

Théorie et codage de l'information

Les codes linéaires

- Chapitre 6 -

PRINCIPE

Définition d'un code linéaire

Soient p un nombre premier et s est un entier positif. Il existe un unique corps de taille $q = p^s$, noté \mathbf{F}_q . L'ensemble $(\mathbf{F}_q)^n$ de tous les n -uples formés d'éléments de \mathbf{F}_q est un espace vectoriel sur \mathbf{F}_q .

Définition 1. \mathcal{L} est un **code linéaire** si \mathcal{L} est un sous-espace vectoriel de $(\mathbf{F}_q)^n$. On dit que \mathcal{L} est un $[n, k]$ -code si $\dim(\mathcal{L}) = k$. Si la distance minimale de \mathcal{L} est d , on parle de $[n, k, d]$ -code.

▷ Attention aux notations "(.)-code" et "[.]-code" !

PRINCIPE

Poids d'un code linéaire

Définition 2. Le poids $\omega(x)$ du mot x de $(\mathbf{F}_q)^n$ est le nombre de composantes non nulles de x .

Le poids minimal $\omega(\mathcal{L})$ du code \mathcal{L} est le minimum des poids de tous les vecteurs non nuls de \mathcal{L} .

Exemple

$$\omega(1101) = 3$$

$$\mathcal{L} = \{00000, 10111, 11010, 01101\} \longrightarrow \omega(\mathcal{L}) = 3$$

PRINCIPE

Poids d'un code linéaire

Définition 3. Si x et y sont deux mots binaires, on appelle intersection de x et y , notée $x \cap y$, l'élément défini par $(x \cap y)(i) = 1$ si $x(i) = y(i) = 1$, et 0 sinon.

En considérant les définitions ci-dessus, on peut démontrer les relations suivantes :

$$\forall x, y \in (\mathbf{F}_q)^n, d_{Ham}(x, y) = \omega(x - y),$$

$$\forall x, y \in (\mathbf{F}_2)^n, d_{Ham}(x, y) = \omega(x) + \omega(y) - 2\omega(x \cap y).$$

Théorème 1. Soit \mathcal{L} un code linéaire. On a : $d(\mathcal{L}) = \omega(\mathcal{L})$.

PRINCIPE

Codes détecteurs et correcteurs linéaires

En utilisant ce qui a été vu au cours précédent, on a immédiatement que :

Théorème 2. *En utilisant la règle de décodage par distance minimale, un code linéaire peut détecter jusqu'à t erreurs, avec $t = \omega(\mathcal{L}) - 1$.*

De plus, il en corrige jusqu'à t' , avec $\omega(\mathcal{L}) = 2t' + 1$ ou $2t' + 2$.

Exemple

$\mathcal{L} = \{00000, 10111, 11010, 01101\}$ est 2-détecteur et 1-correcteur.

MATRICE GÉNÉRATRICE

Définition

Définition 4. Soit \mathcal{L} un $[n, k]$ -code. Une matrice \mathbf{G} de dimension $(k \times n)$ dont les lignes forment une base de \mathcal{L} est une **matrice génératrice** de \mathcal{L} . On a alors :

$$\mathcal{L} = \{x\mathbf{G} \mid x \in (\mathbf{F}_q)^k\}.$$

Exemple

La matrice génératrice suivante permet d'encoder les mots de $(\mathbf{F}_2)^3$.

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

MATRICE GÉNÉRATRICE

Forme systématique

Définition 5. Un $[n, k]$ -code \mathcal{L} est dit sous **forme systématique** s'il existe k positions i_1, \dots, i_k telles que, par restriction des mots du code à ces k positions, on obtient les q^k mots q -aires possibles de longueur k .

Exemple

Le code $\mathcal{C} = \{0000, 0110, 1001, 1010\}$ est systématique sur les positions 1 et 3 :

00 \longrightarrow 0000

01 \longrightarrow 0110

10 \longrightarrow 1001

11 \longrightarrow 1010

MATRICE GÉNÉRATRICE

Forme standard

Théorème 3. *Tout $[n, k]$ -code linéaire a une matrice génératrice de la forme $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{A})$, dite **standard**, où \mathbf{I}_k désigne la matrice unité de dimension k .*

Exemple

La matrice génératrice suivante est sous forme standard.

$$\mathbf{G} = \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right).$$

DUAL D'UN CODE LINÉAIRE

Définition

Définition 6. Soit \mathcal{L} un $[n, k]$ -code.

$$\mathcal{L}^\perp = \{x \in (\mathbf{F}_q)^n \mid x.c = 0, \forall c \in \mathcal{L}\}$$

est appelé **code dual** de \mathcal{L} .

Théorème 4. Soit \mathcal{L} un $[n, k]$ -code linéaire et \mathcal{L}^\perp son dual. On a :

1. $\mathcal{L}^\perp = \{x \in (\mathbf{F}_q)^n \mid x\mathbf{G}^\top = 0\}$ où \mathbf{G} est génératrice de \mathcal{L} ;
2. \mathcal{L}^\perp est un $[n, n - k]$ -code linéaire ;
3. $\mathcal{L}^{\perp\perp} = \mathcal{L}$.

DUAL D'UN CODE LINÉAIRE

Matrice de test

Soit \mathcal{L} un code linéaire de matrice génératrice $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{A})$ de dimension $(k \times n)$.
On pose $\mathbf{H} = (-\mathbf{A}^\top \mid \mathbf{I}_{n-k})$.

Définition 7. *La matrice \mathbf{H} est dite matrice de contrôle ou matrice de test du code linéaire \mathcal{L} .*

On montre aisément que \mathbf{H} est une matrice génératrice de \mathcal{L}^\perp . Cette matrice sera utilisée par la suite pour le décodage.

Théorème 5. *Soit \mathcal{L} un code linéaire ayant \mathbf{H} comme matrice de contrôle. Il existe un mot de poids ω si, et seulement si, il existe ω colonnes de \mathbf{H} linéairement dépendantes.*

EXEMPLE

Matrice génératrice

On construit un $[6, 3]$ -code linéaire binaire en choisissant trois vecteurs linéairement indépendants de $(\mathbf{F}_2)^6$.

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

On obtient l'ensemble des mots du code \mathcal{L} en calculant tous les produits $x\mathbf{G}$ avec $x \in (\mathbf{F}_2)^3$.

EXEMPLE

Mots du code

Les mots du code \mathcal{L} ainsi que leur poids sont donnés par :

x	\mathcal{L}	ω
000	000000	0
001	110110	4
010	011101	4
011	101011	4
100	100101	3
101	010011	3
110	111000	3
111	001110	3

→ \mathcal{L} est 2-détecteur et 1-correcteur.

EXEMPLE

Matrice de contrôle

On écrit la matrice \mathbf{G} sous forme standard (pivot de Gauss) :

$$\mathbf{G} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \cdots \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{array} \right) = (\mathbf{I}_3 \mid \mathbf{A}).$$

La matrice de contrôle $\mathbf{H} = (-\mathbf{A}^\top \mid \mathbf{I}_3)$ de \mathcal{L} est donnée par :

$$\mathbf{H} = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right).$$

EXEMPLE

Mots du code dual

A partir de \mathbf{H} , on obtient les mots du code dual \mathcal{L}^\perp :

x	\mathcal{L}^\perp	ω
000	000000	0
001	110001	3
010	011010	3
011	101011	4
100	101100	3
101	011101	4
110	110110	4
111	000111	3

DÉCODAGE PAR TABLEAU STANDARD

Notion de syndrome

Nous allons à présent exposer une règle de décodage à distance minimale reposant sur un tableau, dit *tableau standard*.

Définition 8. Soit \mathcal{L} un $[n, k]$ -code linéaire de matrice de contrôle \mathbf{H} . Soit x un élément de $(\mathbf{F}_q)^n$. le mot $x\mathbf{H}^\top$ est appelé syndrome de x .

Cette définition permet d'introduire l'application linéaire s suivante :

$$\begin{aligned} s : (\mathbf{F}_q)^n &\longrightarrow (\mathbf{F}_q)^{n-k} \\ x &\longrightarrow s(x) = x\mathbf{H}^\top \end{aligned}$$

Soit x un élément de \mathcal{L} . On note que $s(x) = 0$, ce qui signifie que

$$\mathcal{L} = \ker(s).$$

DÉCODAGE PAR TABLEAU STANDARD

Classes d'équivalence

On définit la classe de x , notée C_x ou $x + \mathcal{L}$, par :

$$C_x = \{x + I \mid I \in \mathcal{L}\}.$$

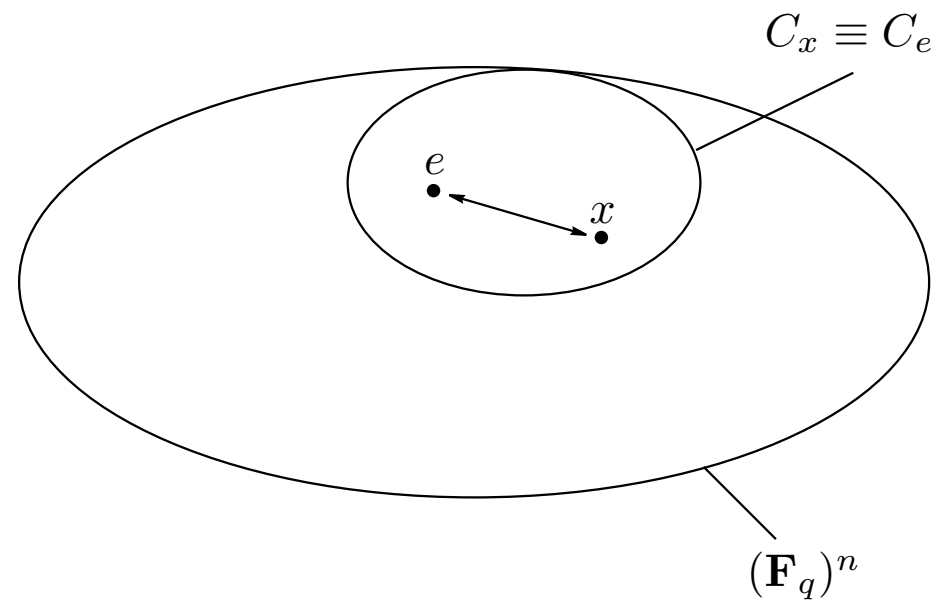
Théorème 6. *L'ensemble des classes C_x , $x \in (\mathbf{F}_q)^n$, forme une partition de $(\mathbf{F}_q)^n$.*

Théorème 7. *Soit \mathcal{L} un $[n, k]$ -code. Les éléments x et y de $(\mathbf{F}_q)^n$ ont le même syndrome si et seulement si ils définissent la même classe.*

DÉCODAGE PAR TABLEAU STANDARD

Construction du tableau

Soit x le mot reçu. Le décodage par distance minimale requiert que l'on décode x par le mot de code c pour lequel $e = x - c$ a le poids le plus faible. Étant donné que c varie dans \mathcal{L} , e est un élément de C_x . Le vecteur d'erreur e a donc le même syndrome que x .



DÉCODAGE PAR TABLEAU STANDARD

Construction du tableau

Mode de construction

1. La première ligne comporte les mots de \mathcal{L} .
2. La ligne j est constituée des $e_j + \mathcal{L}$, où e_j est un mot sélectionné de plus petit poids ne se trouvant pas dans les $j - 1$ lignes précédentes.
3. Le processus est itéré jusqu'à ce que tous les mots de $(\mathbf{F}_q)^n$ soient représentés dans le tableau.

0	c_1	c_2	...	c_m		
e_1	$c_1 + e_1$	$c_2 + e_1$...	$c_m + e_1$	→	$e_1 + \mathcal{L}$
e_2	$c_1 + e_2$	$c_2 + e_2$...	$c_m + e_2$	→	$e_2 + \mathcal{L}$
...	
e_s	$c_1 + e_s$	$c_2 + e_s$...	$c_m + e_s$	→	$e_s + \mathcal{L}$

DÉCODAGE PAR TABLEAU STANDARD

Exemple

Soit \mathcal{L} un $[4, 2]$ -code linéaire binaire défini par la matrice génératrice

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

On a donc :

x	\mathcal{L}	ω
00	0000	0
01	0100	1
10	1101	3
11	1001	2

DÉCODAGE PAR TABLEAU STANDARD

Exemple

Ceci nous conduit au tableau standard suivant :

0000	0100	1101	1001
1000	1100	0101	0001
0010	0110	1111	1011
1010	1110	0111	0011

Inconvénient : espace mémoire occupé (16 Go pour un [32,6]-code!).

⇒ **décodage par syndrome**

DÉCODAGE PAR SYNDROME

Principe

Les mots d'une même ligne du tableau standard ont même syndrome.

⇒ **construction d'une table de représentants et de syndromes**

0	c_1	c_2	...	c_m	$s(e_i)$
e_1	$c_1 + e_1$	$c_2 + e_1$...	$c_m + e_1$	$e_1 \mathbf{H}^\top$
e_2	$c_1 + e_2$	$c_2 + e_2$...	$c_m + e_2$	$e_2 \mathbf{H}^\top$
...
e_s	$c_1 + e_s$	$c_2 + e_s$...	$c_m + e_s$	$e_s \mathbf{H}^\top$

Le calcul du syndrome d'un mot reçu x désigne son représentant e_i . Le décodage s'effectue en calculant $x - e_i$.

DÉCODAGE PAR SYNDROME

Exemple

On considère le code linéaire \mathcal{L} défini par la matrice génératrice \mathbf{G} :

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \rightsquigarrow \mathbf{G}' = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \rightsquigarrow \mathbf{H} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Représentant e_i	Syndrome $e_i \mathbf{H}^\top$
0000	00
1000	01
0010	10
1010	11

DÉCODAGE PAR SYNDROME

Exemple

On suppose que $x = 1110$ a été reçu. Le calcul de son syndrome donne :

$$x\mathbf{H}^\top = \begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}^\top = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

Le représentant de la classe est donc 1010 et l'on obtient

$$c = 1110 - 1010 = 0100.$$