Théorème de Carathéodory

Leçons : 126 ; 151

Références : GOURDON, *Analyse* (p.54)

Prérequis :

- définition de Conv(A)
- le rang sur $\mathbb{Q}$ est le même que le rang sur $\mathbb{R}$
- théorème du rang

Introduction :

On va montrer un théorème qui nous permet de majorer le nombre de points utiles pour exprimer n'importe quel point d'un espace de dimension finie. Par exemple, dans $\mathbb{R}^n$, on pourra exprimer tous les points par seulement n+1 points. Ensuite, on utilisera ce théorème pour montrer l'existence de solutions d'un système diophantien.

Théorème 1. Soit A une partie de $\mathbb{R}^n$, on a

$$\mathrm{Conv}(A) = \left\{ \sum_{i=0}^{n+1} \lambda_i a_i, \text{ avec } a_i \in A, \lambda_i \geqslant 0, \sum_i \lambda_i = 1 \right\}$$

Démonstration.
L'inclusion $\boxed{\supset}$ est claire par définition de Conv(A).

Pour l'inclusion $\boxed{\subset}$, on prend $x \in \mathrm{Conv}(A)$, donc on peut écrire x sous la forme

$$x = \sum_{i=0}^{p} \lambda_i a_i \text{ avec } a_i \in A, \lambda_i \geqslant 0, \sum_i \lambda_i = 1$$

où p est l'entier minimal tel que l'on puisse écrire x sous cette forme.

Raisonnons par l'absurde, supposons que $p \geqslant n+2$. On pose

$$\phi : \left\{ \begin{array}{ccc} \mathbb{R}^p & \to & \mathbb{R}^n \times \mathbb{R} \\ (\nu_i)_1^p & \mapsto & \left(\sum_i \nu_i a_i, \sum_i \nu_i\right) \end{array} \right.$$

L'application ϕ est linéaire, donc on peut appliquer le théorème du rang :

$$p = \dim \mathbb{R}^p = \dim \ker \phi + \text{ rg } \phi$$

Or

$$\text{rg } \phi \leqslant n+1$$

Donc on a

$$\dim \ker \phi \geqslant p - n - 1 \geqslant 1$$

Ainsi on peut choisir $(\alpha_i)_1^p \in \ker \phi \setminus \{0\}$.

On pose

$$\mu_i(t) = \lambda_i + t\alpha_i \text{ pour tout } i \in \{1, \dots, p\} \text{ et pour tout } t \in \mathbb{R}$$

Ainsi on a

$$\sum_{i=1}^{p} \mu_i(t) = \underbrace{\sum_{i=1}^{p} \lambda_i}_{=1} + t \underbrace{\sum_{i=1}^{p} \alpha_i}_{=0} = 1$$

et

$$\sum_{i=1}^{p} \mu_i(t) a_i = \underbrace{\sum_{i=1}^{p} \lambda_i a_i}_{=x} + t \underbrace{\sum_{i=1}^{p} \alpha_i a_i}_{=0} = x$$

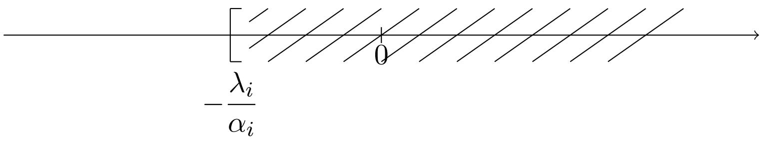
Donc maintenant il faut trouver $\tau \in \mathbb{R}$ tel que

(i) $\mu_i(\tau) \geqslant 0$ pour tout $i \in \{1, \dots, p\}$ (pour coller à la définition des éléments de l'enveloppe convexe)

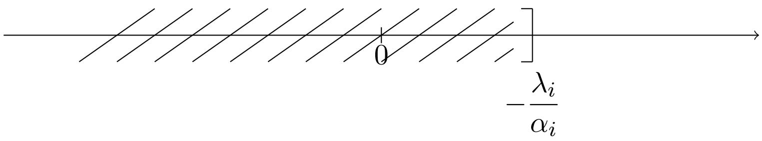
(ii) il existe i_0 tel que $\mu_{i_0}(\tau) = 0$ (pour donner une contradiction avec l'aspect p minimal)

Pour avoir $\mu_i(t) \geqslant 0$

— Si $\alpha_i = 0$, $\mu_i(t) = \lambda_i \geqslant 0$.
— Si $\alpha_i > 0$, il faut que $t \geqslant -\frac{\lambda_i}{\alpha_i}$[1] avec $-\frac{\lambda_i}{\alpha_i} \leqslant 0$



— Si $\alpha_i < 0$, il faut que $t \leqslant -\frac{\lambda_i}{\alpha_i}$[2] avec $-\frac{\lambda_i}{\alpha_i} \geqslant 0$



Comme les α_i sont non tous nuls[3], il existe au moins un α_i non nul et comme $\sum_i \alpha_i = 0$, il existe k et j tels que $\alpha_k > 0$ et $\alpha_j < 0$.

Ainsi on peut poser

$$\tau = \max\{-\frac{\lambda_i}{\alpha_i}, i \in \{1, \dots, p\}, \alpha_i > 0\}$$

On a donc pour tout $i \in \{1, \dots, p\}$, $\mu_i(\tau) \geqslant 0$[4]. On a donc le point (i).

La valeur τ est atteinte en un certain i_0. Or, comme $\tau = -\dfrac{\lambda_{i_0}}{\alpha_{i_0}}$, on a $\mu_{i_0}(\tau) = \lambda_{i_0} + \tau\alpha_{i_0} = 0$. Cela montre le point (ii).

Donc

$$\sum_{i=1}^{p} \mu_i(\tau) \alpha_i = \sum_{i \neq i_0} \mu_i(\tau) \alpha_i = x$$

Ainsi, on a exprimé x avec $p-1$ points, cela contredit donc la minimalité de p.

On a donc prouvé que $p \leqslant n+1$. Ce qui conclut la preuve du théorème. $\square$

1. $\lambda_i + t\alpha_i \geqslant 0 \iff t\alpha_i \geqslant -\lambda_i \iff t \geqslant -\frac{\lambda_i}{\alpha_i}$
2. $\lambda_i + t\alpha_i \geqslant 0 \iff t\alpha_i \geqslant -\lambda_i \iff t \leqslant -\frac{\lambda_i}{\alpha_i}$
3. car les α_i sont dans $\ker \phi$ privé de $\{0\}$
4. car par définition de τ on a le résultat pour les $\alpha_i > 0$, on a $\tau \leqslant 0$ et si $\alpha_i < 0$, on a $-\dfrac{\lambda_i}{\alpha_i} \geqslant 0 \geqslant \tau$

Application 1. Soit $A \in \mathcal{M}_{m,n}(\mathbb{Z})$.
Il y a équivalence entre les deux assertions suivantes :

(i) le système diophantien $Ax = 0$ admet une solution non nulle $x \in \mathbb{N}^n$

(ii) $0_{\mathbb{R}^m} \in \mathrm{Conv}(A_1, \dots, A_n)$ où $A_1, \dots, A_n$ sont les colonnes de la matrice A.

Démonstration. $\boxed{(i) \Longrightarrow (ii)}$ Soit $x \in \mathbb{N}^n$ une solution de $Ax = 0$, d'où

$$0_{\mathbb{R}^m} = \sum_{i=1}^{n} A_i x_i$$

Comme, pour tout $i \in \{1, \dots, n\}$, on a $x_i \geqslant 0$ et x est non nul, on peut prendre

$$\lambda_i = \frac{x_i}{\displaystyle\sum_j x_j}$$

On a donc

$$0_{\mathbb{R}^m} = \sum_{i=1}^{n} \lambda_i A_i$$

avec $\lambda_i \geqslant 0$ et $\sum_i \lambda_i = 1$, donc $0_{\mathbb{R}^m} \in \mathrm{Conv}(A)$.

$\boxed{(ii) \Longrightarrow (i)}$ Soit l minimal tel que $0_{\mathbb{R}^m}$ soit dans l'enveloppe convexe de l colonnes de A. On a alors

$$0_{\mathbb{R}^m} = \sum_{j=1}^{l} \lambda_j A_{i_j} \tag{1}$$

avec $\lambda_j \geqslant 0$ et $i_j \in \{1, \dots, n\}$ distincts.

On note $r = \mathrm{rg}_{\mathbb{R}}\left(A_{i_1}, \dots, A_{i_l}\right)$.

Alors, comme $A_{i_1}, \dots, A_{i_l}$ sont liées par (**??**), on a $\boxed{r < l}$.

De plus, $0_{\mathbb{R}^m} \in \mathrm{Vect}\left(A_{i_1}, \dots, A_{i_l}\right)$ qui est un espace vectoriel de dimension r sur $\mathbb{R}$, donc on a besoin d'au plus $r+1$ éléments de $\{A_{i_1}, \dots, A_{i_l}\}$ pour exprimer $0_{\mathbb{R}^m}$ par le théorème de Carathéodory. Ainsi, par minimalité de l, on a $\boxed{l \leqslant r+1}$. On peut donc en conclure que

$$l = r + 1.$$

On pose

$$\widetilde{A} = \begin{pmatrix} A_{i_1} & \dots & A_{i_l} \end{pmatrix} \in \mathcal{M}_{n,r+1}(\mathbb{Q})$$

qui est une matrice de rang r sur $\mathbb{Q}$. Ainsi, elle est de noyau non réduit à $\{0\}$ puisque le théorème du rang nous donne

$$r + 1 = \dim \ker \widetilde{A} + r$$

Soit $\tilde{y} \in \mathbb{Q}^{r+1}$ non nul tel que $\widetilde{A}\tilde{y} = 0$. On pose $y \in \mathbb{Q}^n$ non nul tel que

$$\begin{cases} y_i = \tilde{y}_i & \text{si } i \text{ est un des } i_j \\ y_i = 0 & \text{sinon} \end{cases}$$

Ainsi

$$Ay = \sum_{i=1}^{n} y_i A_i = \sum_{i \in (i_j)_{j=1}^{l}} \tilde{y_i} \widetilde{A_i} = \widetilde{A}\tilde{y} = 0_{\mathbb{R}^m}$$

Comme $y \in \mathbb{Q}^n$, quitte à multiplier y par un entier suffisamment grand, on peut prendre un $y \in \mathbb{Z}^n$ (toujours non nul).

On veut maintenant montrer que l'on peut trouver $y \in \mathbb{N}^n$. On sait que

$$(\lambda_1, \ldots, \lambda_{r+1}) \in \mathbb{R}_+^{r+1} \cap \ker_{\mathbb{R}} \widetilde{A}$$

Comme le rang ne dépend pas du corps de base, c'est aussi un élément de $\ker_{\mathbb{Q}} \widetilde{A}$. Or on sait que $\dim \ker_{\mathbb{Q}} \widetilde{A} = 1$, donc y et λ sont liés. Il existe donc $\alpha \in \mathbb{R}$ tel que $\alpha y = \lambda$ et on sait que tous les λ_i sont positifs (par définition des λ_i dans (**??**)). Donc tous les y_i sont de même signe, ainsi, quitte à multiplier par -1, on les prend positifs. On vient donc de montrer qu'il existe $y \in \mathbb{N}^n$ non nul tel que $Ay = 0$. □

Remarques :

Je ne vois aucun exemple où on aurait besoin de cette application pour gagner du temps par rapport au temps qu'on prendrait à le faire à la main. Si tu en vois, n'hésite pas à m'envoyer un mail [5].

Astuces de l'agrégatif :

Il faut bien comprendre pourquoi

$$\mathrm{Conv}(A) = \left\{ \sum_{i=0}^{N} \lambda_i a_i, \text{ avec } N \in \mathbb{N}, a_i \in A, \lambda_i \geqslant 0, \sum_i \lambda_i = 1 \right\}$$

et ensuite ce qui est fort dans ce théorème c'est qu'on borne ce N par $n+1$ où n est la dimension de l'espace vectoriel dans lequel on est.

Questions possibles :

- Trouver un convexe de $\mathbb{R}^n$, qui n'est pas engendré par $n+1$ points.

  Réponse : le cercle

5. pierre.le-barbenchon@ens-rennes.fr