Modélisation et résolutions numérique et symbolique de problèmes via les logiciels Maple et MATLAB (MODEL)

Cours n°3 : Codes cycliques

Stef Graillat & Mohab Safey El Din

Université Pierre et Marie Curie (Paris 6)

Résumé du cours précédent

- Codes correcteurs d'erreurs (définition)
- Distance de Hamming et poids de Hamming
- Capacité de détection/correction
- Codes linéaires
- Matrice génératrice, mise sous forme systématique, matrice de contrôle
- Syndromes et erreurs
- Décodage (naif)

Syndrome

Définition 1

Soit $\mathcal{C}$ un code $[n;k]_q$ et $H \in \mathbb{F}_q^{(n-k)\times n}$ sa matrice de contrôle. Le syndrome d'un mot $x \in \mathbb{F}_q^n$ est le mot de longueur $n-k$

$$s = x \cdot H^T.$$

Nous avons

$$\mathcal{C} = \{x \in \mathbb{F}_q^n \mid x \cdot H^T = 0\}$$

Les mots du code sont donc les mots dont le syndrome est nul.
Le syndrome définit un isomorphisme du quotient $\mathbb{F}_q^n/\mathcal{C}$ sur $\mathbb{F}_q^{n-k}$.
Principe du décodage : *Si un syndrome est non nul, on corrige le mot reçu r en appliquant le principe de vraisemblance : on soustrait à r un mot de poids minimum dans sa classe modulo $\mathcal{C}$ (i.e. un mot de poids minimum ayant même syndrome que r).*

Décodage par syndrome

Décode par syndrome

Soit $\mathcal{C}$ un code $[n; k; d]_q$, et $H \in \mathbb{F}_q^{(n-k)\times k}$ une matrice de contrôle. On transmet $c = u + e$, avec $u \in \mathcal{C}$ et une erreur $e \in \mathbb{F}_q^n$

- On calcule
$$s = c \cdot H^T = u \cdot H^T + e \cdot H^T = e \cdot H^T.$$
- On cherche ensuite $e' \in \mathbb{F}_q^n$ de poids minimum tel que :
$$s = e' \cdot H^T.$$
- On décode $c - e' \in \mathcal{C}$. Nous avons $c - e' = u$, si $w_H(e) \leq t$, avec t la capacité de correction du code.

Décodage par syndrome

- Calculer les syndromes s_i des mots e_i de poids inférieurs à la capacité de correction t du code.
- On les stocke dans une table d'association $[(s_i, e_i)]_{i=1}^{i=t}$, les syndromes corrigeables.
- Si on reçoit un mot $c \in \mathbb{F}_q^n$ de syndrome $s = c \cdot H^T$ non nul, il faut regarder si s est corrigeable.
- soit $s \in \mathbb{F}_q^{n-k}$ un syndrome corrigeable et $e \in \mathbb{F}_q^n$ le mot de poids minimum produisant s. Décoder $c = u - e$.

Complexité du décodage : exponentielle en n (c'est trop !) car nécessite de calculer *a priori* la distance minimale.

Contourner la barrière de complexité

$\rightsquigarrow$ controler par construction la distance minimale (codes de Hamming, codes de Reed-Muller, codes poinconnés, codes de Golay).

- Constructions spécifiques, ad hoc
- voir la bibliographie

$\rightsquigarrow$ rajouter un peu de structure algébrique

- la complexité exponentielle provient d'un facteur combinatoire
- idée : quotienter nos constructions de codes par un phénomène de nature combinatoire (par exemple : invariance par permutations de lettres)
- contrainte : garder une structure de codes linéaires (cadre agréable et efficace pour calculer, outils logiciels, etc.)

Contourner la barrière de complexité

$\rightsquigarrow$ controler par construction la distance minimale (codes de Hamming, codes de Reed-Muller, codes poinconnés, codes de Golay).

- Constructions spécifiques, ad hoc
- voir la bibliographie

$\rightsquigarrow$ rajouter un peu de structure algébrique

- la complexité exponentielle provient d'un facteur combinatoire
- idée : quotienter nos constructions de codes par un phénomène de nature combinatoire (par exemple : invariance par permutations de lettres)
- contrainte : garder une structure de codes linéaires (cadre agréable et efficace pour calculer, outils logiciels, etc.)

Codes de Hamming

Propriété 1

Un code linéaire $\mathcal{C}$ a une distance minimale d ssi sa matrice de parité H a d colonnes dépendantes et aucun ensemble d'au plus $d-1$ colonnes dépendantes.

Définition 2 (Code de Hamming)

Pour $r \in \mathbb{N}$, $r \neq 0$, on construit une matrice $\mathcal{H}_r$ à $2^r - 1$ lignes et r colonnes, dont les lignes sont les éléments non nuls de $\mathbb{F}_2^r$. On appelle code de Hamming binaire d'ordre r le code admettant $\mathcal{H}_r$ comme matrice de parité.

Propriété 2

Un code de Hamming binaire d'ordre r est de longueur $2^r - 1$, de dimension $2^r - r - 1$ et de distance minimale 3.

Codes de Reed-Muller

Les codes de Reed-Muller sont des codes binaires de longueur 2^m indicés par un paramètre $\leq r \leq m$. Un tel code est noté $\mathcal{R}(r, m)$.

Définition 3

Les codes de Reed-Muller sont définis par récurrence comme

$$\mathcal{R}(r, m) = \{(u|u+v) \mid u \in \mathcal{R}(r, m-1), v \in \mathcal{R}(r-1, m-1)\}$$

où $\mathcal{R}(0, m) = (1, \ldots, 1)$ (longueur=2^m) et $\mathcal{R}(m, m) = \mathbb{F}_2^{2^m}$.

Si on note $G(r, m)$ une matrice génératrice de $\mathcal{R}(r, m)$ on a

$$G(r, m) = \begin{bmatrix} G(r, m-1) & G(r, m-1) \\ 0 & G(r-1, m-1) \end{bmatrix}$$

Codes de Reed-Muller (suite)

Propriété 3

Soit $0 \leq r \leq m$ deux entiers.

1. *Pour $0 \leq i \leq j \leq m$, on a $\mathcal{R}(i, m) \subset \mathcal{R}(j, m)$.*
2. *La dimension de $\mathcal{R}(r, m)$ est $\sum_{i=0}^{r} \binom{m}{i}$*
3. *La distance minimale de $\mathcal{R}(r, m)$ est 2^{m-r}.*

Codes cycliques

Définition 4

Un code $\mathcal{C}$ est dit cyclique si pour tout mot $c = (c_0, \ldots, c_{n-1})$ de $\mathcal{C}$, le mot $c' = (c_{n-1}, c_0, c_1, \ldots, c_{n-2})$ appartient à $\mathcal{C}$.

On considérera dans la suite des codes **cycliques linéaires**.

Pour mieux appréhender les codes cycliques, on adopte un point de vue polynomial.

Si $c = (c_0, \ldots, c_{n-1})$ est un mot de $\mathbb{F}_q^n$, on lui associe $c(X) = c_0 + c_1 X + \cdots + c_{n-1} X^{n-1}$.

Propriété 4

Si $\mathcal{C}$ est cyclique et $c \in \mathcal{C}$, alors $Xc(X) \mod (X^n - 1)$ est dans $\mathcal{C}$.

Codes cycliques

Définition 4

Un code $\mathcal{C}$ est dit cyclique si pour tout mot $c = (c_0, \ldots, c_{n-1})$ de $\mathcal{C}$, le mot $c' = (c_{n-1}, c_0, c_1, \ldots, c_{n-2})$ appartient à $\mathcal{C}$.

On considérera dans la suite des codes **cycliques linéaires**.
Pour mieux appréhender les codes cycliques, on adopte un point de vue polynomial.
Si $c = (c_0, \ldots, c_{n-1})$ est un mot de $\mathbb{F}_q^n$, on lui associe $c(X) = c_0 + c_1 X + \cdots + c_{n-1} X^{n-1}$.

Propriété 4

Si $\mathcal{C}$ est cyclique et $c \in \mathcal{C}$, alors $Xc(X) \mod (X^n - 1)$ est dans $\mathcal{C}$.

Rappels sur les idéaux

$\rightsquigarrow$ les codes cycliques linéaires sont des idéaux de $\mathbb{F}_q[X]/\langle X^n - 1\rangle$.
Soit A un anneau commutatif et $I \subset A$. On dit que I est un idéal de A ssi

1. pour tout x, y dans I, $x + y \in I$
2. pour tout $a \in A$ et $x \in I$, $a.x \in I$

Un idéal I est principal si il existe $g \in I$ tel que $I = \{ag, a \in A\}$.

Propriété 5

Tous les idéaux de $\mathbb{F}_q[X]$ sont principaux. Tous les idéaux de $\mathbb{F}_q[X]/\langle X^n - 1\rangle$ sont principaux.

Polynôme générateur d'un code cyclique

$\rightsquigarrow$ on peut définir un polynôme générateur g pour les codes cycliques linéaires.

1. $g(X)$ divise $c(X)$ dans $\mathbb{F}_q[X]$ pour tout $c \in \mathcal{C}$
2. $g(X)$ divise dans $X^n - 1$ dans $\mathbb{F}_q[X]$
3. $\deg(g) = n - k$.

$\rightsquigarrow$ étudier les codes cycliques, c'est étudier les diviseurs de $X^n - 1$ dans $\mathbb{F}_q[X]$ $\rightarrow$ propriétés des corps finis.

Polynôme générateur d'un code cyclique

$\rightsquigarrow$ on peut définir un polynôme générateur g pour les codes cycliques linéaires.

1. $g(X)$ divise $c(X)$ dans $\mathbb{F}_q[X]$ pour tout $c \in \mathcal{C}$
2. $g(X)$ divise dans $X^n - 1$ dans $\mathbb{F}_q[X]$
3. $\deg(g) = n - k$.

$\rightsquigarrow$ étudier les codes cycliques, c'est étudier les diviseurs de $X^n - 1$ dans $\mathbb{F}_q[X]$ $\rightarrow$ propriétés des corps finis.

S. Graillat & M. Safey (Univ. Paris 6) MODEL (cours n°2) 13 / 17

Matrice génératrice

Soit $g = g_0 + g_1 X + \cdots + g_{n-k} X^{n-k}$ le générateur d'un code cyclique linéaire $\mathcal{C}$.

Propriété 6

L'ensemble des mots de $\mathcal{C}$ est l'ensemble $\{g(X)a(X) \mid \deg(a(X)) \leq k-1\}$. Le code $\mathcal{C}$ a pour dimension k et pour matrice génératrice la matrice

$$\begin{bmatrix} g_0 & g_1 & \cdots & g_{n-k} & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k} & 0 & \cdots & 0 \\ & & \cdots & & \cdots & & \cdots & \\ 0 & 0 & 0 & \cdots & 0 & g_0 & \cdots & g_{n-k} \end{bmatrix}$$

Mise sous forme systématique

Propriété 7

Si $\mathcal{C}$ est un code cyclique de $\mathbb{F}_q^n$, alors le code dual $\mathcal{C}^\perp$ est aussi cyclique.

Théorème 1

Si $\mathcal{C}$ est un code de longueur n sur $\mathbb{F}_q$ de dimension k de polynôme générateur g, alors $\mathcal{C}^\perp$ est aussi un code cyclique de polynôme générateur $g^\perp = X^{n-k}\frac{h(1/X)}{h_0}$ où $h = \frac{X^n-1}{g}$.

$\rightsquigarrow$ on en déduit la matrice de contrôle !

Remarque : h est obtenu en étudiant le quotient de la division de $X^n - 1$ par g (rappel : g divise $X^n - 1$).

$\rightsquigarrow$ étudier les codes cycliques, c'est étudier les diviseurs de $X^n - 1$ dans $\mathbb{F}_q[X]$ $\rightarrow$ besoin d'en savoir plus sur les propriétés des corps finis et les racines de l'unité.

Mise sous forme systématique

Propriété 7

Si $\mathcal{C}$ est un code cyclique de $\mathbb{F}_q^n$, alors le code dual $\mathcal{C}^\perp$ est aussi cyclique.

Théorème 1

Si $\mathcal{C}$ est un code de longueur n sur $\mathbb{F}_q$ de dimension k de polynôme générateur g, alors $\mathcal{C}^\perp$ est aussi un code cyclique de polynôme générateur $g^\perp = X^{n-k}\frac{h(1/X)}{h_0}$ où $h = \frac{X^n-1}{g}$.

$\rightsquigarrow$ on en déduit la matrice de contrôle !

Remarque : h est obtenu en étudiant le quotient de la division de $X^n - 1$ par g (rappel : g divise $X^n - 1$).

$\rightsquigarrow$ étudier les codes cycliques, c'est étudier les diviseurs de $X^n - 1$ dans $\mathbb{F}_q[X]$ $\rightarrow$ besoin d'en savoir plus sur les propriétés des corps finis et les racines de l'unité.

Décodage

1. Les codes cycliques sont linéaires, on peut donc appliquer une procédure de décodage standard $\rightsquigarrow$ pas vraiment de meilleure complexité !
2. On peut chercher à utiliser le caractère cyclique de ces codes. Calculatoirement, cela implique de tirer profit de la structure de la matrice génératrice (et de la matrice de contrôle).

- ☐ Propriétés structurelles des corps finis et racines de l'unité $\rightsquigarrow$ propriétés sur la distance minimale
- ☐ contrôle de la distance minimale + structure des matrices $\rightsquigarrow$ décodage en temps polynomial (basé sur l'algorithme d'Euclide étendu)

Exemple : Codes BCH (pas dans ce cours)

Conclusion

- Techniques de codage/décodage
- distance et poids de Hamming, capacité de détection et correction d'erreurs
- Codes linéaires, matrice génératrice, matrice de contrôle
- complexité du décodage
- construction de codes cycliques

$\rightsquigarrow$ Algèbre linéaire (dans un corps fini) joue un rôle essentiel (codage décodage mais aussi en cryptologie/sécurité)
$\rightsquigarrow$ Corps fini $\Rightarrow$ Calcul Formel (expérimentations à faire en Maple)
$\rightsquigarrow$ Les strcutures polynomiales semblent permettre d'aller plus loin...