Subject : Intro. to Info. Theory

No. : 1

Date : 10 / 22 / 2019

*

$$\begin{tikzcd}
 & & H_\delta(\mathcal{X}^N) \arrow[r, dotted, "\text{block}"'] & \approx N H(\mathcal{X}) \\
H_0 & & (\mathcal{X}, p) \arrow[ll, "0"] \arrow[u, "\delta"'] \arrow[dl, "\delta"'] \arrow[dr, "\text{var. length symbol (Huffman)}"] & \\
 & H_\delta(\mathcal{X}) & & \approx H(\mathcal{X})
\end{tikzcd}$$

"δ" : "Some" wrongly coded , others perfectly correct

* what if "all" slightly wrongly coded ?

$\{1, 2, \cdots, 100\} \xrightarrow{\text{quantization}} \{5, 10, 15, \cdots, 95\}$

$\mathcal{X}$ $\qquad$ $\mathcal{Y}$

$H_0 = \log 100$ $\qquad$ $H_0 = \log 19$

$\longrightarrow \{20, 40, \cdots, 80\}$

$H_0 = \log 4$

* bits $\downarrow$, error $\uparrow$

distortion

$d(x, y)$

e.g. $d(x, y) = (x-y)^2$

$d(x, y) = [\![x \neq y]\!]$

randomized quantization

$\mathcal{X}$ \ $\mathcal{Y}$	1	3	5
1	1	0	0
2	$\frac{1}{2}$	$\frac{1}{2}$	0
3	0	1	0
4	0	$\frac{1}{2}$	$\frac{1}{2}$
5	0	0	1
6	0	0	1

deterministic quantization

$\mathcal{X}$ \ $\mathcal{Y}$	1	3	5
1	•		
2		•	
3		•	
4			•
5			•
6			•

given $P(x)$, different quantization $P(y|x) \Rightarrow$ different $P(y)$

$\Rightarrow$ different $H(Y)$

* optimal (smallest) $H(Y)$ = rate r if symbol coding on Y
under distortion constraint ?

$$\underbrace{\sum_x \sum_y p(x)\, p(y|x)\, d(x,y)}_{\text{average distortion}} \leq \square$$

*

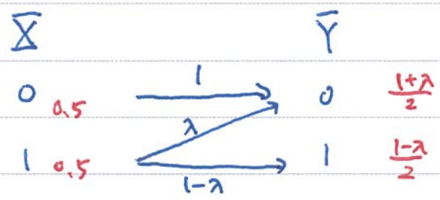
	$\bar{Y}$		$\bar{X}$		$\bar{Y}$	
0.6	0 0		0 0 (0.1)		0 0	0.3
	0 0	$\Leftarrow$	0 1 (0.2)	$\Rightarrow$	0 0	
	0 0		1 0 (0.3)		1 1	0.7
0.4	1 1		1 1 (0.4)		1 1	

rate $\mathcal{H}(0.6)$ — rate $\mathcal{H}(0.7)$

distortion (sq. err.) 0.8 — distortion 0.5

distortion (0/1 error) 0.5 — distortion 0.5

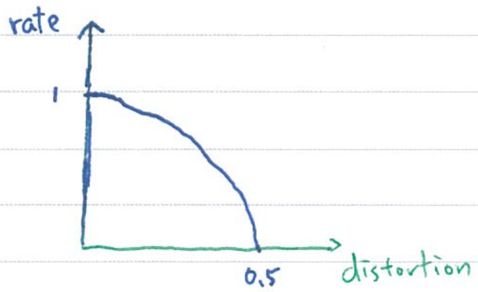
*



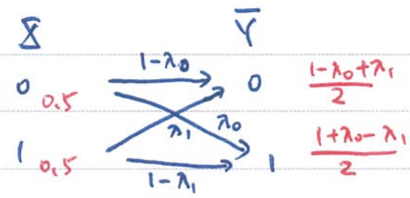
$p(y|x)$

rate $= H(Y) = \mathcal{H}\left(\frac{1-\lambda}{2}\right)$

distortion $= \frac{1}{2}\lambda$



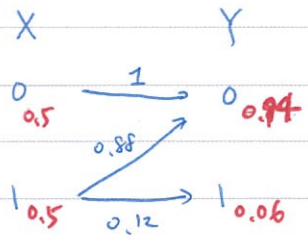
Can we do better? No if (randomized scalar quantizati



distortion $= \frac{\lambda_0+\lambda_1}{2}$

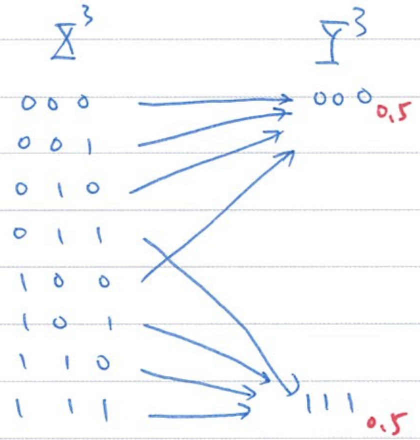
next : analyze rate-distortion tradeoff in general

* can we do better with vector quantization ?



rate $= \frac{1}{3}$

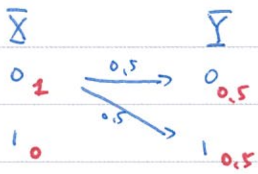
distortion $= 0.44$



rate $= \frac{1}{3} = \frac{\#(Y^3)}{3}$

avg. distortion $= \frac{1}{8} \cdot 6 / 3 = 0.25$

* why ? typical in Y^N can be redundant
what $H(Y)$ means
to quantize X^N



typical in X^N : 0 0 0 0 0 --- 0 , need "0" bit

typical in Y^N : 0 1 1 0 1 0 0 1 , need N bits in total, 1 bit in average