SPÉCIALE MP* : DEVOIR SURVEILLÉ, ÉTUDE D'UN CODAGE

La qualité de la rédaction et de la présentation, la clarté et la précision des raisonnements constitueront un élément important pour l'appréciation des copies.

Notations et objectif du problème

Dans tout le problème E et F sont deux espaces vectoriels sur un corps $\mathbb{K}$ de dimensions respectives k et n avec $0 < k < n$; $\mathcal{B}_E = (e_1, e_2, \ldots, e_k)$ est une base de E et $\mathcal{B}_F = (f_1, f_2, \ldots, f_n)$ est une base de F. On suppose donnée une application linéaire <u>injective</u> g de E dans F et on pose $C = g(E)$. Une telle application g est appelée *code* : en effet, les composantes $x_1, x_2, \ldots, x_n$ de $x = g(u)$ sont les symboles successifs du message codé. Le message codé est transmis et on recueille un élément y de F, dont les composantes $y_1, y_2, \ldots, y_n$ sont les symboles successifs du message reçu. Le message reçu y peut différer de x à cause d'erreurs de transmission affectant certaines composantes. Le but du problème est d'étudier comment on peut reconstituer x puis u à partir de y.

I - Résultats généraux sur les codes

I.1. *Distance minimale d'un code.*

Soit W la fonction de F dans $\mathbb{R}$ qui à tout élément de F fait correspondre le nombre de composantes non nulles de cet élément dans la base $\mathcal{B}_F$.

a. Montrer que la fonction Δ de $F \times F$ dans $\mathbb{R}$ définie par

$$\Delta(x, y) = W(x - y)$$

est une distance sur F
(i.e. $\Delta(x, y) = 0 \Leftrightarrow x = y$, $\Delta(x, y) = \Delta(y, x)$,
$\Delta(x, y) \leqslant \Delta(x, z) + \Delta(z, y)$ pour tout triplet (x, y, z) de F^3).

b. Le nombre d défini par $d = \min\limits_{u \in C \setminus \{0\}} W(u)$ est appelé distance minimale du code.

Soit t la partie entière de $\dfrac{d-1}{2}$, montrer que pour tout élément y de F, il existe *au plus* un élément x de C tel que $\Delta(x, y) \leqslant t$.

I.2. *Applications de contrôle.*

Soit h une application linéaire de F dans un espace vectoriel F'. On suppose que le noyau de h est C. On dit alors que h est une application de contrôle pour le code g.

a. Déterminer, en fonction de n et k, le rang de la famille $(h(f_i))_{i \in [\![1,n]\!]}$.

b. Soient x un élément non nul de C et J l'ensemble des indices i pour lesquels la composante de x sur f_i est non nulle. Montrer que la famille $(h(f_i))_{i \in J}$ est liée

c. Montrer qu'il existe une famille de d vecteurs extraite de la famille $(h(f_i))_{i \in [\![1,n]\!]}$ qui est liée.

d. Montrer enfin que $d \leqslant n - k + 1$.

I.3. *Un exemple d'application de contrôle.*

a. Montrer qu'il existe une partie P de l'ensemble $[\![1, n]\!]$ ayant $n - k$ éléments telle que la famille $(f_i)_{i \in P}$ soit base d'un supplémentaire D de C dans F. Quitte à permuter éventuellement les éléments de la base $\mathcal{B}_F$, on supposera désormais que la partie P ainsi trouvée est la partie $[\![k + 1, n]\!]$.

b. Dans ces conditions, montrer qu'il existe une base $\mathcal{B}'_E$ de E telle que la matrice de l'application linéaire g dans les bases $\mathcal{B}'_E$ et $\mathcal{B}_F$ soit de la forme $G = \begin{pmatrix} I_k \\ A \end{pmatrix}$ où I_k désigne la matrice identité d'ordre k et A est une matrice à k colonnes et $n-k$ lignes.

c. Soit h l'application linéaire de F dans D qui à tout élément x associe la projection de x sur D parallèlement à C. Montrer que h est une application de contrôle pour g et déterminer la matrice H de h lorsque F est muni de la base $\mathcal{B}_F$ et D de la base $(f_{k+1}, \ldots, f_n)$.

d. *Correction d'un message.*
On suppose que, dans la transmission d'un message x, il se produit t erreurs au plus. Soit y un message reçu. On calcule $z = h(y)$. Prouver qu'il existe un élément e de F tel que $h(e) = z$ et $W(e) \leqslant t$. Montrer que $x = y - e$ est l'unique élément de C dont la distance à y est minimum. Ainsi x est le message transmis.

I.4. *Exemple d'un code.*
On suppose que $\mathbb{K}$ est le corps à 2 éléments $\{0, 1\}$.
E est l'espace $\mathbb{K}^4$, F l'espace $\mathbb{K}^7$, ces deux espaces étant muni de leurs bases canoniques respectives. Enfin g est l'application linéaire de E dans F dont la matrice dans les bases ci-dessus est : $\begin{pmatrix} I_4 \\ A \end{pmatrix}$ où $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$.

a. Calculer la matrice H de l'application linéaire h introduite au **3.c**.
b. Déterminer la valeur de la distance minimale d.
c. Soit $y = (1, 0, 1, 0, 0, 1, 0)$. Calculer $h(y)$.
En conclure qu'il existe un élément x de C et un seul tel que $\Delta(x, y) \leqslant 1$. Quel est cet élément x? Quel est l'élément u de $\mathbb{K}^4$ tel que $x = g(u)$?

II - Etude d'une famille de codes

Dans cette partie, le corps de base est le corps à deux éléments $\{0, 1\}$, on fera attention au fait que $\forall x \in \mathbb{K}$, $2x = 0$, $x = -x$. A $x \in \{0, 1\} = \mathbb{K}$ on fait correspondre $\bar{x} \in \mathbb{R}$ tel que $\bar{0} = 0$ et $\bar{1} = 1$.

On désigne par m un entier strictement positif et on pose $k = m + 1$. On note E_k l'espace vectoriel $\mathbb{K}^k$ et F_k l'espace vectoriel de toutes les fonctions définies sur $\mathbb{K}^m$ à valeurs dans $\mathbb{K}$.

II.1. *Construction d'une base de F_k par codage binaire.*

a. Quel est le nombre d'éléments de E_k et de F_k ? Quelle est la dimension de F_k ?

b. Montrer que l'application α de $\mathbb{K}^m$ dans $\{0, 1, \ldots, 2^m - 1\}$ qui à tout $(w_1, w_2, \ldots, w_m)$ de $\mathbb{K}^m$ fait correspondre $\sum\limits_{i=1}^{m} \bar{w}_i 2^{i-1}$ est bijective.

c. Pour tout $(w_1, w_2, \ldots, w_m)$ de $\mathbb{K}^m$ notons $\operatorname{supp}(w_1, w_2, \ldots, w_m)$ l'ensemble des indices $i \in [\![1, m]\!]$ tels que $w_i \neq 0$. Montrer que l'application β de $\{0, 1, \ldots, 2^m - 1\}$ dans l'ensemble des parties de $\{1, 2, \ldots, m\}$ qui, à j écrit sous la forme $\sum\limits_{i=1}^{m} w_i 2^{i-1}$, associe $P_j = \operatorname{supp}(w_1, w_2, \ldots, w_m)$ est bijective.

d. Pour tout élément j de $\{0, 1, 2, \ldots, 2^m - 1\}$ définissons la fonction f_j de $\mathbb{K}^m$ dans $\mathbb{K}$ par :

$$f_j(v_1, v_2, \ldots, v_m) = \begin{cases} 1 \text{ si } \operatorname{supp}(v_1, v_2, \ldots, v_m) = P_j \\ 0 \text{ sinon.} \end{cases}$$

Montrer que la famille $\mathcal{B}_{F_k} = (f_0, f_1, \ldots, f_{2^m-1})$ est une base de F_k.

e. Soit U la fonction de $\mathbb{K}^m$ dans $\mathbb{K}$ constante et égale à 1. Pour tout j, on note V_j la fonction de $\mathbb{K}^m$ dans $\mathbb{K}$ définie par $V_j(v_1, v_2, \dots, v_m) = v_j$ et $\overline{V}_j = U - V_j$.
Exprimer f_j en fonction de $\prod_{i \in P_j} V_i$ et de $\prod_{i \notin P_j} \overline{V}_i$.

f. Montrer que dans la base $\mathcal{B}_{F_k}$ la composante sur f_j d'un élément f de F_k s'écrit sous la forme $f(u)$ où u est un élément de $\mathbb{K}^m$ que l'on déterminera.
Quelles sont les composantes de U sur la base $\mathcal{B}_{F_k}$?
Détailler l'écriture des composantes des V_i lorsque $k = 2, 3, 4$ (on pourra mettre les résultats sous forme de vecteurs lignes).
Donner alors l'écriture des V_i dans la base $\mathcal{B}_{F_k}$ dans le cas général (ici, il est conseillé de donner une écriture précise de V_i en l'exprimant en fonction des f_j).

II.2. *Définition d'une famille de codes.*
On munit désormais E_k de sa base canonique et F_k de la base construite en **1.d**.
On désigne par g_k l'application linéaire de E_k dans F_k définie par :

$$g_k(u_0, u_1, \dots, u_m) = u_0 U + u_1 V_1 + \cdots + u_m V_m.$$

a. Montrer que g_k est une application linéaire injective.
Nous noterons C_k l'image de $g_k(E_k)$ et d_k la distance minimale correspondante.

b. Pour $k = 2, 3$ déterminer tous les éléments de C_k (ceux-ci seront donnés par leurs composantes dans la base $\mathcal{B}_{F_k}$). Déterminer d_2 et d_3.

c. Supposons maintenant $k > 2$.
Comparer les composantes de $g_{k-1}(u_0, u_1, \dots, u_{m-1})$ dans la base $\mathcal{B}_{F_{k-1}}$ avec les 2^{m-1} premières et dernières composantes de $g_k(u_0, u_1, \dots, u_{m-1}, 0)$ dans la base $\mathcal{B}_{F_k}$.
Montrer que $g_k(u_0, u_1, \dots, u_{m-1}, 1)$ a exactement 2^{m-1} composantes non nulles.
En déduire alors par récurrence la valeur de d_k.

II.3. *Introduction d'une structure euclidienne.*
Soit A_k le $\mathbb{R}$-espace vectoriel des fonctions définies sur $\mathbb{K}^m$ à valeurs dans $\mathbb{R}$, muni du produit scalaire qui à tout couple (φ, ψ) d'éléments de A_k associe

$$(\varphi | \psi) = \sum_{w \in \mathbb{K}^m} \varphi(w)\psi(w).$$

D'autre part, pour tout couple (v, w) d'éléments de $\mathbb{K}^m$, où $v = (v_1, v_2, \dots, v_m)$ et $w = (w_1, w_2, \dots, w_m)$, posons $< v, w >= \sum_{i=1}^{m} v_i w_i$ où la somme est calculée dans $\mathbb{K}$.
Pour tout élément v de $\mathbb{K}^m$, définissons la fonction χ_v de $\mathbb{K}^m$ dans $\mathbb{R}$ par :

$$\chi_v(w) = (-1)^{<v,w>}.$$

a. Soit χ un élément de A_k tel que $\chi(0) = 1$ et tel que pour tout couple (w, w') d'éléments de $\mathbb{K}^m$ on ait

$$\chi(w + w') = \chi(w)\chi(w').$$

Montrer que $\chi(w)^2 = 1$ pour tout w. En déduire l'existence d'un unique élément v de $\mathbb{K}^m$ tel que $\chi = \chi_v$.

b. Montrer que, si v est un élément non nul de $\mathbb{K}^m$, il existe un élément s de $\mathbb{K}^m$ tel que $\chi_v(s) = -1$. En déduire que $(\chi_v | \chi_0) = 0$.

c. Prouver finalement que la famille $\left(\frac{1}{2^{m/2}}\chi_v\right)_{v \in \mathbb{K}^m}$ est une base orthonormale de l'espace euclidien A_k.

d. *Transformation de Fourier sur* $\mathbb{K}^m$;
Pour tout élément φ de A_k, on note $\widehat{\varphi}$ l'élément de A_k qui, à tout élément v de $\mathbb{K}^m$ associe $\widehat{\varphi}(v) = (\varphi|\chi_v)$.
Etablir que, pour tout élément φ de A_k :

$$\varphi = \frac{1}{2^m} \sum_{v \in \mathbb{K}^m} \widehat{\varphi}(v)\chi_v. \tag{1}$$

$$(\widehat{\varphi}|\widehat{\varphi}) = 2^m(\varphi|\varphi). \tag{2}$$

$$\widehat{\widehat{\varphi}} = 2^m\varphi. \tag{3}$$

Soit enfin $\mathcal{F}$ l'endomorphisme qui, à tout élément φ associe $\dfrac{1}{2^{m/2}}\widehat{\varphi}$. Enoncer les propriétés de $\mathcal{F}$ traduisant les relations (2) et (3). Déterminer l'image par $\mathcal{F}$ de la base $\left(\dfrac{1}{2^{m/2}}\chi_v\right)$.

II.4. *Application au décodage.*
A toute fonction f de F_k on associe la fonction $\varphi_f\ : \mathbb{K}^m \to \mathbb{R}$ définie par $\varphi_f(w) = (-1)^{f(w)}$.

a. Montrer que, pour tout $i \in [\![1, m]\!]$, il existe un élément unique v de $\mathbb{K}^m$ que l'on déterminera tel que $\varphi_{V_i} = \chi_v$. En déduire $\widehat{\varphi_{V_i}}$. Déterminer enfin $\widehat{\varphi_U}$.

b. Montrer que, pour tout élément $v = (v_1, v_2, \dots, v_m)$ de $\mathbb{K}^m$ et tout élément f de F_k, $\widehat{\varphi}_f(v)$ est égal au nombre de composantes nulles de la fonction $f + \sum\limits_{i=1}^{m} v_iV_i$ sur la base $\mathcal{B}_{F_k}$ diminué du nombre de ses composantes non nulles sur cette base.
En conclure que :

$$\Delta\left(f, \sum_{i=1}^{m} v_iV_i\right) = \frac{1}{2}(2^m - \widehat{\varphi_f}(v)).$$

Par un calcul analogue, montrer que :

$$\Delta\left(f, U + \sum_{i=1}^{m} v_iV_i\right) = \frac{1}{2}(2^m + \widehat{\varphi_f}(v)).$$

c. Dans le cadre du **2** , prenons l'exemple où $m = 3$.
Soit f la fonction qui sur la base $(f_0, f_1, \dots, f_7)$ admet (0,0,1,1,1,1,1,0) pour composantes.
Calculer pour chaque élément v de $\mathbb{K}^3$ la valeur de $\widehat{\varphi_f}(v)$.
En déduire l'élément de C_4 le plus proche de f au sens de la distance Δ.
Déterminer l'élément u de E tel que $g_4(u) = x$.