ACCQ204

Cours 2

Enseignant: Aslan Tchamkerten *Crédit: Rita Ibrahim & Wenceslas Godel*

Codes linéaires

1 Le besoin de structure

Code sur un alphabet $[q] = \{1, 2, \ldots, q\}$
$C : [q]^k \longrightarrow [q]^n$
La mise en mémoire requiert: $n \times q^k$, prohibitif!
Idée: Imposer de la structure sur C pour limiter la mémoire.

Definition 1 *Soit $\mathbb{F}_q$ un corps.*
C est un code linéaire si c'est un sous-espace vectoriel de $\mathbb{F}_q^n$. On le note $[n, k, d]_q$.

Remarque À partir de maintenant tous les codes que l'on va étudier seront des codes linéaires.

2 Représentation G, H

Definition 2 *Le rang d'une matrice $\in \mathbb{F}_q^{k \times n}$ est le nombre maximal de lignes (ou colonnes) indépendantes.*

Definition 3 *Une matrice est dite de rang plein si son rang est $min(k, n)$.*

Theorem 4 *Si $S \subset \mathbb{F}_q^n$ est un sous espace-linéaire*

1. *$|S| = q^k, k \geq 1$, k étant la dimension de S.*

2. $\exists v_1, v_2, \ldots, v_k \in S$ *appelés base de* S *tq.* $\forall x \in S$,
$x = \sum_1^k a_i v_i = (a_1, a_2, \ldots, a_k) \cdot G$
avec $G = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix}$ *appelée matrice "génératrice" de* S.

3. $\exists H$ *matrice* $\in \mathbb{F}_q^{(n-k)\times n}$ *de rang plein tq.* $H \cdot x^T = 0, \forall x \in S$
H *étant la matrice de parité.*

4. $G \perp H \Leftrightarrow G \cdot H^T = 0$.

Lemma 5 *Soit* $k \leq n$ G *une matrice* $k \times n$ *génératrice de* S_1, H *une matrice de parité de dimension* $(n-k) \times n$ *du sous-espace* S_2 *tq.* $G \cdot H^T = 0$. G *et* H *sont supposées de rang plein.*
Alors $S_1 = S_2$.

Preuve

1. $S_1 \subseteq S_2$
$c \in S_1 \Rightarrow \exists y \in \mathbb{F}_q^n$ tq. $c = y \cdot G \Rightarrow c \cdot H^T = y \cdot G \cdot H^T = 0$ car $G \cdot H^T = 0$
$\Rightarrow c \in S_2$.

2. $S_2 \subseteq S_1$
H est de rang plein $\Rightarrow dim\,(Ker\,(H)) = n - dim\,(Im\,(H))$. Or $Ker\,(H) = S_2$ et $dim\,(Im\,(H)) = n-k$. Donc $dim\,(Ker\,(H)) = k \Rightarrow dim\,(S_2) = k$.
De plus G de rang plein $\Rightarrow dim\,(S_1) = k$.
$\stackrel{1}{\rightarrow} S_1 = S_2$.

■

Conséquence 6 *Tout code linéaire* $[n, k, d]_q$ *peut-être représenté avec:*
$min\,(n \cdot k, (n-k) \cdot n) = \mathcal{O}\,(n^2) \neq exp\,(\mathcal{O}\,(n))!$

Complexité du codage: $C\,(m) = m \cdot G$, où m est un vecteur-ligne de taille k et G une matrice de taille $k \times n$.
La complexité est en $\mathcal{O}\,(k \times n)$.

3 Distance minimale d'un code linéaire

Proposition 7 *Pour un code* $C\,[n,k,d]_q$, $d = \min_{c\neq 0\,\in C}\ wt(c)$.

Preuve $d \triangleq \min_{x,y\in C,\,x\neq y}\ \Delta\,(x,y) = \min_{x,y\in C,\,x\neq y}\ \Delta\,(x-y,0)$
Or $\Delta\,(x-y,0) = wt\,(x-y)$.
Donc $d = \min_{c\in C,\,c\neq 0}\ wt\,(c)$.
En effet, la borne inférieure de la quantité $\Delta\,(x-y,0)$ est atteinte car si l'on prend deux éléments, on peut toujours arriver à obtenir c. Par exemple, on choisit $x = c, y = 0$. ■

Proposition 8 *Pour un code* $[n,k,d]_q$ *de matrice de parité H, d est le nombre minimal de colonnes linéairement dépendantes (preuve:* $Hc^T = 0$ *pour tout élément dans le code et donc le c de poids minimal correspondra au nombre minimal de colonnes linéairement dépendantes).*

4 Code de Hamming

Definition 9 *Pour tout entier* $r \geq 2$ *un code de Hamming (binaire) a pour matrice de parité* H_r *telle que :*

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

où la $i^{ème}$ *colonne est la représentation de i en binaire (* $1 \leq i \leq 2^r - 1$ *) sur r bits. Donc* $r = n - k$, *i.e.,* $k = n - r$.

Proposition 10 *Pour tout* $r \geq 2$ *le code de Hamming a distance minimale égale à* 3.

Preuve Les colonnes de la matrice sont deux à deux indépendantes, donc $d \geq 3$. De plus $H_r^1 + H_r^2 + H_r^3 = 0$ et donc $d = 3$. ■

Observation 11 (Code et borne de Hamming) *Par la borne de Hamming*

$$|C| \cdot \mathrm{Vol}(n, \lfloor \frac{d-1}{2} \rfloor) \leq 2^n$$

Pour $d = 3$ *on a*

$$|C| \leq 2^n \cdot \frac{1}{n+1}$$

car $\mathrm{Vol}(n, 1) = n + 1$. *Il suit que*

$$\log_2(|C|) \leq n - \log_2(n+1).$$

Pour le code de Hamming,

$$n = 2^r - 1 \Rightarrow r = \log_2(n+1)$$

et donc

$$\log_2 |C| = k = n - r = n - \log_2(n+1).$$

On déduis que les codes de Hamming atteignent la borne de Hamming.

Observation 12 *Un code atteignant la borne de Hamming est dit parfait. Il existe d'autres codes parfait, par exemple, le code* $[n, 1, n]_2$*, ainsi que d'autres codes du a Golay.*

4.1 Décodage code de Hamming

1. Algo 1 : MAP. la complexité est en $2^{O(n)}$! (besoin de lister tous les mots codes).

2. Algo 2 : Dans le cas où une erreur se produit au maximum par mot code envoyé on a $y^T = c^T + e^T$ avec

$$e^T = \begin{pmatrix} 0 \\ . \\ 0 \\ 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}$$

Alors

$$Hy = Hc + He = He$$

ce qui correspond à la i^{eme} colonne de H (i étant la position du 1 dans e et donc de l'erreur dans y).

Complexité: $O(n \log_2(n))$ (un seul calcul matriciel à faire).

Remarque: Hy est appelé le syndrôme de y.

5 Codes MDS : maximum distance separable

Rappel : la borne du singleton nous dit que pour tout code

$$d \leq n - k + 1 \Rightarrow r + \delta \leq 1 \Rightarrow r \leq 1 - \delta.$$

Definition 13 *Un code est dit MDS (maximum distance separable) si* $d = n - k + 1$.

Proposition 14 *Si un code est MDS alors tout ensemble de k coordonnées définit les mots code de manière unique.*

Preuve Voir preuve borne de Singleton. ■

Definition 15 *Soit C un code avec* q^k *mots codes sur* $\mathbb{F}_q$ *et de longueur n. Soit J un sous-ensemble de* $\{1, 2, .., n\}$ *de coordonnées. J est un* ensemble d'information *si pour tout mot code, les composantes de J le détermine entièrement.*

Corollary 16 *Pour un code MDS, tout J avec* $|J| = k$ *est un ensemble d'information.*

Conjecture 1 *Tout code linéaire* $[n, k]_q$ *MDS satisfait* $n \leq q+1$ *si* $1 \leq k \leq q$ *sauf si* $q = 2^h$ *et* $k = 3$ *ou* $k = q - 1$*, auquel cas on a* $n \leq q + 2$.