Haladó információ- és kódelmélet előadás

Baran Sándor

2018/19 tanév, 2. félév

Irodalom

- Györfi László, Győri Sándor, Vajda István: *Információ- és kódelmélet.* Typotex, 2010.
- Csiszár Imre, Fritz József: *Információelmélet.* Tankönyvkiadó, 1980.
- Cover, Thomas M. and Thomas, Joy A.: *Elements of Information Theory.* Wiley, 2006.
- Togneri, Roberto and de Silva, Christopher J. S.: *Fundamentals of Information Theory and Coding Design.* Chapman & Hall/CRC, 2006.
- Ash, Robert B.: *Information Theory.* Dover Publications, 1990.

Syllabus, eredmények, információk

`arato.inf.unideb.hu/baran.sandor/mischu.html`

Tartalom

Alapfogalmak

Forrásábécé: $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$, $n \geq 2$, véges halmaz. Elemeit (forrás)betűknek nevezzük. Felfoghatóak úgy, mint egy X diszkrét valószínűségi változó (forrás) lehetséges értékei.

$\mathcal{X}^*$: az $\mathcal{X}$ elemeiből álló véges sorozatok halmaza. $\mathcal{X}^*$ elemeit üzeneteknek, vagy közleményeknek nevezzük.

Kódábécé: $\mathcal{Y} = \{y_1, y_2, \ldots, y_s\}$, $s \geq 2$, véges halmaz. Elemeit kódjeleknek nevezzük.

$\mathcal{Y}^*$: az $\mathcal{Y}$ elemeiből álló véges sorozatok halmaza. $\mathcal{Y}^*$ elemei a kódszavak.

Kód: egy $f : \mathcal{X} \to \mathcal{Y}^*$ függvény. $s = 2$: bináris kód.
Időnként kódként hivatkoznak az f kód $\mathcal{K} = f(\mathcal{X})$ értékkészletére.

Ha az f kód értékkészlete különböző hosszúságú kódszavakból áll, akkor változó hosszúságú kódolásról beszélünk.

Betűnkénti kódolás: egy közlemény kódját az egyes forrásbetűk kódjainak egymás után írásával kapjuk.

Egyértelműen dekódolható kódok

Definíció. *Az $f : \mathcal{X} \to \mathcal{Y}^*$ kód egyértelműen dekódolható, ha tetszőleges $\boldsymbol{u} \in \mathcal{X}^*$, $\boldsymbol{v} \in \mathcal{X}^*$ esetén, ahol $\boldsymbol{u} = u_1 u_2 \dots u_k$, $\boldsymbol{v} = v_1 v_2 \dots v_\ell$ és $\boldsymbol{u} \neq \boldsymbol{v}$, teljesül, hogy $f(u_1)f(u_2)\dots f(u_k) \neq f(v_1)f(v_2)\dots f(v_\ell)$. Minden véges kódjelsorozat legfeljebb egy közlemény kódolásával állhat elő.*

Példák.

1. $\mathcal{X} = \{a, b, c\}$, $\mathcal{Y} = \{0, 1\}$, és $f(a) = 1$, $f(b) = 01$, $f(c) = 10110$.
Az f kódoló függvény invertálható, de a kód nem egyértelműen dekódolható. Pl., $f(c)f(a) = 101101 = f(a)f(b)f(a)f(b)$.

2. $\mathcal{X} = \{a, b, c\}$, $\mathcal{Y} = \{0, 1\}$, és $f(a) = 1$, $f(b) = 10$, $f(c) = 100$.
A kód egyértelműen dekódolható, mivel az 1 mindig egy új kódszó kezdetét jelzi.

3. $\mathcal{X} = \{a, b, c\}$, $\mathcal{Y} = \{0, 1\}$, és $f(a) = 1$, $f(b) = 00$, $f(c) = 01$.
A kód egyértelműen dekódolható.

Prefix kódok

Definíció. *Az f kód prefix, ha a lehetséges kódszavak mind különbözőek és egyik kódszó sem folytatása a másiknak.*

Megjegyzések.

1. Minden prefix kód egyértelműen dekódolható.
2. Állandó hosszúságú kód mindig prefix, ha a kódszavai különbözőek.

Példák.

1. $\mathcal{X}=\{a,b,c\}$, $\mathcal{Y}=\{0,1\}$, és $f(a)=1$, $f(b)=00$, $f(c)=01$.
A kód prefix.

2. $\mathcal{X}=\{a,b,c\}$, $\mathcal{Y}=\{0,1\}$, és $f(a)=1$, $f(b)=10$, $f(c)=100$.
A kód nem prefix, de egyértelműen dekódolható.

3. $\mathcal{X}=\{a,b,c,d,e,f,g\}$, $\mathcal{Y}=\{0,1,2\}$, valamint $f(a)=0$, $f(b)=10$, $f(c)=11$, $f(d)=20$, $f(e)=21$, $f(f)=220$, $f(g)=221$.
A kód prefix.

Kódfák

Minden prefix kód ábrázolható egy fagráffal, ahol az egyes kódszavaknak a gyökértől az egyes levelekig tartó töröttvonalak felelnek meg.

Bináris kód esetén pl. a 0-nak a bináris fa „felfelé" tartó ágai, az 1-nek pedig a „lefelé" mutatók felelnek meg.

Példa. Rajzoljuk fel a megfelelő fagráfokat.

1. $\mathcal{Y} = \{0, 1\}$, $\mathcal{K} = \{0, 100, 1010, 1011, 110, 111\}$.
2. $\mathcal{Y} = \{0, 1, 2\}$, $\mathcal{K} = \{0, 10, 11, 20, 21, 220, 221\}$.

Kódszóhosszak

$|f(x)|$: az $x \in \mathcal{X}$ betű $f(x)$ kódjának a kódszóhossza. $\mathcal{L}$ jelöli egy f kódhoz tartozó kódszóhosszak halmazát.

A kódszóhosszak nem lehetnek tetszőlegesek. Pl., nem létezik olyan 4 kódszóból álló egyértelműen dekódolható bináris kód, melynek a kódszóhosszai $\{1, 2, 2, 2\}$.

McMillan egyenlőtlenség

Tétel (McMillan). *Minden egyértelműen dekódolható $f : \mathcal{X} \to \mathcal{Y}^*$ kódra*

$$\sum_{i=1}^{n} s^{-|f(x_i)|} \leq 1,$$

ahol s a kódábécé elemszáma.

Indoklás. Csak prefix kódra. Legyen $M := max_{1 \leq i \leq n} \left|f(x_i)\right|$. Egészítsük ki a $\mathcal{K} = f(\mathcal{X})$ kódszavait M hosszúságúra minden lehetséges módon. Mivel a $\mathcal{K}$ kódfájában minden csúcsból s él indulhat ki, egy $\left|f(x_i)\right|$ hosszúságú kódszót $s^{M-|f(x_i)|}$ féleképpen egészíthetünk ki. A kiegészítésekkel kapott kódsorozatok száma

$$s^{M-|f(x_1)|} + s^{M-|f(x_2)|} + \cdots + s^{M-|f(x_n)|},$$

ami nem lehet több, mint az összes M hosszú kódsorozat s^M száma, azaz

$$s^{M-|f(x_1)|} + s^{M-|f(x_2)|} + \cdots + s^{M-|f(x_n)|} \leq s^M.$$

Az egyenlőtlenséget s^M-el oszva kapjuk az állítást. □

Példa. Tegyük fel, hogy létezik olyan 4 elemű egyértelműen dekódolható bináris kód, melynek kódszóhosszai $\{1, 2, 2, 2\}$. Erre a kódra $2^{-1} + 3 \cdot 2^{-2} > 1$, ami ellentmond a McMillan egyenlőtlenségnek.

Kraft egyenlőtlenség

Tétel (Kraft). *Ha az $L_1, L_2, \ldots, L_n$ pozitív egész számokra*

$$\sum_{i=1}^{n} s^{-L_i} \leq 1,$$

akkor létezik olyan f prefix kód, melyre

$$\left|f(x_i)\right| = L_i, \qquad i = 1, 2, \ldots, n.$$

Példa. Legyen $s = 2$ és $\mathcal{L} = \{1, 3, 3, 3, 4, 4\}$. Ekkor

$$2^{-1} + 3 \cdot 2^{-3} + 2 \cdot 2^{-4} = 1.$$

Egy lehetséges prefix kód:

$$\mathcal{K} = \{0, 100, 101, 110, 1110, 1111\}.$$

Megjegyzés. A McMillan és Kraft egyenlőtlenségekből következik, hogy minden egyértelműen dekódolható kódhoz létezik vele azonos kódhosszú prefix kód. Így elég, ha egy kódtól a speciálisabb prefix tulajdonságot követeljük meg.

Az információmennyiség mérőszáma I.

Hartley (1928): Az n elemű $\mathcal{X}$ halmaz egyes elemeinek azonosításához

$$I = \log_2 n$$

mennyiségű információra van szükség.

Heurisztika. Ha $n = 2^k$, akkor az $\mathcal{X}$ elemeinek reprezentálásához $k = \log_2 n$ hosszú bináris sorozatokat érdemes használni. Ha $\log_2 n \notin \mathbb{Z}$, akkor a szükséges bináris jegyek száma a $\log_2 n$ utáni első egész. Ha $\mathcal{X}$ elemeiből alkotható m hosszú sorozatokat kell binárisan reprezentálni (ezek száma n^m), akkor olyan k hosszra van szükség, melyre $2^{k-1} < n^m \leq 2^k$. Az $\mathcal{X}$ egy elemére eső bináris jegyek $K = k/m$ számára teljesül, hogy $\log_2 n < K \leq \log_2 n + 1/m$. A $\log_2 n$ alsó határ így tetszőlegesen megközelíthető.

A formula az információ mennyiségét a megadáshoz szükséges állandó hosszúságu bináris sorozatok hosszának alsó határaként definiálja. Az információmennyiség egysége: bit. Egy kételemű halmaz azonosításához 1 bit információ szükséges.

Probléma: Hartley nem veszi figyelembe, hogy az $\mathcal{X}$ elemei esetleg nem egyforma valószínűek.

Az információmennyiség mérőszáma II.

Shannon (1948): Egy $\mathrm{P}(A)$ valószínűségű A esemény bekövetkezése

$$I(A) = \log_2 \frac{1}{\mathrm{P}(A)} = -\log_2 \mathrm{P}(A)$$

mennyiségű információt szolgáltat.

Heurisztika. Követelmények az $I(A)$ információmennyiséggel kapcsolatban.

- Ha $\mathrm{P}(A) \leq \mathrm{P}(B)$, akkor $I(A) \geq I(B)$.
  Következmény: $I(A)$ csak a $\mathrm{P}(A)$ értékétől függ, azaz $I(A) = g\big(\mathrm{P}(A)\big)$.
- Független események együttes bekövetkezése esetén az információ összeadódik, azaz ha $\mathrm{P}(A \cdot B) = \mathrm{P}(A)\mathrm{P}(B)$, akkor $I(A \cdot B) = I(A) + I(B)$. Ez azt jelenti, hogy $g(p \cdot q) = g(p) + g(q), \;\; p, q \in]0, 1]$.
- Ha $\mathrm{P}(A) = 1/2$, akkor $I(A) := 1$, azaz $g(1/2) = 1$.

Tétel. *Ha* $g : [0, 1] \to \mathbb{R}$ *olyan függvény, melyre*

a) $g(p) \geq g(q)$, *ha* $0 < p \leq q \leq 1$;

b) $g(p \cdot q) = g(p) + g(q), \;\; p, q \in]0, 1]$;

c) $g(1/2) = 1$,

akkor

$$g(p) = \log_2 \frac{1}{p}, \qquad p \in]0, 1].$$

Az információmennyiség mérőszáma III.

Kapcsolat a két definíció között:

Ha az $\mathcal{X}$ minden eleme azonos $(1/n)$ valószínűségel következik be, akkor egy elem előfordulása éppen $\log_2 n$ információt szolgáltat.

Megjegyzés. A továbbiakban $a \geq 0$ és $b > 0$ esetén

$$0 \log_2 \frac{0}{a} = 0 \log_2 \frac{a}{0} = 0; \quad b \log_2 \frac{b}{0} = +\infty; \quad b \log_2 \frac{0}{b} = -\infty.$$

X: egy $\mathcal{X}$ elemeit értékként felvevő valószínűségi változó.

$p(x)$: az $x \in \mathcal{X}$ forrásbetű előfordulási valószínűsége, azaz

$$p(x) := \mathrm{P}(X = x), \qquad x \in \mathcal{X}.$$

Az X egy értékéhez tartozó átlagos információmennyiség:

$$\sum_{i=1}^{n} p(x_i) I(X = x_i) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) = \mathrm{E}\big(-\log_2 p(X)\big).$$

Entrópia, átlagos kódszóhossz

Definíció. *Az* $\mathcal{X} = \{x_1, x_2, \cdots, x_n\}$ *értékkészletű* X *valószínűségi változó entrópiája*

$$H(X) := \mathrm{E}\big(-\log_2 p(X)\big) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i).$$

Megjegyzés. Ugyanezzel a formulával definiáljuk a

$$\mathcal{P} := \big\{p(x_1), p(x_2), \ldots, p(x_n)\big\}$$

eloszlású $\mathcal{X}$ forrásábécé $H(\mathcal{X})$ entrópiáját, azaz

$$H(\mathcal{X}) := -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i).$$

Definíció. *Egy* $f\colon \mathcal{X} \to \mathcal{Y}^*$ *kód átlagos kódszóhossza*

$$\mathrm{E}(f) := \mathrm{E}\big|f(X)\big| = \sum_{i=1}^{n} p(x_i)\big|f(x_i)\big|.$$

Példák

1. $\mathcal{X}=\{a,b,c\}$, $\mathcal{Y}=\{0,1\}$, és $f(a)=1$, $f(b)=00$, $f(c)=01$.
Valószínűségek: $p(a)=0.6$, $p(b)=0.3$, $p(c)=0.1$.
Rövidebb felírás: $s=2$, $\mathcal{K}=\{1,00,01\}$, $\mathcal{P}=\{0.6,0.3,0.1\}$.

$$H(X) = -0.6\cdot\log_2 0.6 - 0.3\cdot\log_2 0.3 - 0.1\cdot\log_2 0.1 \approx 1.295;$$
$$\mathrm{E}(f) = 0.6\cdot 1 + 0.3\cdot 2 + 0.1\cdot 2 = 1.4$$

2. $s=2$, $\mathcal{L}=\{1,3,3,3,4,4\}$, $\mathcal{P}=\left\{\frac{1}{2},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{16},\frac{1}{16}\right\}$.

$$H(X) = -\frac{1}{2}\cdot\log_2\frac{1}{2} - 3\cdot\frac{1}{8}\cdot\log_2\frac{1}{8} - 2\cdot\frac{1}{16}\cdot\log_2\frac{1}{16} = 2.125;$$
$$\mathrm{E}(f) = \frac{1}{2}\cdot 1 + 3\cdot\frac{1}{8}\cdot 3 + 2\cdot\frac{1}{16}\cdot 4 = 2.125.$$

Cél: az átlagos kódszóhossz alsó határának meghatározása, mivel annál jobb egy kód, minél kisebb az átlagos kódszóhossza.

Keressük azt az f kódot, mely minimalizálja az

$$\mathrm{E}(f)=\sum_{i=1}^{n} p(x_i)\left|f(x_i)\right| \text{ függvényt a } \sum_{i=1}^{n} s^{-|f(x_i)|}\le 1 \text{ feltétel mellett.}$$

Diszkrét alaplemma

Lemma. *Ha a* $p_i \geq 0,\ q_i > 0,\ i = 1, 2, \ldots, n,$ *valós számokra*

$$\sum_{i=1}^{n} p_i = 1 \quad \text{és} \quad \sum_{i=1}^{n} q_i = 1,$$

azaz $\{p_1, p_2, \ldots, p_n\}$ *és* $\{q_1, q_2, \ldots, q_n\}$ *diszkrét valószínűségi eloszlást alkot, akkor*

$$-\sum_{i=1}^{n} p_i \log_2 p_i \leq -\sum_{i=1}^{n} p_i \log_2 q_i$$

és egyenlőség pontosan akkor áll fenn, ha $p_i = q_i,\ i = 1, 2, \ldots, n.$

Megjegyzés. A lemma állítása azzal ekvivalens, hogy

$$\sum_{i=1}^{n} p_i \log_2 \frac{q_i}{p_i} \leq 0,$$

és az állítás tetszőleges alapú logaritmusra teljesül.

Az átlagos kódszóhossz határai

Tétel (Shannon). *Tetszőleges egyértelműen dekódolható* $f\colon \mathcal{X} \to \mathcal{Y}^*$ *kódra*

$$\mathsf{E}(f) = \sum_{i=1}^{n} p(x_i)\big|f(x_i)\big| \geq -\sum_{i=1}^{n} p(x_i)\log_s p(x_i) = \frac{H(\mathcal{X})}{\log_2 s},$$

egyenlőség pedig pontosan akkor áll fenn, ha $p(x_i) = s^{-|f(x_i)|}$, $i = 1, 2, \ldots, n$.

Ha $p(x_i) = s^{-L_i}$, *ahol* $L_i \in \mathbb{N}$, *akkor létezik olyan* f *prefix kód, melyre* $\big|f(x_i)\big| = L_i,\ i = 1, 2, \ldots, n$, *és*

$$\mathsf{E}(f) = \frac{H(\mathcal{X})}{\log_2 s}.$$

Tetszőleges eloszású $\mathcal{X}$ *forrásábécé esetén létezik olyan* $f\colon \mathcal{X} \to \mathcal{Y}^*$ *prefix kód, melyre*

$$\mathsf{E}(f) \leq \frac{H(\mathcal{X})}{\log_2 s} + 1.$$

Indoklás

A tétel első állítása a következővel ekvivalens:

$$H(\mathcal{X}) \leq -\sum_{i=1}^{n} p(x_i) \log_2 s^{-|f(x_i)|}.$$

Alkalmazzuk a diszkrét alaplemmát a

$$p_i := p(x_i), \qquad q_i := \frac{s^{-|f(x_i)|}}{\sum_{j=1}^{n} s^{-|f(x_j)|}}, \quad i = 1, 2, \ldots, n,$$

eloszlásokra.

$$\begin{aligned} H(\mathcal{X}) &= -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \leq -\sum_{i=1}^{n} p(x_i) \log_2 \frac{s^{-|f(x_i)|}}{\sum_{j=1}^{n} s^{-|f(x_j)|}} \\ &= -\sum_{i=1}^{n} p(x_i) \log_2 s^{-|f(x_i)|} + \log_2 \left(\sum_{j=1}^{n} s^{-|f(x_j)|} \right) \quad \text{(McMillan)} \\ &\leq -\sum_{i=1}^{n} p(x_i) \log_2 s^{-|f(x_i)|} + \log_2 1 = -\sum_{i=1}^{n} p(x_i) \log_2 s^{-|f(x_i)|}. \end{aligned}$$

Ha $p(x_i) = s^{-L_i}$, akkor $\sum_{i=1}^{n} s^{-L_i} = 1$. Kraft egyenlőtlenség miatt létezik olyan f prefix kód, melynek kódszóhosszai $L_1, L_2, \ldots, L_n$ és erre a kódra $\mathrm{E}(f) = \frac{H(\mathcal{X})}{\log_2 s}$.

Indoklás

Tetszőleges $\mathcal{X}$ forrásábécé esetén legyen

$$L_i := \lceil -\log_s p(x_i) \rceil, \qquad \text{ahol} \qquad \lceil a \rceil := \min\{n \in \mathbb{Z},\ n \geq a\}.$$

Ekkor

$$-\log_s p(x_i) \leq L_i < -\log_s p(x_i) + 1,$$

tehát $s^{-L_i} \leq p(x_i)$, azaz $\sum_{i=1}^{n} s^{-L_i} \leq \sum_{i=1}^{n} p(x_i) = 1$.

Kraft egyenlőtlenség miatt létezik olyan f prefix kód, melynek kódszóhosszai $L_1, L_2, \ldots, L_n$, és erre a kódra

$$\mathrm{E}(f) = \sum_{i=1}^{n} p(x_i) L_i \leq \sum_{i=1}^{n} p(x_i)\big(-\log_s p(x_i) + 1\big) = 1 + \frac{H(\mathcal{X})}{\log_2 s}.$$

□

Blokkonkénti kódolás

A forrásüzeneteket m hosszúságú blokkokra vágjuk és ezeket a blokkokat kódoljuk.

Formális definíció: egy $f\colon \mathcal{X}^m \to \mathcal{Y}^*$ leképezés.

Olyan, mintha egy új, $\widehat{\mathcal{X}} := \mathcal{X}^m$ forrásábécét kódolnánk.

Forrás blokk: $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$ véletlen vektor. Eloszlása:

$$p(\boldsymbol{x}) = p(x_1, x_2, \ldots, x_m) = \mathrm{P}(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m).$$

Entrópia:

$$H(\boldsymbol{X}) = -\sum_{\boldsymbol{x}\in\mathcal{X}^m} p(\boldsymbol{x}) \log_2 p(\boldsymbol{x}).$$

Ha $X_1, X_2, \ldots, X_m$ független, akkor $H(\boldsymbol{X}) = \sum_{i=1}^{m} H(X_i)$.

Ha $X_1, X_2, \ldots, X_m$ még azonos eloszlású is, akkor $H(\boldsymbol{X}) = mH(X_1)$.

Betűnkénti átlagos kódhossz

Egy m-dimenziós $\boldsymbol{X}$ forrás $f\colon \mathcal{X}^m \to \mathcal{Y}^*$ kódjának betűnkénti átlagos kódhossza

$$\frac{1}{m}\mathsf{E}\big|f(\boldsymbol{X})\big| = \frac{1}{m}\sum_{\boldsymbol{x}\in\mathcal{X}^m} p(\boldsymbol{x})\big|f(\boldsymbol{x})\big|.$$

Shannon tétel:

$$\mathsf{E}\big|f(\boldsymbol{X})\big| \geq \frac{H(\boldsymbol{X})}{\log_2 s}.$$

Következmény. *Ha $X_1, \ldots, X_m$ független, az X-szel azonos eloszlású valószínűségi változók, akkor létezik olyan $f\colon \mathcal{X}^m \to \mathcal{Y}^*$ prefix kód, hogy*

$$\frac{1}{m}\mathsf{E}\big|f(\boldsymbol{X})\big| < \frac{H(X)}{\log_2 s} + \frac{1}{m}.$$

Optimális kódok

Bináris eset: $s = 2$.

Tétel. *Ha az $f : \mathcal{X} \to \{0,1\}^*$ prefix kód optimális, és $\mathcal{X}$ elemei úgy vannak indexelve, hogy $p(x_1) \geq p(x_2) \geq \cdots \geq p(x_n) > 0$, akkor feltehető, hogy f-re teljesül az alábbi három tulajdonság.*

a) $\big|f(x_1)\big| \leq \big|f(x_2)\big| \leq \cdots \leq \big|f(x_n)\big|$, *azaz nagyobb valószínűségekhez rövidebb kódszavak tartoznak.*

b) $\big|f(x_{n-1})\big| = \big|f(x_n)\big|$, *vagyis a két legkisebb valószínűségű forrásbetűhöz tartozó kódszó azonos hosszúságú.*

c) *Az $f(x_{n-1})$ és az $f(x_n)$ kódszavak csak az utolsó bitjükben különböznek.*

Heurisztika. a) Ha $p(x_k) > p(x_j)$ és $|f(x_k)| > |f(x_j)|$, akkor x_j és x_k kódsza- vát felcserélve egy az eredetinél rövidebb átlagos kódszóhosszú kódot kapunk. Az eredeti így nem lehet optimális.

b) Ha $|f(x_{n-1})| < |f(x_n)|$, akkor $f(x_n)$ utolsó bitjét levágva egy az eredetinél rövidebb átlagos kódszóhosszú, ugyancsak prefix kódot kapunk. Az eredeti így nem lehet optimális.

c) Ha létezik olyan $f(x_i)$ kódszó, hogy $f(x_i)$ és $f(x_n)$ csak az utolsó bitben különböznek, akkor a korábbiak alapján $|f(x_i)| = |f(x_{n-1})| = |f(x_n)|$. Ha $i \neq n-1$, akkor cseréljük fel x_i és x_{n-1} kódját.

Bináris Huffman-kód

Tétel. *Tegyük fel, hogy az* $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ *forrásábécé elemei úgy vannak indexelve, hogy* $p(x_1) \geq p(x_2) \geq \cdots \geq p(x_n) > 0$, *és tekintsük azt az* $\widetilde{\mathcal{X}} = \{x_1, x_2, \ldots, x_{n-2}, \widetilde{x}_{n-1}\}$ *forrásábécét, ahol az* $\widetilde{x}_{n-1}$ *szimbólumot az* x_{n-1} *és* x_n *forrásbetűk összevonásával kapjuk és* $p(\widetilde{x}_{n-1}) = p(x_{n-1}) + p(x_n)$.

Ha az új $\{p(x_1), p(x_2), \ldots, p(x_{n-2}), p(x_{n-1}) + p(x_n)\}$ *eloszláshoz ismerünk egy* g *optimális bináris prefix kódot, akkor az eredeti* $\{p(x_1), p(x_2), \ldots, p(x_n)\}$ *eloszlás egy optimális* f *prefix kódját kapjuk, ha a* $g(\widetilde{x}_{n-1})$ *kódszót kiegészítjük egy nullával és egy egyessel, a többi kódszót pedig változatlanul hagyjuk.*

Példa. Adjuk meg az alábbi valószínűségi eloszlásokhoz tartozó bináris Huffman-kódokat. Vizsgáljuk meg az átlagos kódszóhossznak az elméleti alsó korláttól való eltérését.

1. $\mathcal{P}_1 = \{0.68, 0.17, 0.04, 0.04, 0.03, 0.03, 0.01\}$.
2. $\mathcal{P}_2 = \{0.49, 0.14, 0.14, 0.07, 0.07, 0.04, 0.02, 0.02, 0.01\}$.
3. $\mathcal{P}_3 = \{0.15, 0.15, 0.14, 0.14, 0.14, 0.14, 0.14\}$.

Bináris Shannon-Fano-kód

Tegyük fel, hogy az $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ forrásábécé elemei úgy vannak indexelve, hogy $p(x_1) \geq p(x_2) \geq \cdots \geq p(x_n) > 0$. Legyen

$$L_i := \lceil -\log_s p(x_i) \rceil, \qquad \text{ahol} \qquad \lceil a \rceil := \min\{n \in \mathbb{Z},\ n \geq a\},$$

és

$$w_1 := 0, \qquad w_i := \sum_{\ell=1}^{i-1} p(x_\ell), \quad i = 2, 3, \ldots, n.$$

Az x_i forrásbetű $f(x_i)$ kódja legyen $\lfloor 2^{L_i} w_i \rfloor$ bináris alakja L_i hosszon ábrázolva, ahol $\lfloor a \rfloor := \max\{n \in \mathbb{Z},\ n \leq a\}$.

Tétel. *A bináris Shannon-Fano kód prefix és az átlagos kódszóhosszára teljesül* $\mathrm{E}(f) \leq H(\mathcal{X}) + 1$.

Példa. Adjuk meg az alábbi valószínűségi eloszlásokhoz tartozó bináris Shannon-Fano-kódokat.

1. $\mathcal{P}_1 = \{0.68, 0.17, 0.04, 0.04, 0.03, 0.03, 0.01\}$.
2. $\mathcal{P}_2 = \{0.49, 0.14, 0.14, 0.07, 0.07, 0.04, 0.02, 0.02, 0.01\}$.
3. $\mathcal{P}_3 = \{0.15, 0.15, 0.14, 0.14, 0.14, 0.14, 0.14\}$.

Az entrópia tulajdonságai

$X \in \mathcal{X}$, $Y \in \mathcal{Y}$: diszkrét valószínűségi változók.

Tétel.

a) *Ha az X valószínűségi változó n különböző értéket vehet fel pozitív valószínűséggel, akkor*

$$0 \leq H(X) \leq \log_2 n.$$

A bal oldalon egyenlőség pontosan akkor áll fenn, ha X egy valószínűséggel konstans, a jobb oldalon pedig pontosan akkor, ha X eloszlása egyenletes, azaz $p(x_i) = \frac{1}{n}$, $i = 1, 2, \ldots, n$.

b) *Az X és Y diszkrét valószínűségi változókra*

$$H(X, Y) \leq H(X) + H(Y),$$

az egyenlőség pedig pontosan akkor teljesül, ha X és Y független.

c) *Az X tetszőleges $g(X)$ függvényére*

$$H\big(g(X)\big) \leq H(X),$$

az egyenlőség szükséges és elégséges feltétele pedig a g invertálhatósága.

Feltételes entrópia

$$p(x) := \mathrm{P}(X=x); \quad p(y) := \mathrm{P}(Y=y); \quad p(x,y) := \mathrm{P}(X=x,\, Y=y);$$

$$p(x|y) := \mathrm{P}(X=x|\, Y=y) = \frac{p(x,y)}{p(y)};$$

$$p(y|x) := \mathrm{P}(Y=y|X=x) = \frac{p(y,x)}{p(x)}.$$

Definíció. *Az X-nek az $Y=y$ feltétellel vett feltételes entrópiája*

$$H(X|\, Y=y) := -\sum_{x\in\mathcal{X}} p(x|y)\log_2 p(x|y).$$

Az X-nek az Y feltétellel vett feltételes entrópiája

$$H(X|\, Y) := \sum_{y\in\mathcal{Y}} p(y) H(X|\, Y=y) = -\sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} p(x,y)\log_2 p(x|y).$$

A feltételes entrópia tulajdonságai

Tétel. *Legyenek X, Y, és Z véges értékkészletű valószínűségi változók. Ekkor*

a)

$$H(X,Y) = H(Y) + H(X|Y) = H(X) + H(Y|X).$$

b)

$$0 \leq H(X|Y) \leq H(X).$$

A bal oldalon egyenlőség pontosan akkor áll fenn, ha X egy valószínűséggel az Y függvénye, a jobb oldalon pedig pontosan akkor, ha X és Y független.

c)

$$H(X|Z,Y) \leq H(X|Z),$$

egyenlőség pedig akkor és csak akkor áll fenn, ha

$$p(x|z,y) = p(x|z)$$

minden olyan x, y, z-re, amelyre $p(x,y,z) > 0$.

A feltételes entrópia tulajdonságai

d) *Az Y valószínűségi változó minden f függvényére*

$$H(X|Y) \leq H\big(X\big|f(Y)\big),$$

egyenlőség pedig pontosan akkor áll fenn, ha minden rögzített z-re

$$p(x|y) = \mathrm{P}\big(X = x\big|f(Y) = z\big)$$

minden olyan x-re és y-ra, amelyre $f(y) = z$ és $p(y) > 0$.

e) *Az $X_1, X_2, \ldots, X_n$ valószínűségi változók együttes entrópiájára*

$$H(X_1, X_2, \ldots, X_n) = H(X_1) + H(X_2|X_1) + H(X_3|X_2, X_1) + \cdots$$

$$\cdots + H(X_n|X_{n-1}, \ldots, X_1).$$

Infomációforrások

$\mathbb{X}$: információforrás, az $X_1, X_2, \ldots$ valószínűségi változók végtelen sorozata. A forrás az i-edik időpontban az X_i értéket veszi fel.

Mindegyik valószínűségi változó ugyanazon $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ forrásábécéből veszi fel az értékeit.

Az $\mathbb{X}$ forrás emlékezetnélküli, ha az $X_1, X_2, \ldots$ valószínűségi változók függetlenek.

Az $\mathbb{X}$ forrás stacionárius, ha az $X_1, X_2, \ldots$ sorozat stacionárius, azaz minden pozitív n-re és k-ra az $X_1, X_2 \ldots, X_n$ véletlen vektor és az $X_{k+1}, X_{k+2}, \ldots, X_{k+n}$ együttes eloszlása megegyezik.

Az $\mathbb{X}$ forrás ergodikus, ha tetszőleges $f(x_1, \ldots, x_k)$ függvényre

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(X_i, \ldots, X_{i+k-1}) = \mathrm{E} f(X_1, \ldots, X_k) \quad \text{egy valószínűséggel,}$$

ha a határérték létezik.

Adott egy f kódoló $\mathcal{Y} = \{y_1, y_2, \ldots, y_s\}$ kódábécével. Egyértelműen dekódolható. Blokk kódolás $k \geq 1$ blokkhosszal.

Cél: az egy forrásbetűre jutó L átlagos kódszóhossz minimalizálása.

Infomációforrások változó szóhosszú kódolása

$\mathbb{X}$ emlékezetnélküli és stacionárius, betűnkénti kódolás. Egy k hosszúságú üzenet kódjára

$$L = \frac{1}{k}\mathsf{E}\Big(\big|f(X_1)\big| + \cdots + \big|f(X_k)\big|\Big) = \mathsf{E}\big|f(X_1)\big|.$$

Shannon tétel: $$\mathsf{E}\big|f(X_1)\big| \geq \frac{H(X_1)}{\log_2 s}.$$

Létezik f prefix kód: $$\mathsf{E}\big|f(X_1)\big| < \frac{H(X_1)}{\log_2 s} + 1.$$

Blokkonkénti kódolás $f\colon \mathcal{X}^k \to \mathcal{Y}^*$ kódolóval.

$$L = \frac{1}{k}\mathsf{E}\Big(\big|f(X_1,\ldots,X_k)\big|\Big) \geq \frac{1}{k}\frac{H(X_1,\ldots,X_k)}{\log_2 s} \underset{\text{függetlenség}}{=} \frac{H(X_1)}{\log_2 s}.$$

Minden k-re létezik olyan L betűnkénti átlagos kódszóhosszú $f\colon \mathcal{X}^k \to \mathcal{Y}^*$ prefix kód, melyre

$$L < \frac{H(X_1)}{\log_2 s} + \frac{1}{k}.$$

Általánosabb infomációforrások kódolása

Definíció. *Az* $\mathbb{X} = X_1, X_2, \ldots$ *forrás forrásentrópiája a*

$$H(\mathbb{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)$$

mennyiség, amennyiben a határérték létezik.

Tétel. *Ha az* $\mathbb{X} = X_1, X_2, \ldots$ *forrás stacionárius, akkor létezik az entrópiája, és*

$$H(\mathbb{X}) = \lim_{n\to\infty} H(X_n|X_1, X_2, \ldots, X_{n-1}).$$

Tétel. *Ha az* $\mathbb{X} = X_1, X_2, \ldots$ *stacionárius forrást blokkonként kódoljuk az* $f : \mathcal{X}^k \to \mathcal{Y}^*$ *egyértelműen dekódolható kóddal, akkor a kód* L *betűnkénti átlagos kódszóhosszára mindig teljesül az*

$$L \geq \frac{H(\mathbb{X})}{\log_2 s}$$

egyenlőtlenség. A k *blokhosszt elég nagyra választva létezik olyan* f *kód, amelynek* L *betűnkénti átlagos kódszóhossza tetszőlegesen megközelíti a fenti alsó korlátot.*

Univerzális forráskódolás

Adatátvitel költségei az eddig vizsgált (blokk)kódoknál:

- Állandó költség: pl. a forrásszimbólumok gyakoriságai.
- Változó költség: az üzenet kódszavai.

Elméletileg végtelen hosszú forrásokat vizsgálunk. Az állandó költség ekkor fajlagosan nullához tart.

A gyakorlatban a források véges hosszúságuak. Az állandó költség esetleg magasabb lehet, mint az üzenet kódszavainak összhossza.

Adaptív kód: az aktuális forrásszimbólumot az azt megelőző szimbólumok alapján kódoljuk.

Példák:

- Adaptív Hufffman-kód;
- Lempel-Ziv algoritmusok (LZ77, LZ78, LZW).

Az LZ77 algoritmus

Abraham Lempel és Jakov Ziv (1977)

A forrászimbólumokon egy h_a hosszúságú csúszóablakot mozgatunk.

A csúszóablak részei:

- keresőpuffer: a legutóbb kódolt h_k darab forrásszimbólumot tartalmazza;
- előretekintő puffer: a következő h_e darab kódolandó szimbólumot tartalmazza.

Példa. Forrásszöveg

$$\ldots cabracadabrarrarrad \ldots$$

Csúszóablak: $h_a := 13,\ h_k := 7,\ h_e := 6.$

c a b r a c a	d a b r a r	r a r r a d

Az LZ77 algoritmus

Kódolás:

1. Egy hátrafelé mutató mutatóval a kódoló megkeresi a keresőpufferben az előretekintő puffer első szimbólumával megegyező szimbólumokat.
2. Megvizsgálja, a kapott pozíciókkal kezdődően a keresőpufferben lévő szimbólumok milyen hosszan egyeznek meg az előretekintő puffer szimbólumaival.
3. A talált szimbólumok közül kiválasztja azt, ahol a leghosszabb az egyezés.
4. Átküldi a $\langle t, h, c \rangle$ hármast.

   t: a keresőpufferben megtalált szimbólum távolsága az előretekintő puffertől. Ha nincs találat a keresőpufferben, $t = 0$.

   h: a kereső- és az előretekintő puffer egyező szimbólumainak legnagyobb hosszúsága. Ha nincs találat a keresőpufferben, $h = 0$.

   c: az első, az előretekintő pufferben levő nem egyező karakter kódszava.

Példa

Csúszóablak: $h_a := 13,\ h_k := 7,\ h_e := 6.$

c a b r a c a	**d** a b r a r	r a r r a d

d nincs a keresőpufferben. Átküldendő: $\langle 0, 0, f(d)\rangle$.

c	a b r a c a d	**a** b r a r r	a r r a d

a a keresőpufferrben: $t = 2,\ h = 1,\ t = 4,\ h = 1$ és $t = 7,\ h = 4$.
Leghosszabb egyezés: $t = 7,\ h = 4$. Átküldendő: $\langle 7, 4, f(r)\rangle$

c a b r a c	a d a b r a r	**r** a r r a d

r a keresőpufferrben: $t = 1,\ h = 1$ és $t = 3,\ h = 5$.
Leghosszabb egyezés: $t = 3,\ h = 5$. Átküldendő: $\langle 3, 5, f(d)\rangle$

Jellemzők

A $\langle t, h, c \rangle$ kódolásához állandó kódhosszú kód esetén

$$\lceil \log_2 h_k \rceil + \lceil \log_2 h_e \rceil + \lceil \log_2 n \rceil$$

bit szükséges, ahol n a forrásábécé mérete.

Az eljárás hatékonysága aszimptotikusan ($h_k, h_e \to \infty$) tart az optimális algoritmuséhoz, amihez viszont kell a forrás eloszlása is.

Stacionárius és ergodikus forrás esetén az átlagos kódszóhossz $h_k, h_e \to \infty$ esetén konvergál $\frac{H(\mathbb{X})}{\log_2 s}$-hez.

Hatékonyságot növelő módosítások, pl.

- Változó hosszúságú kódok $\langle t, h, c \rangle$ tömörítéséhez. Pl. adaptív Huffman kódolás.
- Duál formátum: $\langle t, h \rangle$ vagy $\langle c \rangle$ továbbítódik. Jelzőbittel azonosítja a formátumokat.
  LZSS – Lempel-Ziv-Storer-Szymanski
- Változtatható méretű pufferek.

Alkalmazások: pkzip, arj

Az LZ78 algoritmus

A kódoló és a dekódoló szótárt épít az előzőleg előfordult sorozatokból.

1. A kódoló megkeresi a forrásszimbólumok aktuális pozíciójától kezdődő leghosszabb egyezést a szótárban.
2. Átküldi az $\langle i, c \rangle$ párt.

   i: az egyező karaktersorozat szótárbeli indexe;

   c: az első nem egyező karakter kódja.

   Ha nem talál egyezést a szótárban, a $\langle 0, c \rangle$ párt küldi át.
3. A szótárba felveszi az i indexű karaktersorozat és a c konkatenációjával kapott stringet. Van eof szimbólum is.

Stacionárius és ergodikus forrás esetén az egy betűre jutó átlagos kódszóhossz konvergál $\frac{H(\mathbb{X})}{\log_2 s}$-hez.

Probléma: a szótár folyamatosan, korlát nélkül növekszik.

Megoldás: egy idő után fix szótár használata, vagy a ritkán használt, illetve felesleges bejegyzések eltávolítása.

Példa

Kódolandó szöveg:

$$dabbacdabbacdabbacdabbacdeecdeecdee$$

a kódoló kimenete	szótár index	szótár bejegyzés	a kódoló kimenete	szótár index	szótár bejegyzés
$\langle 0, f(d)\rangle$	1	d	$\langle 4, f(c)\rangle$	10	bac
$\langle 0, f(a)\rangle$	2	a	$\langle 9, f(b)\rangle$	11	$dabb$
$\langle 0, f(b)\rangle$	3	b	$\langle 8, f(d)\rangle$	12	acd
$\langle 3, f(a)\rangle$	4	ba	$\langle 0, f(e)\rangle$	13	e
$\langle 0, f(c)\rangle$	5	c	$\langle 13, f(c)\rangle$	14	ec
$\langle 1, f(a)\rangle$	6	da	$\langle 1, f(e)\rangle$	15	de
$\langle 3, f(b)\rangle$	7	bb	$\langle 14, f(d)\rangle$	16	ecd
$\langle 2, f(c)\rangle$	8	ac	$\langle 13, f(e)\rangle$	17	ee
$\langle 6, f(b)\rangle$	9	dab			

Az LZW algoritmus

Az LZ78 továbbfejlesztése. Terry Welch (1984)

Az LZ78 $\langle i, c \rangle$ párjából csak az i indexet kell átküldeni. A szótárban szerepelnie kell a teljes forrásábécének.

1. A kódoló az aktuális pozíciótól addig olvassa be a forrásszimbólumokat egy p pufferbe, míg a beolvasott sorozat szerepel a szótárban.
   Legyen c az első karakter, amelyre pc nincs a szótárban.
2. Átküldi a p sorozat indexét.
3. Felveszi a szótárba a pc sorozatot és a c karaktertől folytatja az eljárást.

Alkalmazás: a Unix rendszer compress parancsa, a GIF formátum.

Adaptív szótárméret. Compress esetén 512 bejegyzés, ha megtelik 1024, stb. A felső határ beállítható, maximum 2^{16} bejegyzésig.

Példa

Kódolandó szöveg:

dabbacdabbacdabbacdabbacdeecdeecdee

index	bejegyzés	output	index	bejegyzés	output
1	*a*		14	*acd*	10
2	*b*		15	*dabb*	12
3	*c*		16	*bac*	9
4	*d*		17	*cda*	11
5	*e*		18	*abb*	7
6	*da*	4	19	*bacd*	16
7	*ab*	1	20	*de*	4
8	*bb*	2	21	*ee*	5
9	*ba*	2	22	*ec*	5
10	*ac*	1	23	*cde*	11
11	*cd*	3	24	*eec*	21
12	*dab*	6	25	*cdee*	23
13	*bba*	8			5

Feladat

Kódoljuk az alábbi szövegeket

a) *abbabbabbbaababa*.

b) „bed spreaders spread spreads on beds”. A space és az eof külön karakter.

c) „az ipafai papnak fapipaja van a papi fapipa” A space és az eof külön karakter.

- LZ77 algoritmussal $h_k = 7,\ h_e = 6$ paraméterekkel;
- LZ78 algoritmussal;
- LZW algoritmussal.

Problémafelvetés

Gyakorlati probémák esetén gyakran kell engedni abból, hogy a kódolt üzenet egyértelműen visszaállítható legyen.

Elvárás: a dekódolt üzenet az eredetit „hűen”, de nem feltétlenül pontosan adja vissza.

Példák.

1. Beszéd digitális tárolása/továbbítása: folytonos jelből mintavételezéssel véges értékkészletű jelet kapunk.
2. MPEG-1 vagy MPEG-2 Audio Layer III (MP3) tömörítés (1995): kihagyja a zene/hang azon részeit, amik kívül esnek az átlagember által hallható tartományon. 128 kbit/s mellett az mp3 file mérete 1/11-e a wav méretének.

Blokkból-blokkba kódokat vizsgálunk: a forrásüzenet állandó hosszúságú blokkjait állandó hosszúságú kódokkal kódolja.

Az üzenetek és kódjaik között van egy hűségmérték. Azt méri, az adott kódszó milyen mértékben reprezentálja az eredeti üzenetet.

Állandó hosszúságú blokk-kódok

Változó szóhosszú forráskódolás problémája: ha egy kódszó meghibásodik, gond lehet az utána következő összes további kódszó dekódolásával. Fix hosszú kódoknál nincs ilyen probléma.

$\mathcal{X}$, $\mathcal{Y}$: n-elemű forrás-, illetve s-elemű kódábéce.

Az $f\colon \mathcal{X}^k \to \mathcal{Y}^m$ kód egyértelműen dekódolható, ha

$$n^k \leq s^m.$$

Betűnkénti átlagos kódszóhossz:

$$L = \frac{m}{k} \geq \frac{\log_2 n}{\log_2 s}.$$

A feltétel elégséges is az egyértelműen dekódolható $f\colon \mathcal{X}^k \to \mathcal{Y}^m$ kód létezésére.

A forrás entrópiája csak akkor $\log_2 n$, ha emlékezet nélküli, stacionárius és egyenletes eloszlású. Állandó hosszú kódszavakkal nem tudjuk tetszőlegesen közelíteni a forrásentrópiát, bármekkora is a k blokkhossz.

Forráskódolás előírt hibavalószínűséggel

$\mathbb{X} = X_1, X_2, \ldots$: staconárius forrás.

k hosszúságú üzenetek egy $B \in \mathcal{X}^k$ halmazának valószínűsége:

$$\mathrm{P}(B) := \mathrm{P}\big((X_1, X_2, \ldots, X_k) \in B\big) = \sum_{\boldsymbol{x} \in B} p(\boldsymbol{x}),$$

ahol

$$p(\boldsymbol{x}) := \mathrm{P}(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k), \quad \boldsymbol{x} = (x_1, x_2, \ldots, x_k) \in \mathcal{X}^k.$$

Stacionaritás:

$$\mathrm{P}(B) = \mathrm{P}\big((X_{n+1}, X_{n+2}, \ldots, X_{n+k}) \in B\big), \quad n = 1, 2, \ldots.$$

Definíció. *Az $\mathbb{X}$ stacionárius forrás $f : \mathcal{X}^k \to \mathcal{Y}^m$ kódját akkor nevezzük ε-hibával dekódolhatónk $(0 < \varepsilon < 1)$, ha létezik olyan $f' : \mathcal{Y}^m \to \mathcal{X}^k$ dekódoló függvény, hogy a hibás dekódolás valószínűsége legfeljebb ε, azaz*

$$\mathrm{P}\Big(f'\big(f(X_1, X_2, \ldots, X_k)\big) \neq (X_1, X_2, \ldots, X_k)\Big) \leq \varepsilon.$$

Az ε-hibával való dekódolhatóság feltétele

Megjegyzés. Az $f\colon \mathcal{X}^k \to \mathcal{Y}^m$ pontosan akkor dekódolható ε-hibával, ha f a $\mathcal{X}^k$ egy $1-\varepsilon$-nál nagyobb valószínűségű B részhalmazát invertálhatóan képezi le. Így akkor és csak akkor létezik $f\colon \mathcal{X}^k \to \mathcal{Y}^m$ ε-hibával dekódolható kód, ha létezik olyan $B \subset \mathcal{X}^k$, melyre $\mathrm{P}(B) > 1-\varepsilon$ és $|B| \le s^m$ ($|B|$: a B halmaz számossága).

Keresendő egy olyan minimális számosságú $B \subset \mathcal{X}^k$ üzenethalmaz, amelyre $\mathrm{P}(B) > 1-\varepsilon$. Feltétel az m kódszóhosszra: $s^{m-1} < |B| \le s^m$.

B üzeneteit kölcsönösen egyértelműen kódolhatjuk m hosszú kódszavakkal, a többi üzenetet meg akárhogyan. A kapott ε-hibával dekódolható $f\colon \mathcal{X}^k \to \mathcal{Y}^m$ kód optimális, azaz, ha $g\colon \mathcal{X}^k \to \mathcal{Y}^{m'}$ ε-hibával dekódolható, akkor $m \le m'$.

Indexeljük $\mathcal{X}^k$ üzeneteit csökkenő valószínűségek szerint:

$$p(\boldsymbol{x}_1) \ge p(\boldsymbol{x}_2) \ge \ldots \ge p(\boldsymbol{x}_\ell), \qquad \ell = n^k.$$

$N(k,\varepsilon)$: az az index, melyre

$$\sum_{i=1}^{N(k,\varepsilon)} p(\boldsymbol{x}_i) > 1-\varepsilon, \qquad \sum_{i=1}^{N(k,\varepsilon)-1} p(\boldsymbol{x}_i) \le 1-\varepsilon.$$

A keresett minimális elemszámú üzenethalmaz: $B_{k,\varepsilon} := \bigcup_{i=1}^{N(k,\varepsilon)} \{\boldsymbol{x}_i\}$.

Információstabilis források

$f\colon \mathcal{X}^k \to \mathcal{Y}^m$: ε-hibával dekódolható kód.

$N(k,\varepsilon)$: az egyértelműen dekódolható üzenetek száma.

$$N(k,\varepsilon) \le s^m.$$

Betűnkénti átlagos kódszóhossz:

$$L = \frac{m}{k} \ge \frac{\frac{1}{k}\log_2 N(k,\varepsilon)}{\log_2 s}.$$

Keresendő: $\lim_{k\to\infty} \frac{1}{k}\log_2 N(k,\varepsilon)$.

Definíció. *Az $\mathbb{X} = X_1, X_2, \ldots$ stacionárius forrást információstabilisnak nevezzük, ha minden $\delta > 0$ esetén*

$$\lim_{k\to\infty} \mathrm{P}\left(\left| -\frac{1}{k}\log_2 p(X_1, X_2, \ldots, X_k) - H(\mathbb{X}) \right| > \delta \right) = 0$$

azaz az $Y_k := -\frac{1}{k}\log_2 p(X_1, X_2, \ldots, X_k)$, $k = 1, 2, \ldots$, sorozat sztochasztikusan tart a $H(\mathbb{X})$ forrásentrópiához, ha $k \to \infty$.

Információstabilis források tulajdonságai

Ha $\mathbb{X}$ információstabilis, akkor elég nagy k esetén van olyan $A \subset \mathcal{X}^k$, hogy $\mathrm{P}(A) \approx 1$ és $\boldsymbol{x} \in A$ esetén

$$-\frac{1}{k} \log_2 p(\boldsymbol{x}) \approx H(\mathbb{X}), \quad \text{azaz} \quad p(\boldsymbol{x}) \approx 2^{-kH(\mathbb{X})}.$$

Emellett

$$|A| \approx 2^{kH(\mathbb{X})}.$$

Megjegyzés. Igazolható, hogy stacionárius és ergodikus források információstabilisak. Ezek speciális esetei a stacionárius emlékezetnélküli források, amik így szintén információstabilisak.

Tétel. *Ha az $\mathbb{X}$ stacionárius forrás információstabilis, akkor minden $0 < \varepsilon < 1$ esetén*

$$\lim_{k \to \infty} \frac{1}{k} \log_2 N(k, \varepsilon) = H(\mathbb{X}).$$

Az ε-hibavalószínűségű kódolás tétele

Tétel. *Legyen az $\mathbb{X}$ stacionárius forrás információstabilis. Ekkor, ha az $\mathbb{X}$ forrás k-hosszú blokkjait ε-hibával $(0 < \varepsilon < 1)$ kódoljuk állandó m_k hosszú kódszavakkal, akkor a kódok bármely ilyen tulajdonságú sorozatára*

$$\liminf_{k\to\infty} \frac{m_k}{k} \geq \frac{H(\mathbb{X})}{\log_2 s}.$$

Másrészt, tetszőleges $0 < \varepsilon < 1$ hibavalószínűséghez és $\delta > 0$ számhoz elég nagy k esetén mindig létezik olyan $f\colon \mathcal{X}^k \to \mathcal{Y}^{m_k}$ ε-hibával dekódolható kód, hogy

$$L = \frac{m_k}{k} < \frac{H(\mathbb{X})}{\log_2 s} + \delta.$$

$\frac{H(\mathbb{X})}{\log_2 s}$ a betűnkénti átlagos kódszóhossza alsó határa, amit elég nagy k blokk hossz esetén tetszőlegesen meg tudunk közelíteni.

Jelsebesség

Egy $f : \mathcal{X}^k \to \mathcal{Y}^m$ kód jelsebessége:

$$R := \frac{m}{k} \log_2 s.$$

Megadja, hogy $N = s^m$ darab kódszó bináris reprezentációjánál forrásbetűnként átlagosan hány bitet használunk fel.

Magyarázat. Ha $|\mathcal{X}| = n$, $|\mathcal{Y}| = s$ és k hosszú forrás blokkokat kódolunk m hosszan, akkor $n^k \approx s^m$, azaz $\log_2 n \approx \frac{m}{k} \log_2 s$. Egy n elemű forrás állandó kódszóhosszú bináris kódolásánál az egy betűre eső átlagos kódszóhossz éppen $\log_2 n$, azaz átlagosan ennyi bitet használunk fel.

Az R jelsebességű k hosszú blokkokat kódoló kódok közül az a legkisebb hibával dekódolható, amelyik az $\mathcal{X}^k$ első $N = 2^{kR} (= s^m)$ legnagyobb valószínűségű elemét kódolja egyértelműen. Ha $\mathcal{X}^k$ üzeneteit valószínűségeik szerint csökkenő sorrendben indexeljük, akkor a legjobb kód hibavalószínűsége:

$$P_e(k, R) = \sum_{i \geq 2^{kR}} p(\boldsymbol{x}_i).$$

Jelsebesség és forrásentrópia.

Tétel. *Ha az $\mathbb{X}$ stacionárius forrás információstabilis, akkor a legfeljebb R jelsebességű, k hosszú blokkokat állandó szóhosszon kódoló, legkisebb hibával dekódolható kód hibavalószínűségére igaz, hogy*

$$\lim_{k\to\infty} P_e(k,R) = 0, \qquad \text{ha } R > H(\mathbb{X}),$$

és

$$\lim_{k\to\infty} P_e(k,R) = 1, \qquad \text{ha } R < H(\mathbb{X}).$$

$R > H(\mathbb{X})$ esetén elég hosszú blokkokat használva a dekódolás hibavalószínűsége tetszőlegesen kicsivé tehető.

$R < H(\mathbb{X})$ esetén a blokkhossz növelésével a kód használhatatlanná válik, mivel a dekódolás hibavalószínűsége egyhez tart.

Megjegyzés. Emlékezetnélküli és stacionárius $\mathbb{X}$ forrás esetén igazolható, hogy ha $R > H(\mathbb{X})$ akkor $P_e(k,R)$, ha pedig $R < H(\mathbb{X})$, akkor $1 - P_e(k,R)$ tart exponenciális rendben nullához.

Kvantálás

$\mathbb{X} = X_1, X_2, \ldots$: stacionárius forrás, $X_i \in \mathbb{R}$ abszolút folytonos.

$\mathcal{Q} : \mathbb{R} \to \mathbb{R}$: véges értékkészletű függvény, kvantáló.

$\mathcal{Q}(X_1), \mathcal{Q}(X_2), \ldots$: az $\mathbb{X}$ skalár kvantáltja, diszkrét valószínűségi változó sorozat. $k = 1$ hosszú blokkokat kódoló forráskód.

A kvantálás hűségének mértéke egy n hosszú blokkra:

$$D(\mathcal{Q}) := \frac{1}{n}\mathsf{E}\left(\sum_{i=1}^{n}\left(X_i - \mathcal{Q}(X_i)\right)^2\right) \underset{\text{stacionaritás}}{=} \mathsf{E}\left(X - \mathcal{Q}(X)\right)^2.$$

$D(\mathcal{Q})$: a $\mathcal{Q}$ kvantáló négyzetes torzíása.

X: az $X_1, X_2, \ldots...$ közös eloszlásával megegyező eloszlású valószínűségi változó.

$\{x_1, x_2, \ldots, x_N\}$: a $\mathcal{Q}$ kvantáló értékkészlete. Elemei a kvantálási szintek.

Kvantálási tartományok: $\mathcal{B}_i := \{x \in \mathbb{R} : \mathcal{Q}(x) = x_i\}$, $i = 1, 2, \ldots, N$.

A definíciók diszkrét $\mathbb{X}$ forrás esetén is érvényesek.

Optimális kvantáló

Adott $\{x_1, x_2, \ldots, x_N\}$ kvantálási szintek esetén a legkisebb négyzetes torzítású $\mathcal{Q}$ kvantáló tartományai:

$$\mathcal{B}_i = \{x : |x - x_i| \leq |x - x_j|,\ j = 1, 2, \ldots N\}.$$

Egyenlőség esetén egy adott x-et a legkisebb indexű tartományhoz rendeljük. Ez a legközelebbi szomszéd feltétel. Csak ilyen tulajdonságú kvantálókkal dolgozunk.
Ha $x_1 < x_2 < \cdots < x_n$, akkor a kvantálási tartományok határai:

$$y_i = \frac{x_i + x_{i+1}}{2}, \quad i = 1, 2, \ldots, N-1, \qquad \text{azaz}$$

$$\mathcal{B}_1 =]-\infty, y_1\,], \quad \mathcal{B}_i =]y_{i-1}, y_i], \ i = 1, 2, \ldots, N-1, \quad \mathcal{B}_N =]y_{N-1}, \infty[.$$

$f(x)$: az $\mathbb{X}$ stacionárius forráshoz tartozó sűrűségfüggvény.

Egy adott $\mathcal{B}_i$ tartományhoz tartozó optimális kvantálási szint a $\mathcal{B}_i$ súlypontja:

$$x_i = \frac{\int_{\mathcal{B}_i} x f(x) \mathrm{d}x}{\int_{\mathcal{B}_i} f(x) \mathrm{d}x} = \mathrm{E}\big(X \big| X \in \mathcal{B}_i\big).$$

Egyenletes kvantáló

A

$$\mathcal{Q}(x) = x_i, \qquad \text{ha } x \in \mathcal{B}_i, \quad i = 1, 2, \ldots, N,$$

kvantáló négyzetes torzítása:

$$D(\mathcal{Q}) = \int_{-\infty}^{\infty} \big(x - \mathcal{Q}(x)\big)^2 f(x)\mathrm{d}x = \sum_{i=1}^{N} \int_{\mathcal{B}_i} (x - x_i)^2 f(x)\mathrm{d}x.$$

$[-A, A]$: az X értékkészlete, azaz $f(x) = 0$, ha $x \notin [-A, A]$.

Az N-szintű egyenletes kvantáló alakja:

$$\mathcal{Q}_N(x) = -A + (2i - 1)\frac{A}{N}, \qquad \text{ha}$$

$$-A + 2(i-1)\frac{A}{N} < x \leq -A + 2i\frac{A}{N}, \quad i = 1, 2, \ldots, N.$$

Magyarázat. Az $\mathcal{Q}_N$ kvantáló tartományait a $[-A, A]$ intervallum N egyenlő részre való osztásával kapjuk. A szintek az intervallumok felezőpontjai.

Az egyenletes kvantáló torzítása

Tétel. *Legyen az X abszolút folytonos $[-A, A]$ értékkészlettel és $f(x)$ sűrűségfüggvénnyel. Ekkor az X-et N szinten egyenletesen kvantáló $\mathcal{Q}_N$ kvantáló torzítására*

$$\lim_{N\to\infty} \left(\frac{N}{2A}\right)^2 D(\mathcal{Q}_N) = \frac{1}{12}.$$

Egy kvantálási intervallum hossza $q_N = \frac{2A}{N}$.

Egy q_N hosszúságú intervallumon vett egyenletes eloszlás szórásnégyzete $\frac{q_N^2}{12}$.

Megjegyzés. Ha N elég nagy, $D(\mathcal{Q}_N) \approx \frac{q_N^2}{12}$.

Egy N szintű kvantálónak, mint forráskódnak, a jelsebessége $R = \log_2 N$. A kvantálás egy $\mathcal{Q}(X_1), \mathcal{Q}(X_2), \ldots$ diszkrét stacionárius forrást eredményez. Ezt változó szóhosszal kódolva az átlagos kódszóhossz alsó határa a $H\big(\mathcal{Q}(X_1)\big)$ forrásentrópia.

Differenciális entrópia

Definíció. *Legyen X egy abszolút folytonos valószínűségi változó $f(x)$ eloszlásfüggvénnyel, melynek tartója $\mathcal{S} \subseteq \mathbb{R}$, azaz $f(x) = 0$, ha $x \notin \mathcal{S}$. Ekkor az X (vagy f) differenciális entrópiája*

$$H(X) = H(f) := -\int_{\mathcal{S}} f(x) \log_2 f(x) \mathrm{d}x,$$

amennyiben az integrál létezik.

Példák.

1. Legyen $X \sim \mathcal{U}(a, b)$. Ekkor $H(X) = \log_2(b - a)$ és tetszőleges $[a, b]$ tartójú abszolút folytonos Y esetén $H(Y) \leq H(X)$.

2. Legyen $X \sim \mathcal{E}xp(\lambda)$. Ekkor $H(X) = \log_2 \mathrm{e} - \log_2 \lambda$ és tetszőleges nemnegatív abszolút folytonos Y esetén, melyre $\mathrm{E}Y \leq \frac{1}{\lambda} = \mathrm{E}X$ teljesül, hogy $H(Y) \leq H(X)$.

3. Legyen $X \sim \mathcal{N}(\mu, \sigma)$. Ekkor $H(X) = \frac{1}{2} \log_2 \left(2\pi \mathrm{e} \sigma^2\right)$ és tetszőleges abszolút folytonos Y esetén, melyre $\mathrm{Var}(Y) \leq \sigma^2 = \mathrm{Var}(X)$ teljesül, hogy $H(Y) \leq H(X)$.

Az egyenletes kvantáló entrópiája

Tétel. *Legyen az X abszolút folytonos $[-A, A]$ értékkészlettel és $f(x)$ sűrűségfüggvénnyel és tegyük fel, hogy létezik a $H(f)$ differenciális entrópia. Ekkor az X N-szintű egyenletes kvantálásának $H\big(\mathcal{Q}_N(X)\big)$ entrópiájára*

$$\lim_{N\to\infty}\left(H\big(\mathcal{Q}_N(X)\big)+\log_2\left(\frac{2A}{N}\right)\right)=H(f).$$

Megjegyzés. Ha N elég nagy, $H\big(\mathcal{Q}_N(X)\big)\approx H(f)-\log_2 q_N$.

Következmény. Ha N elég nagy,

$$H\big(\mathcal{Q}_N(X)\big)\approx H(f)-\log_2\sqrt{12D(\mathcal{Q}_N)}.$$

Magyarázat. A torzítás határértéke

$$\lim_{N\to\infty}\left(\frac{N}{2A}\right)^2 D(\mathcal{Q}_N)=\frac{1}{12}.$$

Összevetve a tétel állításával:

$$\lim_{N\to\infty}\left(H\big(\mathcal{Q}_N(X)\big)+\log_2\sqrt{12D(\mathcal{Q}_N)}\right)=H(f).$$

Nem egyenletes kvantálók

Alapelv: a nagy valószínűségű tartományokon rövidebb kvantálási intervallumokat használunk.

Cél: adott X valószínűségű változóhoz határozzuk meg úgy az $x_1 < x_2 < \ldots < x_N$ kvantálási szinteket és a $\mathcal{Q}$ függvényt, hogy a $D(\mathcal{Q})$ minimális legyen.

Az optimális kvantáló eleget tesz az alábbi két szükséges feltételnek (együtt Lloyd-Max feltétel).

1. Legközelebbi szomszéd feltétel:

$$\left|x - \mathcal{Q}(x)\right| = \min_{1 \le i \le N} |x - x_i|, \qquad \forall x \in \mathbb{R}.$$

2. Súlypont feltétel:

   Minden x_j kvantálási szint megegyezik azon X_i minták átlagával (feltételes várható értékével), amelyeket erre a szintre kvantáltunk ($\mathcal{Q}(X_i) = x_j$).

Egy a fenti feltételeket kielégítő kvantálót Lloyd-Max kvantálónak nevezünk.

Példa

Nem minden Lloyd-Max kvantáló optimális.

Legyen X eloszlása egyenletes az $\{1, 2, 3, 4\}$ halmazon. A lehetséges 2-szintű Lloyd-Max kvantálók:

$$\mathcal{Q}_1(1) = 1; \qquad \mathcal{Q}_1(2) = \mathcal{Q}_1(3) = \mathcal{Q}_1(4) = 3;$$

$$\mathcal{Q}_2(4) = 4; \qquad \mathcal{Q}_2(1) = \mathcal{Q}_2(2) = \mathcal{Q}_2(3) = 2;$$

$$\mathcal{Q}_3(1) = \mathcal{Q}_3(2) = 1.5; \quad \mathcal{Q}_3(3) = \mathcal{Q}_3(4) = 3.5.$$

A négyzetes torzítások:

$$D(\mathcal{Q}_1) = D(\mathcal{Q}_2) = 0.5, \qquad D(\mathcal{Q}_3) = 0.25.$$

Csak a $\mathcal{Q}_3$ kvantáló optimális.

Lloyd-Max feltétel abszolút folytonos forrásokra

X: abszolút folytonos valószínűségi változó f sűrűségfüggvénnyel.

Lloyd-Max feltétel:

1. Legközelebbi szomszéd feltétel:

$$y_0 = -\infty, \quad y_i = \frac{x_i + x_{i+1}}{2}, \quad i = 1, 2, \ldots, N-1, \quad y_N = \infty,$$

   ahol y_{i-1} és y_i az x_i-hez tartozó kvantálási intervallum határai.

2. Súlypont feltétel:

$$x_i = \frac{\int_{y_{i-1}}^{y_i} x f(x) \mathrm{d}x}{\int_{y_{i-1}}^{y_i} f(x) \mathrm{d}x}, \qquad i = 1, 2, \ldots, N.$$

Tétel (Fleischer, 1964). *Legyen az $f(x)$ logaritmikusan konkáv, azaz $\log f(x)$ konkáv. Ekkor $f(x)$-re egyetlen N-szintű Lloyd-Max kvantáló létezik, így ez egyben optimális kvantáló is $f(x)$-re.*

Példa. Ha $X \sim \mathcal{U}(a, b)$, akkor a $\mathcal{Q}_N$ N-szintű egyenletes kvantáló az $[a, b]$ intervallumon kielégíti a Lloyd-Max feltételt $f(x)$-re. Mivel $f(x) = (b-a)^{-1}$, $x \in [a, b]$, logaritmikusan konkáv, $\mathcal{Q}_N$ az egyetlen optimális N-szintű kvantáló az egyenletes eloszlásra.

Lloyd-Max algoritmus

Keressük az optimális x_i kvantálási szinteket és a kapcsolódó $\mathcal{B}_i =]y_{i-1}, y_i]$ tartományokat.

Stuart P. Lloyd, Bell Laboratories, 1957 (publikálva: 1982);

Joel Max, General Telephone and Electronics Lab., Waltham, 1960.

Algoritmus

1. Válasszuk meg a kiinduló $x_1 < x_2 < \cdots < x_N$ kvantálási szinteket.
2. Határozzuk meg az y_i intervallumhatárokat a legközelebbi szomszéd feltétel szerint, azaz $y_i = \frac{x_i + x_{i+1}}{2}$, $i = 1, 2, \ldots, N-1$.
3. Az $y_0 = -\infty$ és $y_N = \infty$ választás mellett optimalizáljuk a kvantálót a súlypont feltétellel, meghatározva az új kvantálási szinteket.
4. Vizsgáljuk meg a négyzetes torzítás változását. Amennyiben egy előre megadott küszöb alatt van, álljunk meg, egyébként pedig ismételjük meg a 2. és 3. lépést.

Kompanderes kvantáló

A gyakorlatban az eszközök általában egyenletes kvantálást valósítanak meg. Ez pl. a széles dinamika tartományú beszéd esetén nem eredményez hatékony kódolást.

A forrást egy szigorúan monoton növekvő kompresszorral a $[-1, 1]$ intervallumba transzformáljuk, majd egyenletesen kvantáljuk. Összességében ez egy nem egyenletes kvantáló.

A kvantált értékeket dekódoljuk, majd az expanderrel (a kompresszor inverze) visszaállítjuk az eredeti dinamika tartományt.

Kompander: a kompresszor és expander együtt.

Alkalmazások:

- Digitális telefonrendszereknél az A/D konverter előtt egy kompresszor, majd a D/A konverter után egy expander.
- Professzionális drótnélküli mikrofonoknál, mivel a mikrofon audiojelének dinamika tartománya jóval szélesebb, mint a rádiójeleké.

Kompanderek a beszédkódolásban

8-bites Pulse Code Modulation (PCM) digitális telefonrendszerek.

Észak-Amerika és Japán: μ-law ($\mu = 255$).

$$G_\mu(x) = \operatorname{sign}(x)\,\frac{\log(1+\mu|x|)}{\log(1+\mu)}, \qquad -1 \le x \le 1.$$

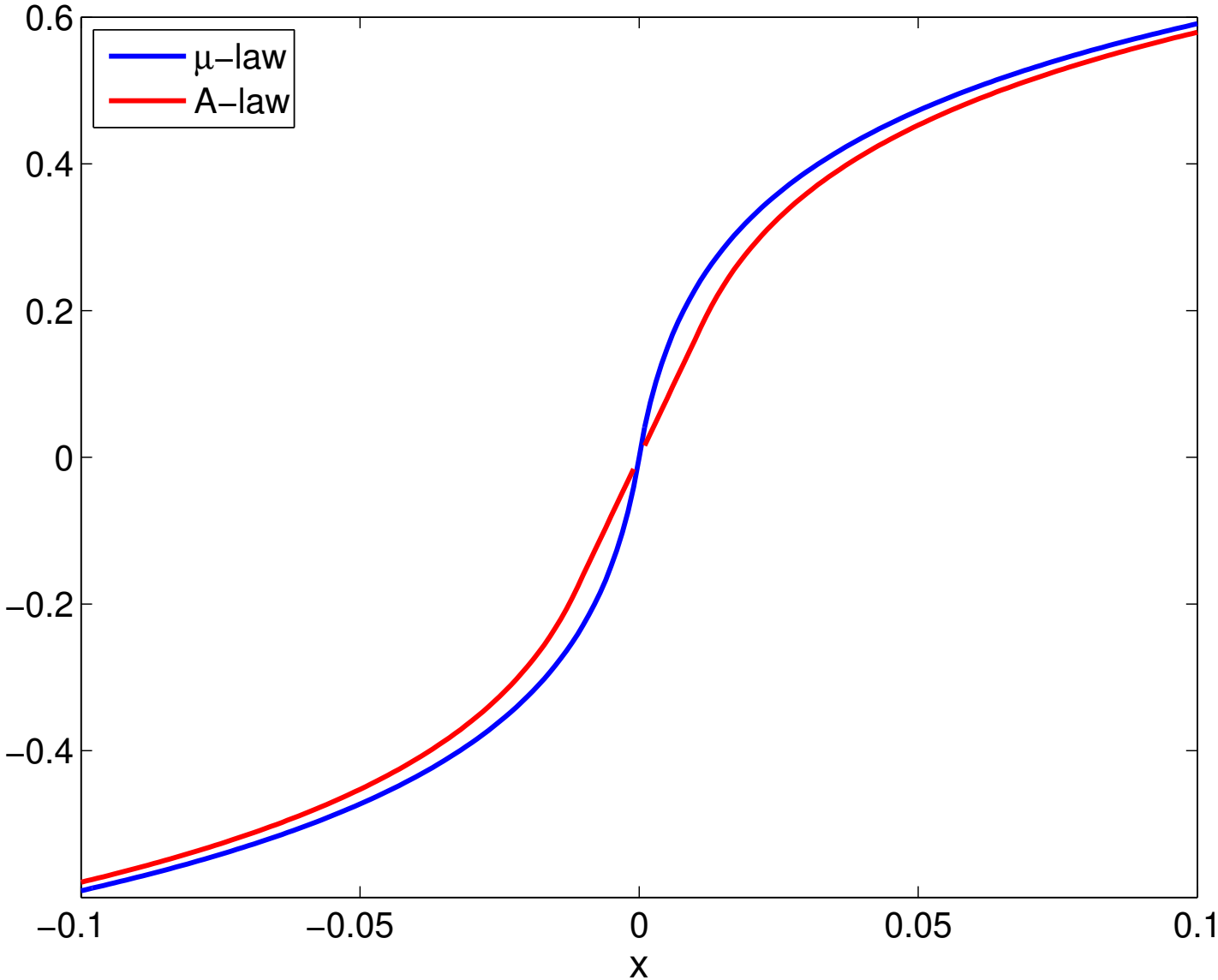
$$G_\mu^{-1}(x) = \operatorname{sign}(x)\,\frac{1}{\mu}\Big((1+\mu)^{|x|}-1\Big), \qquad -1 \le x \le 1.$$

Európa: A-law ($A = 87.7$ vagy $A = 87.6$)

$$G_A(x) = \begin{cases} \operatorname{sign}(x)\,\frac{A|x|}{1+\log A}, & 0 \le |x| < \frac{1}{A}; \\ \operatorname{sign}(x)\,\frac{1+\log|Ax|}{1+\log A}, & \frac{1}{A} \le |x| \le 1. \end{cases}$$

$$G_A^{-1}(x) = \begin{cases} \operatorname{sign}(x)\,\frac{|x|(1+\log A)}{A}, & 0 \le |x| < \frac{1}{1+\log A}; \\ \operatorname{sign}(x)\,\frac{\exp\{|x|(1+\log A)-1\}}{A}, & \frac{1}{1+\log A} \le |x| \le 1. \end{cases}$$

μ-law és A-law



Vektorkvantálás

A forrás kimeneteit többdimenziós eloszlának tekintjük. Így ugyanakkora torzítás mellett jobb tömörítést érhetünk el, különösen, ha az egyes dimenziókhoz tartozó változók korrelálnak.

Példa. Egy színes kép RGB értékeit nem egyenként, hanem egy 3D színtér elemeiként kvantáljuk. 3 skalár kvantálóval a kvantáló tartományok téglatestek, 3D kvantálóval tetszőleges térbeli idomok használhatóak.

$\boldsymbol{X}$: d-dimenziós forrásvektor $f(\boldsymbol{x})$ sűrűségfüggvénnyel.

Kvantáló: $\mathcal{Q} : \mathbb{R}^d \to \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$, $\boldsymbol{x}_i \in \mathbb{R}^d$, $i = 1, 2, \ldots, N$.

Kvantálási tartományok: $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_N$. $\mathbb{R}^d$ egy partíciója, azaz diszjunktak és $\bigcup_{i=1}^{N} \mathcal{B}_i = \mathbb{R}^d$.

$$\mathcal{Q}(\boldsymbol{x}) = \boldsymbol{x}_i, \qquad \text{ha } \boldsymbol{x} \in \mathcal{B}_i, \quad i = 1, 2, \ldots, N.$$

Lloyd-Max feltétel vektorkvantálóra

Négyzetes torzítás:

$$D(\mathcal{Q}) = \frac{1}{d}\mathsf{E}\left\|\boldsymbol{X} - \mathcal{Q}(\boldsymbol{X})\right\|^2 = \frac{1}{d}\sum_{i=1}^{N}\int \|\boldsymbol{x} - \boldsymbol{x}_i\|^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Lloyd-Max feltétel:

1. Legközelebbi szomszéd feltétel: $\mathbb{R}^d$ tartományai Voronoi-tartományok, azaz

$$\mathcal{B}_i = \left\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}_i\| \leq \|\boldsymbol{x} - \boldsymbol{x}_j\|,\ \forall j \neq i\right\}.$$

2. Súlypont feltétel:

$$\boldsymbol{x}_i = \arg\min_{\boldsymbol{y}} \int_{\mathcal{B}_i} \|\boldsymbol{x} - \boldsymbol{y}\|^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

azaz a kimeneti vektorok a kapcsolódó tartományok súlypontjai.

Lloyd-Max algoritmus természetes általánosítása: Linde-Buzo-Gray algoritmus.

Spektrális sűrűség

$X(t),\ t \in \mathbb{R}$: sztochasztikus folyamat. Minden $t \in \mathbb{R}$ időpillanathoz tartozik egy valószínűségi változó.

Definíció. *Az $X(t),\ t \in \mathbb{R}$, sztochasztikus folyamat gyengén stacionárius, ha a várható érték függvénye konstans, az $X(t)$ és $X(s)$ kovarianciái pedig csak a $t - s$ különbségtől függenek. Az*

$$R(\tau) := \mathrm{Cov}\left(X(t+\tau), X(t)\right)$$

függvényt az $X(t)$ gyengén stacionárius folyamat kovarianciafüggvényének, az $R(0)$ értéket pedig a folyamat energiájának nevezzük.

Megjegyzés. Minden stacionárius folyamat gyengén stacionárius.

Definíció. *Legyen $R(\tau)$ az $X(t)$ gyengén stacionárius folyamat kovariancia függvénye. Ha létezik az $R(\tau)$ Fourier transzformáltja, azaz*

$$s(\omega) = \int_{-\infty}^{\infty} R(\tau) \mathrm{e}^{i\omega\tau} \mathrm{d}\tau,$$

akkor az $s(\omega)$ függvényt az $X(t)$ folyamat spektrális sűrűségfüggvényének nevezzük.

Lineáris szűrés

Az általánosítás csorbítása nélkül feltehető, hogy $\mathrm{E}X(t) = 0$.

$h(t)$: négyzetesen integrálható függvény, azaz

$$\int_{-\infty}^{\infty} \big(h(t)\big)^2 \mathrm{d}t < \infty.$$

Definíció. *Az $X(t)$ gyengén stacionárius folyamat lineáris szűrésén az*

$$Y(t) := \int_{-\infty}^{\infty} h(s)X(t-s)\mathrm{d}s$$

folyamatot értjük. A $h(t)$ függvényt a szűrő súlyfüggvényének,

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\mathrm{e}^{i\omega t}\mathrm{d}t$$

Fourier transzformáltját pedig a szűrő átviteli függvényének nevezzük.

Megjegyzés. $Y(t)$ gyengén stacionárius és $\mathrm{E}Y(t) = 0$.

A szűrt folyamat kovarianciája

$s(\omega)$: az $R(\tau)$ kovarianciájú $X(t)$ spektrális sűrűségfüggvénye.

$$s(\omega) = \int_{-\infty}^{\infty} R(\tau) \mathrm{e}^{i\omega\tau} \mathrm{d}\tau \quad \text{és} \quad R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} s(\omega) \mathrm{e}^{-i\omega\tau} \mathrm{d}\omega.$$

$H(\omega)$: a $h(t)$ súlyfüggvény átviteli függvénye.

Az $Y(t)$ szűrt folyamat kovariancia függvénye:

$$\widetilde{R}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} s(\omega) \left| H(\omega) \right|^2 \mathrm{e}^{-i\omega\tau} \mathrm{d}\omega.$$

A szűrt folyamat energiája:

$$\widetilde{R}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} s(\omega) \left| H(\omega) \right|^2 \mathrm{d}\omega.$$

Ideális sávszűrő

Legyen $0 < b < B$. Az ideális sávszűrő átviteli függvénye:

$$H(\omega) = \begin{cases} 1 & \text{ha } |\omega| \in [b, B]; \\ 0 & \text{egyébként.} \end{cases}$$

A kapcsolódó súlyfüggvény:

$$h(t) = \frac{\sin(Bt)}{\pi t} - \frac{\sin(bt)}{\pi t}.$$

Legyen

$$X(t) = A\sin(\omega t + \varphi), \qquad \varphi \sim \mathcal{U}(0, 2\pi).$$

A: amplitúdó; ω: frekvencia; φ: fáziseltolás.

$X(t)$ szűrtje $X(t)$, ha $|\omega| \in [b, B]$ és 0 egyébként.

A sávszűrő energiája:

$$\widetilde{R}(0) = \frac{1}{2\pi} \int_{|\omega| \in [b,B]} s(\omega) \mathrm{d}\omega.$$

Ha $s(\omega)$ folytonos és $B - b$ kicsi (tűszűrő): $\widetilde{R}(0) \approx \frac{1}{\pi}(B - b)s(b)$.

Mintavételezés

$X(t)$: 0 várható értékű gyengén stacionárius folyamat.

$\{X(kT),\ k = 0, \pm 1, \pm 2, \ldots\}$: az $X(t)$ folyamat mintavételezése $T > 0$ mintavételi idővel.

Sinc-függvény:

$$\operatorname{sinc}(t) := \frac{\sin(\pi t)}{\pi t}, \qquad t \in \mathbb{R}.$$

Megjegyzés. $\operatorname{sinc}(0) = 1$ és $\operatorname{sinc}(k) = 0,\ k = 0, \pm 1, \pm 2, \ldots.$

$X(t)$ reprodukciója:

$$\widehat{X}(t) := \sum_{k=-\infty}^{\infty} X(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right), \quad t \in \mathbb{R},$$

ahol

$$\lim_{N\to\infty} \mathrm{E}\left(\widehat{X}(t) - \sum_{k=-N}^{N} X(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right)\right)^2 = 0.$$

Megjegyzés. Ha $t = kT$, akkor $\widehat{X}(t) = X(t)$.

Mintavételi tétel

Tétel. *Ha az $s(\omega)$ spektrális sűrűségfüggvényű 0 várható értékű $X(t)$ gyengén stacionárius folyamat a B sávra korlátozott, vagyis energiája*

$$\widetilde{R}(0) = \frac{1}{2\pi} \int_{-B}^{B} s(\omega) \mathrm{d}\omega,$$

és

$$T < \frac{\pi}{B},$$

akkor minden $t \in \mathbb{R}$ esetén

$$\mathrm{P}\big(\widehat{X}(t) = X(t)\big) = 1.$$

Megjegyzés. Ha $B' = \frac{B}{2\pi}$ a sáv frekvencia, akkor a mintavételi frekvencia legalább a B' duplája kell, hogy legyen.

Példa. A telefonvonalaknál a beszédet 3400 Hz-re sávszűrik és 8000 Hz-el mintavételezik.

CD-minőség esetén 20 kHz-re sávszűrnek és 44100 Hz-el mintavételeznek.

Transzformációs kódolás

1. A kódolandó jelsorozatot diszjunkt blokkokra bontjuk és minden blokkra alkalmazunk egy invertálható transzformációt.
2. Kvantáljuk a transzformált blokkokat.
3. A kvantált értékeket kódoljuk egy bináris kóddal.

$\boldsymbol{x} = (x_0, x_1, \ldots, x_{k-1})^\top$: transzformálandó blokk.

$\boldsymbol{y} = (y_0, y_1, \ldots, y_{k-1})^\top$: transzformált blokk.

$\boldsymbol{A} = (a_{i,j})$: $k \times k$ dimenziós ortonormált transzformációs mátrix.

Transzformáció és inverze:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \quad \text{és} \quad \boldsymbol{x} = \boldsymbol{B}\boldsymbol{y}, \quad \text{ahol} \quad \boldsymbol{B} = \boldsymbol{A}^{-1} = \boldsymbol{A}^\top.$$

Két-dimenziós esetben (képtömörítés) a transzformálandó és transzformált blokkok mátrixok ($\boldsymbol{X}$ és $\boldsymbol{Y}$):

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^\top \quad \text{és} \quad \boldsymbol{X} = \boldsymbol{A}^\top \boldsymbol{Y}\boldsymbol{A}.$$

Példa. JPEG tömörítés esetén 8×8 pixeles blokkokat használnak.

Speciális transzformációk

1. Diszkrét konszinusz transzformáció (DCT) Ahmet, Natarajan, Rao (1974)

Az $\boldsymbol{A}$ mátrix elemei

$$a_{1,j} = \frac{1}{\sqrt{k}}, \quad a_{i,j} = \sqrt{\frac{2}{k}} \cos\left(\frac{(2j-1)(i-1)\pi}{2k}\right), \quad i,j = 1,2,\ldots,k.$$

A legkedveltebb transzformáció. Alkalmazásai: JPEG, MPEG.

2. Diszkrét Walsh-Hadamard transzformáció (DWHT, 1923) Jacques Hadamard, Joseph L. Walsh.

Az $\boldsymbol{A}$ mátrix elemeit rekurzióval számoljuk:

$$\boldsymbol{A}_{2^k} = \begin{bmatrix} \boldsymbol{A}_{2^{k-1}} & \boldsymbol{A}_{2^{k-1}} \\ \boldsymbol{A}_{2^{k-1}} & -\boldsymbol{A}_{2^{k-1}} \end{bmatrix}, \quad \text{ahol} \quad \boldsymbol{A}_1 = 1.$$

$\boldsymbol{A}_{2^k}\boldsymbol{A}_{2^k}^\top = 2^k \boldsymbol{I}_{2^k}$, ezért a transzformáció mátrixa $\frac{1}{\sqrt{2^k}}\boldsymbol{A}_{2^k}$.

Alkalmazásai: JPEG XR (JPEG **ex**tended **r**ange, 2009; Microsoft HD photo) és MPEG-4 AVC vagy H.264 (MPEG-4 Part 10 **A**dvanced **V**ideo **C**oding, 2003; pl. Blue-Ray lemezek).

Részsávos kódolás (SBC: subband coding)

1. A forrást egy szűrőbankon vezetjük át, ami azt frekvenciasávokra bontja (analízis szűrők). Pl. M egyforma széles sávot veszünk.

2. A szűrők kimeneténél mintavételezünk és, hogy szinkronban maradjunk, csökkentjük a minták számát a bemeneti és kimeneti sávszélesség arányában (decimálás, alulmintavételezés). Pl. minden M-edik mintát tartunk meg.

3. Az egyes decimált jeleket külön-külön kódoljuk és továbbítjuk.

4. A dekódolás után a minták közé annyi nullát írunk, hogy visszaállítsuk az eredeti mintaelemszámot (felülmintavételezés).

5. Az egyes mintákat egy szűrőbankon vezetjük keresztül (szintézis szűrők), ami előállítja a kimenő jelet.

Az emberi érzékszervek frekvenciafüggőek. A fontosabb frekvenciákat pontosabban kell rekonstruálni, a kevésbé fontosakat nagyobb torzítással is elég.

Példa. Az emberi hallásnál ha az egyik frekvencia elég hangos, akkor elnyomja (maszkolja) a mellette lévő frekvenciákat.

Példa

A szűrőbank bemenete: $(x_1, x_2, \ldots, x_n)$.

Kimenetek: $(y_1, y_2, \ldots, y_n)$ és $(z_1, z_2, \ldots, z_n)$, ahol $x_0 = 0$ jelöléssel

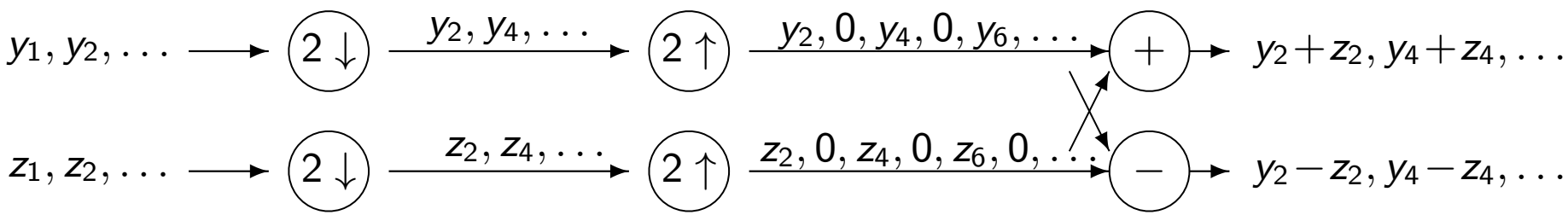
$$y_i = \frac{x_i + x_{i-1}}{2}; \qquad z_i = x_i - y_i = \frac{x_i - x_{i-1}}{2}, \quad i = 1, 2, \ldots, n.$$

Mindkét kimenő jelsorozat simább (kisebb dinamikájú), mint az eredeti, így kisebb torzítással tömöríthető. Megdupláztuk az outputot.

Alulmintavételezés: továbbítsuk csak a páros indexű y_{2i} és z_{2i} jeleket.

Szintézis:

$$x_{2i} = y_{2i} + z_{2i}, \qquad x_{2i-1} = y_{2i} - z_{2i}.$$



Különbségi kódolás

A prediktív kódolás egy speciális esete. Akkor előnyös, ha a szomszédos minták közötti eltérés kicsi, pl. digitális képeken ha nem vagyunk egy él közelében.

Példa. 8 bites intenzitású kép 8 szomszédos pixelének értéke:

$$147,\ 145,\ 141,\ 146,\ 149,\ 147,\ 143,\ 145.$$

Bitenkénti kódolás egyenként 8 biten: 64 bit.

Különbségek (az első kivételével):

$$147,\ -2,\ -4,\ 5,\ 3,\ -2,\ -4,\ 2.$$

A legnagyobb különbség abszolút értéke 5, ami 3 biten kódolható.

A különbségek kódolásához elegendő 4 bit ($3 + 1$ előjel).

8 biten küldjük át a különbségek ábrázolásának hosszát.

Különbségi kódolás hossza: $8 + 8 + 7 \cdot 4 = 44$ bit. 31 % nyereség.

A tömörítés veszteségmentes.

Veszteséges tömörítés, példa

A forrás kimenete:

$$5.4,\ 10.1,\ 7.2,\ 4.6,\ 6.9,\ 12.5,\ 6.2,\ 5.3.$$

A különbségek:

$$5.4,\ 4.7,\ -2.9,\ -2.6,\ 2.3,\ 5.6,\ -6.3,\ -0.9.$$

7 szintű egyenletes kvantáló. A szintek: $-6,\ -4,\ -2,\ 0,\ 2,\ 4,\ 6$.

A kvantált értékek:

$$6,\ 4,\ -2,\ -2,\ 2,\ 6,\ -6,\ 0.$$

A rekonstruált értékek:

$$6,\ 10,\ 8,\ 6,\ 8,\ 14,\ 8,\ 8.$$

A hibák:

$$-0.6,\ 0.1,\ -0.8,\ -1.4,\ -1.1,\ -1.5,\ -1.8,\ -2.7.$$

Hosszabb sorozatok még nagyobb eltéréseket eredményezhetnek.

Kvantálási hibák

$\{x_n\}$, $\{\widehat{x}_n\}$: input, illetve rekonstruált sorozat

$\{d_n\}$: különbség sorozat, $d_n = x_n - x_{n-1}$.

Kvantált különbség sorozat tagjai: $\widehat{d}_n = Q(d_n) = d_n + q_n$.

Rekonstruálás: $\widehat{x}_0 = x_0$ és $\widehat{x}_n = \widehat{x}_{n-1} + \widehat{d}_n$.

$$d_1 = x_1 - x_0; \qquad \widehat{d}_1 = Q(d_1) = d_1 + q_1;$$
$$\widehat{x}_1 = \widehat{x}_0 + \widehat{d}_1 = x_0 + d_1 + q_1 = x_1 + q_1;$$

$$d_2 = x_2 - x_1; \qquad \widehat{d}_2 = Q(d_2) = d_2 + q_2;$$
$$\widehat{x}_2 = \widehat{x}_1 + \widehat{d}_2 = x_1 + q_1 + d_2 + q_2 = x_2 + q_1 + q_2;$$

$$\widehat{x}_n = x_n + \sum_{k=1}^{n} q_k.$$

A kvantálási hibák **kumulálódnak**.

Prediktív kódolás

Az n-edik lépésben a kódoló ismeri az $\widehat{x}_{n-1}$ értékét.

Módosított különbségek: $d_n = x_n - \widehat{x}_{n-1}$.

$$d_1 = x_1 - x_0; \qquad \widehat{d}_1 = Q(d_1) = d_1 + q_1;$$
$$\widehat{x}_1 = \widehat{x}_0 + \widehat{d}_1 = x_0 + d_1 + q_1 = x_1 + q_1;$$
$$d_2 = x_2 - \widehat{x}_1; \qquad \widehat{d}_2 = Q(d_2) = d_2 + q_2;$$
$$\widehat{x}_2 = \widehat{x}_1 + \widehat{d}_2 = \widehat{x}_1 + d_2 + q_2 = x_2 + q_2;$$
$$\widehat{x}_n = x_n + q_n.$$

Cél: minél kisebb d_n különbségek elérése.

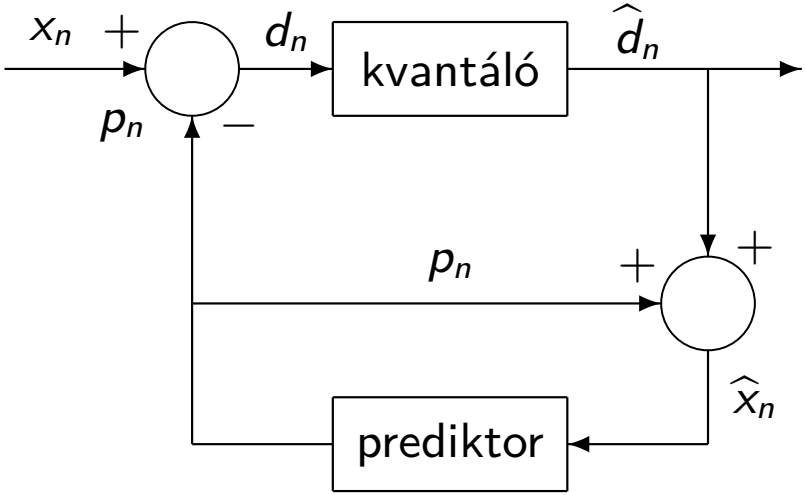
Az x_n értékét a rekonstruált sorozat korábbi értékeinek egy függvényével, a $p_n = f(\widehat{x}_{n-1}, \widehat{x}_{n-2}, \ldots, \widehat{x}_0)$ prediktorral közelítjük.

$$d_n = x_n - p_n = x_n - f(\widehat{x}_{n-1}, \widehat{x}_{n-2}, \ldots, \widehat{x}_0).$$

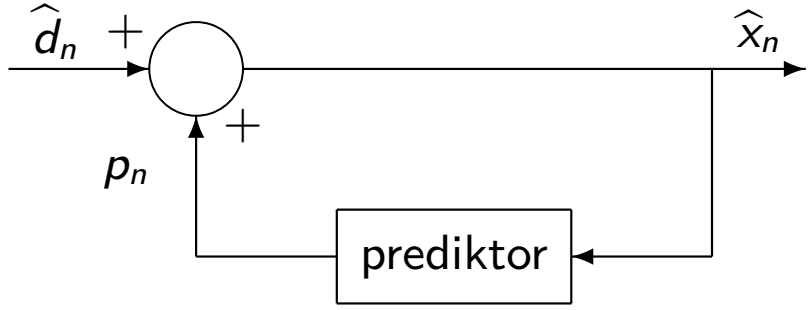
Ez a különbségi impulzuskód moduláció (DPCM: differential pulse code modulation).

Szabadalom: C. Chapin Cutler, Bell Laboratories, 1950.

DPCM



kódoló



dekódoló

A bemeneti jelek idővel változhatnak: adaptív DPCM.

1. Alkalmazkodás a kódoló bemenő x_n jele alapján: előre adaptív módszer. A dekódoló nem ismeri az x_n jelet, az új paramétereket át kell vinni.

2. Alkalmazkodás a kimenő $\widehat{x}_n$ jel alapján: hátra adaptív módszer. Mind a kódoló, mind a dekódoló ismeri.

A kvantáló is lehet adaptív. Előre adaptív eset: az inputot blokkokra bontjuk és az egyes blokkokra optimális kvantáló paramétereit is továbbítjuk.

Jayant-kvantáló

Hátra adaptív kvantáló. Nikil S. Jayant, Bell Laboratories, 1973.

Ha az input érték a kvantáló belső (origóhoz közeli) tartományába esik, finomítsuk a lépésközt, ellenkező esetben növeljük.

Minden kvantálási intervallumhoz tartozik egy szorzótényező, ami a belső részen 1 alatti, azon kívül 1 feletti. A szorzók az origóra szimmetrikusak.

M_k: a k-adik szint szorzótényezője.

Külső tartomány: $M_k > 1$; belső tartomány: $M_k < 1$.

Δ_n: a kvantáló lépésköze az x_n (n-edik) inputnál.

Ha x_{n-1} az $\ell(n-1)$-el jelölt intervallumba esik, akkor

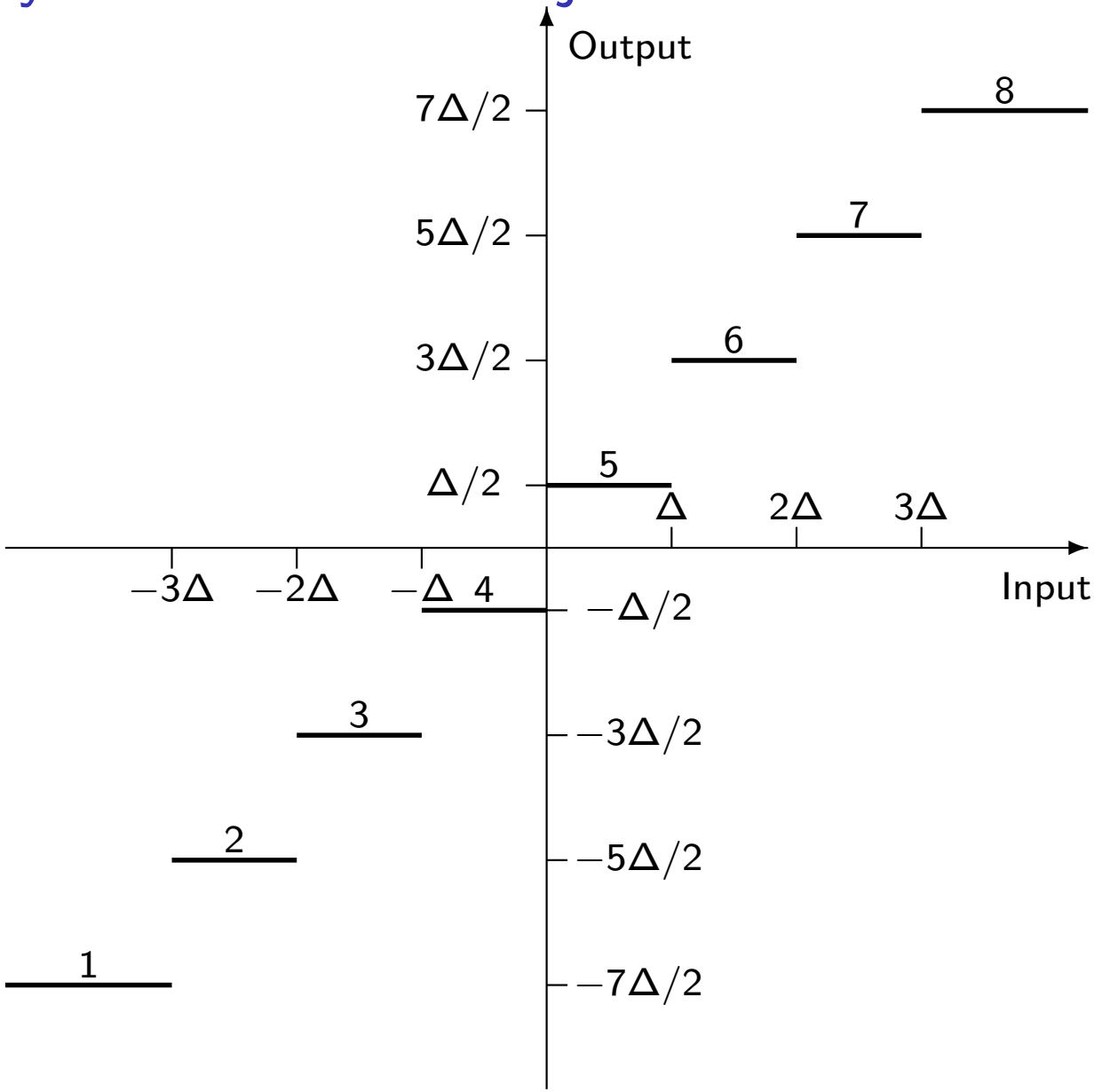
$$\Delta_n = M_{\ell(n-1)}\Delta_{n-1}.$$

A véges pontosság miatt szükséges $\Delta_{\min}$ és $\Delta_{\max}$ megadása.

Példa. Egy 8 szintű (3 bites) kvantáló szorzói:

$M_1 = M_8 = 1.2,\ M_2 = M_7 = 1,\ M_3 = M_6 = 0.9,\ M_4 = M_5 = 0.8.$

3 bites Jayant kvantáló szintjei



A súlyok szimmetrikusak: $M_1 = M_8,\ M_2 = M_7,\ M_3 = M_6,\ M_4 = M_5.$

Példa

- Belső tartomány: $M_4 = M_5 = 0.8,\ M_3 = M_6 = 0.9.$
- Külső tartomány: $M_2 = M_7 = 1,\ M_1 = M_8 = 1.2.$
- Kiinduló lépésköz: $\Delta_0 = 0.5.$
- Input:
  $0.1,\ -0.2,\ 0.2,\ 0.1,\ -0.3,\ 0.1,\ 0.2,\ 0.5,\ 0.9,\ 1.5,\ 1.0,\ 0.9.$

A kvantálás menete:

n	Δ_n	Input	Szint	Output	Hiba	Lépésköz frissítés
0	0.5	0.1	5	0.25	0.15	$\Delta_1 = M_5 \times \Delta_0$
1	0.4	−0.2	4	−0.2	0.0	$\Delta_2 = M_4 \times \Delta_1$
2	0.32	0.2	5	0.16	0.04	$\Delta_3 = M_5 \times \Delta_2$
3	0.256	0.1	5	0.128	0.028	$\Delta_4 = M_5 \times \Delta_3$
4	0.2048	−0.3	3	−0.3072	−0.072	$\Delta_5 = M_3 \times \Delta_4$
5	0.1843	0.1	5	0.0922	−0.0078	$\Delta_6 = M_5 \times \Delta_5$
6	0.1475	0.2	6	0.2212	0.0212	$\Delta_7 = M_6 \times \Delta_6$
7	0.1328	0.5	8	0.4646	−0.0354	$\Delta_8 = M_8 \times \Delta_7$
8	0.1594	0.9	8	0.5578	−0.3422	$\Delta_9 = M_8 \times \Delta_8$
9	0.1913	1.5	8	0.6696	−0.8304	$\Delta_{10} = M_8 \times \Delta_9$
10	0.2296	1.0	8	0.8036	0.1964	$\Delta_{11} = M_8 \times \Delta_{10}$
11	0.2755	0.9	8	0.9643	0.0643	$\Delta_{12} = M_8 \times \Delta_{11}$

Delta moduláció

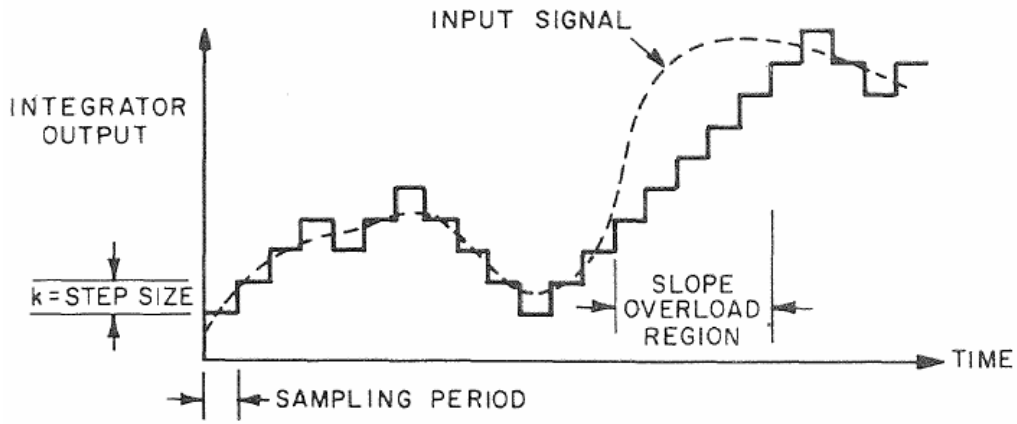
Egy folytonos jelből vett sűrű mintavételezés esetén a szomszédos értékek között kicsi az eltérés.

Delta moduláció (DM): 2 szintű (1 bites) kvantálóval bíró DPCM.

A torzítás csökkentéséhez a mintavételi frekvenciát növelik, akár a sávfrekvencia százszorosára.

Kimeneti értékek: $\pm\Delta$. Rögzített Δ: lineáris delta moduláció.

Probléma: lapos input esetén az output oszcillál, meredeken változó inputot nem tud követni.



Forrás: John Edward Abate: *Linear and adaptive delta modulation.* PhD thesis, Newark College of Engineering, 1967.

Adaptív delta moduláció

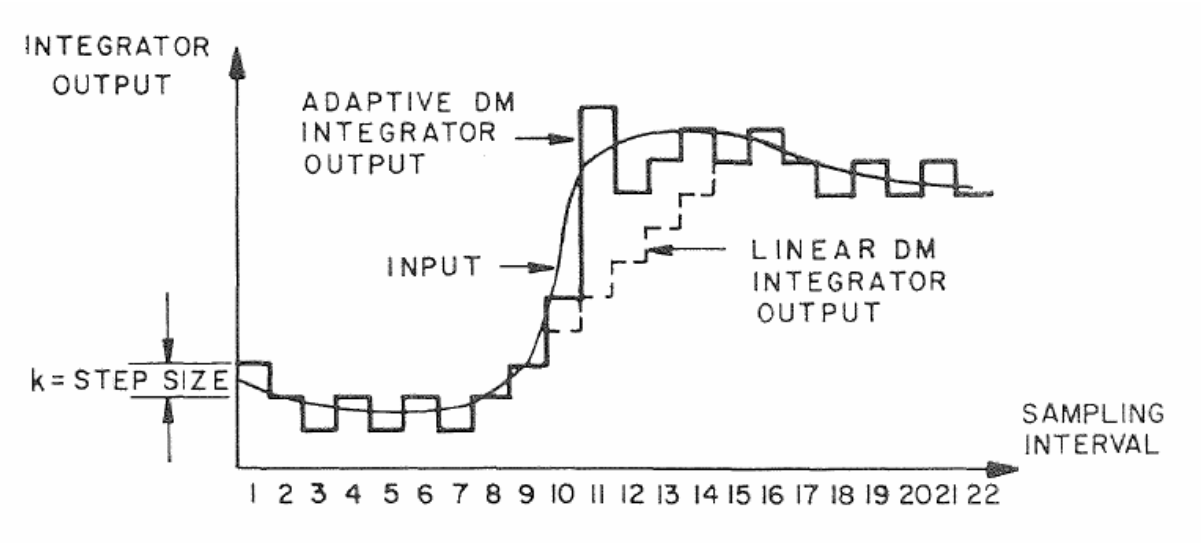
John Edward Abate, AT&T and Bell Laboratories, 1967.

Közel konstans tartomány: kis Δ; gyors változás: nagy Δ lépésköz.

s_n: a DM „lépése” az n-edik időpillanatban, $s_n = \pm\Delta_n$.

Egy lépést figyelő módszer:

$$\Delta_{n+1} = \begin{cases} M_1\Delta_n, & \text{ha } \operatorname{sign} s_n = \operatorname{sign} s_{n-1}; \\ M_2\Delta_n, & \text{ha } \operatorname{sign} s_n \neq \operatorname{sign} s_{n-1}; \end{cases} \qquad 1 < M_1 = \frac{1}{M_2} < 2.$$



Forrás: John Edward Abate: *Linear and adaptive delta modulation.* PhD thesis, Newark College of Engineering, 1967.

Változó meredekségű delta moduláció

Johannes Anton Griefkes, Karel Riemens, Philips, 1970.

Változó meredekségű delta moduláció (CVSD: continuous variable slope delta modulation):

$$\Delta_n = \beta\Delta_{n-1} + \alpha_n\Delta_0.$$

β: konstans, egynél alig kisebb;

$\alpha_n \in \{0, 1\}$: $\alpha_n = 1$, ha a kvantáló legutóbbi K output értékéből J darab azonos előjelű, egyébként $\alpha_n = 0$. Jellemzően: $J = K = 3$.

Mintánként 1 biten kódol, pl. 16 kHz-es mintavétel 16 kbit/s sebességgel kódolódik.

Alkalmazások:

- 16 és 32 kbit/s CVSD: TRI-TAC katonai kommunikációs rend- szer. 16 kbit/s: US Army; 32 kbit/s: US Air Force.
- 64 kbit/s CVSD: bluetooth.

Prediktorok

$p_n = f(\widehat{x}_{n-1}, \widehat{x}_{n-2}, \ldots, \widehat{x}_0)$: a prediktív kódoló (pl. DPCM) prediktora.

Cél: azon optimális f függvény meghatározása, mely minimalizálja a

$$\sigma_d^2 = \mathrm{E}(X_n - p_n)^2$$

négyzetes eltérést. Általánosan a feladat túl bonyolult.

Elég finom kvantálás esetén $\widehat{x}_n \approx X_n$, azaz tekinthető

$$p_n = f(X_{n-1}, X_{n-2}, \ldots, X_0).$$

σ_d^2 minimális, ha

$$f(X_{n-1}, X_{n-2}, \ldots, X_0) = \mathrm{E}(X_n | X_{n-1}, X_{n-2}, \ldots, X_0),$$

de ehhez szükséges a megfelelő feltételes valószínűségek ismerete.

Általában nem ismertek. Normális eloszlású forrás esetén a feltételes várható érték az $X_{n-1}, X_{n-2}, \ldots, X_0$ lineáris függvénye.

Lineáris becslés

N-edrendű lineáris prediktorfüggvény:

$$p_n := \sum_{i=1}^{N} a_i \widehat{x}_{n-i}.$$

Elég finom kvantálás esetén minimalizálandó

$$\sigma_d^2 = \mathrm{E}\left(X_n - \sum_{i=1}^{N} a_i X_{n-i} \right)^2.$$

$R(k) = \mathrm{E}(X_n X_{n+k})$: az X_k nulla várható értékű gyengén stacionárius forrás kovariancia függvénye.

A $\frac{\partial}{\partial a_j}\sigma_d^2 = 0,\ j = 1, 2, \ldots, N$, egyenletekből:

$$\sum_{i=1}^{N} a_i R(i-j) = R(j), \qquad j = 1, 2, \ldots, N.$$

A egyenletrendszer megoldása megadja a prediktor együtthatóit.

Probléma: a megoldandó Wiener-Hopf egyenletrendszer felírásakor feltétel volt a stacionaritás. Ez legfeljebb csak lokálisan teljesül.

Adaptív prediktor

Előre adaptív eset: az inputot blokkokra osztjuk.

Beszédkódolás: 16 ms blokkhossz. 8000 Hz-es mintavételezésnél 128 mintaelem.

Képtömörítés: 8×8 pixeles blokkok.

Az ℓ-edik M hosszú blokk kovarianciájának becslése:

$$R^{(\ell)}(k) = \frac{1}{M - |k|} \sum_{i=(\ell-1)M+1}^{\ell M - |k|} X_i X_{i+|k|}, \qquad R^{(\ell)}(-k) = R^{(\ell)}(k).$$

A bemenetet pufferelni kell, ami késleltetést visz a rendszerbe. A dekódolónak kiegészítő információkra is szüksége van, mert nem ismeri az input jelet.

Hátra adaptív eset: a kódoló kimeneti jelét felhasználva rekurzív formula a

$$d_n^2 = \left(X_n - \sum_{i=1}^{N} a_i \widehat{x}_{n-i} \right)^2$$

minimumának meghatározására.

Hullámformán alapuló beszédkódolás

Referencia módszer: impulzuskód moduláció (PCM: pulse code modulation) codec.

Analóg $300 - 3400$ Hz-es sávszélességű beszédjelből 8000 Hz-el mintavételezünk és 8 bittel kvantáljuk.

Bitsebesség: $8000 \times 8 = \mathbf{64}$ kbit/s.

ITU-T (International Telecommunication Union – Telecommunication Standardization Sector) G.711 távközlési szabvány (1972): PCM kódolás kompanderes kvantálóval (A-law vagy μ-law).

Adaptív DPCM: kihasználja a beszéd különböző mintái közötti korrelációt (Jayant, Bell Laboratories, 1974)

Kvantálás: 5, 4, 3, 2 bit; bitsebesség: **40, 32, 24, 16** kbit/s.

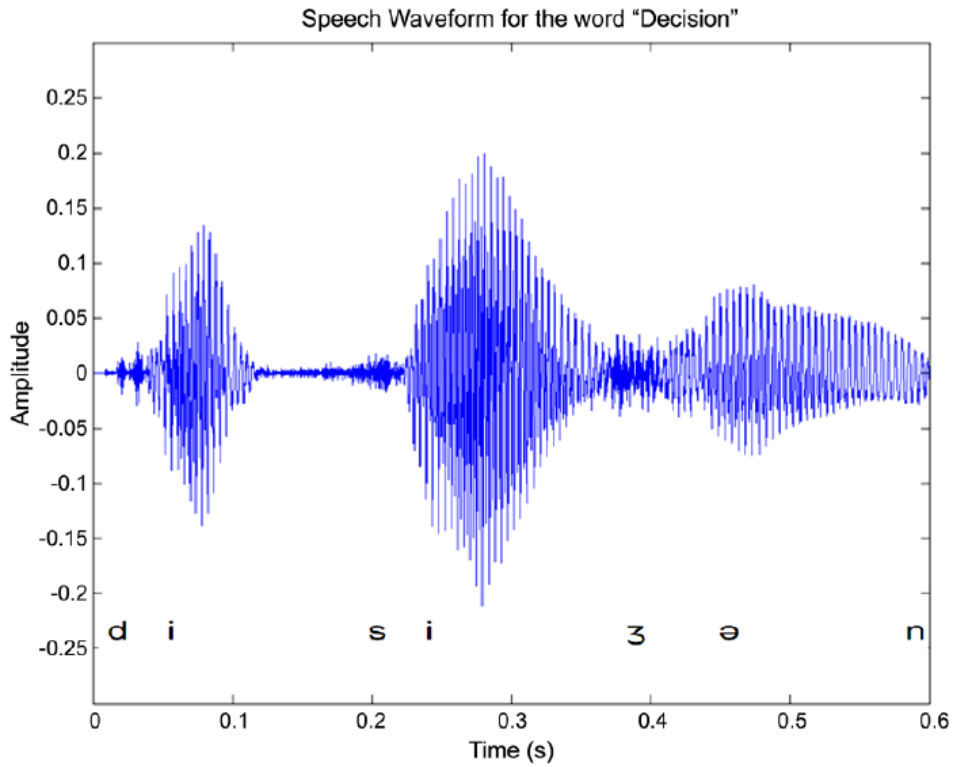
ITU-T G.726 távközlési szabvány (1990): leváltja a korábbi G.721 (32 kbit/s, 1984) és G.723 (24 és 40 kbit/s, 1988) szabványokat.

Legelterjedtebb: **32** kbit/s, standard codec a DECT (digital enhanced cordless telecommunications) telefonkészülékekben (pl. Panasonic KX-TG1100).

Hangképzés

A tüdőből áramló levegő megmozgatja a hangszálakat. A hang a gége, garat, száj- és orrüreg, valamint a nyelv által képzett akadályokon való átjutás során alakul ki (hangképző út).

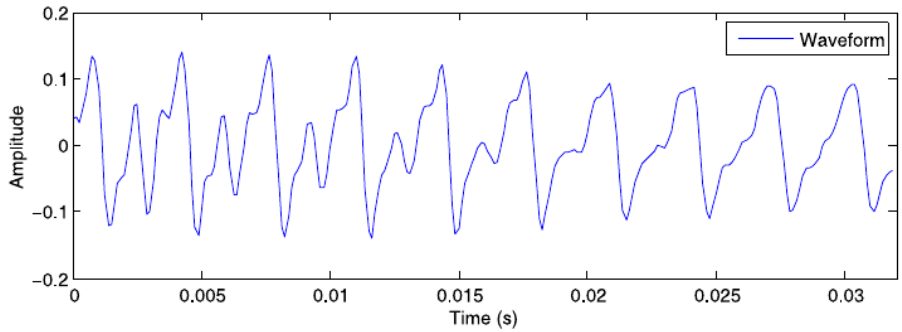
A hangképző út modulálja a hangszálak által képzett zöngét. Mesterséges hangképzésnél egy gerjesztő jelet modulálunk.



A „Decision" szó hullámformája. Forrás: Sun, L., Mkwawa, I.-H., Jammeh, E., Ifeachor, E. *Guide to Voice and Video over IP.* Springer, 2013 (21. old., 2.3 ábra).

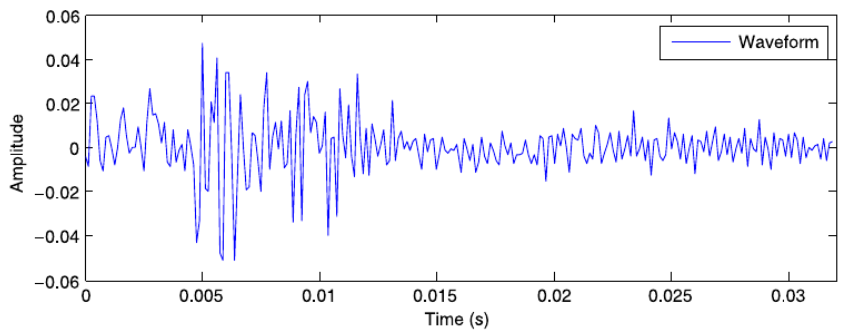
Zöngés és zöngétlen hangok

Zöngés hangok: b, d, dz, dzs, g, gy, j, l, m, n, ny, r, v, z, zs. A hangszalagok egy megadott frekvencián rezegnek (hangmagasság). A beszédminták álperiodikus viselkedést mutatnak.



Zöngés hangminta. Forrás: Sun *et al.* (2013); 22. old., 2.4 ábra.

Zöngétlen hangok: c, cs, f, h, k, p, s, sz, t, ty. A hangszalagok nem rezegnek. A beszédminták zajszerű szerkezetet mutatnak.



Zöngétlen hangminta. Forrás: Sun *et al.* (2013); 23. old., 2.5 ábra.

Paraméteres beszédkódolás

A beszéd rövid, kb 20 ms-os, szegmensekben stacionáriusnak tekinthető, a hangképző út alakja ilyenkor konstans.

Egy-egy stacionárius szegmensben a hangképzű utat egy szűrővel modellezhetjük.

A kódoló analizálja az egyes szegmenseket:

- eldönti zöngés, vagy zöngétlen;
- meghatározza a hangképző szűrő paramétereit;
- megbecsli a gerjesztő jel energiáját, valamint zöngés esetben a hangmagasságát (férfiak: ≈ 125 Hz; nők: ≈ 250 Hz).

A paraméterek binárisan kódolva jutnak el a dekódolónak, ami ezek alapján állítja elő a saját gerjesztő jelét és állítja be a szűrő paramétereit.

- A paraméteres kódolók bonyolultabbak, mint a hullámformán alapulók.
- A hangminőségük lényegesen rosszabb, érthető, de gépi hang jellegű beszédet hallunk.
- Nagyon alacsony bitsebességgel működnek: $\mathbf{1.2 - 4.8}$ kbit/s.

Lineáris prediktív kódolás

LPC: linear predictive coding, Bishnu S. Atal, Bell Labs, 1971.

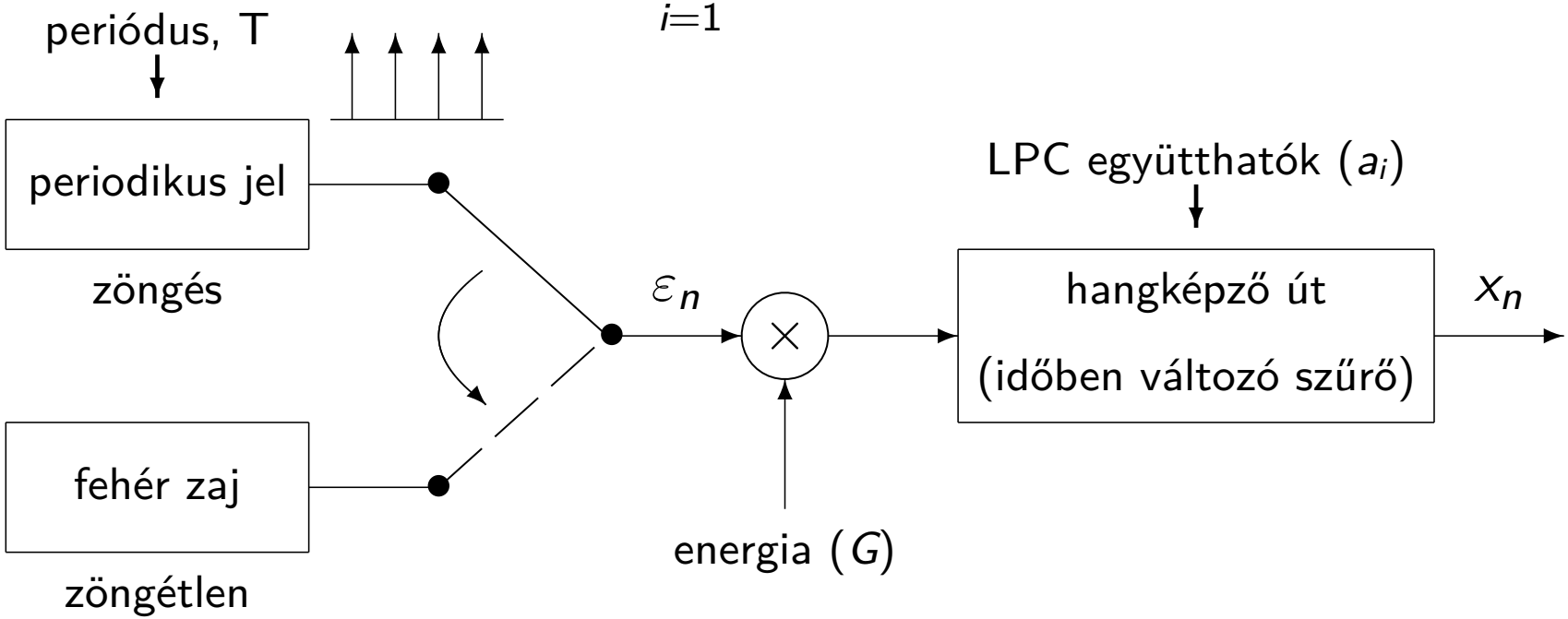
p-lépéses lineáris szűrőt alkalmaz.

ε_n: gerjesztő jel. Zöngés esetben adott frekvenciájú periodikus, zöngétlen esetben fehér zaj (véletlen, független, stacionárius).

G: a szűrő energiája.

x_n: kimenő beszédjel.

$$x_n = \sum_{i=1}^{p} a_i x_{n-i} + G\varepsilon_n.$$



LPC-10 algoritmus

Department of Defence, USA. Federal Standard (FS) 1015 (1984). Főleg titkosított kommunikációnál használatos.

Egy szegmens hossza: 22.5 ms. Továbbítandó információ: **54** bit.

- Gerjesztés típusa (zöngés/zöngétlen): **1** bit.
- Hangmagasság (periódus): **6** bit (logaritmikus karakterisztikájú kompanderes kvantáló).
- Szűrőparaméterek: **41** bit. Érzékeny az 1 közeli együtthatók hibájára. a_1 és a_2 helyett $g_i = (1 + a_i)/(1 - a_i)$, $i = 1, 2$, kvantálódik.
  - Zöngés eset: 10-edrendű prediktív szűrő. Egyenletes kvantáló, g_1, g_2, a_3, a_4: 5 bit; $a_5, \ldots, a_8$: 4 bit; a_9: 3 bit, a_{10}: 2 bit.
  - Zöngétlen eset: 4-edrendű prediktív szűrő. Egyenletes kvantáló, g_1, g_2, a_3, a_4: 5 bit; hibajavítás: 21 bit.
- Energia: **5** bit (logaritmikus karakterisztikájú kompanderes kvantáló).
- Szinkronizáció: **1** bit.

Bitsebesség: 54 bit/22 ms = **2.4** kbit/s. 26.7-szeres kompresszió a PCM-hez képest. Továbbfejlesztett változat: akár **800** bit/s.

Kóddal gerjesztett lineáris prediktív kódoló

Szintézis alapú kódoló (AbS: analysis-by-synthesis): a kódolóban van egy dekódoló is, ami szintetizálja a bemenő beszédjelet. Egy zárt optimalizáló hurokban meghatározza azt a gerjesztő jelet, ami minimalizálja az eredeti és a szintetizált jel közötti hibát. Ennek a paramétereit továbbítja a dekódolónak.

Kóddal gerjesztett lineáris prediktív kódoló (CELP: code excited linear prediction) Manfred R. Schroeder és Bishnu S. Atal, 1985.

Az optimális gerjesztőjelet egy 256 – 1024 elemű kódkönyből választja és annak adatait küldi át a dekódolónak.

Jó minőségű beszédátvitel már **4.8** kbit/s sebességnél.

Lassú a keresés a kódkönyvben, ezért azt részekre bontják.

Eredeti algoritmus (Schroeder és Atal, 1983) Cray-1 szuperszámítógépen (80 MFLOPS; DE HPC: 263.6 TFLOPS): 1s beszéd kódolása 150s idő.

Szabványok: ITU-T G.728 (**16** kbit/s)

Alkalmazás: része a RealAudio és az MPEG-4 Audio formátumnak.

Hangtömörítés

CD minőség: 20 kHz-es sáv, 44100 Hz-es mintavétel, 16 bites egyenletes kvantáló (pl. WAV: waveform audio file format, 1991).

Bitsebesség: $44100 \times 16 \times 2 \approx \mathbf{1400}$ kbit/s ($\times 2$: sztereó).

Hangtömörítésnél figyelembe vett tényezők.

- A hallás legnagyobb érzékenysége $2 - 4$ kHz között van. Kevésbé érzékeny részen ugyanolyan intenzitásu hang halkabbnak tűnik, itt jobban elviseljük a torzítást.
- Adott frekvenciájú nagy intenzitású hang elfedi a vele egyszerre szóló közeli frekvenciákon lévő kisebb intenzitásuakat.
- Adott frekvenciájú nagy intenzitású hang már a bekapcsolása előtt (kb. 2 ms-ig) és kikapcsolása után (kb. 15 ms-ig) is maszkolja a közeli frekvenciákon lévő alacsonyabb intenzitásuakat.

A tömörítendő hangot a frekvenciatartományban analizálják.

Nem minden frekvenciaösszetevőt kell átvinni, és a kódolandókat sem kell azonos pontossággal kvantálni.

MPEG-2 Audio Layer III (MP3) tömörítés, I

Input jel: tömörítetlen PCM audio (pl. WAV file), amit 1152 hosszú blokkokra osztanak (frame).

Két folyamat indul párhuzamosan.

1a. A bemenő minták egy szűrőbankon átvezetve azt 32 egyforma frekvenciasávra bomlanak. 44.1 kHz-es mintavételnél mindegyik sáv $22050/32 = 689$ Hz széles.

1b. Gyors Fourier transzformáció (FFT): a bemenő mintákat az időtartományból átviszi a frekvenciatartományba.

2a. A minták időkorlátos volta miatt a részsávokat ablakolni kell, az MPEG formátum 4 ablakfüggvény típust használ. Ablakolás után az egyes részsávok módosított diszkrét koszinusz transzformáció (MDCT) segítségével 18 további sávra bomlanak. A végeredmény 576 frekvenciaösszetevő.

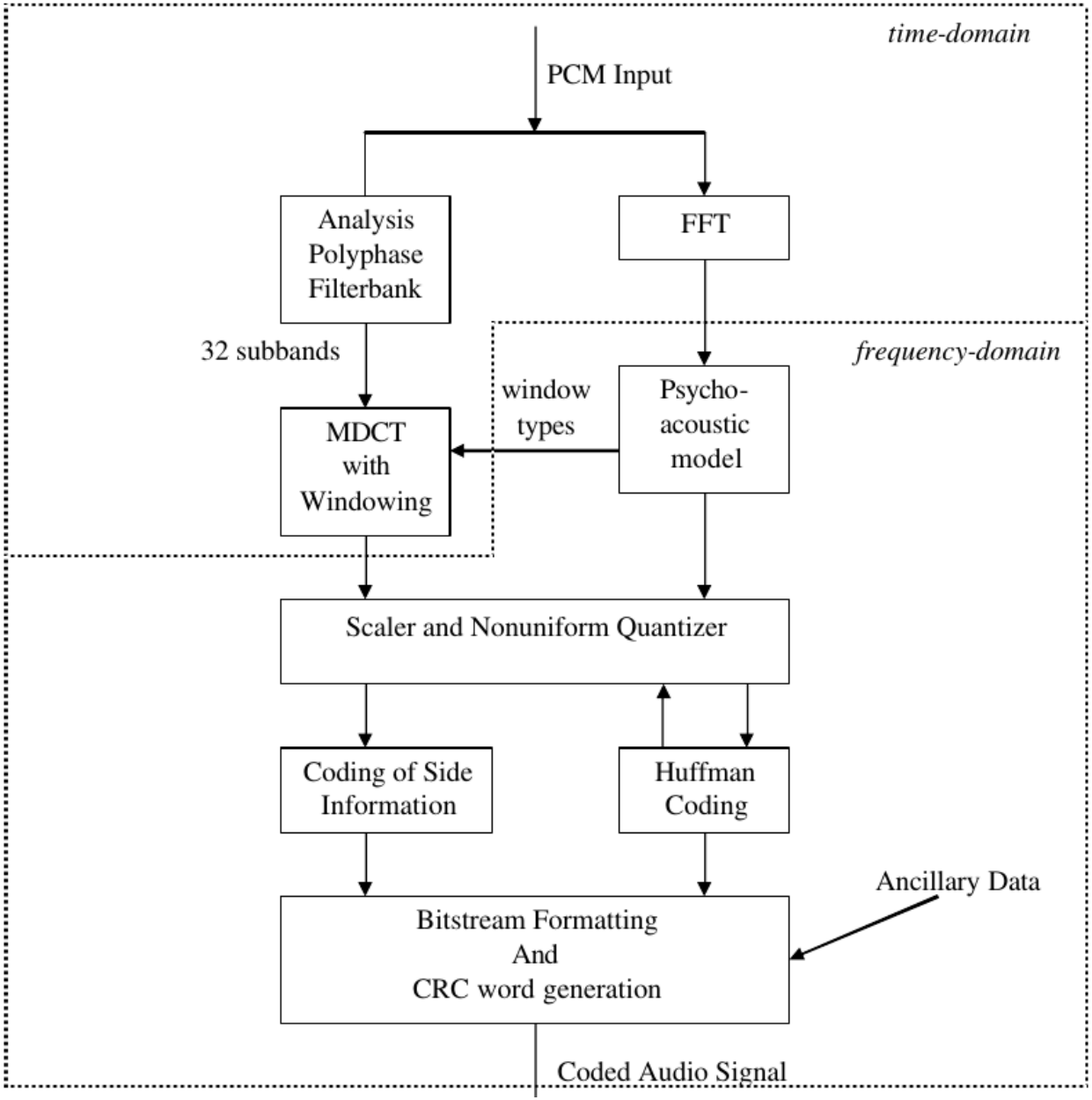
2b. Pszichoakusztikus modell: az emberi hallást szimulálja. Megadja milyen információk hagyhatóak ki a maszkolás miatt, milyen típusú ablakot használjon az MDCT és hogyan kvantálód- janak az egyes frekvenciasávok.

MPEG-2 Audio Layer III (MP3) tömörítés, II

A már a frekvenciatartományban lévő két párhuzamosan futó folyamat összekapcsolódik.

3. A pszichoakusztikus modell információi alapján az 576 frekvenciaösszetevő 22 sávra bomlik és sávonként más-más faktorral skálázva kvantálódik. Itt zajlik a maszkolás.
4. Huffman-kódoló: kódolja a kvantált jeleket. Fix bitsebességnél minden egyes blokkot ugyanannyi bájton kódol. Változó bitsebességnél (VBR) egyes blokkokat rövidebben kódol, a fennmaradó biteket pedig átadja a következő blokknak.
5. Kiegészítő információk kódolása: a tömörítés során kapott és a dekódolónak átadandó információkat kódolja.
6. Multiplexer: a frame fejlécét, CRC (cyclic redundancy code) szavát (hibajavítás), a kiegészítő információkat és a frekvenciaösszetevők kódjait egy továbbítható bitfolyammá rakja össze.

Az MP3 tömörítés vázlata



Forrás: Rassol Raissi, *The theory behind mp3.* www.mp3-tech.org

Az emberi látás

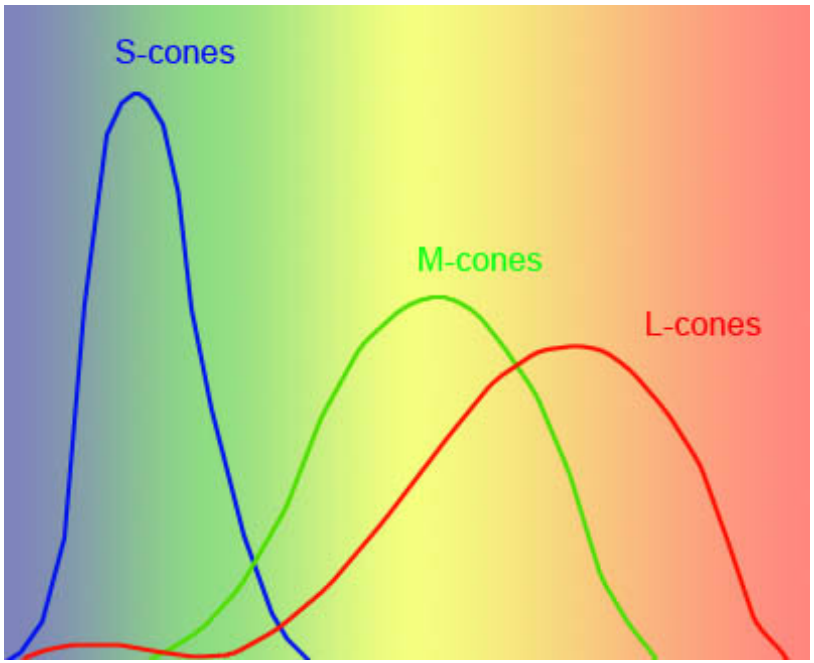
Szín és fényesség érzékelés: pálcikák és csapok.

Pálcikák: perifériás látásnál és kis fényességnél játszanak szerepet.

Csapok: háromféle különböző spektrális érzékenység a látható színtartományban ($\overline{s}(\lambda)$: small; $\overline{m}(\lambda)$: medium; $\overline{\ell}(\lambda)$: large).

Egy $\mathcal{L}(\lambda)$ energiasűrűségű fénysugár által kiváltott inger:

$$(S, M, L)^\top = \int \left(\overline{s}(\lambda), \overline{m}(\lambda), \overline{\ell}(\lambda)\right)^\top \mathcal{L}(\lambda) \mathrm{d}\lambda.$$



Ebből alakul ki a fényesség- és színérzet.

Metamer színek: azonos S, M és L ingert kiváltó spektrumok.

Ezeket nem tudjuk megkülönböztetni.

Színkoordináta-rendszer

Fényesség és színérzet elkülönítése:

$$(X, Y, Z)^\top = \boldsymbol{M}(S, M, L)^\top.$$

Y: fényesség; X és Z: színérzet.

$\boldsymbol{M}$: olyan lineáris transzformáció, ami mindig nemnegatív elemű vektort eredményez.

Gyakorlatban $(X, Y, Z)^\top$ helyett

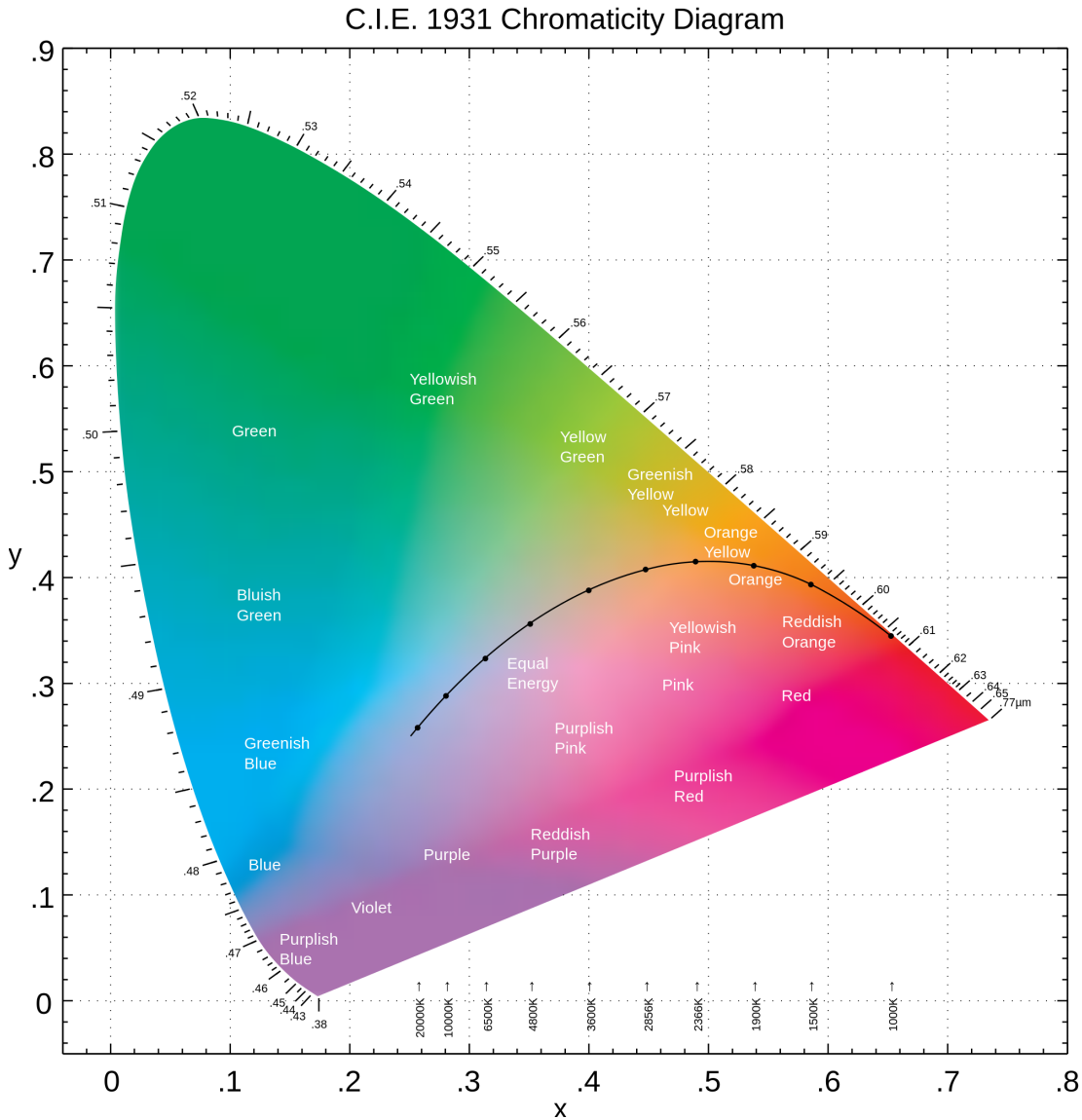
$$(Y, x, y)^\top, \qquad \text{ahol} \quad x = \frac{X}{X+Y+Z}, \quad y = \frac{Y}{X+Y+Z}.$$

Színérzet ábrázolása: (x, y) színkoordináta rendszer.

Ha $\mathcal{L}_i(\lambda)$ spektrumhoz az (x_i, y_i), $i = 1, 2$, pont tartozik, akkor $\mu_1\mathcal{L}_1(\lambda) + \mu_2\mathcal{L}_2(\lambda)$, $\mu_1, \mu_2 > 0$, képe az (x_1, y_1) és (x_2, y_2) közötti szakaszon van.

Monokromatikus spektrum: egyetlen λ_0 összetevőből áll. A monokromatikus spektrumok egy görbét alkotnak. Ez határolja a lehetséges (x, y) értékek halmazát (színpatkó).

Színpatkó



International Commission on Illumination (CIE: Commission internationale de l'éclairage), 1931.

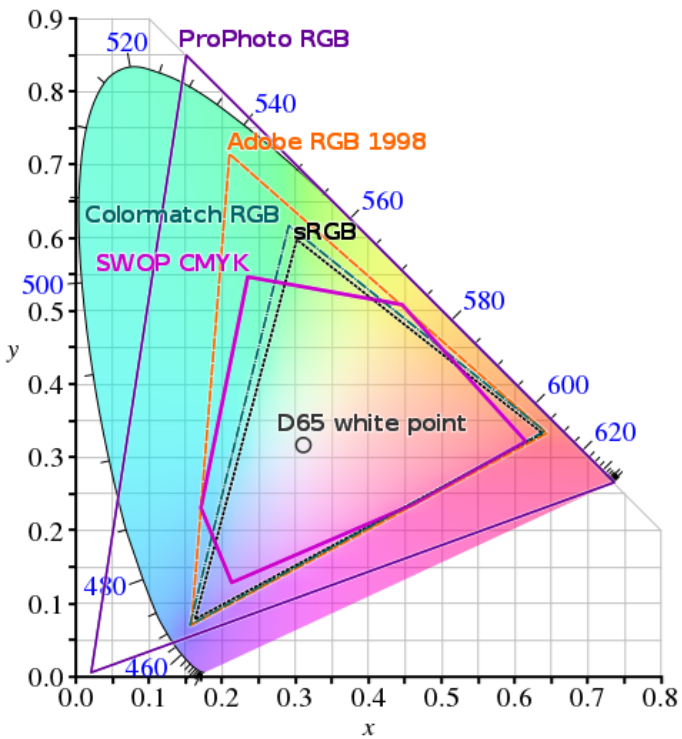
Az RGB színtér

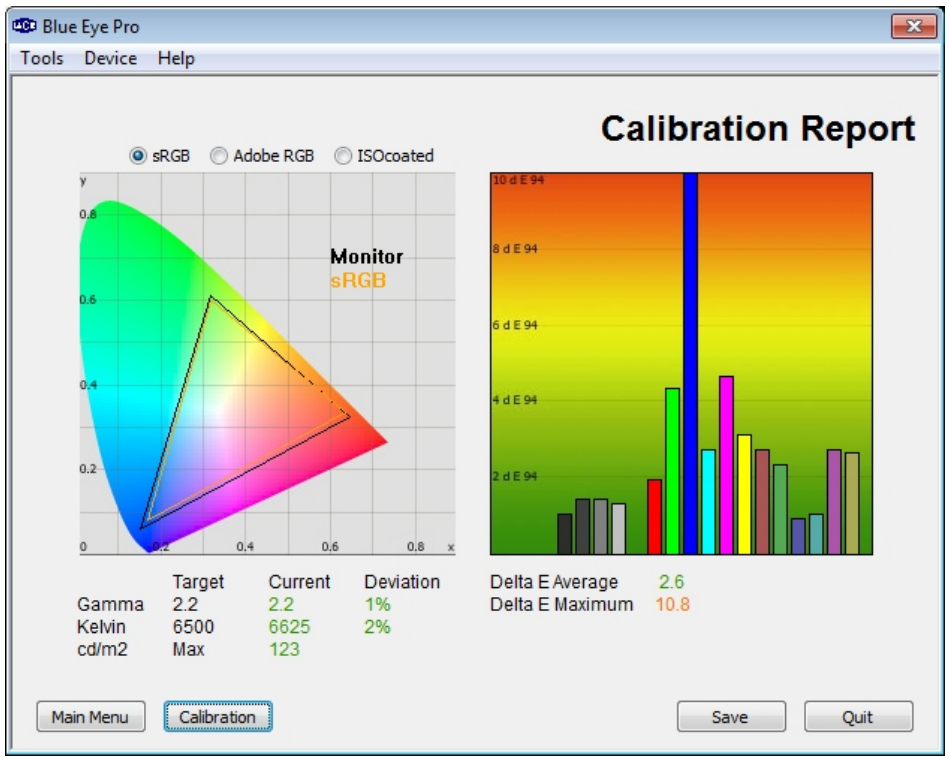
Három szín (R: piros; G: zöld; B: kék) additív keverésével dolgozik.

Az alapszínek által meghatározott háromszög színei ábrázolhatóak.

(R', G', B'): az alapszínek aránya $0 - 255$ egész skálán. 2^{24} szín.

Az (R', G', B') értékek a gamma korrekció után kapott arányok.





Különböző RGB színterek (bal) és az LG 42LB731V TV színkalibrációja (jobb).

Az YCbCr színtér

(Y, x, y): különválasztja a fényességet és a színinformációt.

Probléma: a szem nem egyenletesen érzékel az Y koordinátában.

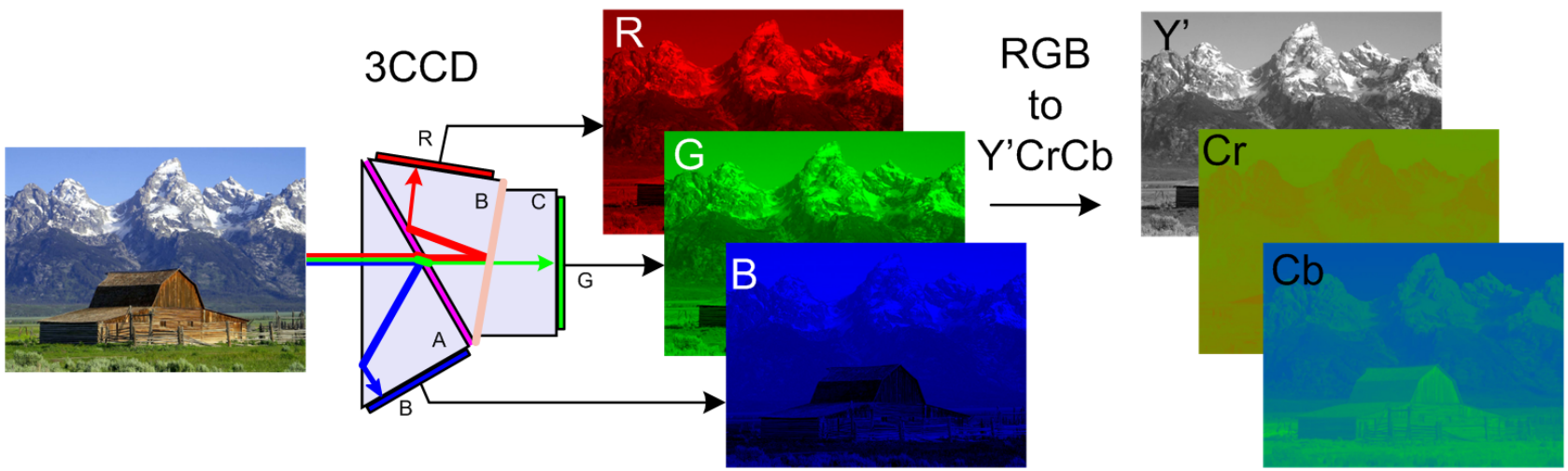
$$\begin{pmatrix} Y' \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \frac{1}{256} \begin{pmatrix} 65.728 & 129.057 & 25.064 \\ -37.945 & -74.494 & 112.439 \\ 112.439 & -94.154 & -18.285 \end{pmatrix} \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix}.$$

Y': luminancia. Egyenletesen érzékeli a szem.

C_b, C_r: krominancia. A kéktől és a pirostól való eltérést adja meg.

Mindhárom koordináta $0 - 255$ egész skálán.

ITU-R BT.601 SDTV szabvány (korábban CCIR 601, 1982).



Graphics Interchange Format (GIF)

Indexelt tárolás: a kép színeit egy palettával írják le. Az egyes képpontok színeinél a palettabeli indexet adják meg.

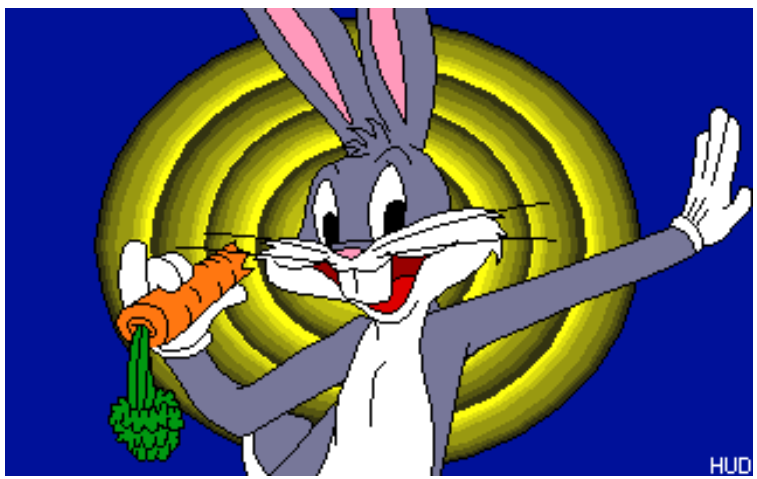
GIF: maximum 8 bites paletta (256 szín) a 24 bites RGB színtérből (3×8 bit). CompuServe, 1987.

Sorfolytonos letapogatás, LZW tömörítés. Veszteségmentes.

Fő alkalmazási terület: kis ikonok, ábrák tömörítése. Jellemzőik:

- Kevés szín.
- Sok a nagy, egyszínű terület és az ismétlődő részlet. Jól tömöríthető az LZW algoritmussal.

Probléma: nem alkalmas pl. fényképek tömörítésére.



Joint Photographic Experts Group (JPEG), I

Veszteséges kódolás (1993). Input kép: YCbCr színtér $8-8$ biten.

A kép a koordináták szerint 3 képsíkra bomlik. Mindegyik sík külön kódolódik. A síkok eltérő felbontása, ha az arány racionális, megengedett. A felbontások legkisebb közös többszörösének megfelelő felbontású kép állítódik vissza. C_b és C_r vízszintes és függőleges felbontása vagy az Y' megfelelő felbontása, vagy annak fele ($4:4:4$: egyforma; $4:2:2$: H feleződik; $4:2:0$: HV feleződik).

1. A képsíkokat 8×8-as négyzetekre bontjuk. Ha a képméret nem osztható 8-cal, a legalsó oszlop/legszélső sor ismétlésével kiegészítjük.
2. A 8×8-as négyzeteket kétdimenziós DCT transzformációval transzformáljuk. A transzformált négyzet elemei: a különböző frekvenciákhoz tartozó felharmonikusok. Bal felső sarok: alacsony frekvenciák, amikre a szem érzékenyebb. $(0,0)$ elem: a 0 frekvencia együtthatója (DC komponens). Négyzetenként gyakran hasonló.
3. A DC komponens különbségi kódolással tömörítődik. Meghatározzuk az előző négyzet DC komponensétől való eltérést.

Joint Photographic Experts Group (JPEG), II

4. A különböző felharmonikusok egyenletesen, de más-más lépésközzel kvantálódnak. Amikre érzékenyebb a szem, azok finomabban.
   Kvantálási tábla: egy az egyes kvantálási lépésközöket megadó mátrix. Függ a tömörítés mértékétől. Nagyobb tömörítési ráta – nagyobb értékek. Pl. 50 %-os tömörítésnél az ajánlott

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 102 & 100 & 103 & 99 \end{bmatrix}$$

Példa

Input mátrix :

$$\begin{bmatrix} 187 & 130 & 113 & 31 & 19 & 125 & 69 & 112 \\ 170 & 52 & 52 & 162 & 207 & 206 & 149 & 51 \\ 188 & 129 & 48 & 160 & 228 & 15 & 185 & 92 \\ 36 & 25 & 26 & 210 & 166 & 217 & 105 & 246 \\ 38 & 149 & 210 & 189 & 198 & 200 & 45 & 30 \\ 114 & 237 & 37 & 222 & 49 & 193 & 168 & 236 \\ 186 & 115 & 251 & 183 & 197 & 22 & 43 & 87 \\ 63 & 112 & 216 & 135 & 47 & 139 & 130 & 22 \end{bmatrix}$$

DCT mátrix :

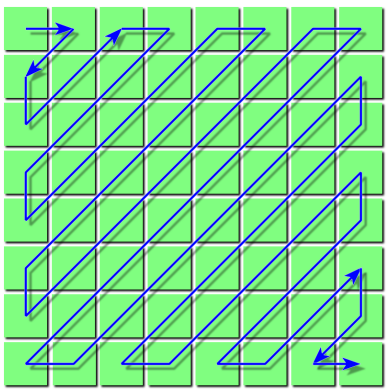
$$\begin{bmatrix} 1021.7 & 16.6 & -104.4 & 24.3 & 43.5 & 21.8 & 0.2 & -57.2 \\ -40.4 & -61.6 & 74.2 & 149.4 & 73.5 & -4.0 & -18.8 & 13.9 \\ -83.2 & 137.5 & 94.3 & 6.4 & -85.3 & 67.3 & 56.8 & 82.1 \\ 29.1 & 47.7 & 74.0 & -32.9 & -56.8 & -61.2 & 52.7 & -100.2 \\ -86.7 & -40.2 & -46.0 & -40.5 & -114.0 & -30.3 & 121.6 & -42.8 \\ -13.0 & -1.0 & 96.6 & -76.6 & 113.9 & -20.8 & 17.1 & 33.1 \\ -2.3 & -20.4 & 157.5 & -26.4 & -49.9 & 7.5 & -102.6 & -72.7 \\ -25.4 & 198.6 & -71.4 & -27.9 & -13.1 & -16.5 & -14.5 & 168.8 \end{bmatrix}$$

Kvantált DCT :

$$\begin{bmatrix} 1024 & 22 & -100 & 32 & 48 & 40 & 0 & -61 \\ -36 & -60 & 70 & 152 & 78 & 0 & 0 & 0 \\ -84 & 143 & 96 & 0 & -80 & 57 & 69 & 56 \\ 28 & 51 & 66 & -29 & -51 & -87 & 80 & -124 \\ -90 & -44 & -37 & -56 & -136 & 0 & 103 & -77 \\ -24 & 0 & 110 & -64 & 81 & 0 & 0 & 0 \\ 0 & 0 & 156 & 0 & 0 & 0 & -120 & -101 \\ 0 & 184 & -95 & 0 & 0 & 0 & 0 & 198 \end{bmatrix}$$

Joint Photographic Experts Group (JPEG), III

5. A különbségi DC együtthatót és a kvantált DCT értékeket cikk-cakk elrendezés szerint sorbarendezzük.
6. A kapott sorozatot részsorozatokra bontjuk. Minden részsorozat eleje 0-kból áll és egy nem 0 elem zárja. Egy részsorozat futamhossz kódja $\big(\{n, s\}, \nu\big)$.
   n: a részsorozat elején lévő 0-futam hossza;
   s: a sorozat végi nem 0 elem kódolásának bithossza;
   ν: a sorozat végi nem 0 elem előjeles bináris alakja.
7. Az $\{n, s\}$ párokat statikus Huffmann kóddal, vagy aritmetikai kódolóval kódoljuk és a kódjaik után felsoroljuk a ν értékek bináris kódjait.



Példa.

$81,\ 0,\ 0,\ 0,\ 0,\ 0,\ -6,\ 0,\ 0,\ 0,\ -12,\ 0,\ 0,\ \ldots$

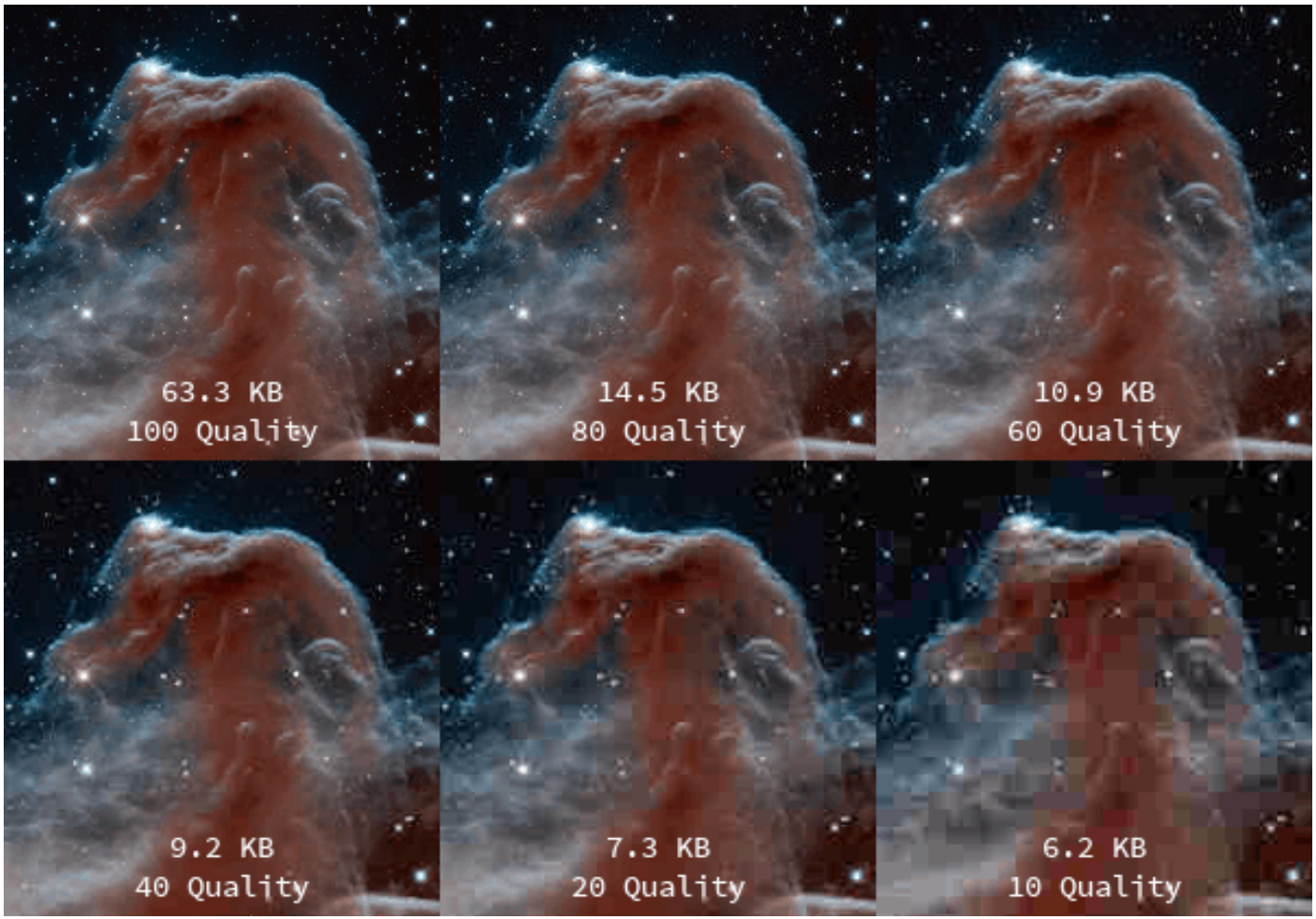
A kód: $\big(\{0, 8\}, 01010001\big);\ \big(\{5, 4\}, 1001\big);$
$\big(\{3, 5\}, 10011\big).$

A Huffman-kód után: 0101000110011 0011.

Tulajdonságok

Szabványok: ISO/IEC 10918, ITU-T T.81, T.83, T.84, T.86

Hatékony, ha nincsenek éles határok. Túl nagy tömörítésnél a kvantálás miatt erős képzaj az éleknél.



Forrás: http://www.gimp.org/tutorials/GIMP_Quickies/

Veszteségmentes JPEG

Kiegészítés a JPEG formátumhoz. Joint Photographic Experts Group, 1993.

Kétdimenziós prediktív kódolás (DPCM).

Sorfolytonos letapogatás. Az (i, j) képpont $X_{i,j}$ értékének $\widehat{X}_{i,j}$ predikciója $X_{i-1,j}$, $X_{i,j-1}$ és $X_{i-1,j-1}$ alapján.

Nyolcféle predikciós séma.

$$0: \quad \widehat{X}_{i,j} = 0; \qquad 4: \quad \widehat{X}_{i,j} = X_{i-1,j} + X_{i,j-1} - X_{i-1,j-1};$$

$$1: \quad \widehat{X}_{i,j} = X_{i-1,j}; \qquad 5: \quad \widehat{X}_{i,j} = X_{i,j-1} + \frac{X_{i-1,j} - X_{i-1,j-1}}{2};$$

$$2: \quad \widehat{X}_{i,j} = X_{i,j-1}; \qquad 6: \quad \widehat{X}_{i,j} = X_{i-1,j} + \frac{X_{i,j-1} - X_{i-1,j-1}}{2};$$

$$3: \quad \widehat{X}_{i,j} = X_{i-1,j-1}; \qquad 7: \quad \widehat{X}_{i,j} = \frac{X_{i-1,j} + X_{i,j-1}}{2}.$$

Bármelyik séma használható, de egy adott képhez csak egyféle.

Adaptív aritmetikai, vagy Huffman kódolás.

Kompressziós arány prediktív esetben (nem 0 séma): kb. $2:1$.

Moving Picture Experts Group (MPEG), I

International Organization for Standardization (ISO) és International Electrotechnical Commission (IEC) munkacsoport.
Szabványok: MPEG-1, MPEG-2, MPEG-4, MPEG-7, MPEG-21.

Veszteséges videotömörítés (1993). MPEG-I részei:

videoszekvencia $\longrightarrow$ képcsoportok $\longrightarrow$ képek (frame) $\longrightarrow$ makroblokkok $\longrightarrow$ blokkok

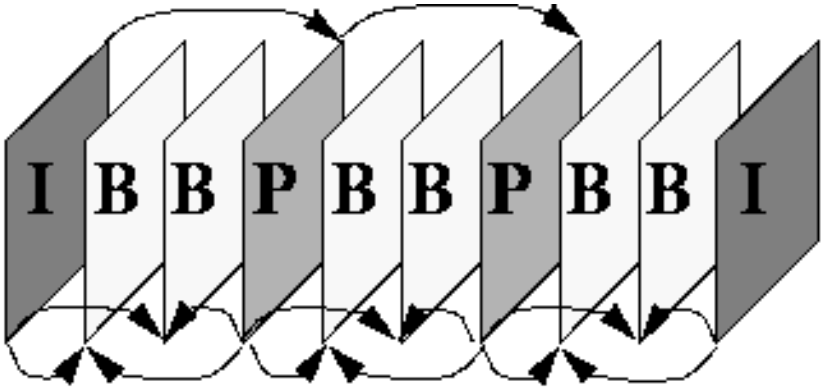
Képcsoport (GOP: group of pictures): függetlenül kódolt egység.

Háromféle képtípus:

- I-típus (intra frame): önálló kép, JPEG tömörítés;
- P-típus (predictive coded frame): az előző I- vagy P-típusú kép segítségével tömörítődik;
- B-típus (bidirectionally predictive coded frame): az előző és/vagy következő I- vagy P-típusú kép segítségével tömörítődik.

Makroblokk: egy kép 16×16 képpont méretű része. 6 darab 8×8 elemű blokkból áll: 4 luminancia (Y') és $1-1$ krominancia (C_r, C_b).
P- és B-típusú képeket makroblokkonként tömörítünk.

Sorrendek



A B-típusú képek kódolásnál előre mozognak, hogy a szomszédos I- és P-típusok után legyenek.

Egyszerűsíti a pufferelést.

Lejátszási sorrend:	0	1	2	3	4	5	6	7	8	9
Képtípus:	I	B	B	P	B	B	P	B	B	I
Sorrend az adatfolyamban:	0	2	3	1	5	6	4	8	9	7

Szokásos sorrend: egy I-típusú képre két P-típusú épül, köztük két B-típusú.

I-típusú képek: önmagukban kódolódnak. Referenciaként szolgálnak a kereséshez.

A P- és B-típusú képek tömörítése a makroblokkonkénti mozgásbecslésen alapszik.

Predikciók

P-típusú képek: minden makroblokkhoz kikeressük az előző I- vagy P-típusú kép leginkább hasonlító makroblokkját (referencia blokk). Csak a távolságot és az irányt (mozgásvektor), valamint a referencia blokktól való eltérést kódoljuk.

Mozgásvektor: Huffmann; eltérés: JPEG-hez hasonló kódolás.
Nincs referencia: makroblokk sima JPEG kódolása.

B-típusú képek: az előző és a rákövetkező I- vagy P-típusú képben keresünk egyezést. Ha mindkét oldalon van egyezés, az átlaguktól való eltérést és a két mozgásvektort kódoljuk. Egyoldali egyezésnél a P-típushoz hasonlóan kódolunk.

MPEG-I tömörítés pl. 356×260 képpont és 24 bites szín esetén.

Típus	Méret	Arány
I	18 Kb	7 : 1
P	6 Kb	20 : 1
B	2.5 Kb	50 : 1
Átlag	4.6 Kb	27 : 1

Video bitsebesség, 30 kép/s: 1.2 Mbit/s; audioval: 1.45 Mbit/s.

Kölcsönös információ

Definíció. *Az X és Y diszkrét valószínűségi változók kölcsönös információján az*

$$I(X;Y) := H(X) + H(Y) - H(X,Y)$$

mennyiséget értjük.

Megjegyzés. A kölcsönös információ szimmetrikus, és

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = I(Y;X).$$

Megjegyzés.

$$\begin{aligned} I(X;Y) &= \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \\ &= \sum_{x,y} p(x,y) \log_2 \frac{p(x|y)}{p(x)} = \sum_{x,y} p(x,y) \log_2 \frac{p(y|x)}{p(y)}. \end{aligned}$$

A kölcsönös információ tulajdonságai

Tétel. *Legyenek X és Y diszkrét valószínűségi változók.*

a) $I(X;Y) \geq 0$ *és* $I(X;Y)$ *pontosan akkor* 0, *ha* X *és* Y *függ- etlen.*

b) $I(X;X) = H(X)$.

c) $I(X;Y) \leq H(X)$ *és* $I(X,Y) \leq H(Y)$.

d) *Az* X *és* Y *bármely* g *és* h *függvényére*

$$I(X;Y) \geq I\big(g(X);h(Y)\big).$$

e) *A következő három állítás ekvivalens:*

i) $I(X;Y) = H(X)$;

ii) $H(X|Y) = 0$;

iii) *létezik olyan* $g : \mathbb{R} \to \mathbb{R}$, *hogy* $\mathrm{P}\big(X = g(Y)\big) = 1$.

Betűnkénti torzítás

$\mathbb{X}$: információforrás.

Az $X_1X_2,\ldots X_k$ blokkot egy $Y_1Y_2\ldots Y_k$ blokkal reprezentáljuk, ahol $X_i\in\mathcal{X},\ Y_i\in\mathcal{Y},\ i=1,2,\ldots,k$.

$d:\mathcal{X}\times\mathcal{Y}\to\mathbb{R}^+$: torzítási mérték.

$d(x,y)$: annak a torzításnak a mértéke, amit az okoz, hogy az $x\in\mathcal{X}$ forrásbetűt az $y\in\mathcal{Y}$ szimbólummal reprezentáljuk.

Az $\boldsymbol{x}=x_1x_2\ldots x_k\in\mathcal{X}^k$ és $\boldsymbol{y}=y_1y_2\ldots y_k\in\mathcal{Y}^k$ blokkok közötti torzítás, amit betűnkénti torzításnak nevezük:

$$d(\boldsymbol{x},\boldsymbol{y})=\frac{1}{k}\sum_{i=1}^{k}d(x_i,y_i).$$

Reprezentálja az $\mathbb{X}$ forrás $\boldsymbol{X}=(X_1,X_2,\ldots,X_k)$ blokkját az $\boldsymbol{Y}=(Y_1,Y_2,\ldots,Y_k)$ blokk.

A betűnkénti átlagos torzítás:

$$\mathsf{E}(d)=\mathsf{E}\big(d(\boldsymbol{X},\boldsymbol{Y})\big)=\sum_{\boldsymbol{x}\in\mathcal{X}^k}\sum_{\boldsymbol{y}\in\mathcal{Y}^k}p(\boldsymbol{x},\boldsymbol{y})d(\boldsymbol{x},\boldsymbol{y}).$$

Példák

1. Legyen $\mathcal{X} = \mathcal{Y}$ és d a Hamming-torzítás:

$$d(x, y) := \begin{cases} 0, & \text{ha } x = y; \\ 1, & \text{ha } x \neq y. \end{cases}$$

Mivel

$$\mathsf{E}\big(d(X_i, Y_i)\big) = \mathsf{P}(X_i \neq Y_i), \qquad i = 1, 2, \ldots, k,$$

a betűnkénti átlagos torzítás

$$\mathsf{E}\big(d(\boldsymbol{X}, \boldsymbol{Y})\big) = \frac{1}{k} \sum_{i=1}^{k} \mathsf{E}\big(d(X_i, Y_i)\big) = \frac{1}{k} \sum_{i=1}^{k} \mathsf{P}(X_i \neq Y_i).$$

2. Legyen $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ és $d := (x - y)^2$.

$$\mathsf{E}\big(d(\boldsymbol{X}, \boldsymbol{Y})\big) = \mathsf{E}\|\boldsymbol{X} - \boldsymbol{Y}\|^2.$$

Betűnkénti hűségkritérium

Betűnkénti hűségkritérium: a betűnkénti átlagos torzítás nem haladhat meg egy adott felső határt.

Pl. az $Y_1, Y_2, \ldots, Y_k$ blokkot akkor fogadjuk el az $X_1, X_2, \ldots, X_k$ blokk reprodukciójaként, ha egy adott $\delta > 0$ számra

$$\mathsf{E}\left(\frac{1}{k}\sum_{i=1}^{k} d(X_i, Y_i)\right) < \delta.$$

Forrás adott hűségű kódolására forráskódokat használunk.

$g : \mathcal{X}^k \to \mathcal{Y}^k$: az $\mathbb{X}$ forrás k-hosszú blokkjait kódoló forráskódja.

A g forráskód betűnkénti átlagos torzítása:

$$D(g) = \mathsf{E}\Big(d\big(\boldsymbol{X}, g(\boldsymbol{X})\big)\Big) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) d\big(\boldsymbol{x}, g(\boldsymbol{x})\big).$$

M: a $g : \mathcal{X}^k \to \mathcal{Y}^k$ forráskód értékkészletének elemszáma.

A g forráskód jelsebessége:

$$R = \frac{\log_2 M}{k}.$$

R-D függvény

A forráskódolás célja: a forrást adott hűségű, minél kisebb jelsebességű kóddal kódoljuk.

$\mathcal{C} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_M\}$: a g értékkészlete ($\boldsymbol{y}_i \in \mathcal{Y}^k$, $i = 1, 2, \ldots, M$).

A torzítás akkor a legkisebb, ha g a következő:

$$g(\boldsymbol{x}) = \boldsymbol{y}_i \quad \text{ha} \quad d(\boldsymbol{x}, \boldsymbol{y}_i) \leq d(\boldsymbol{x}, \boldsymbol{y}_j), \quad j = 1, 2, \ldots, M.$$

$\boldsymbol{x} \in \mathcal{X}^k$ kódja az az $\boldsymbol{y} \in \mathcal{C}$, amelyik legjobban „hasonlít" $\boldsymbol{x}$-re. A jó kód keresése ekvivalens a megfelelő $\mathcal{C}$ keresésével.

Definíció. *Legyen adva az $\mathcal{Y}$ reprodukciós ábécé és a d torzítási mérték. Ekkor az $\mathbb{X}$ emlékezetnélküli stacionárius forrás $\delta \geq 0$ számokra értelmezett R-D függvénye (rate-distorsion)*

$$R(\delta) := \min\Big\{ I(X;Y) : \mathrm{E}\big(d(X,Y)\big) \leq \delta \Big\},$$

ahol a minimumot az összes olyan (X, Y) valószínűségi változópár fölött kell venni, ahol X eloszlása megegyezik a forrás X_i elemeinek közös eloszlásával, Y pedig az $\mathcal{Y}$ halmazból veszi az értékeit. Ha nincs olyan Y, melyre $\mathrm{E}\big(d(X,Y)\big) \leq \delta$, akkor $R(\delta) := \infty$.

Az R-D függvény tulajdonságai

Legyen

$$\delta_{\min} := \sum_x p(x) \min_y d(x,y).$$

Tetszőleges Y-ra

$$\mathrm{E}\big(d(X,Y)\big) = \sum_{x,y} p(x,y) d(x,y) \geq \sum_{x,y} p(x,y) \min_y d(x,y) = \delta_{\min}.$$

Megjegyzés. Az $R(\delta)$ pontosan akkor véges, ha $\delta \geq \delta_{\min}$.

Lemma. *Ha $\delta \geq \delta_{\min}$, akkor $R(\delta)$ a δ monoton csökkenő és konvex függvénye.*

Lemma. *Amennyiben az $\mathbb{X}$ emlékezetnélküli stacionárius forrás $\boldsymbol{X} = (X_1, X_2, \ldots, X_k)$ blokkja és az $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_k)$ véletlen vektor kielégíti a*

$$\mathrm{E}\left(\frac{1}{k}\sum_{i=1}^{k} d(X_i, Y_i)\right) < \delta$$

hűségkritériumot, akkor $I(\boldsymbol{X}; \boldsymbol{Y}) \geq kR(\delta)$.

A forráskódolási tétel megfordítása

Tétel. *Ha az $\mathbb{X}$ emlékezetnélküli és stacionárius forrás M különböző kódszót használó $g : \mathcal{X}^k \to \mathcal{Y}^k$ forráskódja valamely $\delta \geq \delta_{\min}$ számra kielégíti a*

$$D(g) \leq \delta$$

hűségkritériumot, akkor a kód $R = \frac{\log_2 M}{k}$ jelsebességére

$$R \geq R(\delta).$$

Egy δ-nál kisebb torzítású forráskód jelsebessége nem lehet kisebb, mint $R(\delta)$.

Indoklás. A kölcsönös információ és az entrópia tulajdonságai miatt

$$I\big(\boldsymbol{X}; g(\boldsymbol{X})\big) \leq H\big(g(\boldsymbol{X})\big) \leq \log_2 M.$$

Másrészt

$$I\big(\boldsymbol{X}; g(\boldsymbol{X})\big) \geq kR(\delta),$$

amiből adódik, hogy

$$\log_2 M \geq kR(\delta).$$

□

Forráskódolási tétel emlékezetnélküli stacionárius forrásokra

Tétel. *Legyen $R(\delta)$ az $\mathbb{X}$ emlékezetnélküli stacionárius forrás R-D függvénye egy adott $\mathcal{Y}$ reprodukciós ábécére és d torzítási mértékre. Ekkor, ha $\delta \geq \delta_{\min}$, akkor minden $\delta' > \delta$ és $R' > R(\delta)$ esetén elég nagy k-ra létezik egy, az $\mathbb{X}$ forrás k-hosszú blokkjait M darab kódszóval kódoló g forráskód, amelyre*

$$M \leq 2^{kR'}, \quad \text{azaz} \quad \frac{\log_2 M}{k} \leq R',$$

és

$$D(g) < \delta'.$$

Elég hosszú blokkokat használva a kódsebesség $R(\delta)$ alsó határa tetszőlegesen megközelíthető a δ-hoz tetszőlegesen közeli torzítású forráskóddal.

Példa

Legyen $\mathcal{X} = \mathcal{Y}$ és d a Hamming-torzítás:

$$d(x,y) := \begin{cases} 0, & \text{ha } x = y; \\ 1, & \text{ha } x \neq y. \end{cases}$$

$\mathbb{X}$: emlékezetnélküli és stacionárius forrás adott $p(x) := \mathrm{P}(X_i = x)$ eloszlással. Számítsuk ki az $R(0)$ értékét!

Megoldás. A Hamming torzításra $\mathrm{E}\big(d(X,Y)\big) = \mathrm{P}(X \neq Y)$. A $\delta = 0$ esetben $\mathrm{E}\big(d(X,Y)\big) \leq 0$ ekvivalens a $\mathrm{P}(X = Y) = 1$ feltétellel, azaz Y egy valószínűséggel meghatározza X-et. A kölcsönös információ tulajdonságai miatt $I(X;Y) = H(X) = H(\mathbb{X})$. Így $R(0) := \min\big\{I(X;Y) : \mathrm{E}\big(d(X,Y)\big) \leq 0\big\} = H(\mathbb{X})$.

A forráskódolási tétel alapján tetszőleges $\delta > 0$ számra találhatunk olyan $g : \mathcal{X}^k \to \mathcal{Y}^k$ forráskódot, hogy

$$\frac{1}{k}\sum_{i=1}^{k} \mathrm{P}(X_i \neq Y_i) < \delta, \quad \text{ahol} \quad (Y_1, Y_2, \ldots, Y_k) = g(X_1, X_2, \ldots, X_k),$$

és a kód jelsebessége megközelíti a forrás entrópiáját. A kódoló által használt kódszavak száma: $M \approx 2^{kH(\mathbb{X})}$.

Hamming-távolság

$\boldsymbol{u}$, $\boldsymbol{u'}$: egy $\mathcal{X}$ forrásábécé betűiből alkotott k hosszúságú vektorok.

$\boldsymbol{c}$, $\boldsymbol{v}$: egy $\mathcal{Y}$ kódábécé betűiből alkotott n hosszúságú vektorok.

Kódoló: $f : \mathcal{X}^k \to \mathcal{C} \subseteq \mathcal{Y}^n$ invertálható függvény. Az $\boldsymbol{u} \in \mathcal{X}^k$ üzenetet az $\boldsymbol{c} \in \mathcal{C}$ kódszóba képezi le. $\mathcal{C}$ az $\mathcal{X}^k$ kódja.

Egy $\boldsymbol{c} = (c_1, \ldots, c_n)$ bemeneti és $\boldsymbol{v} = (v_1, \ldots, v_n)$ kimeneti sorozat esetén az m-edik pozíciónál a csatorna hibázott, ha $c_m \neq v_m$.

$d(\boldsymbol{c}, \boldsymbol{v})$: azon i pozíciók száma, ahol $c_i \neq v_i$. A $\boldsymbol{c}$ és $\boldsymbol{v}$ vektorok Hamming-távolsága. A csatorna hibáinak száma: $t = d(\boldsymbol{c}, \boldsymbol{v})$. Ez egyszerű hibázás, mert sem a hiba helye, sem az értéke nem ismert.

Megjegyzés. A $d(\boldsymbol{c}, \boldsymbol{v})$ Hamming-távolság metrika, azaz

a) $d(\boldsymbol{c}, \boldsymbol{v}) \geq 0$ és $d(\boldsymbol{c}, \boldsymbol{v}) = 0$ pontosan akkor, ha $\boldsymbol{c} = \boldsymbol{v}$;
b) $d(\boldsymbol{c}, \boldsymbol{v}) = d(\boldsymbol{v}, \boldsymbol{c})$;
c) $d(\boldsymbol{c}, \boldsymbol{v}) \leq d(\boldsymbol{c}, \boldsymbol{w}) + d(\boldsymbol{w}, \boldsymbol{v})$ (háromszög egyenlőtlenség).

Algebrai dekódolás

Dekódolás: egy $g : \mathcal{Y}^n \to \mathcal{C}$ és az $f^{-1} : \mathcal{C} \to \mathcal{X}^k$ egymás utáni alkalmazása. A csatorna kimeneti jeléhez rendel egy forrásüzenetet.

Az f kódoló egyértelműen meghatározza az f^{-1} függvényt. A továbbiakban dekódoló alatt a g függvényt értjük.

Algebrai dekódoló:

$$g(\boldsymbol{v}) = \boldsymbol{c}', \qquad \text{ha} \qquad d(\boldsymbol{c}', \boldsymbol{v}) = \min_{\boldsymbol{c} \in \mathcal{C}} d(\boldsymbol{c}, \boldsymbol{v}).$$

Ha a minimum feltétel több kódszóra is teljesül, akkor tetszőlegesen válasszuk ki az egyiket.

Cél: minden lehetséges $\boldsymbol{v}$ kimenő jelsorozat esetén a hozzá legközelebbi $\boldsymbol{c}$ kódszó meghatározása anélkül, hogy az összes $d(\boldsymbol{c}, \boldsymbol{v})$ távolságot kiszámítanánk.

Definíció. *Egy $\mathcal{C}$ kód kódtávolsága*

$$d_{min} := \min_{\substack{\boldsymbol{c} \neq \boldsymbol{c}' \\ \boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}}} d(\boldsymbol{c}, \boldsymbol{c}').$$

Hibajelzés, hibajavítás

Hibajelzés: a vevőnél csupán detektálni akarjuk a hibázás tényét.

Tétel. *Egy $d_{\min}$ kódtávolságú kód minden legfeljebb $d_{\min} - 1$ számú egyszerű hibát jelezni tud.*

Magyarázat. Egy $\boldsymbol{v}$ vett kódjelsorozat esetén akkor tudjuk a hibázást észrevenni, ha $\boldsymbol{v}$ nem kódszó. Ez biztosan teljesül, ha a $\boldsymbol{c}$ küldött kódszó esetén

$$d_{\min} > d(\boldsymbol{v}, \boldsymbol{c}),$$

azaz a hibák t számára $d_{\min} > t$. □

Hibajavítás: t egyszerű hiba esetén a vett $\boldsymbol{v}$ szóból egyértelműen visszaállítható a küldött $\boldsymbol{c}$ kódszó.

Tétel. *Egy $d_{\min}$ kódtávolságú kód $\left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$ hibát tud javítani.*

Magyarázat. A $\boldsymbol{c}$ küldött kódszó pontosan akkor állítható egyértelműen vissza, ha tetszőleges $\boldsymbol{c}'$ kódszó esetén

$$d(\boldsymbol{v}, \boldsymbol{c}') > d(\boldsymbol{v}, \boldsymbol{c}). \tag{1}$$

A háromszög egyenlőtlenség miatt $d(\boldsymbol{v}, \boldsymbol{c}') \geq d(\boldsymbol{c}, \boldsymbol{c}') - d(\boldsymbol{v}, \boldsymbol{c})$. Az (1) egyenlőtlenség biztosan teljesül, ha $d(\boldsymbol{c}, \boldsymbol{c}') - d(\boldsymbol{v}, \boldsymbol{c}) > d(\boldsymbol{v}, \boldsymbol{c})$, azaz minden $\boldsymbol{c} \neq \boldsymbol{c}'$ esetén $d(\boldsymbol{v}, \boldsymbol{c}') > d(\boldsymbol{v}, \boldsymbol{c})$, tehát $d_{\min}/2 > d(\boldsymbol{v}, \boldsymbol{c})$. □

Példa

$\mathcal{X} = \mathcal{Y} = \{0,1\},\ k = 2,\ n = 5$. Kódoló ($f$):

$$00 \overset{f}{\mapsto} \boldsymbol{c_1} = 00000 \qquad 01 \overset{f}{\mapsto} \boldsymbol{c_2} = 01101 \qquad 10 \overset{f}{\mapsto} \boldsymbol{c_3} = 10110 \qquad 11 \overset{f}{\mapsto} \boldsymbol{c_4} = 11011$$

Dekódoló (g és f^{-1}):

$$\left.\begin{matrix}00000\\10000\\01000\\00100\\00010\\00001\end{matrix}\right\} \overset{g}{\mapsto} 00000 \overset{f^{-1}}{\mapsto} 00 \qquad \left.\begin{matrix}01101\\11101\\00101\\01001\\01111\\01100\end{matrix}\right\} \overset{g}{\mapsto} 01101 \overset{f^{-1}}{\mapsto} 01 \qquad \left.\begin{matrix}10110\\00110\\11110\\10010\\10100\\10111\end{matrix}\right\} \overset{g}{\mapsto} 10110 \overset{f^{-1}}{\mapsto} 10$$

$$\left.\begin{matrix}11011\\01011\\10011\\11111\\11001\\11010\end{matrix}\right\} \overset{g}{\mapsto} 11011 \overset{f^{-1}}{\mapsto} 11 \qquad \left.\begin{matrix}00011\\01010\\10001\\11000\end{matrix}\right\} \overset{g}{\mapsto} 00000 \overset{f^{-1}}{\mapsto} 00 \qquad \left.\begin{matrix}00111\\01110\\10101\\11100\end{matrix}\right\} \overset{g}{\mapsto} 01101 \overset{f^{-1}}{\mapsto} 01$$

Első négy blokk esetén a távolság 1, utolsó kettőnél 2.

Kódtávolság: $d_{\min} = 3$. 2 hibát tud jelezni, 1 hibát tud javítani.

Példa

$\mathcal{X} = \mathcal{Y} = \{0, 1\}$, $k = 2$, $n = 3$. Kódoló (f):

$u_1\,u_2$		$c_1\,c_2\,c_3$	
0 0	$\mapsto$	0 0 0	$\boldsymbol{c_1}$
0 1	$\mapsto$	0 1 1	$\boldsymbol{c_2}$
1 0	$\mapsto$	1 0 1	$\boldsymbol{c_3}$
1 1	$\mapsto$	1 1 0	$\boldsymbol{c_4}$

Az $u_1 u_2$ forrásüzenet egy paritásbittel lett kiegészítve.

Ha $\boldsymbol{v} = v_1 v_2 v_3 \in \mathcal{Y}^3$ nem kódszó, akkor három kódszóból is megkapható 1 bit megváltoztatásával (1 hibával).

Kódtávolság: $d_{\min} = 2$. 1 hibát tud jelezni, 0 hibát tud javítani.

Törléses hiba

A hiba helyét ismerjük, de az értékén nem.

Tétel. *Egy $d_{\min}$ kódtávolságú kód $d_{\min} - 1$ törléses hibát tud javítani.*

Magyarázat. Mivel a kódszavak legalább $d_{\min}$ pozícióban különböznek, nem fordulhat elő, hogy a $\boldsymbol{c}$ és $\boldsymbol{c}'$ kódszavak ugyanazon, legfeljebb $d_{\min} - 1$ pozíciójának törlésével ugyanazt a szót kapjuk. □

Singleton-korlát

$A_s(n, d_{\min})$: egy s elemű kódábécé elemeiből alkotott d_{min} kódtávolságú n hosszú blokk kód kódszavainak maximális száma.

Singleton-korlát:

$$A_s(n, d_{\min}) \leq s^{n-d_{min}+1}.$$

Magyarázat. Legyen $\mathcal{C}$ egy tetszőleges n hosszú kódszavakból álló $d_{\min}$ kódtávolságú blokk-kód. Ha minden kódszavának töröljük az első $d_{\min} - 1$ betűjét, akkor $n - d_{\min} + 1$ hosszúságú új, egymástól különböző kódszavakat kapunk. Ezek maximális száma $s^{n-d_{min}+1}$. Mivel a $\mathcal{C}$ kódot tetszőlegesen választottuk,

$$|\mathcal{C}| \leq A_s(n, d_{\min}) \leq s^{n-d_{min}+1}.$$ □

(n, k) paraméterű kód: a kódoló k hosszú forrásszegmensekhez rendel n hosszú vektorokat. Ekkor $|\mathcal{C}| = s^k$, a Singleton-korlát:

$$d_{min} \leq n - k + 1.$$

Definíció. *Azt a kódot, melyre a Singleton-korlátban egyenlőség áll, maximális távolságú, vagy MDS (maximum distance separable) kódnak nevezzük.*

Hamming-korlát

t: a kód által javítható hibák száma.

Hamming-korlát:

$$A_s(n, d_{\min}) \leq \frac{s^n}{\sum_{i=0}^{t} \binom{n}{i}(s-1)^i}, \quad \text{ahol} \quad t = \left\lfloor \frac{d_{min}-1}{2} \right\rfloor.$$

Magyarázat. Minden egyes kódszó esetén tekintsünk egy t sugarú gömböt, vagyis azokat az n hosszú kódjelsorozatokat, amik legfeljebb t hibával keletkeznek (Hamming-gömb). $(s-1)^i$ darab olyan sorozat van, ami pontosan i pozícióban tér el egy adott kódszótól, így egy gömb $\sum_{i=0}^{t} \binom{n}{i}(s-1)^i$ sorozatot tartalmaz. A kód akkor tud t hibát javítani, ha a gömbök diszjunktak. A gömbökben lévő sorozatok teljes száma kisebb, mint s^n, azaz

$$A_s(n, d_{\min}) \times \sum_{i=0}^{t} \binom{n}{i}(s-1)^i \leq s^n. \qquad \square$$

Megjegyzés. (n, k) paraméterű kód esetén a Hamming-korlát:

$$\sum_{i=0}^{t} \binom{n}{i}(s-1)^i \leq s^{n-k}.$$

Perfekt kódok

Bináris (n, k) típusú kódok esetén a Hamming korlát:

$$\sum_{i=0}^{t} \binom{n}{i} \leq 2^{n-k}.$$

Definíció. *Az olyan kódokat, ahol a Hamming-korlátban egyenlőség áll, perfekt kódoknak nevezzük.*

Egy bináris (n, k) típusú 1 hibát javítani képes kód perfekt, ha

$$1 + n = 2^{n-k}.$$

Példa.

1. Bináris $(5, 2)$ típusú kód, $d_{\min} = 3$. Nem MDS és nem perfekt.
2. Bináris $(3, 2)$ típusú kód, $d_{\min} = 2$. MDS, de nem perfekt.

További korlátok

Gilbert-Varshamov-korlát:

$$A_s(n, d_{\min}) \geq \frac{s^n}{\sum_{i=0}^{d_{\min}-1} \binom{n}{i}(s-1)^i}.$$

Nevező: $d_{\min} - 1$ sugarú Hamming-gömb elemeinek száma.

Plotkin-korlát: Ha $sd_{\min} > (s-1)n$, akkor

$$A_s(n, d_{\min}) \leq \frac{sd_{\min}}{sd_{\min} - (s-1)n}.$$

Bináris esetben, ha $d_{\min}$ páros

$$A_2(n, d_{\min}) \leq \begin{cases} 2\left\lfloor \frac{d_{\min}}{2d_{\min}-n} \right\rfloor, & \text{ha } n < 2d_{\min}; \\ 4d_{\min}, & \text{ha } n = 2d_{\min}. \end{cases}$$

Ha $d_{\min}$ páratlan

$$A_2(n, d_{\min}) \leq \begin{cases} 2\left\lfloor \frac{d_{\min}+1}{2d_{\min}+1-n} \right\rfloor, & \text{ha } n < 2d_{\min}+1; \\ 4d_{\min}+4, & \text{ha } n = 2d_{\min}+1. \end{cases}$$

Bináris lineáris kódok

Definíció. *Egy $\mathcal{C}$ bináris kód lineáris, ha $\mathcal{C}$ lineáris tér, azaz minden $\boldsymbol{c}, \boldsymbol{c'} \in \mathcal{C}$ esetén $\boldsymbol{c} + \boldsymbol{c'} \in \mathcal{C}$.*

Megjegyzés. Ha a $\mathcal{C}$ kód lineáris, akkor $\mathbf{0} \in \mathcal{C}$.

$\boldsymbol{g_1}, \boldsymbol{g_2}, \ldots, \boldsymbol{g_k} \in \mathcal{C}$: a $\mathcal{C}$ kód egy bázisa, ahol $\boldsymbol{g_i} = (g_{i1}, g_{i2}, \ldots, g_{in})$.

$\boldsymbol{G}$: a $\boldsymbol{g_1}, \boldsymbol{g_2}, \ldots, \boldsymbol{g_k}$ sorvektorokból álló $k \times n$ méretű mátrix.

Tetszőleges $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in \mathcal{C}$ esetén egyértelműen létezik $\boldsymbol{u} = (u_1, u_2, \ldots, u_k)$, hogy

$$\boldsymbol{c} = \sum_{i=1}^{k} u_i \boldsymbol{g_i}, \qquad \textit{azaz} \qquad \boldsymbol{c} = \boldsymbol{uG}.$$

Definíció. *A $\boldsymbol{G}$ mátrixot a $\mathcal{C}$ kód generátormátrixának nevezzük.*

Megjegyzés. A generátormátrix nem egyértelmű.

Definíció. *Egy (n, k) paraméterű lineáris kód szisztematikus, ha minden kódszavára igaz, hogy annak utolsó $n - k$ kódjelét elhagyva éppen a neki megfelelő üzenetet kapjuk.*

Szisztematikus kódok

Szisztematikus kód generátormátrixa egyértelmű: $\boldsymbol{G} = (\boldsymbol{I}_k, \boldsymbol{B})$.

$\boldsymbol{I}_k$: $k \times k$ méretű egységmátrix.

$\boldsymbol{B}$: $k \times (n-k)$ méretű mátrix.

A kódszavak alakja: $\boldsymbol{c} = (u_1, \ldots, u_k, c_{k+1}, \ldots, c_n)$.

Első k tag: üzenetszegmens; utolsó $n-k$ tag: paritásszegmens.

Példák.

1. $00 \mapsto 00000, \quad 01 \mapsto 01101, \quad 10 \mapsto 10110, \quad 11 \mapsto 11011.$

Lineáris kód, $n = 5$, $k = 2$. Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \qquad \text{szisztematikus kód.}$$

2. $00 \mapsto 000, \quad 01 \mapsto 011, \quad 10 \mapsto 101, \quad 11 \mapsto 110.$

Lineáris kód, $n = 3$, $k = 2$. Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad \text{szisztematikus kód.}$$

Paritásellenőrző mátrix

Definíció. *Ha egy $(n-k)\times n$ méretű $\boldsymbol{H}$ mátrixra*

$$\boldsymbol{H}\boldsymbol{c}^\top = \boldsymbol{0}$$

*akkor és csak akkor, ha $\boldsymbol{c}\in\mathcal{C}$, akkor a $\boldsymbol{H}$ mátrixot a $\mathcal{C}$ kód paritás-ellenőrző mátrix*ának *(röviden paritásmátrix*ának*) nevezzük.*

Tétel. *Ha $\boldsymbol{G}$ és $\boldsymbol{H}$ ugyanazon $\mathcal{C}$ kód generátormátrixa, illetve paritásmátrixa, akkor*

$$\boldsymbol{H}\boldsymbol{G}^\top = \boldsymbol{0}.$$

Minden lineáris kódnak van paritásmátrixa.

Szisztematikus kód esetén a generátormátrix alakja: $\boldsymbol{G} = (\boldsymbol{I}_k, \boldsymbol{B})$.

A megfelelő paritásmátrix: $\boldsymbol{H} = (\boldsymbol{B}^\top, \boldsymbol{I}_{n-k})$.

Nem szisztematikus generátormátrixot Gauss-eliminációval olyan alakra hozhatunk, aminél léteznek $i_1, i_2, \ldots, i_k$ egészek, hogy az i_j oszlop a j-edik helyen 1, másutt 0. Ez oszlopcserével már szisztematikussá tehető.

Példák

1. $00 \mapsto 00000, \quad 01 \mapsto 01101, \quad 10 \mapsto 10110, \quad 11 \mapsto 11011.$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \qquad \text{szisztematikus kód.}$$

Paritásmátrix:

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

2. $00 \mapsto 000, \quad 01 \mapsto 011, \quad 10 \mapsto 101, \quad 11 \mapsto 110.$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad \text{szisztematikus kód.}$$

Paritásmátrix:

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

Kód minimális súlya

Definíció. *Egy $\boldsymbol{c}$ vektor súlya a koordinátái között levő nem nulla elemek száma. Jelölése: $w(\boldsymbol{c})$.*

Definíció. *Egy $\mathcal{C}$ kód minimális súlyán a*

$$w_{\min} := \min_{\substack{\boldsymbol{c}\in\mathcal{C}\\ \boldsymbol{c}\neq\boldsymbol{0}}} w(\boldsymbol{c})$$

számot értjük.

Tétel. *Ha $\mathcal{C}$ egy lineáris kód, akkor a kódtávolsága megegyezik a minimális súlyával, azaz*

$$d_{\min} = w_{\min}.$$

Magyarázat.

$$d_{\min} = \min_{\boldsymbol{c}\neq\boldsymbol{c}'} d(\boldsymbol{c},\boldsymbol{c}') = \min_{\boldsymbol{c}\neq\boldsymbol{c}'} w(\boldsymbol{c}-\boldsymbol{c}') = \min_{\boldsymbol{c}''\neq\boldsymbol{0}} w(\boldsymbol{c}'') = w_{\min}. \qquad \square$$

Egy $|\mathcal{C}|$ elemszámú $\mathcal{C}$ kód esetén $d_{\min}$ kiszámításához $|\mathcal{C}|(|\mathcal{C}|-1)/2$, $w_{\min}$ kiszámításához $|\mathcal{C}|-1$ művelet szükséges.

Szindróma dekódolás

$\boldsymbol{H}$: egy $\mathcal{C}$ kód paritásmátrixa.

Definíció. *Az $\boldsymbol{s} = \boldsymbol{e}\boldsymbol{H}^\top$ mennyiséget szindómának nevezzük.*

$\boldsymbol{c}$: leadott kódszó; $\boldsymbol{v}$ vett szó; $\boldsymbol{e} = \boldsymbol{v} - \boldsymbol{c}$: hibavektor.

$$\boldsymbol{H}\boldsymbol{v}^\top = \boldsymbol{H}(\boldsymbol{c}+\boldsymbol{e})^\top = \boldsymbol{H}\boldsymbol{c}^\top + \boldsymbol{H}\boldsymbol{e}^\top = \boldsymbol{H}\boldsymbol{e}^\top.$$

$\boldsymbol{H}\boldsymbol{v}^\top$ értéke nem függ az adott kódszótól, csak a hibavektortól.

Szindróma dekódolás:

- A vett $\boldsymbol{v}$ szóból kiszámítjuk az $\boldsymbol{s}^\top = \boldsymbol{H}\boldsymbol{v}^\top = \boldsymbol{H}\boldsymbol{e}^\top$ szindrómát.
- A szindróma alapján megbecsüljük a hibavektort.
- A becsült hibavektort $\boldsymbol{v}$-ből kivonva megkapjuk a kódszóra vonatkozó becslést.

Egy $\boldsymbol{e}$ hibaminta által generált mellékosztály: $\mathcal{C}_{\boldsymbol{e}} := \{\boldsymbol{e}+\boldsymbol{c} \mid \boldsymbol{c}\in\mathcal{C}\}$.

Egy adott mellékosztály elemeihez azonos szindróma tartozik.

Ha $\boldsymbol{e} = \boldsymbol{e}' + \boldsymbol{c}$, akkor $\mathcal{C}_{\boldsymbol{e}} = \mathcal{C}_{\boldsymbol{e}'}$. $\mathcal{C}_{\boldsymbol{0}} = \mathcal{C}$.

Mellékosztály-vezető: azonos mellékosztályba tartozó hibaminták közül a legkisebb súlyú.

Standard elrendezési táblázat

$\mathcal{C}(n,k)$: egy (n,k) típusú kód.

szindróma	mellékosztály-vezető			
$\boldsymbol{s}^{(0)}$	$\boldsymbol{e}^{(0)}=\mathbf{0}$	$\boldsymbol{c}^{(1)}$	$\cdots$	$\boldsymbol{c}^{(s^k-1)}$
$\boldsymbol{s}^{(1)}$	$\boldsymbol{e}^{(1)}$	$\boldsymbol{c}^{(1)}+\boldsymbol{e}^{(1)}$	$\cdots$	$\boldsymbol{c}^{(s^k-1)}+\boldsymbol{e}^{(1)}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\boldsymbol{s}^{(s^{n-k}-1)}$	$\boldsymbol{e}^{(s^{n-k}-1)}$	$\boldsymbol{c}^{(1)}+\boldsymbol{e}^{(s^{n-k}-1)}$	$\cdots$	$\boldsymbol{c}^{(s^k-1)}+\boldsymbol{e}^{(s^{n-k}-1)}$

mellékosztály elemek

Sorrend: $w\big(\boldsymbol{e}^{(i+1)}\big)\geq w\big(\boldsymbol{e}^{(i)}\big)$, $\boldsymbol{e}^{(0)}=\mathbf{0}$, $i=0,1,\ldots,s^{n-k}-2$.

A táblázat elemei mind különbözőek.

Definíció. *Az $\boldsymbol{e}^{(i)}$, $i=1,2,\ldots,s^{n-k}-1$ mellékosztály-vezetőket javítható hibamintáknak nevezzük.*

Vett $\boldsymbol{v}$ szó szindrómája $\boldsymbol{s}^{(i)}$: a választott kódszó $\widehat{\boldsymbol{c}}=\boldsymbol{v}-\boldsymbol{e}^{(i)}$.

Bináris eset: 2^{n-k} javítható hibavektor a szindrómával címezve.

Példa

$00 \mapsto 00000, \quad 01 \mapsto 01101, \quad 10 \mapsto 10110, \quad 11 \mapsto 11011.$

Paritásmátrix:

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Dekódolási táblázat:

szindróma	javítható hibaminta
000	00000
001	00001
010	00010
011	00011
100	00100
101	01000
110	10000
111	01010

Az egyszeres hibák és a 00011, 01010 hibaminták javíthatóak.

Bináris Hamming-kód

Egy hibát javítani képes lineáris, bináris kód. A hibajavításra r bit használható, azaz $n - k = r$.

Cél: rögzített n és r esetén a legnagyobb k elérése.

Paritásmátrix:

$$\boldsymbol{H} = \left(\boldsymbol{a}_1^\top, \boldsymbol{a}_2^\top, \ldots, \boldsymbol{a}_n^\top\right).$$

Legfeljebb egy hiba esetén a javítható $\boldsymbol{e}$ hibavektor: $\boldsymbol{e} = \boldsymbol{0}$ vagy $\boldsymbol{e} = \boldsymbol{e}_i,\ i = 1, 2, \ldots n$ (az i-edik pozíción 1, egyébként 0).

Az $\boldsymbol{s} = \boldsymbol{e}\boldsymbol{H}^\top$ szindróma: $\boldsymbol{0}$ vagy $\boldsymbol{a}_i,\ i = 1, 2, \ldots, n$.

Az $\boldsymbol{e}$ hiba (és így a $\boldsymbol{c}$ kódszó) pontosan akkor adható meg egyértelműen, ha az $\boldsymbol{a}_i$ vektorok mind különbözőek és egyik sem $\boldsymbol{0}$.

$\boldsymbol{H}$ sorainak száma $n-k=r$: legfeljebb 2^r-1 különböző $\boldsymbol{a}_i \neq \boldsymbol{0}$ lehet.

Rögzített r esetén mind a $2^r - 1$ darab különböző bináris $\boldsymbol{a}_i \neq \boldsymbol{0}$ vektort kihasználjuk, ezek lesznek a $\boldsymbol{H}$ oszlopai.

Kódszavak: a $\boldsymbol{H}\boldsymbol{c}^\top = \boldsymbol{0}$ egyenletrendszer megoldásai.

A bináris Hamming-kód tulajdonságai

Megjegyzés. Az (n, k) paraméterű bináris Hamming-kód esetén $n = 2^{n-k} - 1$, azaz a kód perfekt.

Néhány lehetséges számpár:

$n =$	3	7	15	31	63	127
$k =$	1	4	11	26	57	120

Tétel. *Nincs olyan egy hibát javító bináris kód, amely egy Hamming-kóddal azonos szóhosszúságú, és a hozzá tartozó kódszavak száma nagyobb, mint a megfelelő Hamming kód kódszavainak száma.*

Példa. $(7, 4)$ paraméterű bináris Hamming kód:

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

A teletext egyes részeinek kódolása: $(7, 4)$ paraméterű Hamming-kód paritásbittel kiegészítve.

A test fogalma

Definíció. *Egy G halmazt testnek nevezünk, ha elemei között értelmezve van egy összeadásnak és egy szorzásnak nevezett művelet, amelyeket rendre "$+$" és "$\cdot$" szimbólumokkal jelölünk és G rendelkezik a következő tulajdonságokkal:*

1. *G az összeadásra nézve kommutatív csoport, azaz*
   a) *Minden $\alpha, \beta \in G$ esetén $\alpha + \beta \in G$, valamint $\alpha + \beta = \beta + \alpha$.*
   b) *Minden $\alpha, \beta, \gamma \in G$ esetén $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.*
   c) *Létezik egy 0-val jelölt eleme (nulleleme) G-nek, hogy minden $\alpha \in G$ esetén $0 + \alpha = \alpha + 0 = \alpha$.*
   d) *Minden $\alpha \in G$ esetén létezik $\beta \in G$, hogy $\alpha + \beta = 0$. β-t az α additív inverzének nevezzük és $-\alpha$-val jelöljük.*

2. *$G \setminus \{0\}$ a szorzásra nézve kommutatív csoport, azaz*
   a) *Minden $\alpha, \beta \in G \setminus \{0\}$ esetén $\alpha \cdot \beta \in G \setminus \{0\}$, valamint $\alpha \cdot \beta = \beta \cdot \alpha$.*
   b) *Minden $\alpha, \beta, \gamma \in G \setminus \{0\}$ esetén $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$.*
   c) *Létezik egy 1-gyel jelölt eleme (egységeleme) $G \setminus \{0\}$-nak, hogy minden $\alpha \in G \setminus \{0\}$ esetén $1 \cdot \alpha = \alpha \cdot 1 = \alpha$.*
   d) *Minden $\alpha \in G \setminus \{0\}$ esetén létezik $\beta \in G \setminus \{0\}$, hogy $\alpha \cdot \beta = 1$. β-t az α multiplikatív inverzének nevezzük és α^{-1}-gyel jelöljük.*

3. *Minden $\alpha, \beta, \gamma \in G$ esetén $\alpha \cdot 0 = 0 \cdot \alpha = 0$ és $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$.*

Véges testek

Egy s elemszámú testet véges testnek nevezünk és $GF(s)$-sel jelöljük. s nem lehet tetszőleges.

Tétel. *Egy $GF(s)$ véges test esetén $s = p^m$, azaz s vagy prím, vagy pedig prímhatvány.*

Tétel. *Egy p prímszám esetén a $G = \{0, 1, \ldots, p-1\}$ halmaz a modulo p aritmetikával egy véges testet alkot.*

Lemma. *Minden $0 \neq a \in GF(s)$ esetén $a^{s-1} = 1$.*

Lemma. *Minden $0 \neq a \in GF(s)$ esetén létezik egy legkisebb m természetes szám, melyre $a^m = 1$ és az $a, a^2, \ldots, a^m$ elemek mind különbözőek. Az m számot az a elem rendjének nevezzük, m osztója $s-1$-nek.*

Definíció. *Egy $\alpha \in GF(s)$-t a $GF(s)$ véges test primitív elemének nevezzük, ha α rendje $s-1$.*

Tétel. *Minden $GF(s)$ véges testben létezik primitív elem.*

A $GF(p^m)$ test reprezentációja

A $GF(p^m)$ test $\{0, 1, \ldots, p^m - 1\}$ elemeinek $GF(p)$-beli koordinátájú m hosszú vektorokat, illetve belőlük származtatott legfeljebb $(m-1)$-edfokú polinomokat feleltetünk meg.

Összeadás: koordinátánkénti összeadás a $GF(p)$ aritmetikájával.

Definíció. *A $GF(p)$ feletti, nem nulladfokú $P(x)$ polinomot irreducibilis polinomnak nevezzük, ha nem bontható fel két, nála alacsonyabb fokú $GF(p)$ feletti polinom szorzatára.*

Megjegyzés. Minden véges testben található tetszőleges fokszámú irreducubilis polinom.

Példa. Minden elsőfokú polinom irreducubilis. $GF(2)$ esetén ezek x, $x+1$. Másodfokú irreducibilis polinom: különbözik az $(x+1)^2$, $x(x+1)$, x^2 polinomoktól. Egy ilyen van: x^2+x+1.

fokszám	irreducibilis polinom	fokszám	irreducibilis polinom
2	x^2+x+1	6	x^6+x+1
3	x^3+x+1	7	x^7+x^3+1
4	x^4+x+1	8	$x^8+x^4+x^3+x^2+1$
5	x^5+x^2+1	9	x^9+x^4+1

Aritmetikai a $GF(p^m)$ testben

Tétel. *Legyen p egy prímszám, $m \in \mathbb{N}$, $P(x)$ egy $GF(p)$ feletti m-edfokú irreducibilis polinom és $\mathcal{Q} = \{0, 1, \ldots, p^m - 1\}$. Minden $a \in \mathcal{Q}$ elemnek kölcsönösen egyértelműen feleltessünk meg egy $GF(p)$ feletti legfeljebb $(m-1)$-edfokú $a(x)$ polinomot. Az $a(x)$ és $b(x)$ polinomokkal reprezentált $a, b \in \mathcal{Q}$ esetén $a + b$ az a $c \in \mathcal{Q}$ elem, melynek megfelelő $c(x)$ polinomra*

$$c(x) = a(x) + b(x).$$

$a \cdot b$ az a $d \in \mathcal{Q}$ elem, melynek megfelelő $d(x)$ polinomra

$$d(x) = \big[a(x) \cdot b(x)\big] \mod P(x).$$

Ezzel az aritmetikával $\mathcal{Q}$ egy $GF(p^m)$ test.

$GF(p)$ elemei: nulladfokú polinomok. $GF(p)$ a $GF(p^m)$ részének tekinthető.

Példa

Test: $GF(4)$; irreducibilis polinom: x^2+x+1; vektorhossz: $m=2$.

testelem	vektor	polinom	testelem	vektor	polinom
0	00	0	2	10	x
1	01	1	3	11	$x+1$

$2+3$ megfelelője: $x+(x+1)=(1+1)x+1=1$, azaz $2+3=1$.

$2\cdot 3$ megfelelője:

$$x(x+1)=x^2+x=(x^2+x+1)+1=1 \qquad \text{mod } x^2+x+1,$$

azaz $2\cdot 3=1$.

Művelettáblák:

$+$	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

$\times$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

Példa

Test: $GF(8)$; irreducibilis polinom: x^3+x+1; vektorhossz: $m=3$.

testelem	vektor	polinom	testelem	vektor	polinom
0	000	0	4	100	x^2
1	001	1	5	101	x^2+1
2	010	x	6	110	x^2+x
3	011	$x+1$	7	111	x^2+x+1

Művelettáblák:

$+$	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	0	3	2	5	4	7	6
2	2	3	0	1	6	7	4	5
3	3	2	1	0	7	6	5	4
4	4	5	6	7	0	1	2	3
5	5	4	7	6	1	0	3	2
6	6	7	4	5	2	3	0	1
7	7	6	5	4	3	2	1	0

$\times$	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	3	1	7	5
3	0	3	6	5	7	4	1	2
4	0	4	3	7	6	2	5	1
5	0	5	1	4	2	7	3	6
6	0	6	7	1	5	3	2	4
7	0	7	5	2	1	6	4	3

Általános lineáris kódok

Az $\mathcal{Y}$ kódábécé szimbólumai $GF(s)$ elemei, azaz megfeleltethetőek a $\{0, 1, \ldots s-1\}$ számoknak.

Definíció. *Egy $\mathcal{C}$ kód lineáris, ha a $\mathcal{C}$ halmaz vektortér $GF(s)$ felett, azaz minden $\boldsymbol{c}, \boldsymbol{c'} \in \mathcal{C}$ estén $\boldsymbol{c} + \boldsymbol{c'} \in \mathcal{C}$, valamint tetszőleges $\beta \in GF(s)$-re $\beta\boldsymbol{c} \in \mathcal{C}$.*

Generátormátrix: $k \times n$ méretű lineárisan független sorokból álló $\boldsymbol{G}$ mátrix. Egy $\boldsymbol{u}$ üzenet $\boldsymbol{c}$ kódja $\boldsymbol{c} = \boldsymbol{uG}$.

Paritásmátrix: $(n-k) \times n$ méretű $\boldsymbol{H}$ mátrix. $\boldsymbol{c}$ pontosan akkor kódszó, ha $\boldsymbol{Hc}^\top = \boldsymbol{0}$.

Tétel. *Minden lineáris kódnak van paritásmátrixa.*

Bináris Hamming-kód: egy hibát képes javítani. Elég ismerni a hiba helyét. $\boldsymbol{H}$ oszlopai mind különbözőek és egyik sem $\boldsymbol{0}$.

Nembináris Hamming-kód: egy hibát képes javítani. Nemcsak a hiba helyét, hanem az értékét is ismerni kell.

Nembináris Hammming-kód

Paritásmátrix:

$$\boldsymbol{H} = \left(\boldsymbol{a}_1^\top, \boldsymbol{a}_2^\top, \ldots, \boldsymbol{a}_n^\top\right).$$

Az $\boldsymbol{a}_i \neq \boldsymbol{0}$ vektorok mind különbözőek, az $\boldsymbol{a}_i$ első nem 0 eleme 1.

Javítható hibaminta alakja: $\boldsymbol{e} = (0, \ldots, 0, e_i, 0, \ldots, 0)$, azaz az i-edik pozíción a hiba értéke $e_i \neq 0$.

Szindróma: $\boldsymbol{s} = e_i \boldsymbol{a}_i$, és az e_i hiba a szindróma első nem 0 értéke.

A hiba helye: $\boldsymbol{a}_i = \boldsymbol{s}/e_i$, így a $\boldsymbol{H}$ ismeretében i megadható.

$\boldsymbol{H}$ oszlopainak maximális száma:

$$n = \sum_{j=0}^{n-k-1} s^j = \frac{s^{n-k}-1}{s-1}, \qquad \text{azaz} \qquad 1 + n(s-1) = s^{n-k}.$$

s elemű csatornaábécé feletti 1 hibát javítani képes (n, k) típusú kódra a Hamming korlát:

$$1 + n(s-1) \leq s^{n-k}.$$

Tétel. *A maximális hosszúságú nembináris Hamming-kód perfekt.*

$(n, n-2)$ paraméterű nembináris Hamming-kód

Szisztematikus kód $n-k=2$ hosszú paritásszegmenssel.

$\alpha \neq 0$: $GF(s)$ egy $m \geq 0$ rendű eleme. α primitív: $m = s-1$.

Ha $n \leq m+2$, akkor az $1, \alpha, \alpha^2, \ldots, \alpha^{n-3}$ elemek különbözőek.

Primitív eset: $n = s+1$, a lehetséges legnagyobb érték.

Az $(n, n-2)$ paraméterű nembináris Hamming-kód paritásmátrixa:

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 & 0 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-3} & 0 & 1 \end{pmatrix}.$$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & \cdots & 0 & -1 & -\alpha \\ 0 & 0 & 1 & 0 & \cdots & 0 & -1 & -\alpha^2 \\ \vdots & & & & \ddots & & & \vdots \\ 0 & 0 & & 0 & \cdots & 1 & -1 & -\alpha^{n-3} \end{pmatrix}.$$

1 hibát tud javítani: $d_{\min} \geq 3$. Singleton korlát: $d_{\min} \leq n-k+1 = 3$.

Tétel. *Az $(n, n-2)$ paraméterű nembináris Hamming-kód MDS.*

Példa

$GF(7)$ nem 0 elemei:

elem:	1	2	**3**	4	**5**	6
hatványok:	1	2,4,1	3,2,6,4,5,1	4,2,1	5,4,6,2,3,1	6,1
rend:	1	3	6 (primitív)	3	6 (primitív)	2

A $GF(7)$ feletti $(8,6)$ paraméterű Hamming-kód paritásmátrixa:

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 3 & 2 & 6 & 4 & 5 & 0 & 1 \end{pmatrix}.$$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & -3 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & -6 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -4 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 6 & 6 \\ 0 & 1 & 0 & 0 & 0 & 0 & 6 & 4 \\ 0 & 0 & 1 & 0 & 0 & 0 & 6 & 5 \\ 0 & 0 & 0 & 1 & 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 6 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 6 & 2 \end{pmatrix}.$$

Ciklikus kódok

Definíció. *Egy* $\boldsymbol{c} = (c_0, c_1, \ldots, c_{n-1})$ *vektor ciklikus eltoltja az* $S\boldsymbol{c} = (c_{n-1}, c_0, \ldots, c_{n-2})$ *vektor. S-et a ciklikus eltolás operátorának nevezzük.*

Definíció. *A* $\mathcal{C}$ *kódot ciklikus*nak *nevezzük, ha bármely kódszó ciklikus eltoltja is kódszó.*

Példa.

$$\mathcal{C} = \{0000,\ 1100,\ 0110,\ 0011,\ 1001,\ 1111\}.$$

Nem feltétlenül lineáris, pl. $1001 + 0011 = 1010 \notin \mathcal{C}$.

Definíció. *Rendeljünk polinomot az egyes kódszavakhoz:*

$$\boldsymbol{c} = (c_0, c_1, \ldots, c_{n-1}) \mapsto c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}.$$

A $\boldsymbol{c}$ *kódszónak megfeleltetett* $c(x)$ *polinomot kódszópolinom*nak *(röviden kódpolinom*nak*) nevezzük. A kódszópolinomok halmazát* $\mathcal{C}(x)$*-szel jelöljük.*

Példa.

$1100 \mapsto 1 + x;\ 0110 \mapsto x + x^2;\ 0011 \mapsto x^2 + x^3;\ 1001 \mapsto 1 + x^3.$

Kódszópolinomok tulajdonságai

Lemma. *Legyen $c'(x)$ a $\boldsymbol{c}$ kódszó $S\boldsymbol{c}$ eltoltjához rendelt kódszópolinom. Ekkor*

$$c'(x) = \left[xc(x)\right] \qquad \mod (x^n - 1).$$

Magyarázat. Legyen $c(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$, ekkor

$$\begin{aligned} xc(x) &= c_0 x + c_1 x^2 + \cdots + c_{n-1}x^n = c_0 x + c_1 x^2 + \cdots + c_{n-1}\left(x^n - 1\right) + c_{n-1} \\ &= c_{n-1} + c_0 x + c_1 x^2 + \cdots + c_{n-2}x^{n-1} + c_{n-1}\left(x^n - 1\right) = c'(x) + c_{n-1}\left(x^n - 1\right). \end{aligned}$$ □

Tétel. *Minden (n, k) paraméterű, ciklikus, lineáris $\mathcal{C}$ kódban a nem azonosan nulla kódszópolinomok között egyértelműen létezik egy minimális fokszámú $g(x)$ főpolinom (1 főegyütthatójú polinom). $g(x)$ fokszáma $n-k$, és $\boldsymbol{c} \in \mathcal{C}$ pontosan akkor, ha $g(x)|c(x)$, azaz létezik egy olyan $u(x)$ polinom, hogy $c(x) = g(x)u(x)$.*

Definíció. *A $g(x)$ polinomot a $\mathcal{C}$ kód generátorpolinomjának nevezzük.*

Generátorpolinomok

Tétel. *Minden ciklikus, lineáris kód $g(x)$ generátorpolinomjára*

$$g(x) \mid x^n - 1.$$

Mésrészről, ha egy $g(x)$ főpolinomra $g(x) \mid x^n - 1$, akkor létezik egy lineáris ciklikus kód, melynek $g(x)$ a generátorpolinomja.

$\boldsymbol{u} = (u_0, u_1, \ldots, u_{k-1})$: k hosszúságú üzenet, $u_i \in GF(s)$.

Üzenetpolinom: $u(x) = u_0 + u_1 x + \cdots + u_{k-1} x^{k-1}$.

Generátorpolinom: $g(x) = g_0 + g_1 x + \cdots + g_{n-k-1} x^{n-k-1} + x^{n-k}$.

Egy lineáris ciklikus kód generálható az üzenetpolinom és a generátorpolinom szorzásával. Generátormátrixa ($k \times n$ méretű) a $g(x)$ együtthatóinak eltolásával kapható meg:

$$\boldsymbol{G} = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-k-1} & 1 & 0 & \cdots & 0 & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k-2} & g_{n-k-1} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & g_0 & g_1 & g_2 \cdots & 1 & 0 & \\ 0 & 0 & 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_{n-k-1} & 1 \end{pmatrix}.$$

Példa

$GF(2)$ feletti ciklikus kód:

$$\mathcal{C} = \{\, 0000000,\ 1011100,\ 0101110,\ 0010111,\ 1001011,\ 1100101,\ 1110010,\ 0111001 \}.$$

Az 1011100 kódpolinomja: $1 + x^2 + x^3 + x^4$. Minimális fokszámú, ez a generátorpolinom. Üzenetszegmens hossza: $k = 3$.

Generátormátrix és szisztematikus alakja:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}, \quad \boldsymbol{G}_c = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

A szisztematikus és az eredeti generátormátrix paritásmátrixa:

$$\boldsymbol{H}_c = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{H} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Szisztematikus generálás

Tétel. *Minden lineáris ciklikus kód generálható szisztematikusan.*

Módszer. Legyen $u(x)$ egy legfeljebb $(k-1)$-edfokú üzenetpolinom és

$$c(x) = u(x) - \Big(\big[u(x) x^{n-k} \big] \mod g(x) \Big) x^k.$$

$c(x) = 0 \mod g(x)$, így $c(x)$ kódszó.

Példa. $GF(2)$ feletti $(7,3)$ paraméterű kód, $g(x) = 1 + x^2 + x^3 + x^4$.

$\big[x^{4+i}\big] \mod g(x)$ maradékai, $i = 0, 1, 2$:

$$x^4 = 1 + x^2 + x^3 \mod g(x);$$
$$x^5 = 1 + x + x^2 \mod g(x);$$
$$x^6 = x + x^2 + x^3 \mod g(x).$$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Kódszavak:

0000000, 0010111, 0101110, 0111001, 1001011, 1011100, 1100101, 1110010.

Paritásellenőrző polinom

Definíció. *Egy $g(x)$ generátorpolinomú lineáris, ciklikus kód esetén a*

$$h(x) = \frac{x^n - 1}{g(x)}$$

polinomot paritásellenőrző polinomnak nevezzük.

Tétel. *Egy lineáris, ciklikus kódra $c(x)$ pontosan akkor kódszópolinom, ha*

$$c(x)h(x) = 0 \mod (x^n - 1) \qquad \text{és} \qquad \deg\big(c(x)\big) \leq n - 1.$$

Indoklás. Ha $c(x)$ kódszópolinom, akkor alakja $c(x) = u(x)g(x)$, azaz

$c(x)h(x) = u(x)g(x)h(x) = u(x)(x^n - 1)$, tehát $c(x)h(x) = 0 \mod (x^n - 1)$.

Ha $c(x)h(x) = 0 \mod (x^n - 1)$, akkor $c(x)h(x) = a(x)(x^n - 1)$ alakú ahonnan

$$c(x) = a(x)(x^n - 1)/h(x) = a(x)g(x),$$

tehát $c(x)$ kódpolinom. □

Példa. $GF(2)$ feletti $(7, 3)$ paraméterű kód,

$$g(x) = 1 + x^2 + x^3 + x^4, \qquad h(x) = 1 + x^2 + x^3.$$

Szindrómák

Vett jelsorozat: $\boldsymbol{r} = (r_0, r_1, \ldots, r_{n-1})$.

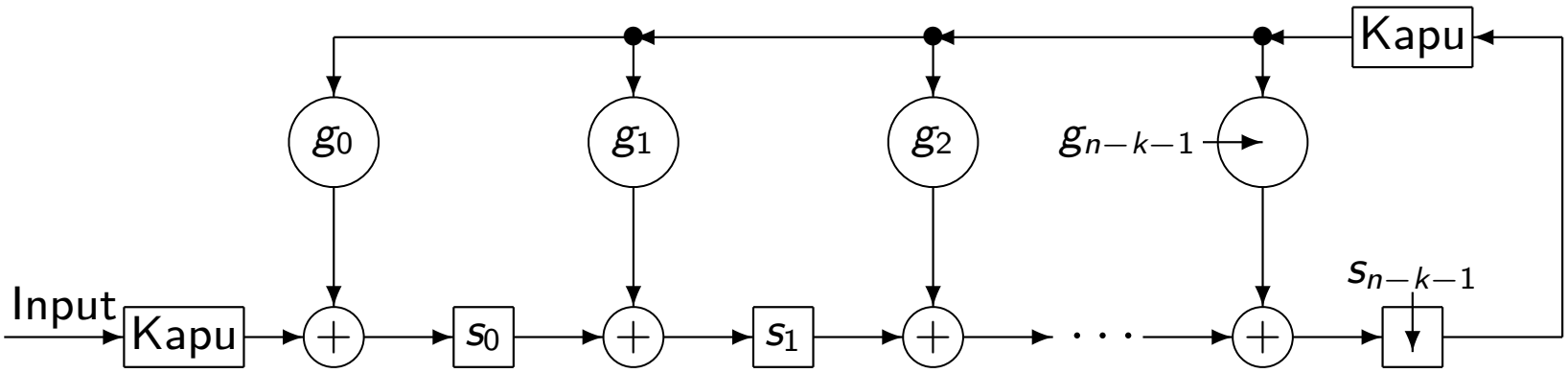
Polinom alak: $r(x) = r_0 + r_1 x + \cdots + r_{n-1} x^{n-1}$.

Szindróma: legfeljebb $n-k-1$-edfokú $s(x)$ polinom, melyre

$$r(x) = a(x)g(x) + s(x).$$

Tétel. *Legyen $s(x)$ az vett $\boldsymbol{r}$ jelsorozat $r(x)$ polinomjához tartozó szindróma. Ekkor az $xs(x)$ polinomnak a $g(x)$ generátorpolinommal való osztásakor kapott $s^{(1)}(x)$ polinom az $\boldsymbol{r}$ jelsorozat $\boldsymbol{r}^{(1)}$ ciklikus eltoltjához tartozó szindróma.*

Következmény. Az $x^i s(x)$ polinomnak a $g(x)$ generátorpolinommal való osztásakor kapott $s^{(i)}(x)$ polinom az $\boldsymbol{r}$ jelsorozat $\boldsymbol{r}^{(i)}$ ciklikus eltoltjához tartozó szindróma.



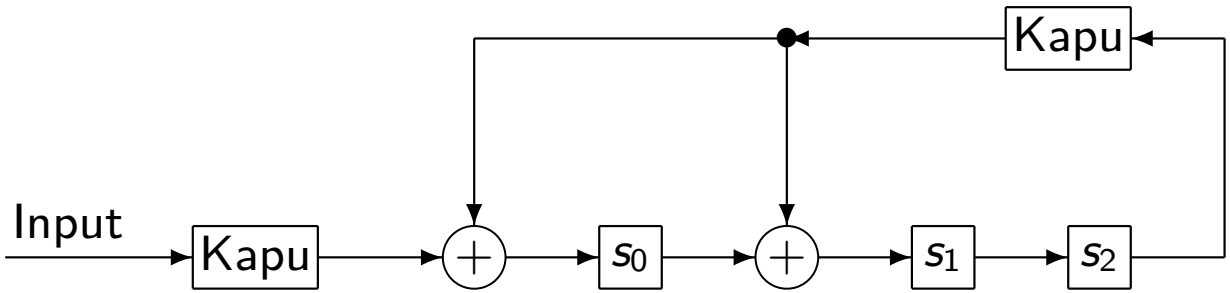
Példa

$GF(2)$ feletti $(7,4)$ paraméterű ciklikus kód, $g(x) = 1 + x + x^3$.

Vett vektor: $\boldsymbol{r} = (0\,0\,1\,0\,1\,1\,0)$; polinom alak: $r(x) = x^2 + x^4 + x^5$.

Szindróma: $\boldsymbol{s} = (1\,0\,1)$; polinom alak: $s(x) = 1 + x^2$.

Szindróma generáló kör:



Shift	Input	A regiszter tartalma	
		0 0 0	(kiinduló állapot)
1	0	0 0 0	
2	1	1 0 0	
3	1	1 1 0	
4	0	0 1 1	
5	1	0 1 1	
6	0	1 1 1	
7	0	1 0 1	($\boldsymbol{s}$ szindróma)
8	-	1 0 0	($\boldsymbol{s}^{(1)}$ szindróma)
9	-	0 1 0	($\boldsymbol{s}^{(2)}$ szindróma)

Bináris Golay-kód

(23, 12) paraméterű lineáris ciklikus kód. Marcel J. E. Golay, 1949.

Generátorpolinom:

$$g(x) = 1 + x^2 + x^4 + x^5 + x^6 + x^{10} + x^{11} \qquad \text{vagy}$$
$$g'(x) = 1 + x + x^5 + x^6 + x^7 + x^9 + x^{11}.$$

Mindkét polinom egy 23 hosszú ciklikus kódot generál, mivel

$$(x-1)g(x)g'(x) = x^{23} - 1.$$

Kódtávolsága 7, azaz 6 egyszerű hibát jelez, 3 egyszerű és 6 törléses hibát javít. Perfekt kód. Jelölés: Golay $[23, 12, 7]_2$.

Kiterjesztett Golay-kód: Golay-kód kiegészítve egy paritásbittel.

Kódtávolság: 8. Jelölés: Golay $[24, 12, 8]_2$.

Alkalmazás: NASA Voyager 1 és 2 űrszonda (1977-), színes fotók továbbítása Golay $[24, 12, 8]_2$ kóddal.

Reed-Solomon-kód

Irving S. Reed és Gustave Solomon, 1960 (MIT)

1. Konstrukció. *Legyenek* $\alpha_0, \alpha_2, \ldots, \alpha_{n-1}$ *a GF(s) különböző elemei* ($n \leq s$) *és* $\boldsymbol{u} = (u_0, u_1, \ldots, u_{k-1})$, $u_i \in GF(s)$, *a k hosszúságú üzenetszegmens*

$$u(x) = u_0 + u_1 x + \cdots + u_{k-1} x^{k-1}$$

üzenetpolinommal. Ekkor a Reed-Solomon-kód $\boldsymbol{u}$ *üzenethez tartozó n hosszú* $\boldsymbol{c}$ *kódszavának a komponensei:*

$$c_0 = u(\alpha_0),\ c_1 = u(\alpha_1),\ c_2 = u(\alpha_2),\ \ldots\ , c_{n-1} = u(\alpha_{n-1}).$$

Jelölés: RS(n,k).

Megjegyzés. A Reed-Solomon-kód lineáris és generátormátrixa:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha_0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_0^{k-1} & \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_{n-1}^{k-1} \end{pmatrix}.$$

Tulajdonságok

Tétel. *Az (n, k) paraméterű Reed-Solomon-kód kódtávolsága $d_{\min} = n - k + 1$, azaz a kód MDS.*

Speciális eset. *Legyen $0 \neq \alpha \in GF(s)$, melynek rendje $m \geq n$, valamint legyenek $\alpha_0 = 1, \alpha_1 = \alpha, \ldots, \alpha_{n-1} = \alpha^{n-1}$. Ekkor a Reed-Solomon-kód generátormátrixa:*

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{k-1} & \alpha^{2(k-1)} & \cdots & \alpha^{(k-1)(n-1)} \end{pmatrix}.$$

Általában $s = 2^r$, gyakori az $s = 2^8 = 256$. Ekkor a Reed-Solomon kód 8-bites szimbólumokat használ, így $8n$ hosszú kódszavakat használó bináris kóddá alakítható.

Népszerű eset: $(255, 223)$ típusú Reed-Solomon kód 8-bites szimbólumokkal. 16 hibát javít.

2. Konstrukció. *Legyen*

$$c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$$

a $\boldsymbol{c} = (c_0, c_1, \ldots, c_{n-1})$ *vektorhoz rendelt polinom, valamint legyen az* α *elem rendje* m*, ahol* $m \geq n$*. A* $\mathcal{C}$ *kódot definiáljuk a következőképpen:*

$$\mathcal{C} = \left\{ \boldsymbol{c} : c(\alpha^i) = 0,\ i = 1, 2, \ldots, n-k \right\}.$$

Megjegyzés. A konstrukció egy olyan kódot eredményez, aminek paritásellenőrző mátrixa

$$\boldsymbol{H} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{n-k} & \alpha^{2(n-k)} & \cdots & \alpha^{(n-k)(n-1)} \end{pmatrix}.$$

Tétel. *Ha* $n = m$ *(nem rövidített Reed-Solomon-kód), akkor a két konstrukció egybeesik.*

Ciklikusság

Tétel. *Ha $n = m$, azaz a Reed-Solomon-kód kódszóhossza megegyezik az α elem rendjével, akkor a kód ciklikus, generátorpolinomja, illetve paritásellenőrző polinomja pedig rendre*

$$g(x) = \prod_{i=1}^{n-k} (x - \alpha^i) \quad \text{és} \quad h(x) = \prod_{n-k+1}^{n} (x - \alpha^i).$$

Alkalmazások, pl.:

- CD lemez: RS$(28, 24)$ (C2 level, lemezhibák, írási hibák) és RS$(32, 28)$ (C1 level, karcolások, ujjlenyomatok);
- DVD lemez: RS$(182, 172)$ (C2 level) és RS$(208, 192)$ (C1 level);
- Digitális földi műsorsugárzás, DVB-T szabvány: RS$(204, 188)$. Generátorpolinom:

$$\begin{aligned} g(x) = {} & x^{16} + 59x^{15} + 13x^{14} + 104x^{13} + 189x^{12} + 68x^{11} + 209x^{10} \\ & + 30x^9 + 8x^8 + 163x^7 + 65x^6 + 41x^5 + 229x^4 + 98x^3 \\ & + 50x^2 + 36x + 59. \end{aligned}$$

Példa

$GF(8)$ feletti RS$(7,4)$ kód: $n = 7,\ k = 4$

$GF(8)$ minden egynél nagyobb eleme primitív.

$\alpha = 2$ hatványai: $2,\ 4,\ 3,\ 6,\ 7,\ 5,\ 1$.

Generátorpolinom:

$$\begin{aligned} g(x) &= (x-\alpha)(x-\alpha^2)(x-\alpha^3) = (x+\alpha)(x+\alpha^2)(x+\alpha^3) \\ &= x^3 + (\alpha+\alpha^2+\alpha^3)x^2 + (\alpha^3+\alpha^4+\alpha^5)x + \alpha^6 \\ &= x^3 + (2+4+3)x^2 + (3+6+7)x + 5 = x^3 + 5x^2 + 2x + 5. \end{aligned}$$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 5 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 5 & 2 & 5 & 1 & 0 & 0 \\ 0 & 0 & 5 & 2 & 5 & 1 & 0 \\ 0 & 0 & 0 & 5 & 2 & 5 & 1 \end{pmatrix}.$$

Bináris generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 101 & 010 & 101 & 001 & 000 & 000 & 000 \\ 000 & 101 & 010 & 101 & 001 & 000 & 000 \\ 000 & 000 & 101 & 010 & 101 & 001 & 000 \\ 000 & 000 & 000 & 101 & 010 & 101 & 001 \end{pmatrix}.$$

Reed-Müller-kód

Irving S. Reed, David E. Müller, 1954

$n = 2^m$, $m \in \mathbb{N}$, $\boldsymbol{v}_0$ egy csupa egyesből álló n hosszú bináris vektor.

$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$: azon $m \times n$ dimenziós bináris mátrix sorai, melynek oszlopait az összes bináris szám m-es alkotja.

$\boldsymbol{v} \otimes \boldsymbol{w}$: a $\boldsymbol{v}$ és $\boldsymbol{w}$ bináris vektorok koordinátánkénti szorzata.

Példa. $m = 2$, $n = 4$, a vizsgálandó mátrix:

$$V_2 := \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

$$\boldsymbol{v}_0 = 1111,\ \boldsymbol{v}_1 = 0011,\ \boldsymbol{v}_2 = 0101,\ \boldsymbol{v}_0 \otimes \boldsymbol{v}_0 = 1111,\ \boldsymbol{v}_0 \otimes \boldsymbol{v}_1 = 0011,$$
$$\boldsymbol{v}_0 \otimes \boldsymbol{v}_2 = 0101,\ \boldsymbol{v}_1 \otimes \boldsymbol{v}_1 = 0011,\ \boldsymbol{v}_1 \otimes \boldsymbol{v}_2 = 0001,\ \boldsymbol{v}_2 \otimes \boldsymbol{v}_2 = 0101.$$

Definíció. *Az r-edrendű $n = 2^m$ kódszóhosszú Reed-Müller-kód egy olyan bináris, lineáris kód, melynek bázisvektorai a $\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ vektorok, valamint ezek legfeljebb r tagú koordinátánkénti szorzatai. Jelölés: RM(r,m).*

Alkalmazás: NASA Mariner 9 űrszonda (1971), fekete-fehér képek továbbítása RM$(1, 5)$ kóddal.

Tulajdonságok

Egy r-edrendű $n = 2^m$ kódszóhosszú Reed-Müller-kód esetén

$$k = 1 + \binom{m}{1} + \cdots + \binom{m}{r},$$

kódtávolsága pedig $d_{\min} = 2^{m-r}$. A Reed-Müller-kód ekvivalens egy paritásbittel kiegészített ciklikus kóddal.

Példa. RM(2, 3) kód:

$$\boldsymbol{v}_0 = 11111111,\ \boldsymbol{v}_1 = 00001111,\ \boldsymbol{v}_2 = 00110011,\ \boldsymbol{v}_3 = 01010101,$$

$$\boldsymbol{v}_1 \otimes \boldsymbol{v}_2 = 00000011,\ \boldsymbol{v}_1 \otimes \boldsymbol{v}_3 = 00000101,\ \boldsymbol{v}_2 \otimes \boldsymbol{v}_3 = 00010001.$$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Dekódolás

Minden egyes üzenetbitre 2^{m-r} független kifejezés írható fel. Ha nincs hiba, ezek mind egyenlőek. Hiba esetén a gyakoribb értéket választjuk helyesnek.

Példa. Az RM$(1,3)$ kód generátormátrixa:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Kódolandó üzenet: $a_0 a_1 a_2 a_3$; kódszó: $b_0 b_1 b_2 b_3 b_4 b_5 b_6 b_7$.

A generátormátrix által meghatározott egyenletek:

$$\begin{array}{cccc} a_0 = b_0 & a_0 + a_2 = b_2 & a_0 + a_1 = b_4 & a_0 + a_1 + a_2 = b_6 \\ a_0 + a_3 = b_1 & a_0 + a_2 + a_3 = b_3 & a_0 + a_1 + a_3 = b_5 & a_0 + a_1 + a_2 + a_3 = b_7. \end{array}$$

Az üzenetbiteket megadó egyenletek:

$$\begin{array}{cccc} a_0 = b_0 & a_1 = b_0 + b_4 & a_2 = b_0 + b_2 & a_3 = b_0 + b_1 \\ a_0 = b_1 + b_6 + b_7 & a_1 = b_2 + b_6 & a_2 = b_1 + b_3 & a_3 = b_2 + b_3 \\ a_0 = b_2 + b_5 + b_7 & a_1 = b_3 + b_7 & a_2 = b_4 + b_6 & a_3 = b_4 + b_5 \\ a_0 = b_3 + b_4 + b_7; & a_1 = b_1 + b_5; & a_2 = b_5 + b_7; & a_3 = b_6 + b_7. \end{array}$$

Bose-Chaudhuri-Hocquenghem (BCH) kód

Raj Chandra Bose és Dwijendra Kumar Ray-Chaudhuri (1960); Alexis Hocquenghem (1959)

Tétel. *Ha az n kódszóhosszú $GF(s)$ feletti $\mathcal{C}$ ciklikus kód $g(x)$ generátorpolinomjának az $\alpha \in GF(s^m)$ elem $d-1$ egymás utáni (különböző) hatványa gyöke, azaz valamely $i_0 \geq 0,\ d > 1$ esetén*

$$g(\alpha^{i_0}) = g(\alpha^{i_0+1}) = \cdots = g(\alpha^{i_0+d-2}) = 0,$$

akkor a kód minimális távolsága $d_{\min} \geq d$.

Definíció. *Az n kódszóhosszú, $n = s^m - 1$, $GF(s)$ feletti kódot t hibát javító BCH-kódnak nevezzük, ha a $g(x)$ generátorpolinomjának gyökei az $\alpha^i \in GF(s^m),\ i = 1, 2, \ldots, 2t$, testelemek.*

Megjegyzés. A tétel alapján $d_{\min} \geq 2t + 1$.

- $t = 1,\ s = 2$: Hamming-kód.
- $m = 1$: Reed-Solomon-kód. t hibát javító R-S-kód egy n-ed-rendű $\alpha \in GF(s)$ segítségével képzett generátorpolinomja:

$$g(x) = (x-1)(x-\alpha)(x-\alpha^2)\cdots(x-\alpha^{2t-1}).$$

Minimálpolinomok

Definíció. *Egy $\alpha \in GF(s^m)$ elem $GF(s)$ feletti minimálpolinomja a legkisebb fokszámú $GF(s)$-beli együtthatókkal rendelkező főpolinom, melynek α gyöke.*

A t hibát javító BCH-kód generátorpolinomja az $\alpha^i \in GF(s^m)$ testelemek $GF(s)$ feletti minimálpolinomjainak legkisebb közös többszöröse.

Példa. $s = 2, m = 3, t = 1$, azaz $n = 2^3 - 1 = 7$.

Legyen $\alpha = 3 \in GF(8)$, $\alpha^2 = 5$. A $GF(8)$-beli műveletek alapján

$$g(x) = x^3 + x^2 + 1 = (x + \alpha)(x + \alpha^2)(x + \alpha^2 + \alpha + 1).$$

Generátormátrix:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

4 üzenet- és 3 paritásbit. Kódtávolság: $d_{\min} = 3$.

Kódfűzés

Egy $\mathcal{C} = \mathcal{C}(n, k)$ kód m-szeres átfűzése egy $\mathcal{C}^m = \mathcal{C}(mn, mk)$ kódot eredményez.

$\boldsymbol{c}^{(i)} = \left(c_0^{(i)}, c_1^{(i)}, \ldots, c_{n-1}^{(i)}\right)$, $i = 1, 2, \ldots, m$: a $\mathcal{C}$ kód kódszavai.

A $\mathcal{C}^m$ átfűzéses kód megfelelő kódszava:

$$\boldsymbol{c} = \left(c_0^{(1)}, c_0^{(2)}, \ldots, c_0^{(m)}, c_1^{(1)}, c_1^{(2)}, \ldots, c_1^{(m)}, \ldots, c_{n-1}^{(1)}, c_{n-1}^{(2)}, \ldots, c_{n-1}^{(m)}\right).$$

A $\boldsymbol{c}^{(i)}$ kódszavakat egy mátrix sorainak tekintjük, majd az mátrix elemeit oszloponként kiolvassuk.

- Lineáris kód átfűzése lineáris kódot eredményez.
- Ha $\mathcal{C}$ kódtávolsága $d_{\min} = d$, akkor $\mathcal{C}^m$ kódtávolsága is d.
- Ciklikus kód átfűzése ciklikus kódot eredményez.

Tétel. *Ha $g(x)$ a $\mathcal{C}$ kód generátorpolinomja, akkor $g\left(x^m\right)$ a $\mathcal{C}^m$ kód generátorpolinomja.*

Az azonos kódtávolság miatt a $\mathcal{C}^m$ ugyanannyi egyszerű, illetve törléses hibát javít, mint a $\mathcal{C}$ kód.

Csomós hibák javítása

Definíció. *A hibavektor egy ℓ hosszúságú szegmense hibacsomó ℓ hosszal, ha a szegmens első és utolsó karaktere nem zéró. Egy kód ℓ hosszúságú hibacsomót javító, ha minden legfeljebb ℓ hosszúságú hibacsomó javítható.*

Tétel. *A $\mathcal{C}^m$ átfűzéses kód $m \cdot t$ hosszúságú hibacsomó javító, ahol t a $\mathcal{C}$ kód hibajavító képessége.*

Tétel. *Egy $\mathcal{C}(n,k)$ lineáris kód ℓ hibajavító képességére teljesül az alábbi Reiger-korlát*

$$\ell \leq \left\lfloor \frac{n-k}{2} \right\rfloor .$$

Definíció. *Azokat a hibajavító kódokat, melyeknél a Reiger-korlátban egyenlőség áll fenn Reiger-optimális kódoknak nevezzük.*

Megjegyzés. Egy MDS kód Reiger-optimális.

Indoklás. MDS kód esetén $d_{min} = n-k+1$, azaz $t = \lfloor (d_{\min}-1)/2 \rfloor = \lfloor (n-k)/2 \rfloor$ egyszerű hibát javít. Mivel $\ell \geq t$, a Reiger-korlátban egyenlőség van. □

Példa

$GF(2)$ feletti

$$g(x) = 1 + x^2 + x^3 + x^4$$

generátorpolinomú $\mathcal{C}(7,3)$ ciklikus kód. Kódtávolság: $d_{\min} = 4$.

Generátor- és paritásmátrixa:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}, \quad \boldsymbol{H} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Az x^i, $i = 1, 2, \ldots 6$, és az $x^i + x^{i+1}$, $i = 0, 1, \ldots, 5$, hibapolinomok szindrómái, azaz a $g(x)$ polinommal való osztáskor kapott maradékai, különbözőek. A nekik megfelelő vektorok javítható hibaminták. Az $\ell = 2$ hosszúságú hibacsomók javíthatóak.

$$\left\lfloor \frac{n-k}{2} \right\rfloor = 2 = \ell : \qquad \text{a kód Reiger-optimális.}$$

Szorzatkódok

$\mathcal{C}(n_i, k_i, d_i)$: d_i kódtávolságú (n_i, k_i) típusú lineáris kódok, $i=1, 2$.

$\mathcal{C}_1 \times \mathcal{C}_2\,(n_1 \cdot n_2, k_1 \cdot k_2, d_1 \cdot d_2)$: szorzatkód. Lineáris.

Elemei: $n_1 \times n_2$ dimenziós mátrixok. Soronként kiolvasva adják a kódszavakat.

Sorok és oszlopok: $\mathcal{C}_1$, illetve $\mathcal{C}_2$ kódbeli kódszavak.

- Üzenethossz: $k_1 \cdot k_2$.
- Kódszóhossz: $n_1 \cdot n_2$.
- Kódtávolság: $d_1 \cdot d_2$.

Szisztematikus $\mathcal{C}_1$ és $\mathcal{C}_2$ esetén az üzenet a bal felső $k_1 \times k_2$ méretű minor.

Ciklikus komponensek esetén létezik olyan elrendezés, hogy a soronkénti kiolvasással kapott kódszavak ciklikusak legyenek.

A szorzatkód képzése

$$\begin{bmatrix} \boldsymbol{c_0^{(0)}} & \boldsymbol{c_1^{(0)}} & \cdots & \boldsymbol{c_{k_1-1}^{(0)}} & c_{k_1}^{(0)} & \cdots & c_{n_1-1}^{(0)} \\ \boldsymbol{c_0^{(1)}} & \boldsymbol{c_1^{(1)}} & \cdots & \boldsymbol{c_{k_1-1}^{(1)}} & c_{k_1}^{(1)} & \cdots & c_{n_1-1}^{(1)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \boldsymbol{c_0^{(k_2-1)}} & \boldsymbol{c_1^{(k_2-1)}} & \cdots & \boldsymbol{c_{k_1-1}^{(k_2-1)}} & c_{k_1}^{(k_2-1)} & \cdots & c_{n_1-1}^{(k_2-1)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ c_0^{(n_2-1)} & c_1^{(n_2-1)} & \cdots & c_{k_1-1}^{(n_2-1)} & c_{k_1}^{(n_2-1)} & \cdots & c_{n_1-1}^{(n_2-1)} \end{bmatrix}$$

Első k_1 oszlop: $\mathcal{C}_2(n_2, k_2, d_2)$ alapján szisztematikus kódolással.

Első k_2 sor: $\mathcal{C}_1(n_1, k_1, d_1)$ alapján szisztematikus kódolással.

Jobb alsó sarok: paritások paritása. Akár az első k_1 sor, akár az első k_2 oszlop paritásaiból képezhető a $\mathcal{C}_2$, illetve $\mathcal{C}_1$ kódszavaival.

Példa. $\mathcal{C} \times \mathcal{C}$ szorzatkód, komponensei egyetlen paritásbittel rendelkező $\mathcal{C}(n, n-1, 2)$ kódok.

A szorzatkód kódtávolsága 4: egy egyszerű hibát javít.

Kaszkád kódok

$\mathcal{C}_1(n_1, k_1, d_1)$: $GF(s)$ test feletti lineáris kód. Belső kód.

$\mathcal{C}_2(N_2, K_2, D_2)$: $GF(s^{k_1})$ test feletti lineáris kód. Külső kód.

A $\mathcal{C}(n_1 N_2, k_1 K_2, d)$ paraméterű $GF(s)$ feletti kaszkád kód képzése:

1. A $k_1 K_2$ hosszú üzenetet k_1 hosszú szegmensekre osztjuk. Egy ilyen szemenst a $\mathcal{C}_2$ kód egyetlen karakternek tekint.
2. A K_2 karakter hosszú $GF(s^{k_1})$ test feletti forrásüzenetből a $\mathcal{C}_2$ kód egy N_2 karakter hosszú kódszót képez. Ez $GF(s)$ felett $k_1 N_2$ karaktert jelent.
3. A $\mathcal{C}_2$ szerinti kódot osszuk fel k_1 hosszú szegmensekre, amikből a $\mathcal{C}_1$ kód egy n_1 karakter hosszú kódszavakat képez. Ez eredményezi a $k_1 K_2$ üzenethez tartozó $n_1 N_2$ hosszú kódot.

A kaszkád kód kódtávolsága: $d \geq d_1 D_2$.

Először a $\mathcal{C}_1$ kódszavait dekódoljuk, majd a $\mathcal{C}_2$ kódszavát.

$\mathcal{C}_2$ elsősorban a csomós, $\mathcal{C}_1$ pedig a véletlen hibákat javítja.

Példa.

$\mathcal{C}_1$: $GF(s)$ feletti $(7, 3)$ paraméterű ciklikus kód. Generátormátrixa:

$$\boldsymbol{G}_1 = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

$\mathcal{C}_2$: $GF(8)$ feletti RS$(7, 4)$ kód. Generátormátrixa:

$$\boldsymbol{G}_2 = \begin{pmatrix} 5 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 5 & 2 & 5 & 1 & 0 & 0 \\ 0 & 0 & 5 & 2 & 5 & 1 & 0 \\ 0 & 0 & 0 & 5 & 2 & 5 & 1 \end{pmatrix}.$$

A kaszkád kód $(49, 12)$ paraméterű. Forrásüzenet, például:

$\boldsymbol{u} = 110\ 111\ 010\ 101$, $GF(8)$ feletti alakban: 6 7 2 5.

$\mathcal{C}_2$ szerinti kódszó:

$\boldsymbol{c} = 3\ 1\ 7\ 5\ 7\ 5\ 5$, bináris alak: 011 001 111 101 111 101 101.

$\mathcal{C}_1$ szerinti kódszavak:

0111001 0010111 1100101 1001011 1100101 1001011 1001011.