2 Hrátky s maticemi

Obecně lze blokový kód délky n definovat jako neprázdnou podmnožinu C množiny $\mathbb{A}^n$, kde $\mathbb{A}$ je nějaká neprázdná konečná množina. $\mathbb{A}$ je zvykem nazývat **abecedou**. Je-li $|\mathbb{A}| = q$, hovoříme o q-árním kódu. Kód je lineární, jestliže $\mathbb{A}$ lze ztotožnit s $\mathbb{F}_q$ tak, aby C bylo lineárním podprostorem. Pokud q není mocnina prvočísla, tak samozřejmě o lineární kód jít nemůže.

Má-li $C \subseteq \mathbb{A}^n$ právě h prvků, hovoříme o (n, h) kódu, případně o $(n, h)_q$ kódu. Lineární kód délky n a dimenze k je tedy $[n, k]_q$ kódem a $(n, q^k)_q$ kódem.

Minimální vzdálenost kódu C je rovna $\min\{d(u, v)\}$, kde $u, v \in C$ a $u \neq v$. Má-li kód C jediný prvek, je jeho minimální vzdálenost rovna $n + 1$. Známe-li minimální vzdálenost d, značíme $(n, h)_q$ kód jako $(n, h, d)_q$ kód. Pro obecně nelineární kódy platí tvrzení 1.3 v nezměněné podobě, na důkazu není třeba nic měnit.

Většinou, avšak ne výhradně, se budeme zabývat kódy lineárními. Nelineární blokové kódy budeme uvažovat zejména tehdy, když budeme chtít vyložit, že určitý lineární kód poskytuje optimální chování, které nelze vylepšit ani při přechodu ke kódu nelineárnímu.

Pro $(n, h)_q$ kód C definujeme **informační poměr** (anglicky rate, jako český jednoslovný ekvivalent budeme používat též slovo **nosnost**) jako zlomek $\frac{\log_q h}{n}$. Pro $[n, k]_q$ kód je nosnost rovna k/n. Informační poměr udává, jaká část zakódované zprávy je potřebná ke sdělení jejího informačního (nekódovaného) obsahu. Cílem teorie kódů je navrhnout kódy, které při zadaném informačním poměru opravují efektivně co nejvíce chyb; případně při zadané schopnosti opravovat maximalizují informační poměr. V praxi se ukazuje jako nutné reagovat na skutečnost, že chyby nebývají rozděleny náhodně, ale vyskytují se ve shlucích. Základem teorie samoopravných kódů jsou však lineární kódy konstruované s předpokladem náhodného výskytu chyby.

Buď F komutativní těleso. Při práci s lineárními kódy má zásadní význam **bodový součin** $u \cdot v = \sum_{i=1}^{n} u_i v_i$ pro vektory $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n) \in F^n$. V případě $F = \mathbb{R}$ jde o součin skalární. Někdy se bodový součin nazývá skalární i nad jinými tělesy, není to však přesné. Je-li například $F = \{0, 1\}$, tak $u \cdot u = 0$ právě když $|u| \equiv 0 \bmod 2$.

Z terminologie skalárního součinu si nicméně vypůjčíme pojem **ortogonální doplněk**. Je-li $M \subseteq F^n$, tak ortogonální doplněk $M^\perp = \{v \in F^n;\ v \cdot u = 0\ \forall u \in M\}$. Lineární obal množiny M budeme značit $\langle M \rangle$.

Lemma 2.1. *Pro každé u, v, $w \in F^n$ je*

- $u \cdot v = v \cdot u$
- $u \cdot (v + w) = u \cdot v + u \cdot w$
- $\forall \lambda \in F$ *je* $u \cdot (\lambda v) = \lambda (u \cdot v)$

Tedy $\cdot$ je symetrická bilineární forma na F^n.

□

Lemma 2.2. *Buď M, $N \subseteq F^n$. Pak*

- *$M^\perp$ je lineární podprostor F^n*
- *$\langle M \rangle^\perp = M^\perp$*
- *$M \subseteq \left(M^\perp\right)^\perp$*
- *z $M \subseteq N$ plyne $M^\perp \supseteq N^\perp$*

□

Lemma 2.3. *Pro M, $N \subseteq F^n$ je $(M \cup N)^\perp = M^\perp \cap N^\perp$. Pokud $0 \in M \cap N$, tak $(M \cup N)^\perp = (M + N)^\perp$, kde $M + N = \{u + v;\, u \in M,\, v \in N\}$.*

□

Důkazy těchto tvrzení není třeba uvádět. Jsou snadné a odpovídají důkazům známým pro skalární součin.

Je-li $u = (\lambda_1, \ldots, \lambda_n) \in F^n$, $u \neq 0$, tak $u^\perp = \{(a_1, \ldots, a_n) \in F^n;\, \sum_{i=1}^n \lambda_i a_i = 0\}$ je zjevně podprostor dimenze $n-1$. Takovým podprostorům se říká **nadrovina**. Je-li $U \subseteq F^n$ podprostor dimenze k a W je nadrovina, tak buď $U \subseteq W$, nebo $U \cap W$ má dimenzi $k-1$.

Buď nyní $0 = U_1 \subset U_1 \subset \ldots \subset U_n = F^n$ řada do sebe vložených podprostorů. Vidíme, že $\dim U_j = j$. Zvolme bázi $e_1, \ldots, e_n$ prostoru F_n tak, aby $U_{i+1} = \langle U_i, e_{i+1} \rangle$, $0 \leq i \leq n-1$. Podle Lemmatu 2.3 je $U_j^\perp = U_{j-1}^\perp \cap e_j^\perp$. Proto má $U_j^\perp$ dimenzi nejvýše o jedničku menší, než $U_{j-1}^\perp$. Máme

$$0 = U_n^\perp \subseteq U_{n-1}^\perp \subseteq \cdots \subseteq U_1^\perp \subseteq U_0^\perp = F^n,$$

přičemž sousední podprostory se liší v dimenzi nanejvýš o jedničku. Jestliže jich je $n+1$, musí být $\dim U_j^\perp = n - j$.

Je-li dán podprostor $V \subseteq F^n$ dimenze k, tak lze jistě sestrojit $0 = U_1 \subset U_1 \subset \ldots \subset U_n = F^n$ tak, aby V bylo rovno některému U_j, $0 \leq j \leq n$. Proto platí

Lemma 2.4. *Ať V je podprostor F^n. Pak $\dim V + \dim V^\perp = n$ a $\left(V^\perp\right)^\perp = V$.*

□

Ať C je $[n,k]_q$ kód. Z lemmatu 2.4 plyne, že $C^\perp$ je $[n, n-k]_q$ kód. Ať H je generující matice $C^\perp$. Označme její řádky $v_1, \ldots, v_{k'}$, kde $k' = n-k$. Víme, že $C^\perp = \langle v_1, \ldots, v_{k'} \rangle$, takže $\left\{u \in F^n; Hu^T = 0\right\} = \{v_1, \ldots, v_{k'}\}^\perp = \left(C^\perp\right)^\perp = C$. Dokázali jsme:

Tvrzení 2.5. *Generující matice doplňkového kódu $C^\perp$ se shodují s prověrkovými maticemi kódu C.*

□

Ze znalosti generující matice není často snadné zjistit minimální váhu kódu. Někdy může pomoci následující pozorování:

Tvrzení 2.6. *Ať C je $[n,k,d]_q$ kód s prověrkovou maticí H. Potom platí, že $d-1$ je rovno největšímu číslu r takovému, že každých r sloupců H je lineárně nezávislých.*

Důkaz. Případ $r=0$ nastává právě tehdy, je-li alespoň jeden ze sloupců matice H nulový. V takovém případě je zjevně $d=1$. Ať $r \geq 1$ a ať kódové slovo $(\lambda_1,\ldots,\lambda_n) \in C$ má právě $t>0$ nenulových hodnot λ_i. Zvolíme-li řádek H, jenž je nenulový v některém sloupci i, pro který platí $\lambda_i \neq 0$, tak vidíme, že odpovídajících t sloupců je lineárně závislých. Proto $t \geq r+1$. Je-li $C=0$, tak $d=n+1$ a $r=n$. Předpokládejme, že $C \neq 0$. Pak lze volit $t=d$, takže $d-1 \geq r$. Současně víme, že existuje vektor $u=(\lambda_1,\ldots,\lambda_n)$ s právě $r+1$ nenulovými hodnotami takový, že $Hu^\mathsf{T}=0$. To znamená, že $u \in C$, a proto $r+1 \geq d$. □

Důsledek 2.7. *Ať C je $[n,k,d]_q$ kód. Potom $d \leq n-k+1$.*

Důkaz. Provérková matice H může obsahovat regulární matici řádu s nanejvýše pro $s=n-k$ (to je počet řádků). Proto $d-1 \leq n-k$. □

Výše uvedená nerovnost se nazývá **Singletonův odhad**. Platí i pro nelineární kódy. Lineární kódy, ve kterých $d=n-k+1$, se nazývají **MDS** (maximum distance separable). Z Důsledku 2.7 okamžitě plyne jejich charakterizace pomocí prověrkové matice:

Důsledek 2.8. *Ať C je $[n,k,d]_q$ kód s prověrkovou maticí H. Kód C je MDS právě když každá její čtvercová podmatice řádu $n-k$ je regulární.*

Zvolme $\ell \geq 2$ a ať $F=\mathbb{F}_q$, $V=F^\ell$ a $V^\# = V\backslash\{0\}$. Na $V^\#$ definujeme ekvivalenci $\sim$ tak, že $u \sim v$ právě když $u=\lambda v$ pro nějaké $\lambda \in F^*$. Vidíme, že $u \sim v$ právě když $\langle u\rangle = \langle v\rangle$. Každý blok $\sim$ má $q-1$ prvků, takže máme $n=\frac{q^\ell-1}{q-1}$ bloků. Z každého z nich vybereme jeden prvek a vybrané prvky označíme $v_1,\ldots,v_n$. Pro $1 \leq i < j \leq n$ platí $\langle v_i\rangle \neq \langle v_j\rangle$ takže v_i a v_j jsou lineárně nezávislé. Ovšem pro $\ell \geq 3$ zjevně existují lineárně závislé vektory v_i, v_j a v_k, kde $1 \leq i < j < k \leq n$. Sestavme prověrkovou matici H tak, aby sloupce byly tvořeny vektory $v_1,\ldots,v_k$. Z Tvrzení 2.6 plyne, že H určuje $[n,\ell,3]$ kód. Všechny takové kódy se nazývají **Hammingovy**. Vidíme také, že v případě binárních kódů je matice H určena až na pořadí sloupců jednoznačně, neboť každý blok $\sim$ obsahuje jediný prvek.

Hammingovy kódy se snadno dekódují. Zabývejme se znovu na chvíli otázkou dekódování $[n,k,d]_q$ kódu C pomocí jeho prověrkové matice H. Ať v je vyslané

kódové slovo, a ať w je slovo přijaté. Rozdílu $e = w - v$ se říka slovo **chybové**. Vidíme, že $Hw^T = He^T$. Jinak řečeno, slovo chybové i přijaté mají stejný syndrom. Je-li $Hw^T \neq 0$, tak pro připadnou opravu hledáme kódové slovo v, aby $|w - v|$ bylo co nejmenší. To znamená, že hledáme e takové, aby $|e|$ bylo co nejmenší a $w - e \in C$. Ovšem $w - e \in C$ právě když $Hw^T = He^T$.

Opravit přijaté slovo na nejbližší kódové slovo tedy znamená pro daný syndrom nalézt slovo nejmenší váhy s týmž syndromem. Je-li takové slovo jediné, je oprava jednoznačná.

Došlo-li k chybě v jediné pozici, řekněme j, tak je syndrom roven j-tému sloupci matice H. Pokud pro binární Hammingův kód vytvoříme H pomocí lexikografického uspořádání tak, že j-tý sloupec je roven $(i_0, \ldots, i_{l-1})$ právě když $j = \sum i_r 2^r$, tak ze syndromu $(e_0, \ldots, e_{l-1}) \neq 0$ dostaneme chybovou pozici okamžitě jako $\sum e_r 2^r$.

Obecně je počet syndormů roven q^{n-k}. Pokud je tato hodnota dostatečně malá, aby ji bylo možno použít jako index tabulky, tak lze ke každému syndromu s nalézt e_s takové, že $s = He_s^\mathsf{T}$ a $|e_s|$ je minimální možná. Algoritmus dokódování se pak redukuje na určení syndromu $s = Hw^\mathsf{T}$, opravu $v = w - e_s$ a nalezení u, že $v = Gu^\mathsf{T}$. Poslední krok je ovšem triviální, je-li generující matice G ve standardním tvaru. Pak je u shodné s prvými k symboly v.

***

Kód C se nazývá **samoortogonální**, jestliže $C \subseteq C^\perp$. Je-li $C = C^\perp$, hovoříme o kódu **samoduálním**. Zvláště důležité jsou binární samoduální kódy s vysokou minimální vahou. Z lemmatu 2.2 vidíme, že lineární kód s generující maticí G je samoortogonální, právě když $u \cdot v = 0$ platí pro libovolné dva řádky u, v matice G.

Pro samoduální $[n, k, d]_q$ kód C musí platit $n = \dim C + \dim C^\perp = 2 \dim C$. Samoduální kódy mohou existovat tedy pouze pro sudou délku. Lze říci, že to jsou samoortogonální kódy maximální možné dimenze (pro samoortogonální kód máme $n = \dim C + \dim C^\perp \geq 2 \dim C$).

Při práci s binárními vektory je zvykem pro $u, v \in \mathbb{F}_2^n$ označit $u \cap v$ ten vektor $w \in \mathbb{F}_2^n$, pro který $w_j = 1$ právě když současně $u_j = 1$ a $v_j = 1$. Je-li tedy $u = i_A$, $v = i_B$, tak $u \cap v = i_{A \cap B}$.

Zásadního významu pro práci s binárními kódy je toto jednoduché pozorování:

Lemma 2.9. *Ať u a v jsou binární vektory délky n. Pak $|u + v| = |u| + |v| - 2|u \cap v|$ a $|u \cap v| \equiv u \cdot v \pmod 2$.*

Důkaz. Pro $A, B \subseteq \{1, \ldots, n\}$ označme na chvíli $A \triangle B$ disjunktní sjednocení $(A \cup B) \setminus (A \cap B)$. Ať $u = i_A$ a $v = i_B$. Máme $u + v = i_{A \triangle B}$, takže dokazovaná rovnost je totéž jako $|A \triangle B| = |A| + |B| - 2|A \cap B|$, což zřejmě platí. Vidíme také, že $u \cdot v$ je součtem tolika jedniček, kolik jich je v $u \cap v$. Proto platí i druhý vztah (Jeho zápisu lze po formální stráce vytknout, že porovnává $u \cdot v$, což je prvek $\mathbb{F}_2$, s celým číslem $|u \cap v|$. Jde však o zápis běžně používaný.) □

Binární kód se C se nazývá **dvojnásobně sudý**, jestliže váha každého kódového slova $u \in C$ je dělitelná čtyřmi.

Lemma 2.10. *Ať C je samoortogonální binární lineární kód s generující maticí G. Ať $u_1, \ldots, u_k$ jsou řádky G. Jestliže je váha každého řádku G dělitelná čtyřmi, tak je kód C dvojnásobně sudý.*

Důkaz. Prvky C jsou právě všechny součty řádků z G. Proto stačí ukázat, že pro libovolná u, $v \in C$ taková, že $|u|$ a $|v|$ jsou dělitelná čtyřmi, platí, že čtyři dělí i $|u+v|$. Z lemmatu 2.9 plyne, že $2|u \cap v| \equiv 0 \bmod 4$ pro všechna u, $v \in C$, neboť C je samoortogonální, a tedy $|u \cap v| \equiv 0 \bmod 2$. Tudíž $|u+v| \equiv |u|+|v| \bmod 4$. □

Ať C je $[n,k,d]_q$ kód s generující maticí G. Kód C je samoortogonální právě když platí $\langle u_1, \ldots, u_k \rangle \subseteq \langle u_1, \ldots, u_k \rangle^\perp = \{u_1, \ldots, u_k\}^\perp$, a to nastává právě když $\{u_1, \ldots, u_k\} \subseteq \{u_1, \ldots, u_k\}^\perp$ (viz Lemma 2.2). Můžeme tedy vyslovit následující lemma, které se často používá společně s Lemmatem 2.10.

Lemma 2.11. *Ať C je $[n,k,d]_q$ kód s generující maticí G. Kód C je samoortogonální právě když $u \cdot v = 0$ platí pro libovolné dva řádky matice G.*

□

Na závěr této kapitoly uvedeme několik pozorování, která se týkají nelineárních kódů.

Lemma 2.12. *Ať C je binární $(2k, 2^k)$ kód. Jestliže $u \cdot v = 0$ pro všechna u, $v \in C$, tak C je lineárním samoduálním kódem.*

Důkaz. Z $C \subseteq C^\perp$ plyne $\langle C \rangle \subseteq C^\perp$ a $\langle C \rangle \subseteq \langle C \rangle^\perp$. Ať $h = \dim C$. Máme $h \leq 2k - h$, a tedy $h \leq k$. Lineární prostor $\langle C \rangle$ tedy obsahuje nejvýše 2^k prvků. Proto musí platit, že $C = \langle C \rangle$. □

Lemma 2.13. *Ať C je binární kód, který obsahuje 0. Předpokládejme, že pro každé $u \in C$ je $u + C$ dvojnásobně sudý lineární kód. Potom $u \cdot v = 0$ pro všechna u, $v \in C$.*

Důkaz. Pro $u, v \in C$ máme $u+v \in u+C$, takže vektory u, v i $u+v$ mají váhu dělitelnou čtyřmi. Lemma proto plyne z Lemmatu 2.9. □

Důsledek 2.14. *Ať C je binární $(2k, 2^k)$ kód takový, že $0 \in C$, a že $u + C$ je dvojnásobně sudý kód pro každé $u \in C$. Potom C je lineární samoduální kód.*

□