Cvičení 5. (kapitola 5.) – Úvod do lineárních kódů

Abeceda kódu $\mathbb{F}_q \sim GF(q)$ – konečné těleso, kde q je mocnina prv.

Množina všech slov nad $GF(q)$ je $(\mathbb{F}_q)^n \sim V(n,q)$ – vektorový prostor

Def: Lineární kód nad $GF(q)$ je podprostor ve $V(n,q)$.

Značení: Je-li C k-dimenzionální podprostor ve $V(n,q)$, pak C se nazývá $[n,k]$-kód, nebo $[n,k,d]$-kód.

Vlastnosti:
- Jestliže C je q-ární $[n,k,d]$-kód, pak C je také q-ární (n, q^k, d)-kód
- Lin. kód vždy obsahuje nulové slovo, $\bar{0}$.
- $dist(C) = \min\{w(\bar{x}) : \bar{x} \in C\} = w(C)$ (váha kódu)

Myšlenka důkazu (podprostora; protože rozdíl je prvkem C; $dist(C) = w(C)$):

$$dist(C) = dist(\bar{x},\bar{y}) \overset{lin}{=} w(\bar{x}-\bar{y}) \geq w(C)$$

$$w(C) = w(\bar{z}) \overset{lin}{=} w(\bar{z}-\bar{0}) = dist(\bar{z},\bar{0}) \geq dist(C)$$

Výhody a nevýhody lin. kódů

Výhody
- Lehce se dá určit $dist(C)$
- $[n,k]$-kód lze určit zadáním pouze k-vektorů báze.
  Pro $[n,k]$-kód generující matice A je $k \times n$ matice s bázovými vektory v řádcích.
- „pěkné" metody kódování a dekódování

Nevýhody
- Potřebujeme-li $q \neq$ od mocniny prvočísla, lin-kód neexistuje
- Někdy, pro dané parametry může existovat i lepší kód, než lineární. Kdy??? (Viz příklad 3.)

Def: Ekvivalence lin. kódů.

C_1, C_2 lineární, $C_1 \sim C_2 \iff C_1$ lze dostat z C_2
- permutací pozic kódu
- násobením symbolů na první pozici nenulovým skalárem

Věta Ekvivalence lin. kódů pomocí gener. matic

C_1, C_2 ← lineární, $C_1 \sim C_2 \Leftrightarrow A_1, A_2$ ← gener. matice

A_1 dostanem z A_2

1. permutací řádků — R_1
2. násobením řádku nenul. skalárem — R_2
3. přičtení nás. jednoho ř. k jinému — R_3

} řádkově ekvivalentní (stejný jen jiná báze)

4. permutací sloupců — S_1
5. násobení sloupce nenul. skalárem — S_2

} ekvivalentní dle definice ekviv.

Věta: Gener. matice lin. kódu může být úpravami 1–5 vždy převedena na „standardní" tvar.

$$G \sim [I_k \mid A]$$

Příklady z přednášky

- kódy C_1, C_2, C_3, C_F, C-opak, jsou lineární
- jejich generující matice, jejich parametry ?
- C_5 – není lineární

Př 1. (i) Ukažte, že E_n je podprostor ve $V(n,2)$

Pro binární abecedu F_2 stačí ukázat jen sčítání.

Důkaz. $\forall \vec{x}, \vec{y} \in E_n$ platí, že $\vec{x}+\vec{y} \overset{?}{\in} E_n$

$\vec{x}$: $w(\vec{x}) = 2k = \sum_{i=1}^{n} x_i$, $k, l \in \mathbb{N}_0$

$\vec{y}$: $w(\vec{y}) = 2l = \sum_{i=1}^{n} y_i$

$dist(\vec{x},\vec{y}) = w(\vec{x}-\vec{y}) =$ pro bin kód $= w(\vec{x}+\vec{y}) = w(\vec{x}) + w(\vec{y}) - 2w(\vec{x}\wedge\vec{y})$

$w(\vec{x}+\vec{y}) = w(\vec{x}) + w(\vec{y}) - 2\,w(\vec{x}\wedge\vec{y})$ Lemma (2.4. v přednášce) – také 2. cvičení

$= 2k + 2l - 2p = 2(k+l-p)$ ← sudé č. $p \in \mathbb{N}_0$

$\Rightarrow \vec{x}+\vec{y} \in E_n$

(ii) Najděte příklad báze pro E_n v prostoru $V(n,2)$.
a určete dimenzi E_n

Ukázali jsme si, že E_n je $(n, 2^{n-1}, 2)$-kód, který vznikne z $V(n-1,2)$ přidáním paritního bitu. Podobně zkusíme sestavit bázi.

Například

$B = ((1,0,0,\overset{b_1}{\dots}0,1),(0,1,0\overset{b_2}{\dots}0,1),(0,0,1,\dots,0,1)\dots(0,0,0\overset{b_{n-1}}{\dots},1,1))$

Je třeba ukázat, že vektory jsou Lin. Nez. a že generují celý podprostor E_n.

Zapíšeme vektory báze do řádků matice

$$\begin{bmatrix} 1&0&0&\dots&0&1\\ 0&1&0&\dots&0&1\\ 0&0&1&\dots&0&1\\ \vdots&&&\ddots&&\vdots\\ 0&0&0&\dots&1&1 \end{bmatrix}$$

Řád matice $n-1 \times n$

Hodnost $n-1$. Matice má zřejmě $n-1$ nezávislých sloupců (řádků)

Zřejmě každý vektor $\bar{x} \in E_n$ lze zapsat jako kombinaci $\bar{b}_1, \dots, \bar{b}_{n-1}$

Je-li na pozicích $1, \dots, n-1$

- lichý počet jedniček bude na poslední pozici 1.
- sudý počet jedniček bude na poslední pozici 0.

$Dim(E_n) = n-1$

Př. 2. $E_n \subset V(n,2)$ a je lineární.

Jaké jsou parametry kódu $C = E_n$?

Zapište generující matici pro E_n v základním tvaru.

E_n je $(n, 2^{n-1}, 2)$-kód. Protože je lineární, je současně i $[n, n-1, 2]$-kód.

Potřebujeme $n-1$ nezávislých vektorů pro generující matici.

Např. může být

$$G = \begin{bmatrix} 1&0&0&\dots&0&1\\ 0&1&0&\dots&0&1\\ 0&&&\ddots&&\vdots\\ \vdots&&&&1&0&1\\ 0&\dots&&&0&1&1 \end{bmatrix}$$

G je $n-1 \times n$ matice tvaru $[I_{n-1} \vdots \bar{1}]$, kde v posl. sloupci jsou jedničky: $\bar{s}_n = \bar{1}$

Př 3. Je lineární $(11, 24, 5)$-kód (kód sestrojený z Hadamardova designu) lineárním kódem?

$24 \neq 2^2$ => není lin. kódem

každý $[n, k, d]$-kód je také (n, q^k, d)-kód

Př ✓ $\bar{b}_1 \quad \bar{b}_2 \quad \bar{b}_3$

3.4.4. Ukažte, že a) trojice vektorů $\{(1,2,0),(2,1,1),(0,2,0)\}$ tvoří množinu lineárně nezávislých vektorů v prostoru $V(4,3)$, b) $\{(1,2,0),(2,1,1),(0,0,2)\}$ není lineárně nezávislá množina vektorů.

a)
$$A = \begin{bmatrix} 1&2&0 \\ 2&1&1 \\ 0&2&0 \end{bmatrix} +r_1 \sim \begin{bmatrix} 1&2&0 \\ 0&0&1 \\ 0&2&0 \end{bmatrix} \sim \begin{bmatrix} 1&2&0 \\ 0&2&0 \\ 0&0&1 \end{bmatrix} \quad h(A)=3 \Rightarrow \text{lin. nezávislé}$$

$$A^T = \begin{bmatrix} 1&2&0 \\ 2&1&2 \\ 0&1&0 \end{bmatrix} +r_1 \sim \begin{bmatrix} 1&2&0 \\ 0&0&2 \\ 0&1&0 \end{bmatrix} \sim \begin{bmatrix} 1&2&0 \\ 0&1&0 \\ 0&0&2 \end{bmatrix} \cdot 2 \sim \begin{bmatrix} 1&2&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}$$

vektory ve sloupcích odpovídá řešení soust $a_1\bar{b}_1 + a_2\bar{b}_2 + a_3\bar{b}_3 = \bar{0}$

b)
$$B = \begin{bmatrix} 1&2&0 \\ 2&1&1 \\ 0&0&2 \end{bmatrix} +r_1 \sim \begin{bmatrix} 1&2&0 \\ 0&0&1 \\ 0&0&2 \end{bmatrix} +r_2 \sim \begin{bmatrix} 1&2&0 \\ 0&0&1 \\ 0&0&0 \end{bmatrix} \Rightarrow h(B)=2 \Rightarrow$$

Lin. závislé

[Jádro lin. zobrazení]

Př 4. Označíme H matici $r \times n$ nad $GF(q)$.

Ukažte, že $C = \{\bar{x} \in V(n,q) : \vec{x} \cdot H^T = \vec{0}\}$ je lineární kód.

[Pozn: Později zavedeme H jako kontrolní matici kódu]

Potř. ukázat, že C je podprostor ve $V(n,q)$

(i) $\forall \bar{x}, \bar{y} \in C, \quad \vec{x} + \vec{y} \overset{?}{\in} C$

$$(\vec{x} + \vec{y}) \cdot H^T = \vec{x} \cdot H^T + \vec{y} \cdot H^T = \vec{0} + \vec{0} = \vec{0} \Rightarrow \vec{x} + \vec{y} \in C$$

(ii) $\forall a \in GF(q) \wedge \vec{x} \in C, \quad a \cdot \vec{x} \overset{?}{\in} C$

$$(a \cdot \vec{x}) \cdot H^T = a \cdot (\vec{x} \cdot H^T) = a \cdot \vec{0} = \vec{0} \Rightarrow a \cdot \vec{x} \in C$$

C je podpr. $V(n,q)$

C je lineární kód ve $V(n,q)$.

Př. 5. (i) Ukažte že, je-li C binární lineární kód, pak kód, který dostaneme přidáním kontrolního paritního symbolu, je také lineární. (Nejedná se jen o kód E_n, jde o obecnější kód)

(ii) Sestavte generující matici binárního $[8,4,4]$-kódu.

z C_F

C_F je $(7,16,3)$-kód

C_1 ← kód ve $V(n,2)$, lineární

$[n, k_1, d_1]$ - kód

C_2 ← kód, který vznikne z C_1 přidáním par. symb.

$[n+1, k_2, d_2]$, $k_1 = k_2$, $d_2 \geq d_1$, $|C_1| = |C_2|$

$$C_2 = \left\{ \hat{\bar{x}} = (x_1, x_2, \dots, x_n, \sum_{i=1}^{n} x_i),\ \bar{x} = (x_1, x_2, \dots, x_n) \in C_1 \right\}$$

Ukažte, že C_2 je lineární.

(1) $\forall\ \hat{\bar{x}}, \hat{\bar{y}} \in C_2$ platí? $\hat{\bar{x}} + \hat{\bar{y}} \in C_2$?

$$\hat{\bar{z}} = \hat{\bar{x}} + \hat{\bar{y}} = \left(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n, \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i \right) =$$

$$= \left(x_1 + y_1, \dots, x_n + y_n, \sum_{i=1}^{n} (x_i + y_i) \right) = \widehat{\bar{x} + \bar{y}} \in C_2$$

Pozn. pro binární lin. kódy stačí ukázat pouze první podmínku, (1) tj. $(\bar{x}, \bar{y}) \in C_1$. Druhá je $a.\bar{x} \in C$, ale a je jen 0 v 1. což je splněno automaticky pokud platí (1)

(ii) $[8,4,4]$ - kód. je $(8,16,4)$ - kód ozn jej C_1

$$\left[\begin{array}{cccc:cccc} 1&0&0&0&0&1&1&1 \\ 0&1&0&0&1&0&1&1 \\ 0&0&1&0&1&1&0&1 \\ 0&0&0&1&1&1&1&0 \end{array}\right]$$

← například takto

každý řádek musí obsahovat (alespoň) právě 4 jedničky. aby $w(c) = 4$

C_F

$$\left[\begin{array}{ccccccc} 1&0&0&0&1&0&1 \\ 0&1&0&0&1&1&1 \\ 0&0&1&0&1&1&0 \\ 0&0&0&1&0&1&1 \end{array}\right] \begin{matrix} 1 \\ 0 \\ 1 \\ 1 \end{matrix} = G \text{ pro } C_2$$

Je kód C_1 stejný jako kód C_2, který vznikne přidáním paritního symbolu k C_F? C_1 je jen ekvivalentní s C_2 – ne stejný.

Př. 6. Dokažte, že v binárním lin. kódu, buď všechna slova mají sudou váhu, nebo přesně polovina má váhu sudou a polovina váhu lichou.

↓ Na cvičení vynecháno.

Označíme E_v – slova sudé váhy
O_d – slova liché váhy pak $E_v \cup O_d = C$.

A je-li $C = E_v$ – všechna slova jsou sudé váhy.

Předpokl., že $O_d \neq \emptyset$, tj. $\exists \bar{y} \in C \wedge \bar{y} \in O_d$.

- $E_v + \bar{y} = \{\bar{x} + \bar{y} \mid \bar{x} \in E_v\} \subseteq O_d$ a tedy
  $|E_v + \bar{y}| = |E_v| \leq |O_d|$
- $O_d + \bar{y} = \{\bar{x} + \bar{y} \mid \bar{x} \in O_d\} \subseteq E_v$
  $|O_d + \bar{y}| = |O_d| \leq |E_v|$

$|E_v| = |O_d| = \frac{1}{2}|C|$

Pozn. $w(\bar{x} + \bar{y}) = w(\bar{x}) + w(\bar{y}) - 2w(\bar{x} \cdot \bar{y}) \Rightarrow$

- váha součtu vektorů liché váhy je sudá
- — " — — jeden sudý a jeden lichý je lichá

Př. 7 Nechť C_1 a C_2 jsou binární lineární kódy, které mají generující matice.

(i) $G_1 = \begin{bmatrix} 1&1&1&1&0 \\ 0&0&1&1&1 \end{bmatrix}$ (ii) $G_2 = \begin{bmatrix} 1&0&0&1&1&0&1 \\ 0&1&0&1&0&1&1 \\ 0&0&1&0&1&1&1 \end{bmatrix}$

zapište všechna slova daných kódů a na základě (Věty 2.2.: $w(C) = dist(C)$) určete jejich min. vzdálenost.

$C_1 = \{ a\bar{v}_1 + b\bar{v}_2 \text{ kde } \bar{v}_1, \bar{v}_2 \in G \text{ a } \forall a, b \in GF(q) \}$

C_1 - binární $[5, 2, 3]$-kód $\Rightarrow |C_1| = 2^2 = 4$

(i) $C_1 = \begin{cases} 00000 \\ 11110 \\ 00111 \\ 11001 \end{cases}$

$\downarrow$

$dist(C) = w(C) = 3$

$a \cdot v_1 + b \cdot v_2$

$\begin{matrix} 00 \\ 10 \\ 01 \\ 11 \end{matrix}$

(ii) C_2 - binární $[7, 3, 4]$-kód $\sim (7, 8, 4)$-kód

$C_2 = \begin{cases} 0000000 & & 000 \\ 1001101 & 1. & 100 \\ 0101011 & 2. & 010 \\ 0010111 & 3. & 001 \\ 1100110 & r_1 + r_2 & 110 \\ 1011010 & r_1 + r_3 & 101 \\ 0111100 & r_2 + r_3 & 011 \\ 1110001 & r_1 + r_2 + r_3 & 111 \end{cases}$

$a\bar{r}_1 + b\bar{r}_2 + c\bar{r}_3$

$\downarrow$

$dist(C) = w(C) = 4$

Př. 8. Nechť C je ternární lineární kód s generující maticí

$\begin{bmatrix} 1&0&1&1 \\ 0&1&1&2 \end{bmatrix}$

Zapište všechna slova kódu C. Určete jeho min. vzdálenost. Rozhodněte, zda je perfektní.

C je ternární $[4, 2, ?]$-kód tj. $(4, 3^2 = 9, ?)$-kód.

$C = \begin{cases} 0000^{00} & 2022^{20} & 2210^{22} \\ 1011^{10} & 0221^{02} & 1202^{12} \\ 0112^{01} & 1120^{11} & 2101^{21} \end{cases}$ $dist(C) = w(C) = 3$

Je perfektní? $M \cdot \underbrace{\left(\binom{4}{0}2^0 + \binom{4}{1}2^1\right)}_{1 \quad + \quad 8} \leq 3^4 \Rightarrow M \leq 9 \Rightarrow$

$\Rightarrow C$ je perfektní.

Př. 9 Nechť C je lineární kód s generující maticí $[I_k | A]$
Ukažte, že jakákoli permutace řádků A vede
k matici generující kód ekvivalentní kódu C.

Ozn C_1 kód s $G_1 = [I_k, A]$
C_2 kód s $G_2 = [I_k, B]$

C_2 je ekviv. C_1, jestliže G_2 lze dostat z G_1
operacemi $R_1 - R_3$ a S_1, S_2 z definice ekvivalence kódů.
Jestliže B vznikne z A lib. permutací řádků
pak matice $[I_k | B]$ generuje kód ekvivalentní
kódu C_1. Proč?

Protože: Nějakou matici $[C | B]$ dostanem z $[I_k | A]$ stejnou
permutací řádků π jako $A \xrightarrow{\pi} B$, tj. $[I_k, A] \xrightarrow{\pi} [C | B]$
Pak inverzní permutací π^{-1} sloupců C dostanem I_k, $C \xrightarrow{\pi^{-1}} I_k$
a tedy $[C | B] \xrightarrow{\pi^{-1}} [I_k | B]$, kde $[I_k, B]$ je gener matice
kódu C_2. Kód C_2 je lineární a ekvivalentní kódu C_1.
Kód C_2 není stejný jako kód C_1, protože permutací různých
sloupců (což sloupce I_k jsou) dostaneme jiný vektorový
podprostor ve $V(n, q)$

Ke každé permutaci existuje perm. inverzní

π cyklický zápis: $(1\,2\,3)(4)$ maticový zápis: $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$

$\pi^3 = id$
$\pi \cdot \pi^2 = id$
$(\pi^{-1}) = \pi^2$
$\pi \cdot \pi^{-1} = id$
$((1\,2\,3)(4)) \cdot (1\,3\,2)(4) = (1)(2)(3)(4) = id$

DU: Nechť C je binární lineární kód s generující maticí

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

$(7, 2^4, 3)$ - kód

Najděte generující matici pro C ve standardním tvaru.

Je kód C stejný jako kód C^F ?

Je kód C ekvivalentní kódu C^F ?