Diszkrét matematika 2

11. előadás
Kódelmélet

Mérai László
merai@inf.elte.hu
https://sites.google.com/view/laszlomerai

Komputeralgebra Tanszék

2023 ősz

Kódtávolság – emlékeztető

Definíció

Egy $\mathcal{C}$ kód kódtávolsága a kódszavak közti minimális távolság: $d(\mathcal{C}) = \min\{d(\mathbf{u}, \mathbf{v}) : \mathbf{u}, \mathbf{v} \in \mathcal{C}, \mathbf{u} \neq \mathbf{v}\}$.

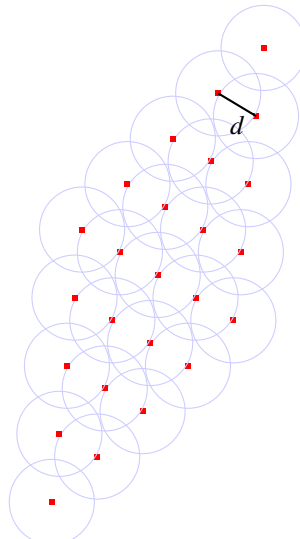
Példa

- Az ismétlő kód ($0 \mapsto 000$, $1 \mapsto 111$) távolsága $d = 3$.
- Paritásbit ($\mathbf{u} \mapsto (u_1, \ldots, u_k, u_1 + u_2 + \cdots + u_k \bmod 2)$) távolsága $d = 2$.

Tétel (Biz.: HF)

Egy $\mathcal{C}$ kód $d = d(\mathcal{C})$ kódtávolsággal:

- $d - 1$ hibát tud jelezni;
- $t = \lfloor (d-1)/2 \rfloor$ hibát tud javítani.



Singleton-korlát

Legyen $\mathcal{C} \subset \Sigma^n$ egy kód $d = d(\mathcal{C})$ minimális távolsággal.

- n minél kisebb, a kódolás annál gazdaságosabb.
- d minél nagyobb, a kód annál több hibát tud jelezni, javítani.
- $\#\mathcal{C}$ minél nagyobb, annál több szót tudunk kódolni.

Tétel (Singleton-korlát)

Egy $\mathcal{C} \subset \Sigma^n$ kód $d = d(\mathcal{C})$ minimális távolság esetén: $\#\mathcal{C} \leq (\#\Sigma)^{n-d+1}$.

Bizonyítás.

- Legyen $\mathcal{C}' \subset \Sigma^{n-d+1}$, amit $\mathcal{C}$ kódszavaiból kapunk az utolsó $d-1$ koordináta eltörlésével.
- Ha $\mathbf{u}, \mathbf{v} \in \mathcal{C}$ ($\mathbf{u} \neq \mathbf{v}$), akkor $d(\mathbf{u}, \mathbf{v}) \geq d$, azaz legalább d pozícióban különböznek. Spec., $\mathbf{u}, \mathbf{v}$ kódok $d-1$ koordináta törlése után is különböznek.
- Azaz $\#\mathcal{C} = \#\mathcal{C}' \leq (\#\Sigma)^{n-d+1}$.

□

Singleton-korlát

Singleton-korlát:
Ha $\mathcal{C} \subset \Sigma^n$ egy kód $d = d(\mathcal{C})$ minimális távolsággal, akkor: $\#\mathcal{C} \leq (\#\Sigma)^{n-d+1}$.

Egy $\mathcal{C}$ kód maximális távolságú (vagy MDS–*maximal distance separable*), ha $\#\mathcal{C} = (\#\Sigma)^{n-d+1}$.

Példa

- Az ismétlő kód
  ($0 \mapsto 000$, $1 \mapsto 111$, $n = 3, \#\mathcal{C} = 2, d = 3, \Sigma = \mathbb{Z}_2$)
  MDS kód: $2 = 2^{3-3+1}$.

- A paritásbit kód
  ($u \mapsto (u_1, \ldots, u_k, u_1 + \cdots + u_k)$, $n = k + 1, \#\mathcal{C} = 2^k, d = 2, \Sigma = \mathbb{Z}_2$)
  MDS kód: $2^k = 2^{k+1-2+1}$.

Később lesz példa további MDS kódokra.

Lineáris kódok

Most algebrai struktúrát vezetünk be a kódokon.

Definíció

Egy $\mathcal{C} \subset \mathbb{F}_q^n$ kód lineáris, ha $\mathcal{C}$ egy lineáris altér $\mathbb{F}_q^n$-ben. Ekkor $k = \dim \mathcal{C}$ a kód dimenziója. Spec. $\#\mathcal{C} = q^k$. Ekkor $\mathcal{C}$ egy (n, k) kód.

Példa

- Az ismétlő kód ($0 \mapsto 000$, $1 \mapsto 111$) egy $(3, 1)$ lineáris kód.
- A paritásbit kód ($u \mapsto (u_1, \ldots, u_k, u_1 + \cdots + u_k \bmod 2)$,) egy $(k + 1, k)$ lineáris kód.

A lineáris kódok kényelmesek.

- **Singleton-korlát**
  Egy $\mathcal{C} \subset \mathbb{F}_q^n$ egy (n, k) lineáris kód $d = d(\mathcal{C})$ minimális távolság.
  Ekkor: $\#\Sigma = \#\mathbb{F}_q = q$, $\#\mathcal{C} = q^k$, így $q^k \leq q^{n-d+1}$, azaz
  $$k \leq n - d + 1.$$

Lineáris kódok

Definíció

Legyen $\mathcal{C} \subset \mathbb{F}_q^n$ egy kód. Az $\mathbf{u} \in \mathcal{C}$ kódszó Hamming-súlya, $w(\mathbf{u}) = \#\{i : u_i \neq 0\}$.

Egy lineáris $\mathcal{C} \subset \mathbb{F}_q^n$ kód súlya: $w(\mathcal{C}) = \min\{w(\mathbf{c}) : \mathbf{c} \in \mathcal{C} \setminus \{0\}\}$

Tétel

Legyen $\mathcal{C}$ egy lineáris kód. Ekkor $d(\mathcal{C}) = w(\mathcal{C})$.

Bizonyítás.

- Minden $\mathcal{C}$ kódra $d(\mathcal{C}) = \min_{\mathbf{u} \neq \mathbf{v}} d(\mathbf{u}, \mathbf{v}) \leq \min_{\mathbf{u} \neq 0} d(\mathbf{u}, 0) = w(\mathcal{C})$ (ahol $\mathbf{u}, \mathbf{v} \in \mathcal{C}$).
- Ha $\mathcal{C}$ lineáris és $d(\mathcal{C}) = d(\mathbf{u}, \mathbf{v})$, akkor $\mathbf{u} - \mathbf{v} \in \mathcal{C}$ és $d(\mathbf{u}, \mathbf{v}) = w(\mathbf{u} - \mathbf{v})$, azaz $d(\mathcal{C}) = w(\mathcal{C})$.

□

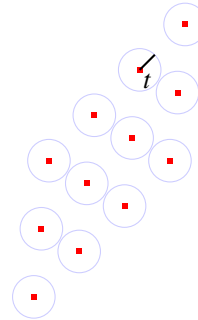
Azaz lineáris kódok esetén a kódtávolság kiszámításához nem kell a $\binom{q^k}{2} \approx q^{2k}/2$ távolságot ellenőrizni, elég a q^k Hamming-súlyt kiszámnolni.

Hamming-korlát

Tétel (Hamming-korlát)

Legyen $\mathcal{C} \subset \mathbb{F}_q^n$ egy lineáris (n,k) kód, mely t hibát tud javítani. Ekkor

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

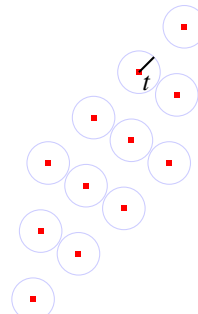


Bizonyítás.

- Tekintsünk egy t sugarú gömböt minden kódszó körül.
- Egy $\mathbf{u}$ kódszó körül azok az $\mathbf{s}$ szavak vannak, melyekre $d(\mathbf{u},\mathbf{s}) \leq t$.
- Ezen $\mathbf{s}$ szavak száma (egy gömb mérete): $G = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$
- $\#\mathcal{C} = q^k$ gömb van, ezek diszjunktak, így $q^k \cdot G \leq q^n$. □

Hamming-korlát

Hamming-korlát: $\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \leq q^{n-k}$.



Egy kód perfekt, ha a Hamming-korlátot egyenlőséggel teljesíti.

Példa

- Az ismétlő kód ($0 \mapsto 000$, $1 \mapsto 111$) perfekt: $1 + 3 = 2^{3-1}$.
- A négyszeres ismétlő kód ($0 \mapsto 0000$, $1 \mapsto 1111$) nem perfekt: $1 + 4 < 2^{4-1}$.
  (Itt $n = 4, k = 1, d = 4, t = 1$.)

A további célunk optimális kódok konstrukciója Singleton- és Hamming-korlát szempontjából tetszőleges n-re.

Generátormátrix

Legyen $\mathcal{C}$ egy lineáris (n, k) kód. Ekkor $\mathcal{C}$ egy altér, így létezik $\mathbf{c}_1, \dots \mathbf{c}_k \in \mathcal{C}$ melyek generálják a $\mathcal{C}$ alteret: $\langle \mathbf{c}_1, \dots, \mathbf{c}_k \rangle = \{a_1\mathbf{c}_1 + \dots + a_k\mathbf{c}_k : a_1, \dots, a_k \in \mathbb{F}_q\} = \mathcal{C}$.

Definíció

Legyen $\mathcal{C}$ egy lineáris (n, k) kód $\mathbf{c}_1, \dots, \mathbf{c}_k$ generátorokkal. Ekkor a $\mathcal{C}$ egy generátormátrixa $G = (\mathbf{c}_1, \dots, \mathbf{c}_k) \in \mathbb{F}_q^{n \times k}$.

Példa

- Az n-szeres ismétléses kód generátormátrixa: $G = (1, 1, \dots, 1)^T = \mathbf{1}^T \in \mathbb{F}_q^{n \times 1}$.
- A paritásbit generátormátrixa:

$$G = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 1 & 1 & \dots & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_k \\ \mathbf{1} \end{pmatrix} \in \mathbb{F}_2^{(k+1) \times k}$$

Generátormátrix

Példa

- Az n-szeres ismétléses kód generátormátrixa: $G = \mathbf{1}^T = (1, 1, \ldots, 1)^T \in \mathbb{F}_q^{n\times 1}$.
- A paritásbit generátormátrixa: $G = \begin{pmatrix} \mathbf{I}_k \\ \mathbf{1} \end{pmatrix} \in \mathbb{F}_2^{(k+1)\times k}$

Megjegyzések

- A generátormátrix nem egyértelmű. Ha $P \in \mathbb{F}_q^{k\times k}$ invertálható, akkor $G \cdot P$ is generátormátrix.
- Az $\mathbf{u} \mapsto G\mathbf{u}$ egy kódolás.

Definíció

Egy $\mathbf{u} \mapsto G\mathbf{u}$ kódolás szisztematikus, ha a kódszavak utolsó $n-k$ elemét elhagyva a kódolandó szót kapjuk, azaz

$$G = \begin{pmatrix} \mathbf{I}_k \\ B \end{pmatrix} \in \mathbb{F}_q^{n\times k}, \quad B \in \mathbb{F}_q^{(n-k)\times k}$$

alakú.

Ellenőrző mátrix

Példa

- Az n-szeres ismétléses kód generátormátrixa: $G = (1, 1, \ldots, 1)^T = \mathbf{1}^T \in \mathbb{F}_q^{n\times 1}$.
  Kapott $\mathbf{w} \in \mathbb{F}_q^n$ szó ellenőrzése: $w_1 \stackrel{?}{=} w_2 \stackrel{?}{=} \ldots \stackrel{?}{=} w_n$
- A paritásbit generátormátrixa: $G = \begin{pmatrix} \mathbf{I}_k \\ \mathbf{1} \end{pmatrix} \in \mathbb{F}_2^{(k+1)\times k}$
  Kapott $w \in \mathbb{F}_2^{k+1}$ szó ellenőrzése: $w_1 + w_2 + \cdots + w_{k+1} \stackrel{?}{=} 0$

Definíció

Legyen $\mathcal{C}$ egy (n, k) kód. Ekkor $\mathcal{C}$ ellenőrző mátrixa az a $H \in \mathbb{F}_q^{(n-k)\times n}$ mátrix melyre $H\mathbf{c} = 0$ pontosan akkor, ha $\mathbf{c} \in \mathcal{C}$.

Példa

- n-szeres ismétléses kód ellenőrző mátrixa: $H = (\mathbf{I}_{n-1}, -\mathbf{1}) \in \mathbb{F}_q^{(n-1)\times n}$
- A paritásbit ellenőrző mátrixa: $H = \mathbf{1} = (1, \ldots, 1) \in \mathbb{F}_2^{1\times(k+1)}$