Diszkrét matematika 2

12. előadás
Kódelmélet

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

Komputeralgebra Tanszék

2023 ősz

Lineáris kódok – emlékeztető

- Legyen $\mathcal{C}$ egy lineáris (n, k) kód, azaz $\mathcal{C}$ egy k dimenziós altér $\mathbb{F}_q^n$-ben.
- Ekkor létezik $\mathbf{c}_1, \dots \mathbf{c}_k \in \mathcal{C}$ melyek generálják a $\mathcal{C}$ alteret: $\{a_1\mathbf{c}_1 + \cdots + a_k\mathbf{c}_k : a_1, \dots, a_k \in \mathbb{F}_q\} = \mathcal{C}$.
- Ekkor a $\mathcal{C}$ egy generátormátrixa $G = (\mathbf{c}_1, \dots, \mathbf{c}_k) \in \mathbb{F}_q^{n\times k}$.
- 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.
- A $\mathcal{C}$ kód ellenőrző mátrixa H, ha $\mathbf{w} \in \mathcal{C} \iff H\mathbf{w} = 0$

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: $H = \mathbf{1} = (1, \dots, 1) \in \mathbb{F}_2^{1\times(k+1)}$

Ellenőrző mátrix

Tétel

Legyen $\mathcal{C}$ egy (n,k) kód, és legyen $G \in \mathbb{F}_q^{n\times k}$ a $\mathcal{C}$ generátormátrixa. Ekkor $H \in \mathbb{F}_q^{(n-k)\times n}$ pontosan akkor a kód ellenőrző mátrixa ha $HG = \mathbf{0} \in \mathbb{F}_q^{(n-k)\times k}$ és $\operatorname{rank} H = n-k$.

Bizonyítás.

- Tekintsük a $\mathbf{u} \mapsto G\mathbf{u}$ kódolást. Ekkor $HG\mathbf{u} = 0$ minden $\mathbf{u} \in \mathbb{F}_q^k$, így $HG = \mathbf{0}$, azaz $\mathcal{C} \subset \operatorname{Ker} H$.
- Megmutatjuk, hogy $\dim \mathcal{C} = \dim \operatorname{Ker} H$.
  Mivel $\dim \mathcal{C} = k$ és $\operatorname{rank} H = n-k = \dim \operatorname{Im} H$, és a dimenziótétel miatt $(\dim \operatorname{Im} H + \dim \operatorname{Ker} H = n)$ $\dim \operatorname{Ker} H = n-(n-k) = k$.
- Mivel $\mathcal{C} \subset \operatorname{Ker} H$ és $\dim \mathcal{C} = \dim \operatorname{Ker} H$, kapjuk $\dim \mathcal{C} = \operatorname{Ker} H$

$\square$

Ellenőrző mátrix

Tétel: Legyen $\mathcal{C}$ egy (n,k) kód, és legyen $G \in \mathbb{F}_q^{n\times k}$ a $\mathcal{C}$ generátormátrixa. Ekkor $H \in \mathbb{F}_q^{(n-k)\times n}$ a kód ellenőrző mátrixa ha $HG = \mathbf{0} \in \mathbb{F}_q^{(n-k)\times k}$ és $\operatorname{rank} H = n-k$.

Tétel

Legyen $\mathcal{C}$ egy (n,k) kód, generátormátrixa $G \in \mathbb{F}_q^{n\times k}$ és tegyük fel, hogy a $\mathbf{u} \mapsto G\mathbf{u}$ kódolás szisztematikus, azaz $G = \begin{pmatrix} \mathbf{I}_k \\ B \end{pmatrix} \in \mathbb{F}_q^{n\times k}$, $B \in \mathbb{F}_q^{(n-k)\times k}$.

Ekkor $H = (-B, \mathbf{I}_{n-k})$.

Bizonyítás. $\operatorname{rank} H = n-k$ és

$$HG = (-B, \mathbf{I}_{n-k}) \cdot \begin{pmatrix} \mathbf{I}_k \\ B \end{pmatrix} = -B + B = \mathbf{0}.$$

□

Szindrómák

Legyen a $\mathcal{C}$ egy (n, k) kód, ellenőrző mátrixa $H \in \mathbb{F}_q^{(n-k)\times n}$.

- Egy kapott $\mathbf{c}$ szó kódszó, ha $H\mathbf{c} = 0$.
- Ha $\mathbf{c} \in \mathcal{C}$ egy kódszó és $\mathbf{e} \in \mathbb{F}_q^n$ egy hibavektor, akkor $\mathbf{w} = \mathbf{c} + \mathbf{e}$ esetén $H\mathbf{w} = H(\mathbf{c} + \mathbf{e}) = H\mathbf{e} = \mathbf{s}$ az $\mathbf{e}$ hibához tartozó szindróma.

Tekintsük a következő táblázatot:

szindróma	mellékosztály vezető	kapott üzenetek		
$\mathbf{s}^{(0)} = \mathbf{0}$	$\mathbf{e}^{(0)} = \mathbf{0}$	$\mathbf{c}^{(1)}$	$\ldots$	$\mathbf{c}^{(q^k)}$
$\mathbf{s}^{(1)} = H\mathbf{e}^{(1)}$	$\mathbf{e}^{(1)}$	$\mathbf{c}^{(1)} + \mathbf{e}^{(1)}$	$\ldots$	$\mathbf{c}^{(q^k)} + \mathbf{e}^{(1)}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{s}^{(q^{n-k}-1)} = H\mathbf{e}^{(q^{n-k}-1)}$	$\mathbf{e}^{(q^{n-k}-1)}$	$\mathbf{c}^{(1)} + \mathbf{e}^{(q^{n-k}-1)}$	$\ldots$	$\mathbf{c}^{(q^k)} + \mathbf{e}^{(q^{n-k}-1)}$
	mellékosztály elemek			

ahol $0 = w(\mathbf{e}^0) \leq w(\mathbf{e}^1) \leq \cdots \leq w(\mathbf{e}^{q^{n-k}-1})$.

Szindrómák

szindróma	mellékosztály vezető	kapott üzenetek		
$\mathbf{s}^{(0)} = \mathbf{0}$	$\mathbf{e}^{(0)} = \mathbf{0}$	$\mathbf{c}^{(1)}$	$\dots$	$\mathbf{c}^{(q^k)}$
$\mathbf{s}^{(1)} = H\mathbf{e}^{(1)}$	$\mathbf{e}^{(1)}$	$\mathbf{c}^{(1)} + \mathbf{e}^{(1)}$	$\dots$	$\mathbf{c}^{(q^k)} + \mathbf{e}^{(1)}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{s}^{(q^{n-k}-1)} = H\mathbf{e}^{(q^{n-k}-1)}$	$\mathbf{e}^{(q^{n-k}-1)}$	$\mathbf{c}^{(1)} + \mathbf{e}^{(q^{n-k}-1)}$	$\dots$	$\mathbf{c}^{(q^k)} + \mathbf{e}^{(q^{n-k}-1)}$
	mellékosztály elemek			

ahol $0 = w(\mathbf{e}^0) \leq w(\mathbf{e}^1) \leq \cdots \leq w(\mathbf{e}^{q^{n-k}-1})$.

Itt minden elem különböző:

- soron belül különbözőek az elemek
- különböző sorok esetén ($j \neq \ell$): $\mathbf{c}^{(i)} + \mathbf{e}^{(j)} = \mathbf{c}^{(k)} + \mathbf{e}^{(\ell)} \Longrightarrow$
  $\mathbf{s}^{(j)} = H\left(\mathbf{e}^{(j)}\right) = H\left(\mathbf{c}^{(i)} + \mathbf{e}^{(j)}\right) = H\left(\mathbf{c}^{(k)} + \mathbf{e}^{(\ell)}\right) = H\left(\mathbf{e}^{(\ell)}\right) = \mathbf{s}^{(\ell)} \Longrightarrow j = \ell$ ↯ $\square$

Szindrómák

szindróma	mellékosztály vezető	kapott üzenetek		
$\mathbf{s}^{(0)} = \mathbf{0}$	$\mathbf{e}^{(0)} = \mathbf{0}$	$\mathbf{c}^{(1)}$	$\dots$	$\mathbf{c}^{(q^k)}$
$\mathbf{s}^{(1)} = H\mathbf{e}^{(1)}$	$\mathbf{e}^{(1)}$	$\mathbf{c}^{(1)} + \mathbf{e}^{(1)}$	$\dots$	$\mathbf{c}^{(q^k)} + \mathbf{e}^{(1)}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{s}^{(q^{n-k}-1)} = H\mathbf{e}^{(q^{n-k}-1)}$	$\mathbf{e}^{(q^{n-k}-1)}$	$\mathbf{c}^{(1)} + \mathbf{e}^{(q^{n-k}-1)}$	$\dots$	$\mathbf{c}^{(q^k)} + \mathbf{e}^{(q^{n-k}-1)}$
	mellékosztály elemek			

Szindrómadekódolás:
Legyen $\mathbf{w}$ a kapott szó. Legyen $\mathbf{s}^{(i)} = H\mathbf{w}$ a hozzá tartozó szindróma és $\mathbf{e}^{(i)}$ a mellékosztály vezető. Ekkor $\mathbf{c} = \mathbf{w} - \mathbf{e}^{(i)}$.

Ha a kód $t = \lfloor (d-1)/2 \rfloor$ hibajavító, akkor elég az első $\sum_{i=0}^{t} \binom{n}{i} q^i$ sort nézni.

Szindrómadekódolás

Példa

Legyen

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{7\times 4} \text{ illetve}$$

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{3\times 7}.$$

szindróma	javítható hibaminták
000	0000000
001	0000001
010	0000010
011	0010000
100	0000100
101	0100000
110	1000000
111	0001000

- A kód súlya $w(\mathcal{C}) = 3$, így $d = 3$ és 1 hibát javít.
- A $\mathbf{c} = (0001111)$ kódszó, $\mathbf{w} = (1001111)$ esetén $H\mathbf{w} = (110)$,így $\mathbf{c} = \mathbf{w} - (1000000)$

Szindrómadekódolás

Példa

Legyen

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \in \mathbb{F}_2^{8\times 4} \text{ illetve}$$

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{4\times 8}.$$

szindróma	javítható hibaminták
0000	00000000
0001	00000001
0010	00000010
0011	00000011
0100	00000100
0101	00000101
0110	00000110
0111	00100000
1000	00001000
1001	00001001
⋮	⋮

- A kód súlya $w(\mathcal{C}) = 4$, így $d = 4$ és 1 hibát javít jól, 2 hibát tippel.

Reed-Solomon kódok

A Reed-Solomon kódok a lineáris kódok leggyakrabban használt családja.

Konstrukció:
Tekintsük az $\mathbb{F}_q$ véges testet, és legyenek az üzenetszavak az $\mathbf{u} \in \mathbb{F}_q^n$ a legfeljebb $k-1$-ed fokú polinomok:

$$\mathbf{u} = (u_0, \ldots, u_{k-1}) \leftrightarrow u(x) = u_0 + u_1 x + \cdots + u_{k-1} x^{k-1} \in \mathbb{F}_q[x], \quad \deg u \leq n-1.$$

Legyenek $\alpha_0, \alpha_1 \ldots, \alpha_{n-1} \in \mathbb{F}_q$ különböző elemek, $n \leq q$.Ekkor a kódolás:

$$c_0 = u(\alpha_0),\ c_1 = u(\alpha_1),\ \ldots,\ c_{n-1} = u(\alpha_{n-1}),$$

és $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{F}_q^n$ a kódszó.

Megjegyzés:

- q tipikusan $q = 2^r$ (implementációs előny)
- r tipikusan nagy, szükséges feltétel: $n \leq 2^r$.

Reed-Solomon kódok

Reed-Solomon kód:

$u(x) = u_0 + u_1 x + \cdots + u_{k-1} x^{k-1} \mapsto \mathbf{c} = \big(u(\alpha_0), u(\alpha_1) \ldots, u(\alpha_{n-1})\big).$

A kód generátormátrixa

$$G = \begin{pmatrix} 1 & \alpha_0 & \alpha_0^2 & \ldots & \alpha_0^{k-1} \\ 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_{n-1} & \alpha_{n-1}^2 & \ldots & \alpha_{n-1}^{k-1} \end{pmatrix} \in \mathbb{F}_q^{n \times k},$$

ugyanis a $G\mathbf{u}$ vektor i-edik koordinátája ($0 \leq i < n$):

$$(1, \alpha_i, \alpha_i^2, \ldots, \alpha_i^{k-1})(u_0, u_1, u_2, \ldots, u_{k-1})^T = u_0 + u_1\alpha_i + u_2\alpha_i^2 + \cdots + u_{k-1}\alpha_i^{k-1} = u(\alpha_i).$$

Reed-Solomon kódok

Reed-Solomon kód:

$u(x) = u_0 + u_1 x + \cdots + u_{k-1}x^{k-1} \mapsto \mathbf{c} = \big(u(\alpha_0), u(\alpha_1) \ldots, u(\alpha_{n-1})\big)$.

Tétel

Az (n,k) paraméterű $RS_q(n,k)$ Reed-Solomon kód kódtávolsága $d = n - k + 1$, azaz a Reed-Solomon kód maximális távolságú (MDS kód).

Bizonyítás.

$$
\begin{aligned}
w(\mathbf{c}) &= \#\{\mathbf{c} \text{ nem } 0 \text{ koordinátái}\} \\
&= n - \#\{\mathbf{c}\ 0 \text{ koordinátái}\} \\
&\geq n - \#\{u(x) \text{ gyökei}\} \\
&\geq n - \deg u(x) \\
&\geq n - k + 1.
\end{aligned}
$$

Mivel a $RS_q(n,k)$ kód lineáris, $w(RS_q(n,k)) = d$.

A Singleton korlát miatt $w(RS_q(n,k)) = d \leq n - k + 1$. $\square$

Reed-Solomon kódok

Megjegyzések

- Gyakori választás $\alpha_i = \alpha^i$, ahol $\alpha \in \mathbb{F}_q$ olyan elem, melyre $\alpha^i \neq 1$ ($1 \leq i < n$, például α generátor $\mathbb{F}_q$-ban).
- CD lemezek esetében $(28, 24)$ paraméterű $\mathbb{F}_{2^8}$ fölötti Reed-Solomon kódot használnak.
- A kódban a Sony és Philips egyeztek meg, a verseny a CD lejátszó készülékek terén zajlott (azaz a dekódoló algoritmus hatékonyságán)