cours 4, le lundi 8 février 2010

Intersection de tribus

Si, sur un même ensemble X, on dispose d'une famille $(\mathcal{F}_i)_{i\in\mathrm{I}}$ de tribus, on peut considérer la famille $\mathcal{F}$ des parties $\mathrm{A} \subset \mathrm{X}$ qui appartiennent à toutes ces tribus. En tant que sous-ensembles $\mathcal{F}_i \subset \mathcal{P}(\mathrm{X})$, cela s'écrit

$$\mathcal{F} = \bigcap_{i\in\mathrm{I}} \mathcal{F}_i.$$

Il est facile de montrer que l'intersection d'une famille de tribus sur un ensemble X est encore une tribu de parties de X.

Tribu engendrée par une classe de parties $\mathcal{C} \subset \mathcal{P}(\mathrm{X})$

Si on considère une classe quelconque $\mathcal{C}$ de parties de X (par exemple : la classe $\mathcal{O}$ de tous les ouverts de $\mathrm{X} = \mathbb{R}$), il existe au moins une tribu qui contient cette classe, à savoir la tribu $\mathcal{P}(\mathrm{X})$, qui est la plus grosse tribu sur X. On peut ensuite considérer (les axiomes de la théorie des ensembles le permettent) l'intersection de toutes les tribus de parties de X qui contiennent la classe donnée $\mathcal{C}$; on obtient ainsi une tribu, qui est clairement la plus petite tribu contenant la classe $\mathcal{C}$. On l'appelle *la tribu engendrée par $\mathcal{C}$*, et on la note $\sigma(\mathcal{C})$.

Si X est un espace topologique avec une famille d'ouverts $\mathcal{O}$, la *tribu borélienne* de X est la tribu $\sigma(\mathcal{O})$ engendrée par les ouverts de X.

Un cas particulièrement important sera la tribu borélienne $\mathcal{B}_{\mathbb{R}}$ de la droite réelle $\mathbb{R}$ (ou bien celle de la droite étendue $\overline{\mathbb{R}}$) ; la mesure de Lebesgue λ est une mesure σ-additive sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$: selon le programme officiel de cet enseignement, son existence est admise.

On peut décrire dans des termes similaires le sous-espace vectoriel engendré par une partie C d'un espace vectoriel E, comme intersection de tous les sous-espaces qui contiennent C. Il s'agit d'une description « par l'extérieur ». Mais dans le cas vectoriel, on dispose aussi d'une description très simple « par l'intérieur », en considérant l'ensemble des combinaisons linéaires de vecteurs de C. On décrit ainsi, en un seul coup, la construction de tous les éléments qui seront, nécessairement, dans tout sous-espace vectoriel contenant C, et, coup de chance, on a ainsi directement un sous-espace vectoriel, qui est donc le plus petit, le sous-espace vectoriel engendré.

La description de $\mathcal{B}_{\mathbb{R}}$ « par l'intérieur » est possible mais délicate : si une tribu contient la classe $\mathcal{O} = \mathcal{C}_0$ des ouverts, elle doit aussi contenir les $\mathrm{G}_\delta = \mathcal{C}_1$, qui sont les intersections dénombrables d'ouverts (une classe qui, entre parenthèses, contient les fermés de $\mathbb{R}$), puis la tribu doit contenir les réunions dénombrables d'ensembles de la classe précédente, formant ainsi les $\mathrm{G}_{\delta\sigma} = \mathcal{C}_2$, puis les $\mathrm{G}_{\delta\sigma\delta} = \mathcal{C}_3$, etc. Le grand malheur est que, même si on poursuit une suite infinie d'opérations donnant les classes successives $\mathcal{C}_0, \ldots, \mathcal{C}_n, \ldots$, obtenant ainsi une classe

$$\mathcal{C}_\omega = \bigcup_{n=0}^{+\infty} \mathcal{C}_n,$$

qui regroupe tous les ensembles obtenus par une itération finie arbitraire des opérations σ et δ, cette classe ne sera toujours pas une tribu : il faut passer à $\mathcal{C}_{\omega+1}$, etc.

Le réconfort, c'est que du point de vue de la mesure, la classe $G_{\delta\sigma}$ suffit : tout ensemble borélien $B \in \mathcal{B}$ peut être encadré par deux ensembles $A_0 \subset B \subset A_1$ qui sont des $G_{\delta\sigma}$ tels que

$$\lambda(B \setminus A_0) = \lambda(A_1 \setminus B) = 0.$$

Découpage en couronnes

Si (B_n) est une suite croissante d'éléments de la tribu $\mathcal{F}$, posons

$$A_0 = B_0 \quad \text{et} \quad A_{n+1} = B_{n+1} \setminus B_n$$

pour tout $n \geq 0$. On voit que les $(A_n) \subset \mathcal{F}$ sont deux à deux disjoints, que $A_0, \ldots, A_n$ est une partition de B_n, donc

$$\mu(B_n) = \sum_{k=0}^{n} \mu(A_k),$$

et on voit que

$$\bigcup_{n \geqslant 0} B_n = \bigcup_{n \geqslant 0} A_n.$$

Il en résulte que : *pour toute suite croissante* (B_n) *d'ensembles de la tribu* $\mathcal{F}$*, on a la* propriété de monotonie

$$\mu\Big(\bigcup_{n \geqslant 0} B_n\Big) = \lim_n \mu(B_n).$$

Cette propriété est le degré 0 du *théorème de convergence monotone*, qui sera vu au prochain cours.

Si (C_n) est une suite quelconque dans $\mathcal{F}$, la suite (B_n) définie par

$$B_n = C_0 \cup \ldots \cup C_n$$

est une suite croissante dans $\mathcal{F}$, qui a la même réunion $\bigcup_n B_n$ que la suite (C_n), donc par monotonie de la mesure

$$\mu\Big(\bigcup_{n \geqslant 0} B_n\Big) = \lim_n \mu(B_n) \leq \lim_n \sum_{k=0}^{n} \mu(C_k) = \sum_{k \geqslant 0} \mu(C_k).$$

Comme $\bigcup_n B_n = \bigcup_{n \geqslant 0} C_n$, on obtient la propriété de *sous-additivité dénombrable*,

$$\mu\Big(\bigcup_{n \geqslant 0} C_n\Big) \leq \sum_{k \geqslant 0} \mu(C_k).$$

Un cas particulier très important est le cas des ensembles de mesure nulle : *si* $(A_i) \subset \mathcal{F}$ *est une famille finie ou dénombrable d'ensembles de mesure nulle, alors leur réunion est de mesure nulle,*

$$\mu\Big(\bigcup_{i \in I} A_i\Big) = 0.$$

Bien sûr, ce principe ne s'applique pas aux réunions qui ne sont pas dénombrables : l'intervalle $[0, 1]$ est de mesure 1 pour λ, et il est réunion des singletons $\{x\}$, $x \in [0, 1]$, qui sont des ensembles de mesure nulle.

Fonctions mesurables réelles

Si f est une fonction sur X à valeurs dans $\overline{\mathbb{R}}$ et c un nombre réel, on notera

$$\{f > c\} = \{x \in \mathrm{X} : f(x) > c\}\,;$$

on notera aussi $\{f \le 1\}$, $\{0 \le f < 2\}$, etc. Si B est un sous-ensemble de $\overline{\mathbb{R}}$ on notera

$$\{f \in \mathrm{B}\} = \{x \in \mathrm{X} : f(x) \in \mathrm{B}\} = f^{-1}(\mathrm{B}).$$

Définition. On dit que f est $\mathcal{F}$-mesurable si $\{f > c\}$ est dans la tribu $\mathcal{F}$ pour tout c réel.

Dans ce cas, les ensembles $\{f \ge c\}$, $\{f < c\}$, $\{f \le c\}$, $\{a \le f < b\}$ sont tous dans la tribu $\mathcal{F}$: par exemple

$$\{f \ge c\} = \bigcap_n \{f > c - 2^{-n}\}$$

est aussi dans $\mathcal{F}$; on obtient $\{f < c\}$ en passant au complémentaire, $\{a \le f < b\}$ en intersectant deux ensembles de l'une des formes précédentes. En fait, on verra plus loin que $\{f \in \mathrm{B}\}$ est dans $\mathcal{F}$ pour tout borélien B de $\overline{\mathbb{R}}$, quand f est $\mathcal{F}$-mesurable.

Exemples.

— Si f est $\mathcal{F}$-mesurable, la fonction $-f : x \in \mathrm{X} \to -f(x)$ est $\mathcal{F}$-mesurable aussi.

— Si f est monotone sur $[a, b]$ et si $\mathcal{B}$ désigne la tribu borélienne de $[a, b]$, alors f est $\mathcal{B}$-mesurable.

— Si $(\mathrm{X}, \mathcal{F})$ est un espace mesurable, toute fonction f de la forme $f = \mathbf{1}_{\mathrm{A}}$, $\mathrm{A} \in \mathcal{F}$, est $\mathcal{F}$-mesurable. En effet, l'ensemble $\{\mathbf{1}_{\mathrm{A}} > c\}$ est égal à X, à A ou à $\emptyset$, selon que $c < 0$, $0 \le c < 1$ ou $c \ge 1$, trois possibilités qui toutes, donnent des ensembles de $\mathcal{F}$.

Plus généralement, une fonction φ sur X est une *fonction $\mathcal{F}$-étagée* s'il existe une partition finie de l'espace X en ensembles $\mathrm{A}_1, \ldots, \mathrm{A}_n$ de $\mathcal{F}$, et des nombres réels a_j tels que

$$\varphi = \sum_{j=1}^{n} a_j \mathbf{1}_{\mathrm{A}_j}.$$

Les fonctions $\mathcal{F}$-étagées sont $\mathcal{F}$-mesurables, puisque

$$\{\varphi > c\} = \bigcup \{\mathrm{A}_j : a_j > c\} \in \mathcal{F},$$

en tant que réunion finie d'ensembles de la tribu.

Remarque : les fonctions en escalier sur $[a, b]$ (de la théorie de l'intégrale de Riemann) sont un cas (très) particulier de fonctions $\mathcal{B}$-étagées.

Opérations sur les fonctions mesurables

Si $(f_i)_{i\in \mathrm{I}}$ est une famille finie ou dénombrable de fonctions $\mathcal{F}$-mesurables à valeurs dans $\overline{\mathbb{R}}$, on voit que la fonction $\sup_{i\in\mathrm{I}} f_i$ est $\mathcal{F}$-mesurable ; en effet, pour tout c réel on a

$$\{\sup_i f_i > c\} = \bigcup_{i\in\mathrm{I}} \{f_i > c\}$$

qui est dans $\mathcal{F}$ comme union dénombrable d'ensembles de $\mathcal{F}$. Si (f_n) est une suite croissante de fonctions $\mathcal{F}$-mesurables, sa limite (simple) est égale à $\sup_n f_n$, donc elle est $\mathcal{F}$-mesurable.

De même, inf dénombrable, limite décroissante, et

$$\limsup f_n = \lim_m \sup_{n\geq m} f_n, \quad \liminf f_n = \lim_m \inf_{n\geq m} f_n,$$

sont des opérations qui préservent les fonctions $\mathcal{F}$-mesurables. Si une suite (f_n) de fonctions $\mathcal{F}$-mesurables tend simplement vers une fonction f, cette fonction est $\mathcal{F}$-mesurable.

Proposition. *La fonction f de* X *dans $[0, +\infty]$ est $\mathcal{F}$-mesurable si et seulement s'il existe une suite croissante (φ_n) de fonctions $\mathcal{F}$-étagées ≥ 0 qui tend simplement vers f.*

Une fonction f de X *dans $\overline{\mathbb{R}}$ est $\mathcal{F}$-mesurable si et seulement s'il existe une suite (φ_n) de fonctions $\mathcal{F}$-étagées qui tend simplement vers f ; on peut supposer que $|\varphi_n| \leq |f|$.*

Preuve. — On pose

$$\varphi_n = \sum_{k=0}^{4^n-1} k2^{-n}\mathbf{1}_{\{k2^{-n}\leq f<(k+1)2^{-n}\}} + 2^n\mathbf{1}_{\{f\geq 2^n\}}.$$

Il s'agit bien d'une fonction $\mathcal{F}$-étagée, puisqu'on a vu que les ensembles de la forme $\{a \leq f < b\}$ sont dans la tribu $\mathcal{F}$. Sur chacun des ensembles

$$\mathrm{A}_{k,n} = \{k2^{-n} \leq f < (k+1)2^{-n}\}, \quad 0 \leq k < 4^n$$

de la partition qui définit φ_n, on voit que la valeur $k2^{-n}$ retenue pour φ_n est un minorant des valeurs de f sur $\mathrm{A}_{k,n}$, et de même sur le dernier ensemble $\{f \geq 2^n\}$; on a donc $\varphi_n \leq f$ pour tout n.

La suite $(\varphi_n(x))$ converge vers $f(x)$ pour tout $x \in \mathrm{X}$; supposons que $f(x) < +\infty$ pour commencer ; on peut trouver un entier n_0 tel que $f(x) < 2^{n_0}$. Pour $n \geq n_0$, il existe un entier k unique tel que $k2^{-n} \leq f(x) < (k+1)2^{-n}$; cet entier est ≥ 0 puisque $f(x) \geq 0$, et $k2^{-n} \leq f(x) < 2^{n_0} \leq 2^n$ montre que $k < 4^n$. On a donc

$$\varphi_n(x) = k2^{-n} > f(x) - 2^{-n}.$$

Si $f(x) = +\infty$, le point x se trouve dans $\{f \geq 2^n\}$ pour tout n, donc $\varphi_n(x) = 2^n$, qui tend bien vers l'infini, valeur de f en x.

Si $k2^{-n} \leq f(x) < (k+1)2^{-n}$ et $0 \leq k < 4^n$, alors $\varphi_n(x) = k2^{-n}$; si on introduit j tel que $2j \leq k \leq 2j+1$ on a aussi $2j2^{-n} \leq f(x) < (2j+2)2^{-n}$, soit $j2^{-(n-1)} \leq f(x) < (j+1)2^{-(n-1)}$, donc $\varphi_{n-1}(x) = j2^{-(n-1)} = 2j2^{-n} \leq \varphi_n(x)$; si $f(x) \geq 2^n$, alors $f(x) \geq 2^{n-1}$ donc $\varphi_{n-1}(x) = 2^{n-1} \leq 2^n = \varphi_n(x)$. La suite (φ_n) est croissante.

II.2. Intégrale

II.2.1. *Intégrale des fonctions étagées positives*

Une fonction $\mathcal{F}$-étagée positive sur X peut se représenter sous la forme

$$\varphi = \sum_{j=1}^{m} a_j \mathbf{1}_{\mathrm{A}_j}$$

où (A_j) est une partition de X en ensembles de $\mathcal{F}$ et où les (a_j) sont des *réels* ≥ 0.

On définit l'intégrale d'une fonction étagée positive par la formule

$$\int_{\mathrm{X}} \varphi \,\mathrm{d}\mu = \sum_{j=1}^{m} a_j \mu(\mathrm{A}_j)\,;$$

la valeur de l'intégrale peut être $+\infty$, dans le cas où un $a_j > 0$ correspond à un ensemble A_j de mesure infinie.

Montrons l'indépendance de l'intégrale par rapport au choix de la représentation de la fonction φ ; si on a une autre représentation,

$$\varphi = \sum_{k=1}^{n} b_k \mathbf{1}_{\mathrm{B}_k},$$

on obtient une partition plus fine de X avec les ensembles

$$\mathrm{C}_{j,k} = \mathrm{A}_j \cap \mathrm{B}_k.$$

On peut définir des réels $c_{j,k} \geq 0$ tels que

$$c_{j,k}\,\mu(\mathrm{C}_{j,k}) = a_j\,\mu(\mathrm{C}_{j,k}) = b_k\,\mu(\mathrm{C}_{j,k})$$

pour tous j, k. Si $\mathrm{C}_{j,k}$ est vide, on prend (par exemple) $c_{j,k} = 0$, et les trois nombres à comparer sont nuls. Si $\mathrm{C}_{j,k}$ n'est pas vide, on y sélectionne un point x ; alors x est dans le morceau A_j de la première partition, donc $\varphi(x) = a_j$, et x est aussi dans le morceau B_k de la deuxième partition, donc $\varphi(x) = b_k$; dans ce cas, on peut prendre $c_{j,k} = a_j = b_k$. Ensuite, on calcule

$$\sum_{j,k} c_{j,k}\,\mu(\mathrm{C}_{j,k}) = \sum_{j=1}^{m} \Bigl(\sum_{k=1}^{n} a_j\,\mu(\mathrm{A}_j \cap \mathrm{B}_k) \Bigr)\,;$$

comme les B_k forment une partition de X, les ensembles $\mathrm{A}_j \cap \mathrm{B}_k$, $k = 1, \ldots, n$, forment une partition finie de A_j et

$$\mu(\mathrm{A}_j) = \sum_{k=1}^{n} \mu(\mathrm{A}_j \cap \mathrm{B}_k).$$

On a donc

$$\sum_{j,k} c_{j,k}\,\mu(\mathrm{C}_{j,k}) = \sum_{j=1}^{m} a_j\,\mu(\mathrm{A}_j)\,;$$

en travaillant dans l'autre direction, on obtient

$$\sum_{j,k} c_{j,k}\,\mu(\mathrm{C}_{j,k}) = \sum_{k=1}^{n} b_k\,\mu(\mathrm{B}_k),$$

ce qui prouve l'indépendance.