Chapitre 6

Mesures produit et théorèmes de Fubini

6.1 Théorèmes de Fubini-Tonelli et Fubini

Définition 6.1. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. On dit que $(X, \mathcal{A}, \mu)$ est *σ-fini* s'il existe une suite de parties $(A_n)_{n\in\mathbb{N}}$ telles que $\mu(A_n) < +\infty$ pour tout n, et $X = \bigcup_n A_n$.

Cette hypothèse est par exemple vérifiée quand $\mu(X) < +\infty$ (donc en particulier quand μ est une mesure de probabilité), quand $X = \mathbb{N}$ muni de la mesure de comptage, ou quand $X = \mathbb{R}^n$ muni de la mesure de Lebesgue.

La méthode de base pour calculer une intégrale d'une fonction de 2 variables est de se ramener à des intégrales de fonctions de 1 variable. Pour cela il nous faut d'abord expliquer comment on peut munir $X \times Y$ d'une structure d'espace mesuré quand X, Y sont tous les deux munis d'une telle structure.

Définition 6.2. Soit $(X, \mathcal{A}, \mu_1)$ et $(Y, \mathcal{B}, \mu_2)$ deux espaces mesurés σ-finis. On note $\mathcal{A} \otimes \mathcal{B}$ la tribu engendrée par les parties de la forme $A \times B$, où $A \in \mathcal{A}$, $B \in \mathcal{B}$; on l'appelle *tribu produit* des tribus $\mathcal{A}$ et $\mathcal{B}$.

Alors il existe une unique mesure ν sur $\mathcal{A} \otimes \mathcal{B}$ vérifiant $\nu(A \times B) = \mu_1(A)\mu_2(B)$ pour tout $A \in \mathcal{A}$ et tout $B \in \mathcal{B}$. Cette mesure est notée $\mu_1 \otimes \mu_2$, et est σ-finie.

On n'essaiera pas de rentrer dans le détail de la construction de cette mesure ; notons que $\mathcal{B}(\mathbb{R}^{n+m}) = \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m)$ et que , si λ_n désigne la mesure de Lebesgue sur $\mathbb{R}^n$, alors on a toujours $\lambda_{n+m} = \lambda_n \otimes \lambda_m$.

La mesure produit $\mu_1 \otimes \mu_2$ étant définie à partir de μ_1 et μ_2, on s'attend à ce qu'il en aille de même de l'intégrale d'une fonction mesurable relativement à $\mu_1 \otimes \mu_2$. Et c'est effectivement le contenu des théorèmes de Fubini.

Théorème 6.3 (Fubini–Tonelli)**.** *Soit $(X_1, \mathcal{T}_1, \mu_1)$ et $(X_2, \mathcal{T}_2, \mu_2)$ deux espaces mesurés σ-finis. Soit $f\colon X_1 \times X_2 \to [0, +\infty]$ une fonction $\mathcal{T}_1 \otimes \mathcal{T}_2$-mesurable. Alors :*

1. *$y \mapsto f(x, y)$ est une fonction mesurable (sur $(X_2, \mathcal{T}_2)$ dans $[0, +\infty]$) pour tout $x \in X_1$, et $x \mapsto \int_{X_2} f(x, y)d\mu_2(y)$ est une fonction mesurable (sur $(X_1, \mathcal{T}_1)$).*
2. *$x \mapsto f(x, y)$ est une fonction mesurable (sur $(X_1, \mathcal{T}_1)$ dans $[0, +\infty]$) pour tout $y \in X_2$, et $y \mapsto \int_{X_1} f(x, y)d\mu_1(x)$ est une fonction mesurable (sur $(X_2, \mathcal{T}_2)$).*
3. *On a*
$$\int_{X_1\times X_2} f(x, y)d\mu_1 \otimes \mu_2(x, y) = \int_{X_1} \left(\int_{X_2} f(x, y)d\mu_2(y)\right) d\mu_1(x) = \int_{X_2} \left(\int_{X_1} f(x, y)d\mu_1(x)\right) d\mu_2(y) .$$

Dans le cas particulier des fonctions boréliennes, on obtient :

Corollaire 6.4 (Fubini-Tonelli pour les fonctions boréliennes)**.** Soit m, p deux entiers et $f\colon \mathbb{R}^m \times \mathbb{R}^p \to [0, +\infty]$ une fonction borélienne. Alors on a

$$\begin{aligned}\int_{\mathbb{R}^{m+p}} f d\lambda_{m+p} &= \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^p} f(x, y)d\lambda_p(y)\right) d\lambda_m(x) \\ &= \int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^m} f(x, y)d\lambda_m(x)\right) d\lambda_p(y)\end{aligned}$$

En pratique, dans les exercices, m et p vaudront le plus souvent 1 ou 2.

Remarque 6.5. Comme dans le cas de $\mathbb{R}$, intégrer une fonction de n variables sur une partie borélienne A de $\mathbb{R}^n$ revient à intégrer la fonction $f \cdot \mathbf{1}_A$ sur $\mathbb{R}^n$, donc l'énoncé ci-dessus peut être utilisé pour calculer les intégrales de fonctions boréliennes définies sur des sous-parties boréliennes de $\mathbb{R}^n$.

Exemple. Calculons l'aire du disque unité $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

Par définition, l'aire d'une partie D est l'intégrale de la fonction caractéristique de D sur $\mathbb{R}^2$; on a donc

$$\begin{aligned}
\text{aire}(D) &= \int_{\mathbb{R}^2} \mathbf{1}_D(x, y) d\lambda_2(x, y) \\
&= \int_{x=-1}^{1} \left(\int_{y=-\sqrt{1-x^2}}^{\sqrt{1-x^2}} d\lambda(y) \right) d\lambda(x) \\
&= \int_{-1}^{1} 2\sqrt{1-x^2} dx \\
&= 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(t) dt \\
&= \pi
\end{aligned}$$

On voit que, dans le calcul ci-dessus, on n'a pas pris en compte les points en dehors de D, puisque $\mathbf{1}_D$ vaut 0 en ces points : en pratique, on détermine dans quel domaine varie x (ici, $[-1, 1]$) puis, *à x fixé*, dans quel domaine varie y, et on calcule les intégrales correspondantes.

Bien sûr, on aurait pu aussi faire le calcul « dans l'autre sens », en intégrant d'abord par rapport à x puis par rapport à y ; ici ça ne changerait rien (la fonction intégrée et le domaine sont symétriques en x, y) mais dans d'autres cas un des calculs peut être beaucoup plus facile que l'autre.

Dans le cas où f n'est pas à valeurs positives, on a d'abord besoin de s'assurer que $|f|$ est intégrable, ce qui peut être fait en appliquant le théorème de Fubini-Tonelli. On peut alors appliquer le théorème suivant qui s'applique à toutes les fonctions intégrables :

Théorème 6.6 (Fubini)**.** *Soit $(X_1, \mathcal{T}_1, \mu_1)$ et $(X_2, \mathcal{T}_2, \mu_2)$ deux espaces mesurés σ-finis. Soit $f : X_1 \times X_2 \to \mathbb{R}$ une fonction* intégrable. *Alors :*

1. *$y \mapsto f(x, y)$ est une fonction intégrable (sur X_2) pour presque tout $x \in X_1$, et $x \mapsto \int_{X_2} f(x, y) d\mu_2(y)$ est une fonction intégrable (sur X_1).*
2. *$x \mapsto f(x, y)$ est une fonction intégrable (sur X_1) pour preseque tout $y \in X_2$, et $y \mapsto \int_{X_1} f(x, y) d\mu_1(x)$ est une fonction intégrable (sur X_2)*
3. *On a*

$$\int_{X_1 \times X_2} f(x, y) d\mu_1 \otimes \mu_2(x, y) = \int_{X_1} \left(\int_{X_2} f(x, y) d\mu_2(y) \right) d\mu_1(x) = \int_{X_2} \left(\int_{X_1} f(x, y) d\mu_1(x) \right) d\mu_2(y) \ .$$

Pour les fonctions boréliennes intégrables, on obtient :

Corollaire 6.7 (Théorème de Fubini pour les fonctions boréliennes)**.** Soit $f : \mathbb{R}^{m+p} \to \mathbb{R}$ une fonction borélienne telle que $|f|$ est intégrable. Alors on a :

$$\begin{aligned}
\int_{\mathbb{R}^{m+p}} f d\lambda_{m+p} &= \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^p} f(x, y) d\lambda_p(y) \right) d\lambda_m(x) \\
&= \int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^m} f(x, y) d\lambda_m(x) \right) d\lambda_p(y)
\end{aligned}$$

Remarque 6.8. Dans le cas particulier d'une fonction continue sur un compact de $\mathbb{R}^{m+p}$, on sait qu'elle est toujours intégrable ; il n'y a donc pas besoin dans ce cas d'utiliser au préalable le théorème de Fubini-Tonelli, par contre il faut bien vérifier que la fonction f est continue, et surtout que le domaine d'intégration est compact...

Pour essayer de se convaincre que le théorème de Fubini est « raisonnable », donnons-en une preuve dans le cas où on intègre une fonction *continue* f sur un rectangle $[a,b]\times[c,d]$. Notons d'abord que sous ces hypothèses $|f|$ est bien intégrable puisqu'elle est bornée et que le domaine d'intégration est de mesure finie. On veut montrer que :

$$\int_a^b \left(\int_c^d f(x,y)dy\right) dx = \int_c^d \left(\int_a^b f(x,y)dx\right) dy \ .$$

Pour cela, on introduit deux fonctions $F, G\colon [c,d]\to\mathbb{R}$ en posant

$$F(t) = \int_a^b \left(\int_c^t f(x,y)dy\right) dx \quad \text{et} \quad G(t) = \int_c^t \left(\int_a^b f(x,y)dx\right) dy \ .$$

Commençons par étudier F ; elle est de la forme $F(t) = \displaystyle\int_a^b h(x,t)dt$, avec $h(x,t) = \displaystyle\int_c^t f(x,y)dy$. A x fixé, la fonction $t\mapsto h(x,t) = \displaystyle\int_c^t f(x,y)dy$ est dérivable, de dérivée $f(x,t)$; de plus cette dérivée est bornée. On peut donc appliquer le théorème de dérivation des intégrales à paramètres pour conclure que F est dérivable sur $[c,d]$, de dérivée $F'(t) = \displaystyle\int_a^b f(x,t)dx$.

Passons à G : le théorème de continuité des intégrales à paramètres nous permet de voir que $g\colon y\mapsto \displaystyle\int_a^b f(x,y)dx$ est une fonction continue sur $[c,d]$. Puisque $G(t) = \displaystyle\int_c^t g(y)dy$, on voit que G est dérivable et que $G'(t) = g(t) = \displaystyle\int_a^b f(x,t)dx$.

On voit donc que F et G sont dérivables, et que $F' = G'$ sur $[c,d]$; de plus on a $F(c) = G(c) = 0$. On en conclut que F et G sont égales sur $[c,d]$, en particulier $F(d) = G(d)$, ce qu'on voulait démontrer.

6.2 Théorème de changement de variables

En pratique, pour calculer une intégrale multiple, on est souvent amené à faire un changement de variables pour se ramener à un domaine plus simple sur lequel appliquer le théorème de Fubini. Toutes les fonctions ne sont pas acceptables pour ce théorème : les fonctions que l'on peut utiliser pour faire un changement de variable sont les *difféomorphismes de classe* $\mathcal{C}^1$.

Définition 6.9. Étant donné U un ouvert de $\mathbb{R}^n$, et $\varphi\colon U\to\mathbb{R}^n$ une fonction de classe $\mathcal{C}^1$, on note $J\varphi(x)$ la matrice jacobienne de φ en un point $x\in U$; rappelons qu'il s'agit de la matrice dont le coefficient sur la i-ième ligne et la j-ième colonne vaut $\dfrac{\partial\varphi_i}{\partial x_j}(x)$.

Définition 6.10. Soit U, V deux ouverts de $\mathbb{R}^n$. Une fonction $\varphi\colon U\to V$ est un difféomorphisme de classe $\mathcal{C}^1$ si :

1. φ est une bijection de U sur V.
2. φ est de classe $\mathcal{C}^1$ sur U, c'est-à-dire que chaque dérivée $\dfrac{\partial\varphi_i}{\partial x_j}$ existe et est continue sur U.
3. La bijection réciproque φ^{-1} est également de classe $\mathcal{C}^1$.

Remarque 6.11. Comme $\varphi^{-1}\circ\varphi(x) = x$ pour tout $x\in U$ par définition, on obtient en appliquant la règle de la chaîne que $J\varphi^{-1}(\varphi(x))J\varphi(x) = I_n$ (la matrice identité) pour tout $x\in U$; en particulier $J\varphi(x)$ doit être inversible pour tout $x\in U$ si φ est un difféomorphisme (ce qui revient à dire que son déterminant est non nul).

Un exemple fondamental de difféomorphisme de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$ est donné par une application linéaire et inversible ; rappelons que si φ est linéaire, sa matrice jacobienne est simplement la matrice de φ dans la base canonique de $\mathbb{R}^n$.

Souvent, pour vérifier qu'une application est un difféomorphisme de classe $\mathcal{C}^1$, on utilise la caractérisation suivante.

Théorème 6.12 (Théorème d'inversion globale)**.** *Soit U un ouvert de $\mathbb{R}^n$, et $\varphi\colon U \to \mathbb{R}^n$ une application de classe $\mathcal{C}^1$. Si φ est injective sur U et $\det(J\varphi(x)) \neq 0$ pour tout $x \in U$, alors $\varphi(U)$ est un ouvert de $\mathbb{R}^n$ et φ est un difféomorphisme de classe $\mathcal{C}^1$ de U sur $\varphi(U)$.*

Remarque 6.13. Toute la difficulté du théorème précédent est dans la démonstration que $\varphi(U)$ est ouvert. Si on suppose que $\varphi(U)$ est ouvert dans l'énoncé ci-dessus, alors le fait que φ soit un difféomorphisme vient simplement du fait qu'on sait calculer « explicitement »l'inverse d'une matrice inversible (via le lien entre inverse et comatrice).

Nous pouvons maintenant énoncer le théorème de changement de variables.

Théorème 6.14 (Théorème de changement de variables)**.** *Soit U, V deux ouverts de $\mathbb{R}^n$, et $\varphi\colon U \to V$ un difféomorphisme de classe $\mathcal{C}^1$. Rappelons qu'on note λ_n la mesure de Lebesgue sur $\mathbb{R}^n$. Alors on a :*

1. *Pour toute partie B borélienne de U, $\lambda_n(\varphi(B)) = \int_B |\det(J\varphi(x))| d\lambda_n(x)$.*
2. *Si $f\colon V \to [0, +\infty]$ est borélienne, alors*
$$\int_V f(x) d\lambda_n(x) = \int_U f \circ \varphi(y) |\det(J\varphi(y))| d\lambda_n(y) \ .$$
3. *Si $f\colon V \to \mathbb{R}$ est intégrable, alors $y \mapsto f \circ \varphi(y)|\det(J\varphi(y))|$ est intégrable sur U et on a*
$$\int_V f(x) d\lambda_n(x) = \int_U f \circ \varphi(y) |\det(J\varphi(y))| d\lambda_n(y) \ .$$

La définition de l'intégrale fait que (3) est une conséquence immédiate de (2) ; et (1) est un cas particulier de (2) appliqué à la fonction caractéristique de $\varphi(B)$. La définition fait aussi qu'il n'est pas difficile de déduire (2) une fois qu'on a démontré (1).

Ce théorème est difficile à démontrer, et on ne va pas essayer de donner une idée de la preuve dans le cas où φ n'est pas une application linéaire. Discutons un peu la preuve de (1) dans le cas où φ est une application linéaire. Étudions quelques cas particuliers :

1. La matrice M de φ dans la base canonique est la matrice diagonale $\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & t & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$. Alors on a $\varphi([0,1]^n) = [0,1] \times [0,1] \times \ldots [0,t] \times [0,1] \times \ \ldots [0,1]$ si $t > 0$, et $[0,1] \times [0,1] \times \ldots [t,0] \times [0,1] \times \ \ldots [0,1]$ si $t < 0$. Dans les deux cas, on a par définition de la mesure de Lebesgue que
$$\lambda_n(\varphi([0,1]^n) = |t| = |\det(M)|\lambda_n([0,1]^n) \ .$$
2. La matrice M est la matrice par blocs $\begin{pmatrix} N & \\ & I_{n-2} \end{pmatrix}$, où $N = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Alors on a $\varphi([0,1]^n) = D \times [0,1]^{n-2}$, où D est le parallélogramme de sommets $(0,0)$, $(0,1)$, (1.1) et $(1,2)$ (faites un dessin !). Par conséquent $\lambda_n(\varphi([0,1]^n) = \text{aire}(D) = 1$, et $\det(M) = \det(N) = 1$. Donc dans ce cas on a bien comme espéré que $\lambda_n(\varphi([0,1]^n) = |\det(M)|\lambda_n(\varphi([0,1]^n)$.
3. La matrice M est une *matrice de permutation*, c'est-à-dire que φ est une application linéaire permutant les vecteurs de base. Alors la déterminant de M vaut ± 1 (ce dont on peut se convaincre par récurrence, en développant le déterminant de M par rapport à la première ligne par exemple, et en utilisant que M a exactement un 1 sur chaque ligne et chaque colonne, et des 0 ailleurs ; ou bien en revenant à la formule définissant le déterminant), et $\varphi([0,1]^n) = [0,1]^n$. On a donc encore $\lambda_n(\varphi([0,1]^n) = |\det(M)|\lambda_n(\varphi([0,1]^n)$

Pour l'instant, on a le résultat désiré pour des matrices de trois types, et $B = [0,1]^n$; par pavage, on peut se convaincre qu'on a alors le résultat désiré pour une matrice d'un de ces trois types et B n'importe quelle partie borélienne de $\mathbb{R}$. Ensuite, un théorème d'algèbre linéaire nous dit que l'on peut écrire toute matrice inversible

M sous la forme $M = M_1 M_2 \dots M_p$, où chaque matrice M_i est d'un des trois types ci-dessus. On peut donc écrire, si φ est une bijection linéaire de matrice M et B une partie borélienne de $\mathbb{R}^n$:

$$\begin{aligned}
\lambda_n(\varphi(B)) &= \lambda_n(M_1 \dots M_p(B)) \\
&= |\det(M_p)| \lambda_n(M_1 \dots M_{p-1}(B)) \\
&= |\det(M_p)| \dots |\det(M_1)| \lambda_n(B) \\
&= |\det(M)| \lambda_n(B).
\end{aligned}$$

Exemple (changement de variables en polaires). On considère l'application $\phi : U =]0, +\infty[\times]0, 2\pi[\to \mathbb{R}^2$ définie par $\phi(r, \theta) = (r \cos\theta, r \sin\theta)$.

Alors, la matrice jacobienne de ϕ est $\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$, de déterminant r.

De plus, ϕ est injective et $\phi(U) = \mathbb{R}^2 \setminus [0, +\infty[= V$.

Ainsi, ϕ est un $\mathcal{C}^1$-difféomorphisme de U sur V. Comme $\lambda_2(\mathbb{R}^2 \setminus V) = 0$, c'est-à-dire $\mathbb{R}^2 \setminus V$ est négligeable, il n'est pas gênant que ϕ ne soit pas un difféomorphisme de U sur $\mathbb{R}^2$ tout entier.

Par exemple, calculons

$$I = \int_D (x+y)^2 dx dy, \text{ où } D = \{(x, y) \colon x^2 + y^2 < 1\}.$$

En utilisant le théorème de changement de variables avec les coordonnées polaires (et le théorème de Fubini), on obtient $\phi^{-1}(D) =]0, 1[\times]0, 2\pi[$ et

$$\begin{aligned}
I &= \int_{\phi^{-1}(D)} (r\cos\theta + r\sin\theta)^2 r dr d\theta \\
&= \int_0^1 \left(\int_0^{2\pi} r^3 (\cos^2\theta + \sin^2\theta + 2\cos\theta\sin\theta) d\theta \right) dr \\
&= \int_0^1 r^3 \left(\int_0^{2\pi} (1 + \sin 2\theta) d\theta \right) dr \\
&= \int_0^1 2\pi r^3 dr \\
&= \frac{\pi}{2}.
\end{aligned}$$