p(x)\right] \\
&= E\left[-\log_2\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\right] \\
&\ldots \quad (cont.)
\end{aligned}
$$

ガウス分布情報源のエントロピー(3)

- エントロピーの算出 (続き)

$$
\begin{aligned}
H(X) &= -\log_2\frac{1}{\sqrt{2\pi\sigma^2}} - E\left[\log_2\left(\exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\right] \\
&= \log_2\sqrt{2\pi\sigma^2} - \log_2\left(\exp\left(-\frac{E\left[x^2\right]}{2\sigma^2}\right)\right) \\
&= \log_2\sqrt{2\pi\sigma^2} - \log_2\left(\exp\left(-\frac{1}{2}\right)\right) \qquad \because E\left[x^2\right] = \sigma^2 \\
&\ldots \quad (cont.)
\end{aligned}
$$

Entropy of Gaussian Source(3)

- Calculation of entropy (contd.)

$$
\begin{aligned}
H(X) &= -\log_2\frac{1}{\sqrt{2\pi\sigma^2}} - E\left[\log_2\left(\exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\right] \\
&= \log_2\sqrt{2\pi\sigma^2} - \log_2\left(\exp\left(-\frac{E\left[x^2\right]}{2\sigma^2}\right)\right) \\
&= \log_2\sqrt{2\pi\sigma^2} - \log_2\left(\exp\left(-\frac{1}{2}\right)\right) \qquad \because E\left[x^2\right] = \sigma^2 \\
&\ldots \quad (cont.)
\end{aligned}
$$

ガウス分布情報源のエントロピー(4)

- エントロピーの算出 (続き)

$$
\begin{aligned}
H(X) &= \log_2\sqrt{2\pi\sigma^2} + \log_2 e^{\frac{1}{2}} \\
&= \log_2\sqrt{2\pi\sigma^2} + \log_2\sqrt{e} \\
&= \log_2\sqrt{2\pi\sigma^2 e}
\end{aligned}
$$

(ビット/サンプル)

Entropy of Gaussian Source(4)

- Calculation of entropy (contd.)

$$
\begin{aligned}
H(X) &= \log_2\sqrt{2\pi\sigma^2} + \log_2 e^{\frac{1}{2}} \\
&= \log_2\sqrt{2\pi\sigma^2} + \log_2\sqrt{e} \\
&= \log_2\sqrt{2\pi\sigma^2 e}
\end{aligned}
$$

(bit/sample)

最大エントロピー定理

- アナログ情報源のエントロピーの最大化
  - 一様分布では $H(X)$ は無限大に発散する
  - 制約を与えた上での最大化に意味がある
- 平均パワーが設定値(Pとする)以下でのエントロピーの最大化問題
  - 拘束条件

$$\int_{-\infty}^{\infty} x^2 p(x)dx \le P$$
$$\int_{-\infty}^{\infty} p(x)dx = 1$$

Maximum Entropy Theorem

- Maximization of entropy of analog source
  - $H(X)$ get infinity for uniform distribution
  - Maximization under constraint has meaning
- Entropy maximization problem when average power is set to less than certain value.
  - constrained condition

$$\int_{-\infty}^{\infty} x^2 p(x)dx \le P$$
$$\int_{-\infty}^{\infty} p(x)dx = 1$$

最大エントロピー定理(2)

- Lagrangeの未定係数λ, μを導入した評価関数 J の最大化

$$J = -p(x)\log_2 p(x) + \lambda x^2 p(x) + \mu p(x)$$

- $p(x)$ で偏微分して 0 とおき, 解の形を仮に求める

$$\frac{\partial J}{\partial p(x)} = -\log_2 p(x) - \log_2 e + \lambda x^2 + \mu = 0$$
$$p(x) \equiv \alpha e^{-\beta x^2}$$

Maximum Entropy Theorem(2)

- Maximize evaluation function J with Lagrange's uncertain coefficients λ, μ

$$J = -p(x)\log_2 p(x) + \lambda x^2 p(x) + \mu p(x)$$

- Take partial derivative by $p(x)$, put 0, obtain a form of p(x)

$$\frac{\partial J}{\partial p(x)} = -\log_2 p(x) - \log_2 e + \lambda x^2 + \mu = 0$$
$$p(x) \equiv \alpha e^{-\beta x^2}$$

最大エントロピー定理(3)

- $p(x)$ を拘束条件に代入し, 定数α, βを求めると

$$\alpha = \frac{1}{\sqrt{2\pi P}} \quad , \quad \beta = \frac{1}{2P}$$

となり, $p(x)$ は

$$p(x) = \frac{1}{\sqrt{2\pi P}} e^{-\frac{x^2}{2P}}$$

となるから, 平均 0,分散 P のガウス分布であることがわかる

Maximum Entropy Theorem(3)

- Obtain constants α, β by inserting $p(x)$ to constrained condition

$$\alpha = \frac{1}{\sqrt{2\pi P}} \quad , \quad \beta = \frac{1}{2P}$$

Thus, $p(x)$ is given by

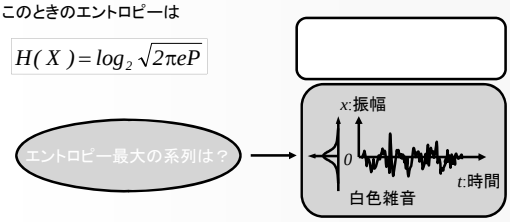
$$p(x) = \frac{1}{\sqrt{2\pi P}} e^{-\frac{x^2}{2P}}$$

We find that it is gaussian with average 0, variance P.

最大エントロピー定理(4)

- 最大エントロピー定理:「平均電力が P 以下の情報源では，エントロピーが最大になるのは，分散が P の白色ガウス情報源である」
- このときのエントロピーは

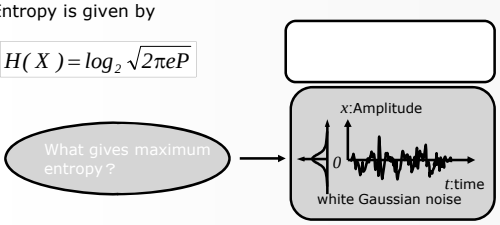
$$H(X) = \log_2 \sqrt{2\pi e P}$$



Maximum Entropy Theorem(4)

- Maximum Entropy Theorem:「When source's average power is less than P, white gaussian source with variance P gives the maximum entropy」
- Entropy is given by

$$H(X) = \log_2 \sqrt{2\pi e P}$$



歪みを許す情報源符号化

- 無歪みの場合
  - 常に $x=y$ であり，復号器の出力 Y の値 y を知れば，完全に情報源出力 X の値 x を知ることができる
  - Y の値を知ったときに X について知る情報量は $H(X)$
- 歪みを許す場合
  - Y の値を知っても，X になおあいまいさが $H(X|Y)$ だけ残る
  - Yの値を知ったとき、X に関して得る平均情報量は、相互情報量 $I(X;Y)$ に相当する（ $I(X;Y)=H(X)-H(X|Y)$ ）

Source Coding with Distortion

- In case of no distortion
  - always $x=y$, we can know value x from source output X by knowing value y of decoder output Y
  - Information to know X is $H(X)$ by knowing value of Y
- In case of allowing distortion
  - Ambiguity $H(X|Y)$ still remains to X even we know value of Y
  - Information of X under known Y is given by $I(X;Y)=H(X)-H(X|Y)$ (Mutual information)

歪みを許す情報源符号化(2)

- 情報量を伝送するのに必要な符号量
  - 無歪みの場合にはエントロピー $H(X)$
  - 歪みを許す場合には $I(X;Y)$
- 問題点
  - 歪みを許す場合には平均符号長 $I(X;Y)$ が同じでも，情報源符号化の方法により，平均歪み d_m が異なる
  - 逆に，平均歪み d_m が同じでも，情報源符号化の方法により，平均符号長 $I(X;Y)$ が異なる

Source Coding with Distortion(2)

- Required bits to transmit information
  - Entropy $H(X)$ in case of no distortion
  - $I(X;Y)$ in case of allowing distortion
- Problem
  - Average distortion d_m differs by source coding method even if average code length $I(X;Y)$ is the same when allowing distortion
  - That is, average code length $I(X;Y)$ differs by source coding method even if average distortion d_m is the same

歪みを許す情報源符号化(3)

- レート歪み関数 (Rate-distortion function)
  - 与えられた値 D に対し，平均歪み d_m が

    $d_m \le D$

    の条件を満たすような情報源符号化を行い，そのときの $I(X;Y)$ の最小値 $R(D)$ をレート歪み関数と呼ぶ

$$R(D) = \min_{d_m \le D} \left(I(X;Y) \right)$$

Source Coding with Distortion(3)

- Rate-distortion function
  - When source coding satisfies the next condition in regard to average distortion d_m to given value D,

    $d_m \le D$

    the minimum value $R(D)$ of $I(X;Y)$ is called rate-distortion function

$$R(D) = \min_{d_m \le D} \left(I(X;Y) \right)$$

歪みを許す情報源符号化(4)

- 歪みを許す場合の情報源符号化定理
  - 平均歪み d_m を D 以下に抑えるという条件の下で，任意の正数 ε に対し，情報源 S を1情報源あたりの平均符号長 L が

    $R(D) \le L < R(D) + \varepsilon$

    となる2元符号へ符号化できる．

Source Coding with Distortion(4)

- Source coding theorem when distortion is allowed
  - Under the condition making average distortion d_m not exceed D, we can encode source S to binary symbol code having average code length L per input in regard to an arbitrary positive number ε.

    $R(D) \le L < R(D) + \varepsilon$

レート歪み関数

- レート歪み関数の計算
  - 平均歪み d_m が D 以下になるという拘束条件の下で，情報源出力 X と復号器出力 Y の間の相互情報量 $I(X;Y)$ を最小化
  - 符号化法を変えれば，$P(y|x)$ が変化する ($P(x|y)$ ではないことに注意)
  - $P(y|x)$ の変化によって，$I(X;Y)$ が変化する
- 最小化の対象 (解法はLagrangeの未定係数法による)

$$I(X;Y) = \sum_x P(x) \sum_y P(y|x) \log_2 \frac{P(y|x)}{P(y)}$$

Rate-distortion Function

- Calculation of rate-distortion function
  - Under the condition of that average distortion d_m does not exceed D, minimize mutual information $I(X;Y)$ between source output X and decoder output Y
  - $P(y|x)$ varies by changing coding method (Note: not $P(x|y)$)
  - $I(X;Y)$ varies by changing $P(y|x)$
- Minimization object (solved by Lagrangian method)

$$I(X;Y) = \sum_x P(x) \sum_y P(y|x) \log_2 \frac{P(y|x)}{P(y)}$$

レート歪み関数(2)

- 拘束条件

$$P(y)=\sum_x P(x)P(y|x)$$
$$d_m=\sum_x P(x)\sum_y P(y|x)d(x,y)\le D$$

ただし

$$P(y|x)\ge 0$$
$$\sum_y P(y|x)=1$$

Rate-distortion Function(2)

- Condition

$$P(y)=\sum_x P(x)P(y|x)$$
$$d_m=\sum_x P(x)\sum_y P(y|x)d(x,y)\le D$$

where

$$P(y|x)\ge 0$$
$$\sum_y P(y|x)=1$$

白色ガウス情報源のレート歪み関数

- 歪み測度の定義

$$d(x,y)=(x-y)^2$$

- 問題設定
  - X の確率密度関数 $p(x)$ が

$$p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}$$

で与えられることが前提

Rate-distortion Function for White Gaussian Source

- Definition of distortion measure

$$d(x,y)=(x-y)^2$$

- Setting problem
  - Let probability density function $p(x)$ from source X can be given by

$$p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}$$

白色ガウス情報源のレート歪み関数(2)

- 平均歪み d_m を D 以下に抑えるという条件

$$d_m=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x)p(y|x)(x-y)^2dxdy\le D$$

の下で，次式を条件付確率密度関数 $P(y|x)$ に関して最小化

$$I(X;Y)=\int_{-\infty}^{\infty}p(x)\left[\int_{-\infty}^{\infty}p(y|x)\log_2\frac{p(y|x)}{p(y)}dy\right]dx$$

ただし，

$$p(y)=\int_{-\infty}^{\infty}p(x)p(y|x)dx$$

(問題設定終わり)

Rate-distortion Function for White Gaussian Source(2)

- Condition to suppress average distortion d_m under D

$$d_m=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x)p(y|x)(x-y)^2dxdy\le D$$

Let us minimize the next equation in regard to the conditional probability function $P(y|x)$.

$$I(X;Y)=\int_{-\infty}^{\infty}p(x)\left[\int_{-\infty}^{\infty}p(y|x)\log_2\frac{p(y|x)}{p(y)}dy\right]dx$$

where,

$$p(y)=\int_{-\infty}^{\infty}p(x)p(y|x)dx$$

(end of setting problem)

白色ガウス情報源のレート歪み関数 (3)

- 解法
  - $I(X;Y)=H(X)-H(X|Y)$ であるから，$I(X;Y)$ の最小化は $H(X|Y)$ の最大化
  - $Z=X-Y$ と置くと，$H(X|Y)=H(Z|Y)$ (差を知っても X を知ることになる) このとき

$$H(Z|Y) \le H(Z)$$

であるから

$$H(X|Y) \le H(Z)$$

Rate-distortion Function for White Gaussian Source(3)

- Solution
  - $I(X;Y)=H(X)-H(X|Y)$ Thus, minimization of $I(X;Y)$ corresponds to maximization of $H(X|Y)$
  - Let us define $Z=X-Y$, $H(X|Y)=H(Z|Y)$ (can know X by knowing difference), we can write

$$H(Z|Y) \le H(Z)$$

Thus,

$$H(X|Y) \le H(Z)$$

白色ガウス情報源のレート歪み関数 (4)

- 解法（続き）
  - 一方，平均歪み d_m は Z^2 の期待値であり，これが D の値を越えないことが拘束条件となる

$$d_m = E\left[Z^2\right] \le D$$

  - この条件の下に，$H(Z)$ が最大となるのは，“最大エントロピー定理”により，Z が平均 0，分散 D の白色ガウス情報源である
  - このときの Z のエントロピー

$$H(Z) = \log_2 \sqrt{2\pi e D}$$

Rate-distortion Function for White Gaussian Source(4)

- Solution (contd.)
  - Average distortion d_m is expected value of Z^2. This value should not exceed D (constrained condition).

$$d_m = E\left[Z^2\right] \le D$$

  - Under this condition, $H(Z)$ is maximized by “Maximum Entropy Theorem” when Z is white gaussian source with average 0, variance D.
  - In this case, entropy of Z is given by

$$H(Z) = \log_2 \sqrt{2\pi e D}$$

白色ガウス情報源のレート歪み関数 (5)

- 解法（続き）
  - 相互情報量 $I(X;Y)$ の最小値は次式で与えられる

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &\ge H(X) - H(Z) \\ &= \log_2 \sqrt{2\pi e \sigma^2} - \log_2 \sqrt{2\pi e D} \\ &= \frac{1}{2} \log_2 \frac{\sigma^2}{D} \qquad (bit/sample) \end{aligned}$$

Rate-distortion Function for White Gaussian Source(5)

- Solution (contd.)
  - The minimum value of mutual information $I(X;Y)$ is given by the next equation.

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &\ge H(X) - H(Z) \\ &= \log_2 \sqrt{2\pi e \sigma^2} - \log_2 \sqrt{2\pi e D} \\ &= \frac{1}{2} \log_2 \frac{\sigma^2}{D} \qquad (bit/sample) \end{aligned}$$

白色ガウス情報源のレート歪み関数 (6)

- 解法（続き）
  - レート歪み関数 $R(D)$ は $I(X;Y)$ の最小値

$$R(D) = \min_{d_m \le D} I(X;Y)$$
$$= \frac{1}{2}\log_2 \frac{\sigma^2}{D} \qquad (bit/sample)$$

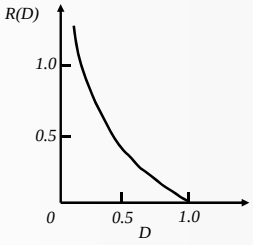
Rate-distortion Function for White Gaussian Source(6)

- Solution (contd.)
  - Rate distortion function $R(D)$ is the minimum value of $I(X;Y)$

$$R(D) = \min_{d_m \le D} I(X;Y)$$
$$= \frac{1}{2}\log_2 \frac{\sigma^2}{D} \qquad (bit/sample)$$

白色ガウス情報源のレート歪み関数 (7)

- R-D グラフ (分散1の場合)
  - 下に凸
  - $R(D)=\infty$ at $D=0$



Rate-distortion Function for White Gaussian Source(7)

- R-D Curve (Variance=1)
  - Convex to downside
  - $R(D)=\infty$ at $D=0$