MT242, Cours n° 5, Lundi 21 Février 2000.

Exercice traité.

On considère la forme quadratique Q définie pour tout vecteur $v = (x, y, z) \in \mathbb{R}^3$ par

$$\mathrm{Q}(v) = 2(x - y + 2z)^2 - 3(y + z)^2.$$

On cherche le rang de Q. Comme on a une combinaison de deux carrés de formes linéaires, on est sûr que le rang est ≤ 2 (cette remarque n'a pas été faite jusqu'ici). Les deux formes linéaires sont

$$\ell_1(v) = x - y + 2z, \quad \ell_2(v) = y + z,$$

et $\mathrm{Q} = 2\ell_1^2 - 3\ell_2^2$. On voit que les deux formes linéaires sont indépendantes en écrivant leurs deux matrices ligne dans la base canonique. On cherche ensuite le noyau de Q. Puisque Q est combinaison linéaire à coefficients non nuls des carrés de deux formes linéaires indépendantes ℓ_1 et ℓ_2, on sait que

$$\mathrm{N}(\mathrm{Q}) = \{v \in \mathbb{R}^3 : \ell_1(v) = \ell_2(v) = 0\}.$$

La résolution donne $\mathrm{N}(\mathrm{Q}) = \{(-3t, -t, t) : t \in \mathbb{R}\}$.

On cherche maintenant une base β-orthogonale, où β désigne la forme polaire de Q,

$$\beta(v, v') = 2(x - y + 2z)(x' - y' + 2z') - 3(y + z)(y' + z'),$$

si $v' = (x', y', z')$. Pour ce faire on commence par compléter ℓ_1, ℓ_2 en une base du dual, ce qui revient à compléter la décomposition de Gauss avec des carrés indépendants à coefficients nuls. On a une infinité de choix. On prendra le plus simple,

$$\mathrm{Q}(v) = 2(x - y + 2z)^2 - 3(y + z)^2 + 0z^2,$$

qui correspond à $\ell_3(v) = z$, qui est bien indépendante des deux premières, comme on le constate en écrivant la matrice 3×3 dont les trois lignes sont les coordonnées des trois formes linéaires (une *ligne* par forme linéaire)

$$\mathrm{L} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

On obtiendra la base β-orthogonale correspondante en écrivant la matrice inverse de cette matrice, et en prenant les trois vecteurs v_1, v_2 et v_3 dont les coordonnées sont dans les colonnes de L^{-1}. Le calcul donne

$$\mathrm{L}^{-1} = \begin{pmatrix} 1 & 1 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

ce qui donne les vecteurs $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$ et $v_3 = (-3, -1, 1)$. Dans la base $\mathbf{v} = (v_1, v_2, v_3)$ on pourra vérifier que la matrice $\mathrm{B} = \mathrm{Mat}(\mathrm{Q}, \mathbf{v})$ est

$$\mathrm{B} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Sous cette forme il est facile d'identifier le noyau : il est engendré par les vecteurs v_j de la base β-orthogonale $\mathbf{v}$ tels que $\mathrm{Q}(v_j) = 0$, donc ici $\mathrm{N}(\mathrm{Q}) = \mathbb{R}v_3$, ce qui redonne bien le même résultat qu'au début de l'exercice...

Simplification des coefficients

Il s'agit ici de simplifier les coefficients des combinaisons linéaires de carrés obtenus dans la méthode de Gauss, ou bien, ce qui revient au même, de simplifier les coefficients que l'on peut obtenir dans la matrice B d'une forme quadratique dans une base β-orthogonale (la matrice B est diagonale dans ce cas ; on s'intéresse aux coefficients diagonaux).

Par exemple, on pourrait écrire la forme quadratique de l'exemple précédent sous la forme

$$\mathrm{Q}(v) = (\sqrt{2}\,x - \sqrt{2}\,y + 2\sqrt{2}\,z)^2 - (\sqrt{3}\,y + \sqrt{3}\,z)^2,$$

et on a maintenant $\mathrm{Q} = m_1^2 - m_2^2$, avec deux formes linéaires indépendantes m_1 et m_2,

$$m_1(v) = \sqrt{2}(x - y + 2z), \quad m_2(v) = \sqrt{3}(y + z).$$

On a un peu simplifié les coefficients de la décomposition de Gauss, en remplaçant 2 et -3 par 1 et -1. Mais si on travaille avec des réels, il est impossible de se débarrasser du signe $-$ qui vient avec le deuxième carré. Résumons :

Si $\mathbb{K} = \mathbb{C}$, toute forme quadratique sur un espace vectoriel E *de dimension finie peut s'écrire sous la forme*

$$\mathrm{Q} = \ell_1^2 + \cdots + \ell_r^2,$$

où $\ell_1, \ldots, \ell_r$ sont des formes linéaires sur E*, linéairement indépendantes. L'entier r est le rang de* Q*. Si β désigne la forme bilinéaire polaire de* Q*, on peut trouver une base β-orthogonale* $\mathbf{v}$ *de* E *telle que la matrice* $\mathrm{B} = \mathrm{Mat}(\mathrm{Q}, \mathbf{v})$ *soit diagonale, avec r coefficients diagonaux égaux à* 1 *et les autres nuls.*

Si $\mathbb{K} = \mathbb{R}$, toute forme quadratique sur un espace vectoriel E *de dimension finie peut s'écrire sous la forme*

$$\mathrm{Q} = \ell_1^2 + \cdots + \ell_s^2 - \ell_{s+1}^2 - \cdots - \ell_{s+t}^2,$$

où $\ell_1, \ldots, \ell_{s+t}$ sont des formes linéaires sur E*, linéairement indépendantes. L'entier $r = s + t$ est le rang de* Q*. Si β désigne la forme bilinéaire polaire de* Q*, on peut trouver une base β-orthogonale* $\mathbf{v}$ *de* E *telle que la matrice* $\mathrm{B} = \mathrm{Mat}(\mathrm{Q}, \mathbf{v})$ *soit diagonale, avec s coefficients diagonaux égaux à* 1*, t coefficients diagonaux égaux à* -1 *et les autres nuls. On verra plus loin que les nombres s et t ne dépendent que de* Q*, et pas de la base β-orthogonale choisie.*

Regardons un cas un peu exotique, celui du corps $\mathbb{K} = \mathbb{Z}/11\,\mathbb{Z}$ à 11 éléments $\overline{0}$, $\overline{1}$, $\overline{2}$, $\overline{3}$, $\overline{4}$, $\overline{5}$, $\overline{6}$, $\overline{7}$, $\overline{8}$, $\overline{9}$, $\overline{10}$. Les scalaires non nuls qui sont des carrés sont $\overline{1}$, $\overline{4}$, $\overline{9}$, $\overline{5}$, $\overline{3}$. Le scalaire $\overline{2}$ n'est pas un carré, et on constate que tout non-carré $x \in \mathbb{K}$ peut s'écrire $x = \overline{2}\,d^2$ pour un certain $d \in \mathbb{K}$ (on pourrait le démontrer).

Si $\mathrm{Q} = \sum_{i=1}^{n} c_i \ell_i^2$ est une forme quadratique sur un espace vectoriel E de dimension n sur ce corps $\mathbb{K}$, avec $c_i \in \mathbb{K}$, on peut remplacer tous les $c_i \ell_i^2$ avec c_i carré par m_i^2, et tous les $c_i \ell_i^2$ avec c_i non carré par $\overline{2} m_i^2$, où $m_i = d_i \ell_i$ est une forme linéaire qui est un multiple non nul de ℓ_i.

Exercice proposé. On peut aller un peu plus loin. Montrer que $\overline{2}\ell_1^2 + \overline{2}\ell_2^2$ peut s'écrire $m_1^2 + m_2^2$, avec m_1, m_2 indépendantes appartenant à $\mathrm{Vect}(\ell_1, \ell_2)$. Il en résulte qu'on peut pousser la simplification des coefficients en disant que s'il y a deux coefficients $\overline{2}$, on peut les remplacer par deux coefficients $\overline{1}$ (mais il faut aussi remplacer les deux formes linéaires dont le carré était multiplié par $\overline{2}$ par deux autres formes linéaires).

Si $\mathbb{K} = \mathbb{Z}/11\,\mathbb{Z}$, toute forme quadratique sur un espace vectoriel E de dimension finie sur $\mathbb{K}$ peut s'écrire sous la forme

$$\mathrm{Q} = c_1 \ell_1^2 + \cdots + c_r \ell_r^2,$$

où $\ell_1, \ldots, \ell_r$ sont des formes linéaires sur E, linéairement indépendantes, et où les coefficients (c_i) sont égaux à $\overline{1}$ ou bien $\overline{2}$, avec au plus **un** *coefficient égal à $\overline{2}$.*

Cas réel : étude du signe

L'étude du signe d'une forme quadratique réelle reviendra en calcul différentiel, dans l'étude des extrema locaux.

Lemme 1.3.2. *Soient Q une forme quadratique sur un espace vectoriel E réel et F un sous-espace vectoriel de E, de dimension finie k.*

On a $\mathrm{Q}(v) \ge 0$ pour tout $v \in \mathrm{F}$ si et seulement s'il existe une base β-orthogonale $v_1, \ldots, v_k$ de F telle que $\mathrm{Q}(v_j) \ge 0$ pour tout $j = 1, \ldots, k$.

On a $\mathrm{Q}(v) > 0$ pour tout $v \in \mathrm{F}\setminus\{0\}$ si et seulement s'il existe une base β-orthogonale $v_1, \ldots, v_k$ de F telle que $\mathrm{Q}(v_j) > 0$ pour tout $j = 1, \ldots, k$.

Démonstration. On va traiter le premier point. Si on suppose que $\mathrm{Q}(v) \ge 0$ pour tout vecteur $v \in \mathrm{F}$, on aura évidemment $\mathrm{Q}(v_j) \ge 0$ pour tous les vecteurs v_j, $j = 1, \ldots, k$ d'une base β-orthogonale quelconque de F.

Inversement, si on a $\mathrm{Q}(v_j) \ge 0$ pour tous les vecteurs v_j, $j = 1, \ldots, k$ d'une base β-orthogonale de F, on considère un vecteur $v \in \mathrm{F}$ quelconque, on écrit $v = x_1 v_1 + \cdots + x_k v_k$, avec $x_i \in \mathbb{R}$, puis

$$\mathrm{Q}(v) = \beta(v, v) = \sum_{i,j=1}^{k} x_i x_j \, \beta(v_i, v_j) = \sum_{i=1}^{k} x_i^2 \, \beta(v_i, v_i) = \sum_{i=1}^{k} x_i^2 \, \mathrm{Q}(v_i) \ge 0.$$

Signature d'une forme quadratique réelle

On considère un espace vectoriel réel E de dimension finie et une forme quadratique Q sur E. On désigne par β la forme bilinéaire polaire de Q.

On dira que Q est de *signature* (s, t) s'il existe une base β-orthogonale $(e_1, \ldots, e_n)$ de E qui contient s vecteurs tels que $\mathrm{Q}(e_i) > 0$ et t vecteurs tels que $\mathrm{Q}(e_i) < 0$.

En termes de décomposition de Gauss, cela équivaut à dire que la décomposition de Gauss $\sum_{i=1}^{n} c_i \ell_i^2$ en combinaison de carrés de n formes linéaires indépendantes contient s termes tels que $c_i > 0$ (*somme* de carrés) et t termes tels que $c_i < 0$ (*différence* de carrés). Cette définition n'a de sens que si on montre que ces deux nombres s et t ne dépendent que de Q, et pas de la base β-orthogonale particulière (ou de la décomposition de Gauss particulière).

Supposons donc que $(e_1, \ldots, e_n)$ soit une base β-orthogonale de E, avec $\beta(e_i, e_i) > 0$ pour $i = 1, \ldots, s$ et $\beta(e_i, e_i) \le 0$ pour $i > s$. D'après le lemme 1.3.2 appliqué à $-\mathrm{Q}$, on voit que $\mathrm{Q}(v) \le 0$ pour tout vecteur $v \in \mathrm{G} = \mathrm{Vect}(e_{s+1}, \ldots, e_n)$. Soit $\mathbf{e}' = (e'_1, \ldots, e'_n)$ une autre base β-orthogonale, et supposons que maintenant $\beta(e'_i, e'_i) > 0$ pour $i = 1, \ldots, s'$, et $\beta(e'_i, e'_i) \le 0$ pour $i > s'$. D'après le lemme 1.3.2, on a $\mathrm{Q}(v) > 0$ pour tout vecteur de $\mathrm{F}' \setminus \{0\}$, où $\mathrm{F}' = \mathrm{Vect}(e'_1, \ldots, e'_{s'})$. Il en résulte que $\mathrm{F}' \cap \mathrm{G} = \{0\}$ (en effet, si $v \in \mathrm{F}'$ et $v \ne 0_{\mathrm{E}}$, on a $\mathrm{Q}(v) > 0$, donc $v \notin \mathrm{G}$). On en déduit que $\dim \mathrm{F}' + \dim \mathrm{G} \le \dim \mathrm{E} = n$, donc $s' + (n - s) \le n$, ce qui donne $s' \le s$. En échangeant les rôles des bases $\mathbf{e}$ et $\mathbf{e}'$, on obtient aussi $s \le s'$, donc $s = s'$. Le raisonnement est le même pour montrer que $t = t'$.

Définition. On dit que la forme quadratique Q définie sur l'espace vectoriel réel de dimension finie E, et de forme polaire β, est *de signature* (s, t) si sa matrice dans une base β-orthogonale contient s coefficients > 0 et t coefficients < 0 sur la diagonale.

Définition. Soit Q une forme quadratique sur un espace vectoriel réel E et soit β sa forme polaire ; on dit que β (ou Q) est *positive* sur E si $\mathrm{Q}(v) \geq 0$ pour tout vecteur $v \in \mathrm{E}$. On dit que β (ou Q) est *définie positive* sur E si $\mathrm{Q}(v) > 0$ pour tout vecteur $v \in \mathrm{E}$ non nul.

Remarque. La forme quadratique Q sur l'espace vectoriel réel E de dimension n est définie positive si et seulement si sa signature est $(n, 0)$.

En effet, si $(e_1, \ldots, e_n)$ est une base β-orthogonale de E, on aura $\mathrm{Q}(e_i) > 0$ pour tout $i = 1, \ldots, n$ lorsque la signature est $(n, 0)$; d'après le lemme 1.3.2 appliqué avec $\mathrm{F} = \mathrm{E}$ il en résulte que Q est définie positive sur E. L'implication inverse est évidente : si Q est définie positive et si $(e_1, \ldots, e_n)$ est une base β-orthogonale, on aura $\mathrm{Q}(e_i) > 0$ pour tout $i = 1, \ldots, n$, donc la signature est $(n, 0)$.

Equivalence de formes quadratiques

Dans ce paragraphe on se replace dans le cas d'un corps $\mathbb{K}$ quelconque (mais tel que $1_{\mathbb{K}} + 1_{\mathbb{K}} \neq 0$).

On dit que deux formes quadratiques Q_1 et Q_2 sur un $\mathbb{K}$-espace vectoriel E sont équivalentes s'il existe un automorphisme u de E tel que $\mathrm{Q}_2(v) = \mathrm{Q}_1(u(v))$ pour tout $v \in \mathrm{E}$. Si $\mathrm{Q}_1 = \sum c_j (\ell_j)^2$, alors $\mathrm{Q}_2 = \sum c_j (\ell_j \circ u)^2$, et les formes $\ell_j \circ u$ sont linéairement indépendantes. Il en résulte que Q_1 et Q_2 ont le même rang.

Lemme 1.3.3. *Si $\ell_1, \ldots, \ell_n$ et $m_1, \ldots, m_n$ sont deux bases du dual E^*, il existe un automorphisme u de E tel que $m_i = \ell_i \circ u$ pour tout $i = 1, \ldots, n$.*

Démonstration. Il existe des vecteurs (v_j) tels que $\ell_i(v_j) = \delta_{i,j}$ et des vecteurs (w_j) tels que $m_i(w_j) = \delta_{i,j}$. On sait que $(v_j)_{j=1}^n$ et $(w_j)_{j=1}^n$ sont deux bases de E. On peut donc définir un endomorphisme u de E en posant $u(w_j) = v_j$ pour $j = 1, \ldots, n$. Comme (v_j) est aussi une base de E, on en déduit que u est inversible. Alors pour chaque $i = 1, \ldots, n$, on a

$$(\ell_i \circ u)(w_j) = \ell_i(u(w_j)) = \ell_i(v_j) = \delta_{i,j} = m_i(w_j)$$

pour tout $j = 1, \ldots, n$, donc $m_i = \ell_i \circ u$ puisque ces deux applications linéaires coïncident sur une base de E.

Proposition. *Si $\ell_1, \ldots, \ell_n$ et $m_1, \ldots, m_n$ sont deux bases du dual E^*, si $\mathrm{Q}_1 = \sum_{i=1}^n c_i \ell_i^2$ et $\mathrm{Q}_2 = \sum_{i=1}^n c_i m_i^2$, avec les mêmes coefficients (c_i), alors Q_1 et Q_2 sont équivalentes.*

Démonstration. D'après le lemme 1.3.3, il existe un automorphisme u de E tel que $m_i = \ell_i \circ u$ pour tout $i = 1, \ldots, n$. il en résulte immédiatement que $\mathrm{Q}_2 = \sum_{i=1}^n c_i m_i^2 = \sum_{i=1}^n c_i (\ell_i \circ u)^2 = \mathrm{Q}_1 \circ u$.

Il en résulte que :

dans le cas complexe, deux formes quadratiques sont équivalentes si et seulement si elles ont le même rang ;

dans le cas réel, deux formes quadratiques sont équivalentes si et seulement si elles ont la même signature.

Formes bilinéaires positives. Inégalité de Cauchy-Schwarz

Soit E un espace vectoriel réel et soit β une forme bilinéaire symétrique positive sur l'espace E ; soient $v, w \in \mathrm{E}$; on a

$$\beta(tv + w, tv + w) \ge 0,$$

pour tout nombre réel t. On a donc un trinôme réel en la variable t qui ne change pas de signe quand t décrit $\mathbb{R}$,

$$t^2\beta(v, v) + 2t\,\beta(v, w) + \beta(w, w) \ge 0,$$

donc son discriminant est ≤ 0, ce qui donne l'inégalité de Cauchy-Schwarz

$$\big(\beta(v, w)\big)^2 \le \beta(v, v)\,\beta(w, w).$$

Proposition. Inégalité de Cauchy-Schwarz. *Soit* E *un espace vectoriel réel et soit* β *une forme bilinéaire symétrique positive sur l'espace* E *; on a*

$$\forall v, w \in \mathrm{E}, \quad \big|\beta(v, w)\big| \le \sqrt{\beta(v, v)}\,\sqrt{\beta(w, w)}.$$

Exemple. Considérons la forme bilinéaire β définie sur l'espace vectoriel réel E des fonctions réelles continues sur $[a, b]$ (avec $a < b$) par la formule

$$\beta(f, g) = \int_a^b f(t)g(t)\,dt$$

pour toutes fonctions réelles continues f et g. Il est clair que cette forme bilinéaire est positive, donc on peut lui appliquer l'inégalité de Cauchy-Schwarz. On obtient ainsi l'inégalité

$$\Big|\int_a^b f(t)g(t)\,dt\Big| \le \Big(\int_a^b f(t)^2\,dt\Big)^{1/2}\Big(\int_a^b g(t)^2\,dt\Big)^{1/2}.$$

1.4. Produit scalaire et normes euclidiennes

Définition. Un *espace euclidien* est un espace vectoriel réel E **de dimension finie** muni d'une forme bilinéaire symétrique β définie positive.

On dit que β définit le *produit scalaire* de l'espace euclidien E, et on notera en général le produit scalaire de deux vecteurs v et w de E par

$$v \,.\, w = \beta(v, w).$$

En termes précis, un espace euclidien n'est pas seulement un espace vectoriel, mais un couple (E, β) d'un espace vectoriel E et d'une forme bilinéaire symétrique β définie sur $\mathrm{E} \times \mathrm{E}$.

Exemple.

L'exemple le plus commun sera $\mathbb{R}^n$ avec son produit scalaire usuel, défini par

$$\beta(v, w) = v \,.\, w = x_1\,y_1 + \cdots + x_n\,y_n$$

pour deux vecteurs quelconques $v = (x_1, \ldots, x_n)$ et $w = (y_1, \ldots, y_n)$ de $\mathbb{R}^n$.

Pour un espace euclidien E fixé, on notera $\|v\| = \sqrt{v \,.\, v}$ et l'inégalité de Cauchy-Schwarz donne dans ce cas

$$\bigl|v \,.\, w\bigr| \le \|v\| \, \|w\|.$$

Définition. On suppose $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et on suppose que E est un espace vectoriel sur $\mathbb{K}$. On dit que la fonction $\mathrm{N} : \mathrm{E} \to [0, +\infty[$ est une norme sur E si elle vérifie les trois propriétés suivantes.

(i) Pour tous $v, w \in \mathrm{E}$, on a $\mathrm{N}(v + w) \le \mathrm{N}(v) + \mathrm{N}(w)$.
(ii) Pour tout $\lambda \in \mathbb{K}$ et tout $v \in \mathrm{E}$, on a $\mathrm{N}(\lambda v) = |\lambda| \, \mathrm{N}(v)$.
(iii) On a $\mathrm{N}(v) = 0$ si et seulement si $v = 0_{\mathrm{E}}$.

Lemme 1.4.1. *Soit* E *un espace euclidien ; l'application* $v \to \|v\|$ *est une norme sur* E.

Démonstration. On voit que $\|v\| = 0$ signifie que $v \,.\, v = \beta(v, v) = 0$, donc $v = 0_{\mathrm{E}}$ puisque β est définie positive. A partir de l'inégalité de Cauchy-Schwarz, on voit facilement que

$$\|v + w\| \le \|v\| + \|w\|$$

(inégalité triangulaire) ; en effet, on obtient en développant le carré scalaire de $v + w$ et en utilisant la symétrie ($v \,.\, w = w \,.\, v$)

$$\|v + w\|^2 = (v + w) \,.\, (v + w) = v \,.\, v + 2\, v \,.\, w + w \,.\, w \le$$

$$v \,.\, v + 2 \sqrt{v \,.\, v} \sqrt{w \,.\, w} + w \,.\, w = (\sqrt{v \,.\, v} + \sqrt{w \,.\, w})^2 = (\|v\| + \|w\|)^2.$$

Base orthonormée

Soit E un espace euclidien ; on dit qu'un système $(v_1, \ldots, v_n)$ de vecteurs de E est une **base orthonormée** de E si c'est une base de E et si les vecteurs (v_i) sont deux à deux orthogonaux et de norme 1,

$$\forall i \ne j, \quad v_i \,.\, v_j = 0 \,; \quad \forall i = 1, \ldots, n, \ \|v_i\| = 1.$$

Théorème. *Tout espace euclidien admet une base orthonormée.*

Démonstration. On a essentiellement déjà vu ce résultat dans ce chapitre.

Cours n° 6, Mercredi 23 Février 2000.

Chapitre 2. Topologie

2.1. Bolzano-Weierstrass

Une *suite* d'éléments d'un ensemble non vide X est une application $\xi : \mathbb{N} \to \mathrm{X}$, c'est à dire une correspondance $n \in \mathbb{N} \to \xi(n) \in \mathrm{X}$. Dans la pratique on utilise rarement la notation de fonction pour représenter une suite. On posera plutôt $x_n = \xi(n)$ pour tout $n \ge 0$ et on parlera de la suite $(x_n)_{n \ge 0}$; on notera parfois simplement $(x_n)_n$ ou (x_n). La variable n est une variable muette qui peut être remplacée par une autre, par exemple $(x_k)_{k \ge 0}$ représente la même suite.

Une sous-suite de la suite (x_n) est une nouvelle suite $k \in \mathbb{N} \to y_k \in \mathrm{X}$ obtenue de la façon suivante : on se donne une suite **strictement croissante** d'entiers $n_0 < n_1 < \ldots < n_k < \ldots$ et on pose $y_k = x_{n_k}$ pour tout $k \geq 0$. Dans la pratique on notera aussi la sous-suite $(y_k)_k$ par $(x_{n_k})_k$.

Exemple. Considérons la suite de nombres réels définie par $x_n = (-1)^n$ pour tout entier $n \geq 0$. Posons $n_k = 2k$ pour tout $k \geq 0$. On obtient ainsi une suite strictement croissante d'entiers, qui permet de considérer la sous-suite $(x_{n_k})_k = (x_{2k})_k$ de la suite donnée. On voit que pour tout $k \geq 0$, on a $x_{2k} = (-1)^{2k} = 1$. On a ainsi trouvé une sous-suite de la suite de départ, qui est une sous-suite convergente (en fait une sous-suite constante) alors que la suite de départ ne converge pas. Le théorème de Bolzano-Weierstrass ira dans cette direction qui consiste à chercher des sous-suites convergentes pour une suite donnée.

Les deux propositions qui suivent sont censées être des rappels de première année. Elles sont données sans démonstration.

Proposition 2.1.1. *Si une suite (x_n) de nombres réels tend vers une limite $\ell \in \mathbb{R}$, toute sous-suite (x_{n_k}) converge aussi vers ℓ. Le résultat est vrai aussi dans le cas d'une limite infinie.*

Proposition 2.1.2. *Si $f : [a,b] \to \mathbb{R}$ est une fonction continue au point $y_0 \in [a,b]$, si (x_n) est une suite de nombres réels de $[a,b]$ qui tend vers y_0, alors la suite $(f(x_n))$ tend vers $f(y_0)$.*

Théorème 2.1.1. Théorème de Bolzano-Weierstrass. *Toute suite bornée de nombres réels admet des sous-suites convergentes vers un nombre réel.*

Ce résultat n'est évidemment pas possible pour une suite non-bornée. Par exemple, la suite $x_n = n$ n'est pas convergente dans $\mathbb{R}$; on peut dire qu'elle converge vers $+\infty$, qui n'est pas un nombre réel. Toute sous-suite de cette suite converge aussi vers $+\infty$, ce qui montre qu'aucune sous-suite ne peut converger vers une limite réelle.

Démonstration. Soit (x_n) une suite bornée de nombres réels, disons $|x_n| \leq \mathrm{M}$ pour tout $n \geq 0$; on va d'abord trouver un nombre y_0 qui soit candidat raisonnable pour être limite d'une sous-suite. Définissons une fonction f sur $\mathbb{R}$ de la façon suivante : supposons donné $y \in \mathbb{R}$; s'il n'y a qu'un nombre fini d'indices n tels que $x_n \leq y$, on pose $f(y) = 0$, sinon, s'il existe une infinité d'indices n tels que $x_n \leq y$, on pose $f(y) = 1$. Il est facile de voir que $f(-\mathrm{M}-1) = 0$ (parce qu'il n'y a aucun terme de la suite tel que $x_n < -\mathrm{M}$) et $f(\mathrm{M}) = 1$ (parce que tous les termes de la suite sont $\leq \mathrm{M}$). De plus la fonction f est croissante, et ne prend que les deux valeurs 0 et 1. Il y a nécessairement un endroit précis où elle change de valeur, c'est à dire un y_0 tel que, pour tout $\varepsilon > 0$, on ait $f(y_0 - \varepsilon) = 0$ et $f(y_0 + \varepsilon) = 1$. On trouve y_0 en disant que y_0 est la borne supérieure de l'ensemble non vide et majoré $\mathrm{A} \subset \mathbb{R}$ formé de tous les réels y tels que $f(y) = 0$. Puisque $f(y_0 - \varepsilon) = 0$ et $f(y_0 + \varepsilon) = 1$, il y a un ensemble infini $\mathrm{B} \subset \mathbb{N}$ d'indices n tels que $x_n \leq y_0 + \varepsilon$, mais seulement un ensemble fini C d'indices n tels que $x_n \leq y_0 - \varepsilon$; on en déduit que l'ensemble différence $\mathrm{B} \setminus \mathrm{C}$ est infini. Pour tout $\varepsilon > 0$, on a donc la propriété

$\mathrm{P}(\varepsilon)$: il existe une infinité d'indices n tels que $y_0 - \varepsilon < x_n \leq y_0 + \varepsilon$.

Montrons maintenant qu'on peut trouver une sous-suite (x_{n_k}) qui converge vers y_0. On applique la propriété $\mathrm{P}(\varepsilon)$ avec des valeurs successives $1, 1/2, 1/4, \ldots$, c'est à dire $\varepsilon = 2^{-k}$ pour $k = 0, 1, \ldots$ En appliquant $\mathrm{P}(1)$ on peut d'abord choisir un entier n_0 tel que $y_0 - 1 < x_{n_0} \leq y_0 + 1$; en supposant que $n_0 < n_1 < \ldots < n_{k-1}$ sont déjà choisis, on sait d'après $\mathrm{P}(2^{-k})$ qu'il existe une infinité d'indices n tels que $y_0 - 2^{-k} < x_n \leq y_0 + 2^{-k}$. On peut donc en trouver au moins un, qu'on choisira comme n_k, qui soit $> n_{k-1}$. On aura donc $|x_{n_k} - y_0| \leq 2^{-k}$ pour tout entier $k \geq 0$, ce qui montre que la sous-suite (x_{n_k}) converge vers y_0.

Rappel des propriétés des fonctions continues sur un fermé borné de $\mathbb{R}$ (avec Bolzano)

Lorsque f est une fonction réelle *minorée* définie sur un ensemble non vide X, on peut considérer la borne inférieure $\inf f(\mathrm{X})$ de l'ensemble $f(\mathrm{X})$ de ses valeurs sur X, et on peut toujours trouver une suite (x_n) de points de X telle que

$$\inf f(\mathrm{X}) = \lim_n f(x_n).$$

En effet, pour tout $n \geq 1$ on peut trouver par définition de la borne inférieure une valeur $f(x_n)$ de f en un certain point $x_n \in \mathrm{X}$ telle que $\inf f(\mathrm{X}) \leq f(x_n) < \inf f(\mathrm{X}) + 1/n$, donc la suite des valeurs $f(x_n)$ converge vers $\inf f(\mathrm{X})$. Mais on ne peut pas en général trouver un point $x \in \mathrm{X}$ tel que $f(x) = \inf f(\mathrm{X})$. Si un tel point x existe, il est clair que $f(x)$ est alors *la plus petite valeur de f* sur l'ensemble X, et on convient de noter dans ce cas

$$f(x) = \min f(\mathrm{X}).$$

(et de façon analogue, on note $f(y) = \max f(\mathrm{X})$ si f atteint son maximum sur X au point $y \in \mathrm{X}$). Nous allons voir que tel est le cas pour les fonctions continues définies sur un intervalle fermé borné non vide de $\mathbb{R}$.

Théorème 2.1.2. *Si f est une fonction réelle définie et continue en tout point d'un intervalle fermé borné $[a, b] \subset \mathbb{R}$, (avec $a < b$), elle est bornée et elle atteint son maximum et son minimum sur $[a, b]$.*

Démonstration. Montrons d'abord que f est bornée sur $[a, b]$. Sinon, on pourrait trouver pour tout entier $n \geq 0$ un point $x_n \in [a, b]$ tel que $|f(x_n)| > n$. Par Bolzano, on trouve une sous-suite (x_{n_k}) convergente vers un certain point $y_0 \in \mathbb{R}$, et on a $y_0 \in [a, b]$ puisque $a \leq x_{n_k} \leq b$ et $x_{n_k} \to y_0$ (passage à la limite des inégalités larges). On devrait alors avoir $|f(x_{n_k})| \to |f(y_0)|$ par continuité de la fonction f au point y_0, mais $|f(x_{n_k})| > n_k$ tend vers l'infini, ce qui est contradictoire. On en déduit que f est bornée sur $[a, b]$.

L'ensemble $f([a, b])$ des valeurs de f sur $[a, b]$ est non vide (puisque $[a, b]$ est non vide) et borné, donc on peut définir sa borne inférieure $m = \inf f([a, b])$ et on peut trouver une suite de points $x_n \in [a, b]$ telle que $(f(x_n))$ converge vers m. D'après le théorème 2.1.1 (de Bolzano-Weierstrass), il existe une sous-suite (x_{n_k}) qui converge vers un point $y_0 \in \mathbb{R}$. Puisque $x_{n_k} \in [a, b]$ pour tout k, on obtient $y_0 \in [a, b]$ comme avant, donc y_0 est bien dans l'ensemble de définition de f. Puisque f est continue au point y_0, on doit avoir $f(y_0) = \lim_k f(x_{n_k}) = m$, c'est à dire que la fonction f atteint son minimum sur $[a, b]$ au point y_0. La même démonstration s'applique pour trouver un point $y_1 \in [a, b]$ où le maximum de f est atteint.

Bolzano-Weierstrass multi-dimensionnel

Passons au cas de $\mathbb{R}^d$, $d \geq 2$. Si on a une suite (v_n) de vecteurs de $\mathbb{R}^d$, avec $v_n = (x_{n,1}, \ldots, x_{n,d})$ pour tout $n \geq 0$, on dit que (v_n) converge vers le vecteur $v = (x_1, \ldots, x_d)$ si chaque coordonnée de v_n converge vers la coordonnée correspondante de v,

$$\forall i = 1, \ldots, d, \quad x_i = \lim_n x_{n,i}.$$

On dit que la suite (v_n) est bornée si toutes les suites coordonnées $(x_{n,i})_n$ sont bornées,

$$\exists \mathrm{M},\ \forall i = 1, \ldots, d,\ \forall n \geq 0, \quad |x_{n,i}| \leq \mathrm{M}.$$

Théorème 2.1.3. Bolzano-Weierstrass vectoriel. *De toute suite bornée (v_n) dans $\mathbb{R}^d$ on peut extraire une sous-suite convergente.*

Démonstration. Donnons-la, pour simplifier, dans le cas de $\mathbb{R}^2$. Ecrivons $v_n = (x_n, y_n) \in \mathbb{R}^2$ pour tout n. Puisque la suite (v_n) est bornée, il en résulte que les deux suites réelles (x_n) et (y_n) sont bornées. D'après Bolzano-Weierstrass en une dimension (théorème 2.1.1), on peut d'abord extraire une sous-suite convergente (x_{n_k}) de la suite (x_n), de limite $x \in \mathbb{R}$. On considère maintenant la suite bornée $(y_{n_k})_k$, et on peut en extraire une sous-suite convergente $(y_{n_{k_j}})_j$, sous-suite de limite $y \in \mathbb{R}$. La sous-suite $(x_{n_{k_j}})_j$ reste convergente vers x, et au total la sous-suite vectorielle $(v_{n_{k_j}})_j = (x_{n_{k_j}}, y_{n_{k_j}})_j$ converge dans $\mathbb{R}^2$ vers $v = (x, y) \in \mathbb{R}^2$.

Equivalence de normes

Définition. On dit que deux normes N_1 et N_2 sur un espace vectoriel réel ou complexe E sont *équivalentes* s'il existe une constante K telle que $\mathrm{N}_2 \leq \mathrm{K}\,\mathrm{N}_1$ et $\mathrm{N}_1 \leq \mathrm{K}\,\mathrm{N}_2$.

Exemples simples de normes sur $\mathbb{R}^d$. Pour tout vecteur $v = (x_1, \ldots, x_d) \in \mathbb{R}^d$, considérons les trois quantités

$$\|v\|_1 = \sum_{i=1}^{d} |x_i|; \quad \|v\|_2 = \Big(\sum_{i=1}^{d} |x_i|^2\Big)^{1/2}; \quad \|v\|_\infty = \max_{1 \leq i \leq d} |x_i|.$$

Il est facile de vérifier que $v \to \|v\|_1$ et $v \to \|v\|_\infty$ sont des normes sur $\mathbb{R}^d$. On a vu dans le paragraphe “espaces euclidiens” que $v \to \|v\|_2$ est une norme.

On peut montrer facilement que ces trois normes sont équivalentes sur $\mathbb{R}^d$. On a

$$\|v\|_\infty \leq \|v\|_2 \leq \|v\|_1; \quad \|v\|_1 \leq d\,\|v\|_\infty.$$