Information Theory & Coding

Sept 15th 2020

Yesterday:

Data Compression

- Source code: $c: \mathcal{U} \mapsto \{0,1\}^*$
- injective code: $u \neq v \Rightarrow c(u) \neq c(v)$
- Kraft Sum $(c) = \sum_{u \in \mathcal{U}} 2^{-\text{length}(c(u))}$
- extending a code c:

$$c^n : \mathcal{U}^n \to \{0,1\}^*$$
$$u_1 \dots u_n \to c(u_1)c(u_2)\dots c(u_n)$$

$$c^* : \mathcal{U}^* \to \{0,1\}^*$$
exactly same

Notations: $\mathcal{U}, \mathcal{V}$ are sets

$\mathcal{U} \times \mathcal{V} : \{(u,v) : u \in \mathcal{U}, v \in \mathcal{V}\}$

$|\mathcal{U}|$ = cardinality of $\mathcal{U}$

$|\mathcal{U} \times \mathcal{V}| = |\mathcal{U}| \cdot |\mathcal{V}|$

$\mathcal{U}^n = \underbrace{\mathcal{U} \times \mathcal{U} \times \dots \times \mathcal{U}}_{n} = \{(u_1 \dots u_n) : u_1 \in \mathcal{U}, u_2 \in \mathcal{U} \dots\}$

$|\mathcal{U}^n| = |\mathcal{U}|^n$

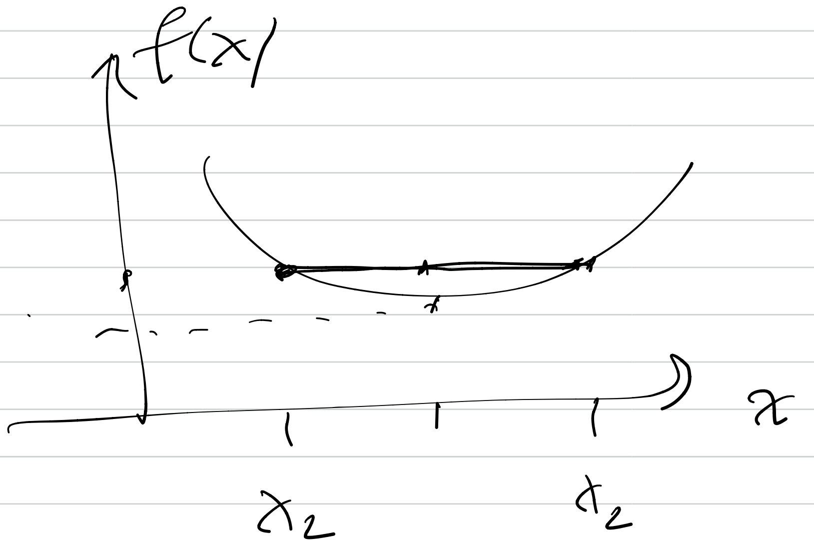
Convex function: (convex $\cup$)

$$f : \underbrace{\text{subset of } \mathbb{R}^n}_{\text{Convex}} \longmapsto \mathbb{R} \quad \text{with the}$$

following property

$$\forall x_1, x_2 \quad f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

$$0 \leq \lambda \leq 1$$



Concave $\cap$ function f is a function where $-f$ is convex.

Fact: Logarithm is a concave function.

Lemma: $\ln x \leq x - 1 \quad \forall x > 0$, equality iff $x = 1$

Pf: Recall Taylor expansions:

$$f(x) = f(x_0) + f'(x_0)(x-1) + \frac{1}{2} f''(\xi)(x-x_0)^2$$

$$\ln x = \underbrace{\ln(1)}_{0} + 1 \cdot (x-1) + \underbrace{\frac{1}{2}\left(\frac{-1}{\xi^2}\right)}_{\leq 0} \underbrace{(x-1)^2}_{\geq 0}$$

$$\leq (x-1)$$

Corollary: if $p_1, p_2, \dots, p_n \geq 0$, $\sum_i p_i = 1$,

$x_1, x_2, \dots, x_n > 0$, then

$$\sum_{i=1}^{n} p_i \log x_i \leq \log \left(\sum_{i=1}^{n} p_i x_i\right).$$

($\equiv$ log is $\frown$)

Pf: note: sufficient to prove for $\log = \ln$.

let $A = \sum_{i=1}^{n} p_i x_i$, so we need to prove

$$\left(\sum_{i=1}^{n} p_i \ln x_i\right) - \ln A \leq 0.$$

$$\equiv \sum_{i=1}^{n} p_i \ln x_i - \sum_{i=1}^{n} p_i \ln A \leq 0$$

$$\equiv \sum_{i=1}^{n} p_i \ln \frac{x_i}{A} \leq 0.$$

But $\sum p_i \ln \frac{x_i}{A} \leq \sum_i p_i \left(\frac{x_i}{A} - 1\right)$

$$= \left(\frac{1}{A} \underbrace{\sum_{i=1}^{n} p_i x_i}_{A}\right) - \underbrace{\left(\sum_{i=1}^{n} p_i\right)}_{=1}$$

$$= 1 - 1$$

$$= \underline{0} \quad //$$

Back to source coding:

Def: suppose $c : \mathcal{U} \to \{0,1\}^*$, $d : \mathcal{V} \to \{0,1\}^*$, let

$c \times d : \mathcal{U} \times \mathcal{V} \to \{0,1\}^*$

$(c \times d)(u,v) = c(u)\,d(v)$: concatenation.

Lemma: $\mathrm{KraftSum}(c \times d) = \mathrm{KraftSum}(c)\,\mathrm{KraftSum}(d)$

Pf: $KS(c \times d)$

$$= \sum_{u,v}' 2^{-\text{length}\,(c\times d)(u,v)}$$

$$= \sum_{u,v}' 2^{-[\text{length}\,c(u) + \text{length}\,d(v)]}$$

$$= \sum_{u,v}' 2^{-\text{length}\,c(u)} \cdot 2^{-\text{length}\,d(v)}$$

$$= \left(\sum_{u}' 2^{-\text{length}\,c(u)}\right)\left(\sum_{v}' 2^{-\text{length}\,d(v)}\right)$$

$$= KS(c) \cdot KS(d) \quad //$$

Recall: c injective $\Rightarrow KS(c) \leq \log(1 + |\mathcal{U}|)$

Thm: c is uniq. decodable $\Rightarrow KS(c) \leq 1$.

Pf: c is uniq decodable $\equiv c^*$ is injective

$\Rightarrow (\forall n,\ c^n$ is injective.$)$

(**Recall**: $c^* : \mathcal{U}^* \to \{0,1\}^*$

$c^*(u_1 \ldots u_n) = c(u_1)\,c(u_2) \ldots c(u_n)$.)

$$\Rightarrow \underset{\substack{\parallel \\ KS(c)^n}}{KS(c^n)} \leq \log_2(1 + |\mathcal{U}^n|) = \log_2(1 + |\mathcal{U}|^n).$$

So: $\forall n$, $KS(c)^n \leq \log_2(1+|\mathcal{U}|^n)$. $(*)$

if $KS(c) > 1 \Rightarrow$ LHS is exponentially $\nearrow$ with n
RHS is linear in n.
$\Rightarrow$ contradiction with $(*)$

$\Rightarrow KS(c) \leq 1$. //

Aside on Taylor expansions:

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(x_0)\frac{(x-x_0)^n}{n!}$$

$$f(x) = f(x_0) + (x-x_0) f'(\xi) \qquad \xi \text{ between } x_0 \,\&\, x$$

$$= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(\xi)(x-x_0)^2$$

$c \times d$ example $\quad \mathcal{U} = \{a, b\}$

$\mathcal{V} = \{\alpha, \beta\}$

$c(a) = 0 \quad c(b) = 11, \quad d(\alpha) = 000 \quad d(\beta) = 1$

$\Rightarrow (c\times d)(a\beta) = 01$

$(c \times d)(b\alpha) = 11000$

So: KS of a U.D. code c is ≤ 1.

Yesterday we also saw: Suppose U is RV taking values in $\mathcal{U}$, and c is a code $: \mathcal{U} \to \{0,1\}^*$

then $E[\text{length } c(\bar{U})] \geq H(U) - \log_2 KS(c)$

$\left(= \sum_u p(u)\,\text{length}(c(u))\right)$

with $H(U) \stackrel{\Delta}{=} \sum_u p(u) \log_2 \frac{1}{p(u)}$

Corrolary: if c is Uniq. Decodable then

$E(\text{length } c(U)) \geq H(U)$.

Def: a code $c: \mathcal{U} \to \{0,1\}^*$ is said to be prefix-free if $u \neq v \Rightarrow$

$c(u)$ is not a prefix of $c(v)$.

Def: a bit-string $a = (a_1 a_2 .. a_m)$ is a prefix of a bit-string $b = (b_1 \dots b_n)$ if (*) $n \geq m$ & $a_i = b_i$ $i = 1 \dots m$.

Observation: c is P.F. $\Rightarrow$ c is Uniq. Decodable

Example: $\mathcal{U} = \{a, b, c\}$

$c: 0, 10, 11$

$c^*(..) = 0|11|0|0|11|10|0$

a c a a c b a

Example $\mathcal{U} = \{a, b, c\}$

$0 \quad 01 \quad 11$

$c^* = 0111111 - \quad 0$

Ans. to Sina: Sardinas–Petersen test (sp??)

Thm: suppose $\ell : \mathcal{U} \to \{0,1,2,3,\dots\}$,

and $\underbrace{\sum_u 2^{-\ell(u)} \le 1}$. Then

there exists a P.F. $c : \mathcal{U} \to \{0,1\}^*$ s.t.

$\forall u$ length $c(u) = \ell(u)$

Example: $\mathcal{U} = \{a, b, c\}$

ℓ: 1, 2, 2 $\qquad 2^{-1} + 2^{-2} + 2^{-2} = 1$ ✓

c: 0, 10, 11

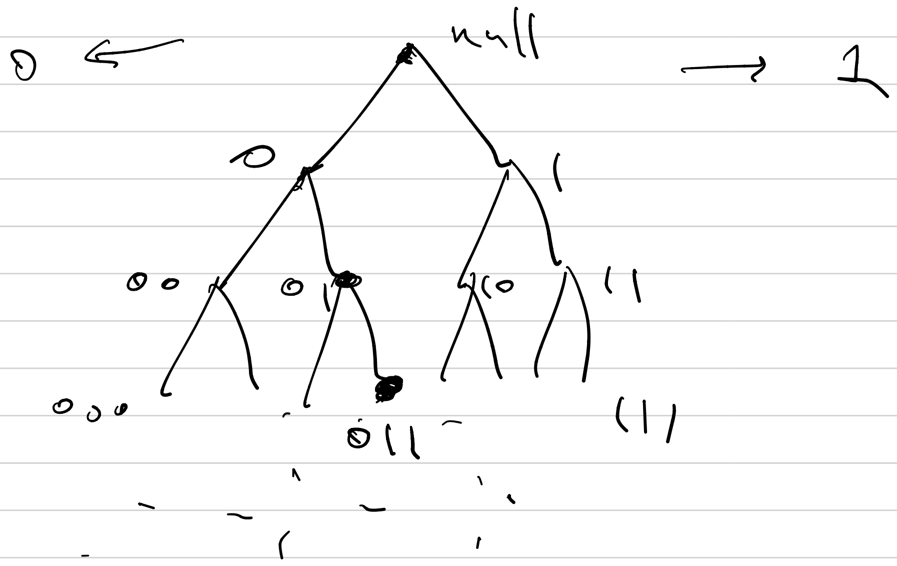
c: 0, 11, 10

In particular: if a code $\tilde{c} : \mathcal{U} \to \{0,1\}^*$

has $KS(\tilde{c}) \le 1 \Rightarrow \exists c$ PF. st

$\forall u$ length$(c(u))$ = length$(\tilde{c}(u))$.

Preliminary: Binary trees.



Pf of the thm: given $\{\ell(u)\}$ let

$\mathcal{U} = \{u_1, \dots, u_K\}$ s.t. $\ell(u_1) \le \ell(u_2) \le \dots \le \ell(u_K)$.

Consider the following procedure:

start with an infinite binary tree, mark all nodes as available.

For $i = 1, \dots K$

assign to u_i $c(u_i)$ = any available string at depth $l(u_i)$ in the tree

mark this string and all its descendants as unavailable.

Print $\{c(u_i) : i = 1 \dots K\}$.

To complete the proof we need to ensure that will succeed. to this end let

$F(i, d)$: fraction of available nodes at depth d just before the i'th step of execution

$F(1, d) = 1 \quad \forall d \geq 0$

$$F(i+1, d) = F(i, d) - \left(2^{d - l(u_i)}\right) 2^{-d} \quad d \geq l(u_i)$$

$$= F(i, d) - 2^{-l(u_i)} \quad d \geq l(u_i)$$

$$F(1, d) = 1 \quad > 0$$

$$F(2, d) = 1 - 2^{-l(u_1)} \quad > 0$$

$$F(3, d) = 1 - 2^{-l(u_1)} - 2^{-l(u_2)} \quad > 0$$

$$F(K, d) = 1 - \underbrace{\left[2^{-l(u_1)} + \dots + 2^{-l(u_{K-1})}\right]}_{< 1 \text{ because } \left(2^{-l(u_1)} \dots + 2^{-l(u_K)}\right) \leq 1} > 0$$

//