Lec. 06

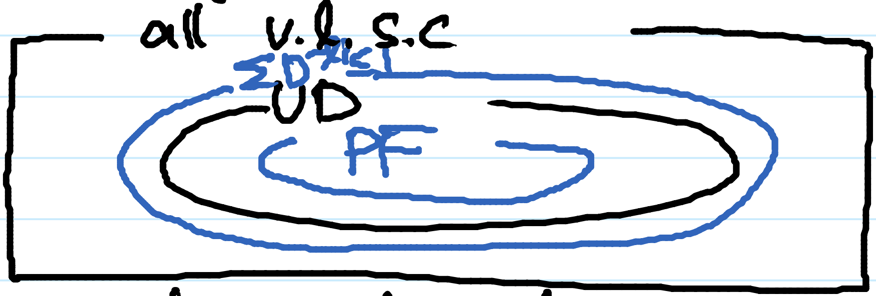
2013년 3월 19일 화요일
오후 1:54

Review

- single-source single-destination
  - data compression
  - " transmission
- zero-error data compression

$$\underset{\mathcal{C}}{\text{minimize}}\ L(\mathcal{C})$$

subject to $\mathcal{C}$ is UD.

- 

- Entropy lower bound
  - Kraft's inequality

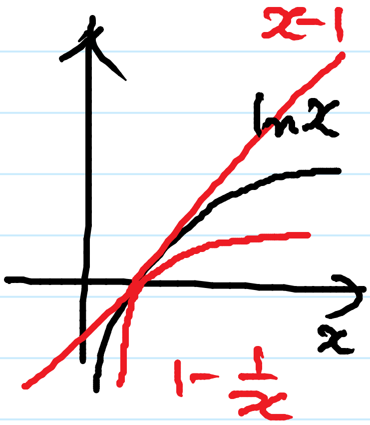
    If $\mathcal{C}$ is a D-ary UD v.l.s.c w/ $(l_i)_{i=1}^{|\mathcal{X}|}$,

    then $$\sum_{i=1}^{|\mathcal{X}|} D^{-l_i} \le 1$$

  - Fundamental Inequality

$$1-\frac{1}{x} \le \ln x \le x-1$$

iff $x=1$ iff $x=1$



the entropy of X

$$H_D(X) \triangleq \sum_{i=1}^{|X|} p_i \log_D \frac{1}{p_i}$$

$$= -\sum_{i=1}^{|X|} p_i \log_D p_i$$

$$= E\left\{ \log_D \frac{1}{P_X(X)} \right\}$$

$$= \sum_{x \in X} \log_D \frac{1}{P_X(x)} P_X(x)$$

representation

$$L(\mathcal{C}) - H_D(X) = \sum_{i=1}^{|X|} p_i l_i - \sum_{i=1}^{|X|} p_i \log_D \frac{1}{p_i}$$

$$= \sum_{i=1}^{|X|} p_i \log_D D^{l_i} p_i \quad \text{iff } p_i = D^{-l_i}, \forall i$$

$$= \sum_{i=1}^{|X|} p_i \frac{\ln D^{l_i} p_i}{\ln D} \geq \sum_{i=1}^{|X|} p_i \frac{1}{\ln D}\left(1 - \frac{1}{D^{l_i} p_i}\right)$$

$$= \frac{1}{\ln D}\left(\sum_{i=1}^{|X|} p_i - \sum_{i=1}^{|X|} \cancel{p_i} \frac{1}{\cancel{p_i}} D^{-l_i} \right)$$

$$\geq \frac{1}{\ln D}(1-1) = 0$$

if $\mathcal{C}$ is UD.

- PF codes

Optimization $\rightarrow$ constraint set partitioning

$(l_i)_{i=1}^{|X|}$

$$\begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_{|X|} \end{bmatrix}$$

In the subset of D-ary UD codes w/ $l_1, l_2, \cdots, l_{|X|}$,
there exists a PF code w/ the same codeword length profile.

proof) $\sum_{i=1}^{|X|} D^{-l_i} \le 1$

$l'_i \leftarrow$ re-ordered codeword lengths.

mathematical induction — basis/basic step
transfinite induction — induction/inductive step
conclusion

existence
uniqueness

Page 03

2013년 3월 19일 화요일

오후 1:54

• $\min_{\mathcal{C} \in \mathcal{BUD}} L(\mathcal{C}) < H_D(X) + 1$

(proof) Given a source w/ $(p_i)_{i=1}^{|X|}$, there exists a D-ary PF code w/ $l_i \triangleq \lceil -\log_D p_i \rceil, \forall i$, because

$$\sum_{i=1}^{|X|} D^{-l_i} = \sum_{i=1}^{|X|} D^{-\lceil -\log_D p_i \rceil} \le \sum_{i=1}^{|X|} p_i = 1.$$

Simplify!

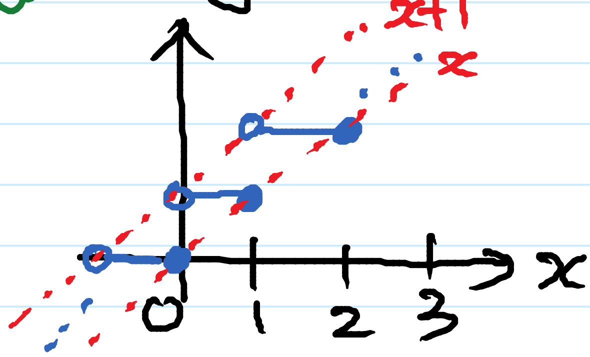
$\lceil x \rceil$ Visualize!

$\log_D p_i - 1 < -\lceil -\log_D p_i \rceil \le \log_D p_i$

$-\log_D p_i \le \lceil -\log_D p_i \rceil < -\log_D p_i + 1$

$y = \lceil x \rceil$

$x \le \lceil x \rceil < x + 1$

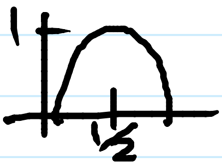


Then,

$$L(\mathcal{C}) = \sum_{i=1}^{|X|} p_i l_i < \sum_{i=1}^{|X|} p_i(-\log_D p_i + 1)$$

$$= H_D(X) + 1.$$

★ $h_b(p) \le 1$ bit.



page 04

2013년 3월 19일 화요일

오후 2:53

- $\boxed{\text{a binary source}} \rightarrow X_i \leftarrow$ i.i.d.

$2^{-1} = \frac{1}{2}$

$2^{-2} =$

$2^{-3} =$

$2^{-4} =$

$\vdots$

X

x	$P_X(x)$
0	p
1	$1-p$

(X_1, X_2)

(x_1, x_2)	$P_{X_1 X_2}(x_1, x_2)$
0 0	p^2
0 1	$p(1-p)$
1 0	$p(1-p)$
1 1	$(1-p)^2$

$$X^{(N)} \triangleq (X_1, \cdots, X_N)$$

$$H_D(X_1 \cdots X_N) \le L(\mathcal{C}^{(N)}) < H_D(X_1, \cdots, X_N) + 1$$

$$N H_D(X) \le L(\mathcal{C}^{(N)}) \le N H_D(X) + 1$$

$$H_D(X) \le \frac{1}{N} L(\mathcal{C}^{(N)}) \le H_D(X) + \frac{1}{N}.$$

- "asymptotically achieve"

$\mathcal{C}$ achieves the entropy lower bound.

vs.

The sequence $\mathcal{C}^{(N)}$ of codes achieves the entropy lower bound as N tends to infinity

Huffman codes

Experience ⇒ Lesson

page 05

2013년 3월 19일 화요일

오후 3:09

- Fixed-length source coding

x	$P_X(X=x)$	$C(x)$
1	P_1	
2	P_2	
$\vdots$	$\vdots$	
$\lvert\mathcal{X}\rvert$	$P_{\lvert\mathcal{X}\rvert}$	

(i) fixed length

(ii) zero-error

$\lceil \log_D |\mathcal{X}| \rceil$

non-zero error

"We will construct a sequence of fixed-length D-ary source code $\mathcal{C}^{(N)}$ such that

(i) codeword symbols / source output $\to R$

(ii) error rate $\to 0$."