IUT Villetaneuse
Université Sorbonne Paris Nord

BUT Informatique - Formation en alternance
Modélisations mathématiques (R5.12)

Année 2024-2025

Table des matières

Chapitre 1

Introduction aux chaînes de Markov

Sommaire

1.1 Propriété de Markov

On considère un espace probabilisé discret $(\Omega, \mathbb{P})$ et une suite infinie de variables aléatoires $X_0, X_1, X_2, \ldots$ définies sur Ω et prenant toutes leurs valeurs dans le même ensemble $\mathcal{E}$.

Définition 1.1

*On dit que la suite $X_0, X_1, X_2, \ldots$ est une **chaîne de Markov** si, pour tout $n \in \mathbb{N}$ et pour tout $(x_0, \ldots, x_n, x_{n+1}) \in \mathcal{E}^{n+2}$, on a*

$$\mathbb{P}\big(\{X_{n+1} = x_{n+1}\}/\{X_n = x_n\} \cap \{X_{n-1} = x_{n-1}\} \cap \cdots \cap \{X_0 = x_0\}\big) = \mathbb{P}(\{X_{n+1} = x_{n+1}\}/\{X_n = x_n\}).$$

*Cette propriété est appelée **propriété de Markov**. On l'écrira plus rapidement*

$$\mathbb{P}(X_{n+1} = x_{n+1}/X_n = x_n,\, X_{n-1} = x_{n-1},\, \cdots, X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1}/X_n = x_n).$$

Intuitivement, la propriété de Markov exprime que, pour prédire l'état du système à l'instant $n+1$, la connaissance de ses états à tous les instants $n, n-1, \cdots, 1, 0$ n'apporte pas plus d'information que la seule connaissance à l'instant n.

L'ensemble $\mathcal{E}$ est appelé **l'espace des états** de la chaîne de Markov. Étant donnés deux états e et e', la probabilité conditionnelle

$$\mathbb{P}(X_{n+1} = e'/X_n = e)$$

est appelée la **probabilité de transition** de l'état e à l'état e' à l'étape n. On la note souvent $p_{e,e'}(n)$. On dira de plus que la chaîne de Markov est **homogène** si la probabilité de transition $p_{e,e'}(n)$ ne dépend pas de n. On note alors simplement $p_{e,e'}$. En pratique, les états sont souvent numérotés, ce qui signifie que l'on a écrit $\mathcal{E} = \{e_1, e_2, \ldots\}$. On peut alors remplacer, sans risque de confusion, les états par leurs numéros dans toutes les notations ; par exemple, on peut écrire $X_n = i$ au lieu de $X_n = e_i$ ou encore $p_{i,j}$ à la place de p_{e_i,e_j}.

Dans toute la suite, on se limitera à des chaînes de Markov homogènes et dont l'espace des états $\mathcal{E}$ est fini.

1.2 Matrice de transition

On considère une chaîne de Markov $X_0, X_1, X_2, \ldots$ dont l'espace des états est $\mathcal{E} = \{e_1, e_2, \ldots, e_k\}$.

Définition 1.2

*La **matrice de transition** de cette chaîne de Markov est la matrice de dimensions $k \times k$ dont le coefficient à la i^{eme} ligne, j^{eme} colonne, est p_{e_i,e_j} (où $1 \leq i, j \leq k$).*

Propriété 1.1

La matrice de transition est telle que

1. *tous ses coefficients sont positifs ou nuls,*
2. *la somme des coefficients d'une même ligne est égale à 1.*

Toute matrice carrée vérifiant les propriétés 1 et 2 ci-dessus est dite **stochastique**.

1.3 Graphe des transitions

Soit $X_0, X_1, X_2, \ldots$ une chaîne de Markov dont l'espace des états est $\mathcal{E}$.

Définition 1.3

*Le **graphe des transitions** de cette chaîne de Markov est le graphe tel que*

1. *l'ensemble de ses sommets est $\mathcal{E}$,*
2. *il existe une arête du sommet e au sommet e' si et seulement si la probabilité de transition $p_{e,e'}$ est non nulle ; dans ce cas, l'arête de e à e' porte l'étiquette $p_{e,e'}$.*

1.4 Transitions en plusieurs étapes

Soit $X_0, X_1, X_2, \ldots$ une chaîne de Markov dont on note $\mathcal{E}$ l'espace des états.

Définition 1.4

*Étant donné un entier $q \geq 1$, la **probabilité de transition en q étapes** (ou q transitions) de l'état e à l'état e' est la probabilité conditionnelle $\mathbb{P}(X_q = e'/X_0 = e)$.*

On la note $p^{(q)}_{e,e'}$ ou simplement $p^{(q)}_{i,j}$ si e est l'état numéro i et e' l'état numéro j.

On obtient comme conséquence de l'homogénéité de la chaîne de Markov :

Propriété 1.2

Pour tous e, e' dans $\mathcal{E}$ et pour pour tous entiers $n, q \geq 1$ on a

$$p^{(q)}_{e,e'} = \mathbb{P}(X_{n+q} = e'/X_n = e).$$

Si les états de la chaîne de Markov sont numérotés, c'est à dire si l'on a écrit $\mathcal{E} = \{e_1, e_2, \ldots, e_k\}$, on peut définir la matrice de transition en q étapes de façon analogue à la matrice de transition, en utilisant les $p^{(q)}_{e_i,e_j}$ au lieu des p_{e_i,e_j}. Précisément :

Définition 1.5

*Pour tout entier $q \geq 1$, la **matrice de transition en q étapes** est la matrice $k \times k$, dont le coefficient à la i^{eme} ligne, j^{eme} colonne, est $p^{(q)}_{e_i,e_j}$ (où $1 \leq i, j \leq k$.)*

Théorème 1.1

Notons M la matrice de transition et $M^{(q)}$ la matrice de transition en q étapes ($q \in \mathbb{N}^*$). Alors on a $M^{(q)} = M^q$. Autrement dit, la matrice de transition en q étapes n'est rien d'autre que la puissance q^{ieme} de la matrice de transition.

1.5 Lois de probabilité des X_n

Soient $X_0, X_1, X_2, \ldots$ une chaîne de Markov dont l'espace des états est $\mathcal{E} = \{e_1, e_2, \cdots, e_k\}$. On désigne par M la matrice de transition et par $M^{(q)}$ la matrice de transition en q étapes.

On s'intéresse à la loi de probabilité de X_n, autrement dit on cherche à connaître la probabilité $\mathbb{P}(X_n = e_i)$ pour tout $i \in \{1, \cdots, k\}$, et ceci pour chaque $n \in \mathbb{N}$.

Définition 1.6

Pour tout $n \in \mathbb{N}$, on pose

$$p^{(n)} = \big(\mathbb{P}(X_n = e_1), \ldots, \mathbb{P}(X_n = e_k)\big),$$

*qui est donc un vecteur de $\mathbb{R}^k$. On pourra aussi regarder $p^{(n)}$ comme une matrice à une ligne et k colonnes. Ce vecteur $p^{(n)}$ est appelé la **distribution de probabilité à la n^{ieme} étape.** Dans le cas $n = 0$, on parle de **distribution initiale**.*

Propriété 1.3

Pour tout $n \in \mathbb{N}$,

- *$p^{(n)}$ a toutes ses coordonnées positives ou nulles, et la somme de ses coordonnées vaut 1,*
- *on a la relation matricielle $p^{(n+1)} = p^{(n)} M$,*
- *on a la relation matricielle $p^{(n)} = p^{(0)} M^{(n)}$.*

1.6 Comportement asymptotique

On considère dans ce paragraphe une chaîne de Markov $X_0, X_1, X_2, \cdots$ dont l'espace des états est $\mathcal{E} = \{e_1, \cdots, e_k\}$. Sa matrice de transition est notée M.

1.6.1 Loi stationnaire

Définition 1.7

*On dit qu'un vecteur $v = (y_1, \cdots, y_k) \in \mathbb{R}^k$ est un **vecteur de probabilité** s'il vérifie*

(a) $\forall i \in \{1, \cdots, k\} \quad y_i \geq 0$;

(b) $\sum_{i=1}^{k} y_i = 1$.

Remarque 1.1 *Se donner un vecteur de probabilité $v = (y_1, \cdots, y_k)$ revient à se donner une loi de probabilité sur l'ensemble $\mathcal{E} = \{e_1, \cdots, e_k\}$. Les vecteurs $p^{(n)}$ du paragraphe 1.5 définissent bien sûr des vecteurs de probabilités en posant $y_i = P(X_n = e_i)$.*

Définition 1.8

*On dit qu'un vecteur de probabilité v est **stationnaire** (ou **fixe**) pour la chaîne de Markov $X_0, X_1, X_2 \cdots$ s'il vérifie l'équation matricielle $v\,M = v$. Dans ce cas, on dit aussi que la loi de probabilité sur $\mathcal{E}$ déterminée par v est une **loi stationnaire**.*

Théorème 1.2

La chaîne de Markov $X_0, X_1, X_2 \cdots$ possède au moins un vecteur de probabilité fixe.

1.6.2 Distribution limite

Définition 1.9

*On dit qu'un vecteur de probabilité v est la **distribution limite** de la chaîne de Markov $X_0, X_1, X_2, \ldots$ si, pour tout vecteur de probabilité w, on a*

$$\lim_{n \to +\infty} w\, M^{(n)} = v;$$

autrement dit, si les vecteurs de distribution $p^{(n)}$ tendent vers v indépendamment de la distribution initiale $p^{(0)}$.

Théorème 1.3

Si la chaîne de Markov $X_0, X_1, X_2 \cdots$ admet une distribution limite v alors v est l'unique vecteur de probabilité stationnaire.

1.6.3 Un critère d'existence d'une distribution limite

Définition 1.10

- *La chaîne de Markov $X_0, X_1, X_2, \cdots$ est dite **irréductible** si son graphe des transitions est fortement connexe ; ceci est équivalent à dire que pour tout $(i,j) \in \{1, \cdots, k\}^2$ il existe $q \in \mathbb{N}$ tel que $p_{i,j}^{(q)} > 0$.*
- *La chaîne de Markov $X_0, X_1, X_2, \cdots$ est dite **régulière** (ou **primitive**) s'il existe un entier $q \geq 1$ tel que tous les coefficients de la matrice de transition en q étapes $M^{(q)}$ sont strictement positifs.*

Remarque 1.2 *Une chaîne de Markov régulière est aussi irréductible mais la réciproque n'est pas vraie.*

Théorème 1.4

Si une chaîne de Markov $X_0, X_1, X_2 \cdots$ est irréductible alors elle admet une distribution limite.

1.7 Exercices

Exercice 1. Un enfant collectionne des figurines qu'il trouve dans ses paquets de biscuits favoris. La série complète comporte quatre figurines. On note X_n le nombre de figurines différentes que possède cet enfant suite à l'achat du n^{ieme} paquet de biscuits.
1) Justifier que cette suite $(X_n)_{n\geq 0}$ est une chaîne Markov dont on précisera l'espace des états.
2) Calculer les probabilités des transitions puis donner la matrice de transition.
3) Tracer le graphe des transitions.
4) On note Y la v.a.r. égale au nombre de paquets qu'il faut acheter pour que l'enfant ait sa collection complète. Pour $i \in \{0,1,2,3\}$, on considère aussi la v.a.r. Y_i égale au nombre de paquets à acheter pour que la collection passe de l'état i à l'état $i+1$.
   a) Quelle sont les lois de probabilité des Y_i ?
   b) Quelle relation a-t-on entre Y et les Y_i ? En déduire $\mathbb{E}(Y)$ et $Var(Y)$.

Exercice 2. Dans un atelier, deux machines fonctionnent indépendamment l'une de l'autre. Chaque machine a pour probabilité p de tomber en panne au cours d'une journée.
I On suppose que si une machine tombe en panne au cours d'une journée, alors un technicien viendra la réparer pendant la nuit suivante, de sorte qu'elle se retrouvera en état de marche le lendemain matin. Cependant, le technicien ne peut pas réparer deux machines défectueuses en une nuit. On note X_n le nombre de machines en panne au début du n^{ieme} jour ($n \in \mathbb{N}$).
1) Justifier que la suite $(X_n)_{n\geq 0}$ est une chaîne de Markov dont on précisera l'espace des états.
2) Calculer les probabilités de transition et donner la matrice de transition M.
3) Former le graphe des transitions.
4) Initialement, toutes les machines sont en état de marche. Calculer la probabilité qu'il y ait

une machine en panne au début du deuxième jour.
II On suppose maintenant que le technicien se déplacera pendant une des deux nuits suivant la panne. La suite $(X_n)_{n\geq 0}$ est-elle encore une chaîne de Markov ?

Exercice 3. La fabrication industrielle d'un certain produit de déroule en deux étapes consécutives. A la fin de chaque étape, les objets fabriqués sont inspectés et trois situations sont possibles :

- l'objet est très défectueux, ce qui arrive avec probabilité 0,1. Dans ce cas l'objet est jeté.
- l'objet est légèrement défectueux, ce qui arrive avec probabilité 0,3. Il passe alors une nouvelle fois par la même étape de fabrication.
- l'objet est en bon état, ce qui arrive avec probabilité 0,6. Il continue alors la chaîne de fabrication s'il en était à la première étape ou bien est commercialisé s'il en était à la deuxième étape.

1) Décrire ce processus de fabrication par une chaîne de Markov à quatre états.
2) Déterminer la matrice de transition M.
3) Tracer le graphe des transitions.
4) Calculer la probabilité qu'un objet soit commercialisé en étant passé une seule fois par chaque étape de fabrication.

Exercice 4. On considère les jets successifs d'un dé non truqué et on note X_n le maximum des nombres obtenus lors des n premiers lancers.
1) Justifier que la suite des X_n définit bien une chaîne de Markov puis donner sa matrice de transition M.
2) Quelle est la distribution $p^{(1)}$?
3) Justifier, au vu de la dernière ligne de M, que cette chaîne de Markov n'est pas régulière.
4) Quelle est la probabilité d'avoir obtenu un résultat au moins égal à 5 avant le quatrième lancer ?

Exercice 5. Partie I. Un vendeur ambulant partage son temps entre trois villes A,B,C selon les modalités suivantes :

- Si un certain jour il se trouve dans la ville A, alors il s'installera le lendemain dans la ville B ou la ville C et son choix sera fait en tirant à pile ou face.
- Si un certain jour il se trouve dans la ville B, alors il jette deux pièces de monnaie et décide pour le lendemain
  - d'aller dans la ville A s'il obtient deux piles,
  - d'aller dans la ville C s'il obtient deux faces,
  - de rester dans la ville B dans les autres cas.
- Si un certain jour il se trouve dans la ville C, alors il s'installera le lendemain dans la ville A ou la ville B et son choix sera fait en tirant à pile ou face.

On note X_n la ville où travaille ce vendeur lors de la n^{ieme} journée d'observation.
1) Justifier que la suite $X_0, X_1, X_2, \cdots$ est une chaîne de Markov dont on précisera l'espace des états.
2) Donner la matrice de transition M puis former le graphe des transitions de cette chaîne de Markov.
3) a) Justifier que cette chaîne de Markov est régulière puis déterminer sa distribution limite.
b) Si la tournée du vendeur a commencé depuis longtemps et si vous voulez le rencontrer, dans quelle ville vous semble-t-il préférable de vous rendre ?
4) On suppose dans cette question que la tournée commence par la ville C. Quelle est la probabilité qu'il ne soit pas de retour dans cette ville trois jours plus tard ?

Partie II. On reprend les hypothèses de la partie I avec la seule différence suivante :

— Le commerce tellement bien dans la ville B que si, un certain jour, le vendeur atteint cette ville alors il s'y installera définitivement.

1) Donner la nouvelle matrice de transition puis justifier que cette nouvelle chaîne de Markov n'est pas régulière.

2) On suppose ici que le vendeur commence sa tournée en choisissant à pile ou face entre les villes A et C. Calculer la probabilité qu' il soit installé dans la ville B trois jours plus tard.

Exercice 6. (homogénéité des transitions d'ordre supérieur)

Soit $X_0, X_1, X_2 \dots$ une chaîne de Markov dont on note E l'espace des états. La probabilité de transition en n étapes ($n \geq 2$) est notée $p_{i,j}^{(n)}$ (où i, j sont des éléments de E) et est définie par $p_{i,j}^{(n)} = \mathbb{P}(X_n = j / X_0 = i)$. Une propriété du cours dit que

$$\forall n \geq 2 \; \forall k \geq 1 \; \forall i \in E \; \forall j \in E \quad p_{i,j}^{(n)} = \mathbb{P}(X_{n+k} = j / X_k = i).$$

Autrement dit, les probabilités des transitions en n étapes ne dépendent que des états considérés et non pas de leur place sur « l'axe du temps ». On propose dans cet exercice de le vérifier dans le cas où $k = 2$.

1) Justifier que $\mathbb{P}(X_{n+2} = j / X_n = i) = \sum_{k \in E} \mathbb{P}(X_{n+2} = j, X_{n+1} = k / X_n = i)$.

2) En déduire que

$$P(X_{n+2} = j / X_n = i) = \sum_{k \in E} \frac{P(X_{n+2} = j / X_{n+1} = k, X_n = i)}{P(X_n = i)} P(X_{n+1} = k, X_n = i).$$

3) Utiliser finalement la propriété de Markov et l'homogénéité pour obtenir

$$\mathbb{P}(X_{n+2} = j / X_n = i) = \sum_{k \in E} \frac{\mathbb{P}(X_2 = j / X_1 = k, X_0 = i)}{\mathbb{P}(X_0 = i)} \mathbb{P}(X_1 = k, X_0 = i)$$

puis conclure.

Exercice 7. On considère une chaîne de Markov $(X_n)_{n \geq 0}$ à deux états.

1) Justifier que la matrice de transition M est de la forme $\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$ où a, b sont deux réels tels que $0 \leq a \leq 1$ et $0 \leq b \leq 1$.

2) On suppose dans cette question que $a + b = 0$ ou $a + b = 2$.

a) Quelles sont alors les seules matrices M possibles ?

b) Vérifier que, dans l'un des deux cas trouvés au a), la chaîne de Markov admet un unique vecteur de probabilité fixe v que l'on explicitera.

c) Etant donné une distribution initiale $p^{(0)}$, rappeler la relation entre $p^{(n)}$, $p^{(0)}$, M^n. A-t-on nécessairement $\lim_{n \to +\infty} p^{(n)} = v$? Que pensez-vous de la réciproque du Théorème 1.3 du cours ?

3) On suppose dans cette question que $0 < a + b < 2$.

a) Montrer par récurrence sur n que

$$\forall n \geq 1 \quad M^n = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}.$$

b) En déduire que la chaîne de Markov admet pour distribution limite le vecteur $v = (\frac{b}{a+b}, \frac{a}{a+b})$.

c) On choisit dans cette question $0 < a < 1$ (par exemple $a = \frac{1}{2}$) et $b = 0$.
i) Combien vaut alors la probabilité de transition $p_{2,2}$ (on dit que l'état 2 est *absorbant*) ?
ii) Vérifier que, pour tout $n \geq 1$, on a $p_{2,1}^{(n)} = 0$. La chaîne de Markov est-elle irréductible ?

Exercice 8. Dans la suite, N est un entier donné ≥ 1.
I On note X_n la position à l'instant n d'une particule qui se promène sur les points $\{0, 1, \ldots, N\}$ de l'axe réel de la manière suivante :

- Si à un instant n elle se trouve sur un point $k \in \{1, \ldots, N-1\}$ alors, à l'instant $n+1$, elle se sera déplacée vers la droite avec probabilité p, vers la gauche avec probabilité q, et sera restée sur place avec probabilité $r = 1 - p - q$.
- Si à un instant n elle se trouve au point 0 alors, à l'instant $n+1$, elle se sera déplacée vers la droite avec probabilité p et sera restée sur place avec probabilité $1 - p$.
- Si à un instant n elle se trouve au point N alors, à l'instant $n+1$, elle se sera déplacée vers la gauche avec probabilité q et sera restée sur place avec probabilité $1 - q$.

1) Justifier que la suite de X_n définit une chaîne de Markov.
2) Quelle est la forme de la matrice de transition ?
3) On choisit ici $N = 2$ et $p = q = \frac{1}{3}$. Calculer la probabilité que cette particule soit revenue à sa place d'origine au bout de trois déplacements.
II Reprendre les questions du **I** en supposant cette fois que la particule reste en O si elle atteint ce point et de même pour le point N. Cette chaîne de Markov est-elle irréductible ?

Exercice 9. Une urne A contient 2 boules blanches tandis qu'une urne B contient 4 boules rouges. A chaque étape du processus, on extrait une boule de chaque urne et l'on échange ces deux boules. On note X_n le nombre de boules rouges dans l'urne A après n échanges.
1) Justifier que la suite $X_0, X_1, X_2, \cdots$ est une chaîne de Markov et donner sa matrice de transition.
2) Quelle est la probabilité qu'il y ait 2 boules rouges dans l'urne A après 3 échanges ?
3) a) Pourquoi peut-on affirmer que cette chaîne de Markov admet une distribution limite ?
b) Quelle est, sur le long terme, la probabilité d'avoir 0,1 ou 2 boule(s) rouge(s) dans l'urne A ?

Exercice 10. (Urnes d'Ehrenfest) On considère deux urnes A et B dans lesquelles sont réparties au total N boules numérotées de 1 à N. On répète l'expérience aléatoire qui consiste à tirer au hasard un nombre entier i entre 1 et N et et à changer d'urne la boule numéro i. On note X_n le nombre de boules présentes dans l'urne A après avoir répété n fois cette expérience.
1) Justifier que la suite $X_0, X_1, X_2, \cdots$ est une chaîne de Markov et préciser sa matrice des transitions.
2) Cette chaîne de Markov est-elle irréductible ?
3) Montrer que sa distribution limite est la loi binomiale $\mathcal{B}(N; 1/2)$.

Chapitre 2

Codes détecteurs et codes correcteurs d'erreurs

Sommaire

2.1 Addition modulo 2

Définition 2.1

- *L'addition modulo 2 sur l'ensemble* $\{0,1\}$ *est l'opération* $\oplus$ *définie par*

$$0 \oplus 0 = 0 \qquad 1 \oplus 0 = 1 \qquad 0 \oplus 1 = 1 \qquad 1 \oplus 1 = 0$$

- *L'addition modulo 2 sur* $\{0,1\}^n$ *(où* $n \in \mathbb{N}^*$*) est l'opération, également notée* $\oplus$*, définie à partir de l'addition* $\oplus$ *sur* $\{0,1\}$ *par*

$$(x_1, \cdots, x_n) \oplus (y_1, \cdots, y_n) = (x_1 \oplus y_1, \cdots, x_n \oplus y_n).$$

Les élements de $\{0,1\}^n$, c'est à dire les n-uplets dont chaque coordonnée vaut 0 ou 1, sont appelés **mots binaires** de **longueur** n. Un mot binaire $(x_1, \cdots, x_n) \in \{0,1\}^n$ est simplement

noté $x_1 \cdots x_n$. On note aussi $\mathbf{0}_n$ (ou $\mathbf{0}$ s'il n'y a pas d'ambiguité sur n) l'élément de $\{0,1\}^n$ dont toutes les coordonnées valent 0, c'est à dire $\mathbf{0}_n = \underbrace{0 \cdots 0}_{n \text{ fois}}$.

Remarque 2.1 *Il y aussi une multiplication modulo 2, qui coïncide en fait avec la multiplication usuelle, donnée par*

$$0 \otimes 0 = 0 \qquad 1 \otimes 0 = 0 \qquad 0 \otimes 1 = 0 \qquad 1 \otimes 1 = 1$$

Propriété 2.1

Pour tous m, m', m'' dans $\{0,1\}^n$ on a :

1. $m \oplus (m' \oplus m'') = (m \oplus m') \oplus m''$ *(associativité de $\oplus$)*
2. $m \oplus m' = m' \oplus m$ *(commutativité de $\oplus$)*
3. $m \oplus \mathbf{0}_n = m$ *($\mathbf{0}_n$ est élement neutre pour $\oplus$)*
4. $m \oplus m = \mathbf{0}_n$

Exemple 2.1 *Dans $\{0,1\}^4$ on a $(1,1,0,1) \oplus (1,0,0,1) = (0,1,0,0)$.*

2.2 Poids des mots et distance de Hamming

Définition 2.2

- *Le **poids** d'un mot binaire $m = x_1 \cdots x_n$ est le nombre de x_i égaux à 1. On le note $w(m)$.*
- *La **distance de Hamming** entre deux mots binaires $m = x_1 \cdots x_n$ et $m' = y_1 \cdots y_n$ de même longueur n, notée $d(m, m')$ est le nombre d'indices i tels que $x_i \neq y_i$. Autrement dit, $d(m, m') = w(m \oplus m')$ et en particulier $d(m, \mathbf{0}_n) = w(m)$.*

Propriété 2.2

Pour tous m, m', m'' dans $\{0,1\}^n$ on a :

1. $d(m, m') \geq 0$
2. $d(m, m') = d(m', m)$
3. $d(m, m') = 0$ *si et seulement si* $m = m'$
4. $d(m, m') \leq d(m, m'') + d(m'', m')$ *(inégalité triangulaire)*
5. $d(m \oplus m", m' \oplus m") = d(m, m')$ *(invariance par translation)*

2.3 Canal de transmission binaire symétrique sans mémoire

Nous considérons la question de la transmission de données binaires, c'est à dire des suites de 0 et de 1 (bits), depuis un émetteur vers un receveur. Ces données sont acheminées via un canal « bruité », ce qui signifie que des 0 peuvent être transformés en 1, et inversement.

$$\underbrace{1\underline{0}11000110101\underline{1}100010\cdots}_{\text{émis}} \xrightarrow{\text{canal bruité}} \underbrace{1\underline{1}11000110101\underline{1}000010\cdots}_{\text{reçu}}$$

Se pose alors la question de la détection des erreurs de transmission et de leur correction par le receveur. Par la suite, nous considérons l'altération d'un bit comme un phénomène aléatoire et nous faisons de plus les hypothèses simplificatrices suivantes :

- le canal est **symétrique** : la probabilité qu'un bit soit altéré ne dépend pas de sa valeur 0 ou 1.
- le canal est **sans mémoire** : les altérations éventuelles des différents bits sont indépendantes les unes des autres et ont toutes la même probabilité $p \in]0,1[$. Dans la pratique, p est beaucoup plus proche de 0 que de 1.

En conséquence, le nombre (aléatoire) d'erreurs survenant lors de la transmission d'un mot binaire de longueur n est modélisé par une loi binomiale $\mathcal{B}(n;p)$.

2.4 Codage par blocs

Le codage par blocs est une technique habituelle pour la transmission de données binaires.

- Pour l'expéditeur de l'information, elle consiste :

1. À découper la suite de 0 et de 1 en blocs de longueur k, où $k \geq 1$ est fixé à l'avance. Par exemple $k = 5$ sur le schéma ci-dessous :

$$\underbrace{10110}_{\text{bloc 1}}\underbrace{00110}_{\text{bloc 2}}\underbrace{10111}_{\text{bloc 3}}\underbrace{00010}_{\text{bloc 4}}\cdots$$

2. À coder chaque bloc, c'est à dire lui ajouter un nombre donné de **bits de contrôle**, qui dépendent du bloc considéré et dont le rôle est de permettre la détection ou la correction d'erreurs de transmission. On obtient ainsi un mot binaire de longueur $n = k + r$, appelé **mot de code**, où r est le nombre de bits de contrôle. C'est ce mot qui est transmis via le canal bruité. Par la suite, on supposera toujours que les bits de contrôle sont ajoutés sur la droite du bloc codé ; on parle alors de **code systématique**. Le schéma suivant illustre le cas $k = 5$ et $r = 2$.

$$\underbrace{x_1\,x_2\,x_3\,x_4\,x_5}_{bloc} \xrightarrow{\text{codage}} \overbrace{x_1\,x_2\,x_3\,x_4\,x_5\;\underbrace{x_6\,x_7}_{\text{contrôle}}}^{\text{mot de code émis}} \xrightarrow{\text{canal bruité}} \overbrace{y_1\,y_2\,y_3\,y_4\,y_5\,y_6\,y_7}^{\text{mot binaire reçu}}$$

- Pour le receveur de l'information :

1. À vérifier que le mot binaire reçu est bien un mot de code (détection d'erreurs) et éventuellement corriger ces erreurs.
2. À extraire les k premiers bits, qui sont ceux portant l'information utile (décodage).

L'ensemble de tous les mots de code est appelé **code** et est noté $\mathcal{C}$. Ainsi $\mathcal{C}$ est un sous-ensemble de $\{0,1\}^n$ contenant 2^k mots. On note $\varphi : \{0,1\}^k \to \{0,1\}^n$ la fonction de codage, c'est à dire la fonction qui à un bloc associe le mot de code correspondant. Dans l'exemple précédent, $\varphi(x_1x_2x_3x_4x_5) = x_1x_2x_3x_4x_5x_6x_7$. La **distance minimale** du code est la plus petite distance de Hamming entre deux mots de code différents ; on la note d. On dit alors que l'on a un code de **type $[n, k, d]$**, ou de **type $[n, k]$** s'il est inutile de préciser d. Les nombres n et k sont appelés respectivement **dimension** et **longueur** du code. Le **rendement** est le nombre k/n.

2.5 Détection d'erreurs

Pour détecter les erreurs de transmission, le receveur regarde si le mot reçu appartient à $\mathcal{C}$. Si ce n'est pas le cas, au moins une erreur sur un bit est survenue.

Théorème 2.1 (*situation de détection certaine*)

Considérons un code de type $[n, k, d]$. Avec le critère ci-dessus, le receveur détecte tous les mots binaires contenant entre 1 et $d-1$ erreurs.

Preuve : Dire que le mot reçu $m' \in \{0,1\}^n$ contient entre 1 et $d-1$ erreurs revient à dire que $m' = m \oplus e$ où m est le mot de code transmis et $e \in \{0,1\}^n$ est un « vecteur d'erreurs » tel que $1 \leq w(e) \leq d-1$. Comme $d(m, m') = d(m, m \oplus e) = d(\mathbf{0}_n, e) = w(e)$, le mot binaire m' n'est pas un mot de code. □

2.6 Correction d'erreurs

Pour corriger le mot reçu m', le receveur le remplace par un mot $m \in \mathcal{C}$ qui minimise la distance de Hamming $d(m, m')$. En particulier il conserve m' si m' est déjà un mot de code.

Remarque 2.2 *On peut montrer que, si p est petit (plus précisément si $p < 1/2$), ce qui toujours le cas en pratique, cette méthode de correction revient à remplacer m' par $m' \oplus e$ où $e \in \{0,1\}^n$ est le « vecteur d'erreurs » le plus probable altérant un mot de code en m', c'est à dire pour lequel $m' = m \oplus e$ avec $m \in \mathcal{C}$. En partant du principe que ce vecteur d'erreurs le plus probable est le véritable vecteur d'erreurs, on écrit $m' = m \oplus e$ avec $m \in \mathcal{C}$ donc $m' \oplus e = (m \oplus e) \oplus e = m \oplus (e \oplus e) = m \oplus \mathbf{0} = m$ avec $d(m, m') = d(m' \oplus e, m') = d(e, \mathbf{0}) = w(e)$, qui est minimale parmi toutes les distances de Hamming entre m' et un mot de code.*

Théorème 2.2 (*situation de correction certaine*)

Considérons un code de type $[n, k, d]$. Avec la méthode ci-dessus, le receveur corrige convenablement tout mot binaire comportant strictement moins de $d/2$ erreurs.

Preuve : Supposons qu'un mot reçu $m' \in \{0,1\}^n$ ne soit pas bien corrigé. Cela signifie que $m' = m \oplus e$ où m est le mot de code émis, $e \in \{0,1\}^n$ est un « vecteur d'erreurs », et que l'on a remplacé m' par un mot de code $\widehat{m} \neq m$ parce que $d(\widehat{m}, m') \leq d(m, m')$. Avec l'inégalité triangulaire, on obtient alors

$$d \leq d(m, \widehat{m}) \leq d(m, m') + d(m', \widehat{m}) \leq 2 \times d(m, m') = 2 \times d(m, m \oplus e) = 2 \times d(\mathbf{0}_n, e) = 2 \times w(e)$$

donc $w(e) \geq d/2$, autrement dit au moins $d/2$ erreurs de transmission sont survenues. □

2.7 Inégalité de Hamming et codes parfaits

Théorème 2.3 (*Inégalité de Hamming*)

Considérons un code $\mathcal{C}$ de type $[n,k,d]$ et notons t le plus grand entier strictement inférieur à $d/2$ (explicitement $t = \frac{d}{2} - \frac{3+(-1)^d}{4}$). Alors on a

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t} \leq 2^{n-k}$$

Preuve : Soient un mot quelconque $m \in \{0,1\}^n$ et p un entier entre 0 et n. Les mots m' vérifiant $d(m,m') = p$ sont les mots de $\{0,1\}^n$ qui diffèrent de m sur exactement p bits. Il y en a donc $\binom{n}{p} = \frac{n!}{p!(n-p)!}$ puisque ce coefficient binomial est le nombre de façons de choisir p bits parmi les n bits de m. En conséquence il y a $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}$ mots $m' \in \{0,1\}^n$ vérifiant $d(m,m') \leq t$. On note $S(m;t)$ l'ensemble de tous ces mots, que l'on nomme *sphere de centre m et de rayon t*. Remarquons maintenant pour $m_1 \neq m_2$ dans $\mathcal{C}$ on a $S(m_1,t) \cap S(m_2,t) = \emptyset$. En effet si un mot m' était commun à ces deux sphères alors on obtiendrait avec l'inégalité triangulaire

$$d(m_1,m_2) \leq d(m_1,m') + d(m',m_2) \leq t + t < d$$

ce qui contredit la définition de d. Comme il y a 2^k sphères de rayon t dont le centre est un mot de $\mathcal{C}$, on en déduit que

$$2^k \times \left(\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t} \right) \leq 2^n$$

En divisant par 2^k on obtient bien

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t} \leq 2^{n-k}$$

□

Remarque 2.3 *Ce résultat permet de majorer t, et donc aussi d, lorsque n et k sont fixés. Il suffit de calculer successivement les termes de la suite croissante $u_0 = \binom{n}{0} = 1$, $u_1 = \binom{n}{0} + \binom{n}{1} = 1 + n$, $u_2 = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} = 1 + n + \frac{n(n-1)}{2}$ etc. Cette suite dépasse 2^{n-k} au plus tard à son $(n+1)^{eme}$ terme car $u_n = \binom{n}{0} + \cdots + \binom{n}{n} = 2^n > 2^{n-k}$. D'après le Théorème 2.3, le plus petit entier x tel que $\binom{n}{0} + \cdots + \binom{n}{x} > 2^{n-k}$ est strictement plus grand que t.*

Définition 2.3

Un code $\mathcal{C}$ de type $[n,k,d]$ est ***parfait*** *si pour tout mot $m' \in \{0,1\}^n$ il existe un mot $m \in \mathcal{C}$ (nécessairement unique) tel que $d(m,m') \leq t$, où t est le plus grand entier strictement inférieur à $d/2$. Ceci est équivalent à dire que*

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t} = 2^{n-k}$$

Ainsi, pour un code parfait de type $[n,k,d]$, il n'y a que deux alternatives lorsqu'un mot m' est reçu :

- ou bien m' contient strictement moins que $d/2$ erreurs, et alors il est bien corrigé ;
- ou bien m' contient $d/2$ erreurs ou plus, et alors il est mal corrigé.

2.8 Codes linéaires

2.8.1 Généralités

Dans ce paragraphe, nous manipulons seulement des matrices à coefficients dans $\{0,1\}$; de plus les opérations sur ces matrices (c'est à dire l'addition, la multiplication par un scalaire dans $\{0,1\}$ et la multiplication des matrices entre elles) sont définies de la même façon que dans le cadre usuel mais en remplaçant l'addition des réels par l'addition $\oplus$ dans $\{0,1\}$ (et la multiplication des réels par la multiplication $\otimes$ dans $\{0,1\}$). Par exemple

$$\begin{pmatrix}1&1\\0&1\end{pmatrix}+\begin{pmatrix}0&1\\1&1\end{pmatrix}=\begin{pmatrix}1&0\\1&0\end{pmatrix}\qquad\text{et}\qquad\begin{pmatrix}1&0&1\\1&1&0\end{pmatrix}\times\begin{pmatrix}1&1\\1&0\\0&1\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=I_2$$

Les propriétés habituelles des opérations sur les matrices (commutativité de l'addition, associativité de l'addition et de la multiplication, distributivité de l'addition sur la multiplication) restent vraies.

Définition 2.4

Un code systématique de type $[n,k,d]$ est ***linéaire*** *quand, pour tout bloc $x_1\cdots x_k\in\{0,1\}^k$, le mot de code $y_1\cdots y_n=\varphi(x_1\cdots x_k)$ se calcule par la formule matricielle*

$$\underbrace{\begin{pmatrix}y_1&\cdots&y_n\end{pmatrix}}_{matrice\,1\times n}=\underbrace{\begin{pmatrix}x_1&\cdots&x_k\end{pmatrix}}_{matrice\,1\times k}\times G$$

où G est une matrice de taille $k\times n$ de la forme $G=(I_k\mid P)$, c'est à dire obtenue en prenant la matrice identité I_k de taille $k\times k$ et en collant sur sa droite une matrice P de taille $k\times(n-k)$. Cette matrice G est appelée la ***matrice génératrice*** *du code.*

Remarque 2.4 *En particulier $\varphi(100\cdots0)$ se lit sur la première ligne de G, $\varphi(010\cdots0)$ sur la deuxième ligne, $\cdots$, $\varphi(00\cdots01)$ sur la dernière ligne. De plus, grâce à la formule matricielle de la définition, il suffit de connaitre ces k mots de code pour pouvoir coder n'importe quel bloc.*

Exemple 2.2 *Pour $k=3$ et $n=5$, on a G de la forme $\begin{pmatrix}1&0&0&x&y\\0&1&0&x'&y'\\0&0&1&x''&y''\end{pmatrix}$ et la fonction de codage $\varphi:\{0,1\}^3\to\{0,1\}^5$ vérifie*

$$\varphi(100)=100xy\qquad\varphi(010)=010x'y'\qquad\varphi(001)=010x''y''$$

où x,y,x',y',x'',y'' sont les bits de contrôle.

Propriété 2.3

Considérons un code systématique linéaire $\mathcal{C}$ de type $[n,k,d]$. Alors on a :

1. *Pour tous blocs b,b' on a $\varphi(b\oplus b')=\varphi(b)\oplus\varphi(b')$.*
2. *On a $\varphi(\mathbf{0}_k)=\mathbf{0}_n$; en particulier $\mathbf{0}_n\in\mathcal{C}$.*
3. *la distance minimale d est égale au poids minimal des mots non nuls de $\mathcal{C}$.*

Preuve : 1. Notons $b = x_1 \cdots x_k$ et $\varphi(b) = y_1 \cdots y_n$. De même $b' = x'_1 \cdots x'_k$ et $\varphi(b') = y'_1 \cdots y'_n$. On a $b \oplus b' = x_1 \oplus x'_1 \cdots x_k \oplus x'_k$ et $\varphi(b \oplus b')$ se lit dans la matrice-ligne $\begin{pmatrix} x_1 \oplus x'_1 & \cdots & x_k \oplus x'_k \end{pmatrix} \times G$ puisque le code est linéaire. Mais on a les égalités matricielles

$$\begin{aligned} \begin{pmatrix} x_1 \oplus x'_1 & \cdots & x_k \oplus x'_k \end{pmatrix} \times G &= \Big(\begin{pmatrix} x_1 & \cdots & x_k \end{pmatrix} + \begin{pmatrix} x'_1 & \cdots & x'_k \end{pmatrix} \Big) \times G \\ &= \begin{pmatrix} x_1 & \cdots & x_k \end{pmatrix} \times G + \begin{pmatrix} x'_1 & \cdots & x'_k \end{pmatrix} \times G \\ &= \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} + \begin{pmatrix} y'_1 & \cdots & y'_n \end{pmatrix} \\ &= \begin{pmatrix} y_1 \oplus y'_1 & \cdots & y_n \oplus y'_n \end{pmatrix} \end{aligned}$$

la première et la dernière égalité étant vraies par définition de la somme de deux matrices (ici des matrices-lignes), la deuxième par distributivité de l'addition $+$ des matrices par rapport à leur multiplication $\times$ et la troisième par linéarité du code. Ainsi $\varphi(b \oplus b') = y_1 \oplus y'_1 \cdots y_n \oplus y'_n = \varphi(b) \oplus \varphi(b')$.
2. On a $\mathbf{0}_k = \mathbf{0}_k \oplus \mathbf{0}_k$ donc on déduit du 1) ci-dessus et de la dernière ligne de la Propriété 2.1 que $\varphi(\mathbf{0}_k) = \varphi(\mathbf{0}_k) \oplus \varphi(\mathbf{0}_k) = \mathbf{0}_n$.
3. On note δ le poids minimal d'un mot de code non nul. Considérons deux mots de code $m \neq m'$ minimisant la distance de Hamming, c'est à dire vérifiant $d(m, m') = d$. Puisque m et m' sont des mots de code, il existe des blocs b et b' dans $\{0,1\}^k$ tels que $\varphi(b) = m$ et $\varphi(b') = m'$. On a donc avec le 1) précédent

$$\mathbf{0}_n \neq m \oplus m' = \varphi(b) \oplus \varphi(b') = \varphi(b \oplus b') \in \mathcal{C}$$

et donc $d = d(m, m') = w(m \oplus m') \geq \delta$. L'inégalité inverse $d \leq \delta$ est vraie aussi car si m'' est un mot de code non nul de poids minimal alors on a $\delta = w(m'') = d(\mathbf{0}_n, m'') \geq d$ puisque $\mathbf{0}_n \in \mathcal{C}$ d'après le 2) . □

2.8.2 Détection et correction par syndromes

Le but de ce paragraphe est de voir que, pour un code systématique linéaire, on peut détecter et corriger des erreurs sans connaitre tous les mots de code.

Définition 2.5

- *La **matrice de contrôle** d'un code sytématique linéaire de type $[n, k, d]$ et de matrice génératrice $G = (I_k \,|\, P)$ est la matrice $H = (P^\intercal \,|\, I_{n-k})$ de taille $(n-k) \times n$, où $P^\intercal$ désigne la transposée de P.*
- *Le **syndrome** d'un mot $m = y_1 \cdots y_n \in \{0,1\}^n$ est le mot $z_1 \cdots z_{n-k} \in \{0,1\}^{n-k}$ qui se lit dans la matrice-ligne $\begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} \times H^\intercal$. Il est noté $\sigma(m)$.*

Exemple 2.3 *Si G est comme dans l'Exemple 2.2, on a*

$$H = \begin{pmatrix} x & x' & x'' & 1 & 0 \\ y & y' & y'' & 0 & 1 \end{pmatrix} \text{ et } H^\intercal = \begin{pmatrix} x & y \\ x' & y' \\ x'' & y'' \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Le syndrome du mot $01010 \in \{0,1\}^5$ *est* $\sigma(01010) = x' \oplus 1\, y' \in \{0,1\}^2$.

Propriété 2.4

Considérons un code systématique linéaire $\mathcal{C}$ de type $[n,k,d]$ et de matrice de contrôle H. Alors pour tous mots m et m' dans $\{0,1\}^n$ on a :

1. *$\sigma(m \oplus m') = \sigma(m) \oplus \sigma(m')$.*
2. *$m \in \mathcal{C}$ si et seulement si $\sigma(m) = \mathbf{0}_{n-k}$.*
3. *$\sigma(m) = \sigma(m')$ si et seulement si il existe $m'' \in \mathcal{C}$ tel que $m' = m \oplus m''$.*

Preuve : 1. Les arguments sont les mêmes que pour le 1) de la Propriété 2.3, avec $\sigma(m)$ au lieu de $\varphi(b)$ et $H^\intercal$ au lieu de G.
2. On se contente de le démontrer pour $n = 5$ et $k = 3$, c'est à dire avec G comme dans l'Exemple 2.2 et donc $H^\intercal$ comme dans l'Exemple 2.3. Pour $m = x_1x_2x_3x_4x_5 \in \{0,1\}^5$ on a

$$\sigma(m) = \mathbf{0}_2 \Leftrightarrow \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{pmatrix} \times \begin{pmatrix} x & y \\ x' & y' \\ x'' & y'' \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = (0\,0) \Leftrightarrow \begin{cases} x_1 \otimes x \oplus x_2 \otimes x' \oplus x_3 \otimes x'' \oplus x_4 = 0 \\ x_1 \otimes y \oplus x_2 \otimes y' \oplus x_3 \otimes y'' \oplus x_5 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_1 \otimes x \oplus x_2 \otimes x' \oplus x_3 \otimes x'' = x_4 \\ x_1 \otimes y \oplus x_2 \otimes y' \oplus x_3 \otimes y'' = x_5 \end{cases} \quad \text{car } x_4 \oplus x_4 = 0 \text{ et } x_5 \oplus x_5 = 0$$

$$\Leftrightarrow \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & x & y \\ 0 & 1 & 0 & x' & y' \\ 0 & 0 & 1 & x'' & y'' \end{pmatrix}$$

$$\Leftrightarrow m \in \mathcal{C}.$$

3. On a

$$\sigma(m) = \sigma(m') \Leftrightarrow \sigma(m) \oplus \sigma(m') = \mathbf{0}_{n-k} \Leftrightarrow \sigma(m \oplus m') = \mathbf{0}_{n-k} \Leftrightarrow m \oplus m' \in \mathcal{C}$$
$$\Leftrightarrow \text{il existe } m'' \in \mathcal{C} \text{ tel que } m' = m \oplus m''$$

la première équivalence étant due au fait que $\sigma(m') \oplus \sigma(m') = \mathbf{0}_{n-k}$, la deuxième et la troisième sont donnés par le 1) et le 2) précédents et la dernière est le fait que $m \oplus m' = m''$ si et seulement si $m' = m \oplus m''$ parce que $m \oplus m = \mathbf{0}_n$. □

En combinant le 2) de la Propriété 2.4 avec le Théorème 2.1, on voit que le receveur détecte de façon certaine tout mot contenant entre 1 et $d-1$ erreurs simplement en calculant son syndrome.

Théorème 2.4 (*correction par syndromes*)

Soit un code systématique linéaire $\mathcal{C}$ de type $[n,k,d]$ et de matrice de contrôle H. Formons une table à deux colonnes et 2^{n-k} lignes de la façon suivante :

• on place dans la colonne de gauche les 2^{n-k} mots de $\{0,1\}^{n-k}$;

• on parcours les mots dans $\{0,1\}^n$ par ordre de poids croissant et l'on calcule pour chaque mot m rencontré son syndrome $\sigma(m) \in \{0,1\}^{n-k}$.

- si la case de la table à droite de $\sigma(m)$ est vide, on y inscrit m.

- sinon, on passe au mot suivant.

Une fois cette table remplie, tout mot binaire reçu $m' \in \{0,1\}^n$ contenant strictement moins de $d/2$ erreurs est bien corrigé en le remplaçant par $m \oplus m'$, où m est le mot inscrit dans la table à droite de $\sigma(m')$.

Preuve : Pour $m \in \{0,1\}^n$, considérons l'ensemble $S(m) = \{m' \in \{0,1\}^n \,|\, \sigma(m) = \sigma(m')\}$, autrement dit l'ensemble des mots $m' \in \{0,1\}^n$ ayant le même syndrome que m (y compris m lui-même). Le 3) de la Propriété 2.4 nous dit que les mots dans $S(m)$ sont ceux qui s'obtiennent en additionnant à m un mot de $\mathcal{C}$. De plus, quand on additionne à m deux mots différents de $\mathcal{C}$ alors on obtient deux mots différents de $S(m)$ car $m \oplus m_1'' = m \oplus m_2''$ si et seulement si $m_1'' = m_2''$, à nouveau parce que $m \oplus m = \mathbf{0}_n$. On obtient donc que $S(m)$ contient le même nombre d'éléments que $\mathcal{C}$, à savoir 2^k. Maintenant remarquons que la réunion de tous ces ensembles $S(m)$ est égale à $\{0,1\}^n$ (car $m \in S(m)$) et que deux ensembles $S(m_1)$ et $S(m_2)$ sont ou bien égaux ou bien disjoints, selon que $\sigma(m_1) = \sigma(m_2)$ ou que $\sigma(m_1) \neq \sigma(m_2)$. En conséquence, il y a $2^n/2^k = 2^{n-k}$ ensembles $S(m)$ différents et donc aussi 2^{n-k} valeurs différentes pour les syndromes. Ainsi chaque mot de $\{0,1\}^{n-k}$ est le syndrome de 2^k mots de $\{0,1\}^n$ exactement. En particulier toutes les lignes de la deuxième colonne du tableau seront remplies.

Supposons maintenant que $m' \in \{0,1\}^n$ soit un mot reçu et que $m \in \{0,1\}^n$ soit le mot inscrit dans la table à droite de $\sigma(m')$. Par construction de la table, ceci signifie que $\sigma(m) = \sigma(m')$ et que m a un poids minimal parmi les éléments de $S(m')$. On corrige m' en le remplaçant par $m \oplus m'$ qui est bien un mot de code puisque, comme on l'a déjà vu, le 1) et le 2) de la Propriété 2.4 donnent

$$\sigma(m) = \sigma(m') \Leftrightarrow \sigma(m) \oplus \sigma(m') = \mathbf{0}_n \Leftrightarrow \sigma(m \oplus m') = \mathbf{0}_n \Leftrightarrow m \oplus m' \in \mathcal{C}.$$

De plus $d(m', m \oplus m') = d(\mathbf{0}_n, m) = w(m)$. On sait aussi avec le 3) de la Propriété 2.4 que pour tout autre mot $\widehat{m} \in \mathcal{C}$ on a $m' \oplus \widehat{m} \in S(m')$ donc $w(m) \leq w(m' \oplus \widehat{m}) = d(m', \widehat{m})$ ce qui montre que $m \oplus m'$ est bien un mode de code dont la distance à m' est minimale. On conclut en appliquant le Théorème 2.2. □.

2.9 Exercices

Rappel : dans les exercices qui suivent, on suppose les mots de codes acheminés à travers un canal de transmission binaire symétrique et sans mémoire. La probabilité qu'un bit soit mal transmis est notée p.

Exercice 1. Prouver la Propriété 2.1 du cours, en commençant par le cas $n = 1$.

Exercice 2. Dans $\{0,1\}^5$, on considère les mots binaires $m = 10101$, $m' = 11010$ et $m'' = 01010$. Donner les poids de ces mots ainsi que les distances de Hamming $d(m, m')$, $d(m, m'')$, $d(m', m'')$.

Exercice 3. On considère le code systématique $\mathcal{C} \subset \{0,1\}^6$ contenant les mots suivants :

000000	001110	010101	011100	100011	101010	110001	111111

1) Donner le type $[n, k, d]$ de ce code ainsi que son rendement.
2) a) Combien d'erreurs peut-on détecter de façon certaine ?
b) Quand un mot reçu contient une erreur, peut-on la corriger de façon certaine ?
3) Le code $\mathcal{C}$ est-il linéaire ?
4) On suppose que l'on a reçu le mot $m' = 010011$.
a) m' contient-il (au moins) une erreur ? Le cas échéant, comment corriger m' ?
b) Écrire la liste de tous les vecteurs d'erreurs possibles, c'est à dire les vecteurs $e \in \{0,1\}^6$ tels que $m' = m \oplus e$ avec $m \in \mathcal{C}$.
c) Exprimer en fonction de p la probabilité de chaque vecteur d'erreurs possible. Donner une formule générale exprimant cette probabilité en fonction du poids du vecteur considéré.

d) Montrer que, pour $p < 1/2$, la probabilité d'un vecteur d'erreurs e décroît quand le poids $w(e)$ croît.

e) En déduire le ou les vecteurs d'erreurs les plus probables. Pourquoi ceci légitime-t-il la correction faite au a) ?

5) Reprendre la question 4) avec $m' = 101101$.

6) Le code $\mathcal{C}$ est-il parfait ?

Exercice 4. On code les blocs de $\{0,1\}^2$ avec l'application de codage $\varphi : \{0,1\}^2 \to \{0,1\}^5$ définie par $\varphi(ab) = ab\bar{a}\bar{b}c$ où $\bar{0} = 1$, $\bar{1} = 0$ et de plus $c = 0$ si $w(ab)$ est pair, $c = 1$ si $w(ab)$ est impair.

1) Donner le type $[n, k, d]$ de ce code ainsi que son rendement.

2) Avec ce code, combien d'erreurs sont détectées de façon certaine ? Combien sont corrigées de façon certaine ?

3) Ce code est-il linéaire ?

4) On suppose que l'on a reçu le mot $m' = 10001$.

a) m' comporte-t-il (au moins) une erreur ? Le cas échéant, comment est-il corrigé ?

b) Écrire la liste de tous les vecteurs d'erreurs possibles, c'est à dire les vecteurs $e \in \{0,1\}^5$ tels que $m' = m \oplus e$ avec $m \in \mathcal{C}$.

c) On suppose $p < 1/2$. En procédant comme à l'exercice précédent, trouver le ou les vecteurs d'erreurs les plus probables. Pourquoi ceci légitime-t-il la correction faite au a) ?

5) Reprendre la question 4) avec $m' = 11111$.

6) Ce code est-il parfait ?

Exercice 5. (bit de parité)

On considère le code systématique de type [4,3,d] dont la fonction de codage $\varphi : \{0,1\}^3 \to \{0,1\}^4$ est définie par $\varphi(x_1x_2x_3) = x_1x_2x_3x_4$ où x_4 est choisi de telle façon qu'il y ait un nombre pair de 1 parmi $x_1x_2x_3x_4$.

1) Quel est le rendement de ce code ? Quelle est la valeur de d ?

2) Combien d'erreurs peut-on détecter de façon certaine ? Combien peut-on en corriger de façon certaine ?

3) Exprimer x_4 en fonction de x_1,x_2 et x_3. En déduire que ce code est linéaire.

4) a) Pour que le receveur ne s'aperçoive pas qu'un mot reçu m' est erroné, combien m' doit-il comporter d'erreurs ? Calculer la probabilité que ceci arrive.

b) En déduire la probabilité qu'un mot reçu ne soit pas détecté sachant qu'il est erroné.

c) Calculer la probabilité du b) quand $p = 0,01$.

Exercice 6. (codage par répétition)

On considère le code systématique de type [3,1,d] dont la fonction de codage $\varphi : \{0,1\} \to \{0,1\}^3$ est $\varphi(0) = 000$ et $\varphi(1) = 111$.

1) Quel est le rendement de ce code ? Quelle est la valeur de d ?

2) Ce code est-il linéaire ?

3) Combien d'erreurs ce code peut-il détecter de façon certaine ? Combien peut-il en corriger de façon certaine ?

4) a) Pour que le receveur ne s'aperçoive pas qu'un mot reçu m' est erroné, combien m' doit-il comporter d'erreurs ? Calculer la probabilité que ceci arrive.

b) En déduire la probabilité qu'un mot reçu ne soit pas détecté sachant qu'il est erroné.

c) Calculer la probabilité du b) quand $p = 0,01$.

5) Calculer la probabilité qu'un mot reçu soit mal corrigé sachant qu'il est erroné.

Exercice 7. On considère le code systématique linéaire $\mathcal{C}$ de matrice génératrice $G = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix}$

1) Donner le type $[n, k, d]$ de ce code.
2) Combien d'erreurs ce code détecte-t-il de façon certaine ? Combien en corrige-t-il de façon certaine ?
3) Corriger les messages reçus 11011 et 11010.
4) Ce code est-il parfait ?

Exercice 8. (code de Hamming de type [7,4,3])

On considère le code systématique linéaire $\mathcal{C}$ de matrice génératrice $G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$

1) Vérifier que ce code est de type [7,4,3].
2) Ce code est-il parfait ?
3) En utilisant les syndromes, détecter les messages reçus erronés parmi 1101001, 1011100 et 1110010 puis les corriger.

Exercice 9. Pour un code (systématique) de type [5,2,d], quelle majoration de d obtient-on en utilisant l'inégalité de Hamming ? Donner un exemple avec la plus grande valeur de d possible.

Exercice 10. Le but de cet exercice est de démontrer que, pour un code linéaire $\mathcal{C}$, ou bien tous les mots de code se terminent par 0, ou bien exactement la moitié d'entre eux se terminent par 0. Par la suite, on note $\mathcal{C}_0$ l'ensemble des mots de code se terminant par 0 et $\mathcal{C}_1$ l'ensemble des mots de code se terminant par 1.
1) Justifier que $\mathcal{C}_0 \neq \emptyset$.
2) On suppose dans cette question que $\mathcal{C}_1 \neq \emptyset$, autrement dit qu'il existe au moins un mot de code se terminant par 1, et on choisit $\widehat{m} \in \mathcal{C}_1$.

a) Justifier que, pour tout $m \in \mathcal{C}_0$, on a $m \oplus \widehat{m} \in \mathcal{C}_1$. On définit donc une fonction $f : \mathcal{C}_0 \to \mathcal{C}_1$ en posant $f(m) = m \oplus \widehat{m}$ pour tout $m \in \mathcal{C}_0$.

b) Montrer que f est une bijection, c'est à dire qu'elle est injective et bijective.

c) Conclure.