

## Cours 3 Matrice génératrice

- À propos de la linéarité

On se donne deux entiers  $k \leq n$ . On appelle traditionnellement  $k$  la dimension du code et  $n$  la longueur du code. La fonction de codage  $\varphi$  est une application de  $(\mathbb{F}_2)^k$  à valeurs dans  $(\mathbb{F}_2)^n$ . Pour tout message  $a \in (\mathbb{F}_2)^k$ , on est capable de déterminer un mot envoyé  $u = \varphi(a) \in (\mathbb{F}_2)^n$ . L'ensemble  $\mathcal{C} = \{\varphi(a), a \in (\mathbb{F}_2)^k\} \subset (\mathbb{F}_2)^n$  de tous les mots envoyés est appelé le code ; c'est l'ensemble image de l'application  $\varphi$ .

L'addition dans  $\mathbb{F}_2$  s'étend sans difficulté aux mots de  $(\mathbb{F}_2)^k$  ou de  $(\mathbb{F}_2)^n$ , appelés aussi "vecteurs" ; il suffit de faire l'addition composante par composante. Si  $a = (a_1, \dots, a_k) \in (\mathbb{F}_2)^k$  et  $b = (b_1, \dots, b_k) \in (\mathbb{F}_2)^k$ , alors  $a + b = (a_1 + b_1, \dots, a_k + b_k)$ . On a bien entendu une définition analogue si on remplace  $(\mathbb{F}_2)^k$  par  $(\mathbb{F}_2)^n$ .

On peut aussi définir la multiplication d'un scalaire  $\lambda \in \mathbb{F}_2$  par un vecteur

$a = (a_1, \dots, a_k) \in (\mathbb{F}_2)^k$  en multipliant toutes les composantes de  $a$  par le nombre  $\lambda$  :

$$\lambda \cdot (a_1, \dots, a_k) = (\lambda a_1, \dots, \lambda a_k).$$

Muni de l'addition "+" et de la multiplication par un scalaire ".", l'ensemble  $((\mathbb{F}_2)^k, +, \cdot)$  constitue une structure mathématique avec deux opérations, deux lois de composition. Nous ne développerons pas ses propriétés dans le cadre de ce cours.

La fonction de codage  $\varphi : (\mathbb{F}_2)^k \rightarrow (\mathbb{F}_2)^n$  est linéaire si et seulement si

$\forall a, b \in (\mathbb{F}_2)^k, \varphi(a + b) = \varphi(a) + \varphi(b)$  et  $\forall \lambda \in \mathbb{F}_2, \forall a \in (\mathbb{F}_2)^k, \varphi(\lambda a) = \lambda \varphi(a)$ . En particulier,  $\varphi(0_{(\mathbb{F}_2)^k}) = 0_{(\mathbb{F}_2)^n}$  et le vecteur zéro (de  $(\mathbb{F}_2)^n$ ) est toujours un mot du code :  $0 \in \mathcal{C}$ .

- Vecteurs de base

On peut décomposer un vecteur (ou un mot !)  $a \in (\mathbb{F}_2)^k$  à l'aide des "vecteurs de base"  $e_j$  définis par  $e_j = (0, \dots, 0, 1, 0, \dots, 0)$  avec l'unique nombre "1" à la  $j^{\text{o}}$  position. On vérifie (à l'aide d'une preuve par récurrence) que le vecteur ligne  $a = (a_1, \dots, a_k)$  s'écrit aussi sous la forme  $a = \sum_{j=1}^k a_j e_j$ . Alors la linéarité de l'application  $\varphi$  entraîne

$\varphi(a) = \varphi(\sum_{j=1}^k a_j e_j) = \sum_{j=1}^k \varphi(a_j e_j) = \sum_{j=1}^k a_j \varphi(e_j)$ . On note  $c_j = \varphi(e_j)$  le mot du code issu du codage du  $j^{\text{o}}$  vecteur de base. C'est un élément de  $(\mathbb{F}_2)^n$  que nous décomposons sur la base  $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  de  $(\mathbb{F}_2)^n$ . Le vecteur ligne  $\varepsilon_\ell$  est défini par

$\varepsilon_\ell = (0, \dots, 0, 1, 0, \dots, 0) \in (\mathbb{F}_2)^n$ , avec le nombre "1" en  $\ell^{\text{o}}$  position.

Attention aux dimensions ! Les vecteurs  $e_j \in (\mathbb{F}_2)^k$  sont des messages alors que les vecteurs  $\varepsilon_\ell \in (\mathbb{F}_2)^n$  sont des mots envoyés, de  $n$  bits.

Si  $c_j = (g_{j1}, g_{j2}, \dots, g_{jn})$ , on peut aussi écrire  $c_j = \sum_{\ell=1}^n g_{j\ell} \varepsilon_\ell$ . Alors le calcul de  $\varphi(a)$  se poursuit de la façon suivante :

$$\varphi(a) = \sum_{j=1}^k a_j \varphi(e_j) = \sum_{j=1}^k a_j c_j = \sum_{j=1}^k a_j \left[ \sum_{\ell=1}^n g_{j\ell} \varepsilon_\ell \right] = \sum_{\ell=1}^n \left[ \sum_{j=1}^k a_j g_{j\ell} \right] \varepsilon_\ell.$$

À l'aide du tableau de nombres  $g_{j\ell}$  pour  $1 \leq j \leq k$  et  $1 \leq \ell \leq n$ , on est capable d'expliciter le codage  $u = \varphi(a)$  d'un message quelconque  $a \in (\mathbb{F}_2)^k$ . Le mot du code se décompose sur les vecteurs  $\varepsilon_\ell$  et  $u = \sum_{\ell=1}^n u_\ell \varepsilon_\ell$ . De plus, le  $\ell^o$  bit du mot envoyé est une fonction linéaire des bits du message initial :  $u_\ell = \sum_{j=1}^k a_j g_{j\ell}$ .

- Matrice génératrice

La matrice génératrice, en général notée  $G$ , est un tableau à  $k$  lignes et  $n$  colonnes qui contient à l'intersection de la  $j^o$  ligne et de la  $\ell^o$  colonne le nombre  $g_{j\ell} \in \mathbb{F}_2$  :  $G = (g_{j\ell})_{1 \leq j \leq k, 1 \leq \ell \leq n}$ . C'est simplement la liste des mots du code obtenue en codant les vecteurs de base  $e_j \in (\mathbb{F}_2)^k$  :  $c_j = \varphi(e_j) = \sum_{\ell=1}^n g_{j\ell} \varepsilon_\ell$ .

- Exemple : duplication de bit

Cet exemple a été introduit au chapitre précédent. On a dans ce cas  $k = 1$  et  $n = 2$ . Le code est linéaire, ce qui est cohérent avec le choix  $\varphi(0) = 00$ . On garde en mémoire simplement la valeur  $\varphi(1) = 11$ . On stocke ces données dans une matrice  $G$  à une ligne et deux colonnes :  $G = (1 \ 1)$ . On a ici  $G \in \mathcal{M}_{1,2}(\mathbb{F}_2)$ .

- Exemple : double duplication de bit

On a maintenant  $k = 1$  et  $n = 3$  : chaque bit est dupliqué deux fois avant envoi dans le canal de transmission. La fonction de codage s'écrit dans ce cas  $\varphi(0) = 000$ ,  $\varphi(1) = 111$ . On se contente de considérer les divers bits qui constituent  $\varphi(1)$  pour former la matrice génératrice :  $G = (1 \ 1 \ 1) \in \mathcal{M}_{1,3}(\mathbb{F}_2)$ .

- Exemple : duplication de deux bits

Cet exemple a également été introduit au chapitre précédent. On a dans ce cas  $k = 2$  et  $n = 4$ . La fonction de codage  $\varphi$  peut être décrite complètement par les relations  $\varphi(00) = 0000$ ,  $\varphi(10) = 1010$ ,  $\varphi(01) = 0101$  et  $\varphi(11) = 1111$ . Cette fonction de codage est bien linéaire. D'une part, pour tout  $a \in (\mathbb{F}_2)^2$ , on a  $\varphi(0.a) = 0$ .  $\varphi(a) = 0000 \in (\mathbb{F}_2)^4$ . D'autre part, on a aussi  $\varphi(1.a) = 1$ .  $\varphi(a)$  et surtout  $1111 = 1010 + 0101 = \varphi(10) + \varphi(01)$ . Avec des notations plus formelles, on a donc  $\varphi((1,0) + (0,1)) = \varphi(1,0) + \varphi(0,1)$ . On détermine la matrice génératrice  $G \in \mathcal{M}_{2,4}(\mathbb{F}_2)$  en reportant les mots  $c_1 = \varphi(1, 0)$  et  $c_2 = \varphi(0, 1)$ . D'où

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \in \mathcal{M}_{2,4}(\mathbb{F}_2).$$

- Exemple : contrôle de parité

Un message de trois bits se code avec quatre bits *via* l'ajout d'un "bit de parité". On a donc  $k = 3$  et  $n = 4$ . La fonction de codage associée  $\varphi$  peut s'écrire

$\varphi(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3, \alpha_1 + \alpha_2 + \alpha_3)$  avec  $\alpha_j \in \mathbb{F}_2$ . Cette fonction de codage est bien linéaire et nous incitons le lecteur à le vérifier. La matrice génératrice n'utilise que trois mots :  $c_1 = \varphi(100) = 1001$ ,  $c_2 = \varphi(010) = 0101$  et  $c_3 = \varphi(001) = 0011$ . On a donc

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in \mathcal{M}_{3,4}(\mathbb{F}_2).$$

- Codage linéaire et multiplication d'une matrice ligne par la matrice génératrice

Nous avons vu plus haut que le message  $a = \sum_{j=1}^k \alpha_j e_j$  avec  $\alpha_j \in \mathbb{F}_2$  se code en un mot  $u = \sum_{\ell=1}^n u_\ell \varepsilon_\ell \in (\mathbb{F}_2)^n$  avec de plus la relation  $u_\ell = \sum_{j=1}^k \alpha_j g_{j\ell}$ . On identifie un vecteur ligne avec une matrice à une ligne. La dernière relation est donc le produit de la ligne  $a$  par la matrice génératrice  $G$ :  $u = aG$ , avec  $u \in \mathcal{M}_{1,n}(\mathbb{F}_2)$ ,  $a \in \mathcal{M}_{1,k}(\mathbb{F}_2)$  et  $G \in \mathcal{M}_{k,n}(\mathbb{F}_2)$ .

La relation  $u = aG$  permet de calculer le codage du mot envoyé  $u$  à partir du message  $a$  et de la matrice génératrice  $G$ .

On a par exemple pour la duplication de deux bits,  $\varphi(11) = (11) \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} = (1111)$ .

Cette écriture ne fait que traduire la linéarité. Dans le cas plus général d'un vecteur ligne à deux composantes et d'une matrice de deux lignes et quatre colonnes, on a

$(\alpha, \beta) \begin{pmatrix} a & b & c & d \\ e & f & g & h \end{pmatrix} = (\alpha a + \beta e, \alpha b + \beta f, \alpha c + \beta g, \alpha d + \beta h)$ , en ajoutant ici des séparateurs sous forme de virgule afin de faciliter la lecture.

Pour le contrôle de parité, on a typiquement  $\varphi(110) = (110) \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = (1100)$ . Rap-

pelons que pour un message de trois bits arbitraires  $(\alpha, \beta, \gamma)$  et une matrice de trois lignes et quatre colonnes, on a

$(\alpha, \beta, \gamma) \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & \ell \end{pmatrix} = (\alpha a + \beta e + \gamma i, \alpha b + \beta f + \gamma j, \alpha c + \beta g + \gamma k, \alpha d + \beta h + \gamma \ell)$ .

## Exercices

- Transposée de matrice

On se donne deux entiers  $n$  et  $k$  de sorte que  $n \geq k \geq 1$ . On note  $I_k$  la matrice identité comportant  $k$  lignes et  $k$  colonnes. On se donne aussi une matrice comportant  $k$  lignes et  $n$  colonnes que l'on peut écrire sous forme de blocs  $G = (I_k P)$ .

- Quel est l'ordre de la matrice  $P$ ?
- Que vaut  $G^t$ , transposée de la matrice  $G$ ?
- Dans le cas où  $G$  est la matrice génératrice du contrôle de parité pour lequel on a  $k = 3$  et  $n = 4$ , expliciter la matrice  $P$  et comparer le résultat de la question b) à un calcul direct.

- Matrices deux par deux dans  $\mathbb{F}_2$

On se donne  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{F}_2)$ .

- Combien y a-t-il de telles matrices ?

On pose  $\tilde{A} = \begin{pmatrix} d & b \\ c & a \end{pmatrix}$ .

- Montrer que  $A\tilde{A} = \tilde{A}A = (ad + bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ .

- Montrer que si  $ad + bc$  est non nul, la matrice  $A$  est inversible.

- d) Donner la liste de toutes les matrices carrées d'ordre deux à coefficients dans  $\mathbb{F}_2 \equiv \{0, 1\}$  telles que  $ad + bc \neq 0$ .
- e) Donner aussi la liste des inverses de ces matrices inversibles.
- f) Si  $ad + bc = 0$ , quelle est la valeur des produits  $A\tilde{A}$  et  $\tilde{A}A$  ?
- g) En déduire que si  $ad + bc = 0$ , la matrice  $A$  n'est pas inversible.
- h) Combien y a-t-il de matrices carrées d'ordre deux non inversibles à coefficients dans  $\mathbb{F}_2$  ?
- i) Donner la liste des matrices non inversibles dans  $\mathcal{M}_2(\mathbb{F}_2)$ .

• Probabilité de non détection

On complète un message  $a \in (\mathbb{F}_2)^3$  par un bit de parité pour obtenir un mot envoyé  $u \in (\mathbb{F}_2)^4$ . On a  $u_j = a_j$  pour  $1 \leq j \leq 3$  et  $\sum_{j=1}^4 u_j = 0$ . On suppose que la probabilité de mauvaise transmission du canal est  $p$  avec  $0 < p \ll 1$ . Pour fixer les idées, on peut par exemple supposer  $p = 10^{-3}$ .

- a) Quelles sont les différentes probabilités pour qu'un message envoyé  $u \in (\mathbb{F}_2)^4$  soit mal transmis par le canal de communication ?
- b) Quelles est la probabilité qu'une mauvaise transmission donne lieu à une détection de la défaillance ?
- c) Quelles est la probabilité conjointe à la fois de mauvaise transmission et de non détection d'une défaillance pour ce canal de communication ?
- d) Quelle est la probabilité pour qu'un message mal transmis ne soit pas détecté par ce processus d'ajout d'un bit de parité ?

• Associativité du produit des matrices carrées

On se donne un entier  $n \geq 1$  et trois matrices carrées  $A, B$  et  $C$  d'ordre  $n$ ; on a donc  $A, B, C \in \mathcal{M}_n(\mathbb{F}_2)$ .

- a) Que vaut  $(AB)_{ik}$  en fonction des coefficients des matrices  $A$  et  $B$  ?
- b) Que vaut  $((AB)C)_{iq}$  en fonction des coefficients des matrices  $A, B$  et  $C$  ?
- c) Que vaut  $(BC)_{jq}$  en fonction des coefficients des matrices  $B$  et  $C$  ?
- d) Calculer  $(A(BC))_{iq}$  en fonction des coefficients des matrices  $A, B$  et  $C$ .
- e) Déduire des questions précédentes la propriété d'associativité :  $A(BC) = (AB)C$ . Dans un produit de matrices, on peut placer les parenthèses où l'on veut !