DISTANCES ET TOPOLOGIES SUR LES TRIBUS ET LES FILTRATIONS

Ce texte est simplement le fruit de quelques réflexions échangées par courrier électronique. Il contient certainement des choses bien connues que seule notre flemme d'aller chercher le bon livre nous empêche de découvrir dans la littérature.

1. Distances entre sous-tribus

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. La première distance dont on peut vouloir munir cet espace porte sur les événements : si A, $B \in \mathcal{A}$, on pose

$$\Delta(A,B) = \mathbb{P}(A \triangle B)^{1/2}$$

où $A \triangle B$ désigne la différence symétrique des événements A et B. L'exposant $1/2$ est justifié par l'identité

$$\mathbb{P}(A \triangle B)^{1/2} = \mathbb{E}[\mathbb{1}_A + \mathbb{1}_B - 2\mathbb{1}_A\mathbb{1}_B]^{1/2} = \mathbb{E}\big[(\mathbb{1}_A - \mathbb{1}_B)^2\big]^{1/2}, \tag{1}$$

qui permet de voir que Δ satisfait l'inégalité triangulaire.[1] Cependant l'application Δ n'est pas une distance sur $\mathcal{A}$ car elle ne vérifie pas la condition de séparation et n'est donc qu'une pseudo-distance. C'est en revanche une distance sur $\mathcal{A}/\mathcal{N}(\mathcal{A},\mathbb{P})$ — $\mathcal{N}(\mathcal{A},\mathbb{P})$ est l'ensemble des négligeables de la tribu $\mathcal{A}$ — l'ensemble des classes d'équivalence de $\mathcal{R}_{\mathbb{P}}$ définie sur $\mathcal{A} \times \mathcal{A}$ par

$$A \,\mathcal{R}_{\mathbb{P}}\, B \iff \mathbb{P}(A \triangle B) = 0.$$

On remarque que $A \,\mathcal{R}_{\mathbb{P}}\, B$ si et seulement si les tribus $\sigma(A) \vee \mathcal{N}(\mathcal{A},\mathbb{P})$ et $\sigma(B) \vee \mathcal{N}(\mathcal{A},\mathbb{P})$ sont égales.

Considérons deux sous-tribus $\mathcal{B}$ et $\mathcal{C}$. La pseudo-distance entre ces deux tribus est naturellement définie par

$$\Delta(\mathcal{B},\mathcal{C}) = \max\Big(\sup_{B\in\mathcal{B}} \inf_{C\in\mathcal{C}} \mathbb{P}(B\triangle C)^{1/2},\ \sup_{C\in\mathcal{C}} \inf_{B\in\mathcal{B}} \mathbb{P}(B\triangle C)^{1/2}\Big)$$

et qui devient une distance lorsqu'on passe aux classes d'équivalences modulo les négligeables de $(\Omega, \mathcal{A}, \mathbb{P})$. Comme précédemment, $\Delta(\mathcal{B},\mathcal{C}) = 0$ si et seulement si $\mathcal{B} \vee \mathcal{N}(\mathcal{A},\mathbb{P})$ et $\mathcal{C} \vee \mathcal{N}(\mathcal{A},\mathbb{P})$ sont égales. C'est donc aussi une distance pour l'ensemble des sous-tribus complètes de $\mathcal{A}$, c'est-à-dire sur l'ensemble des sous-tribus contenant $\mathcal{N}(\mathcal{A},\mathbb{P})$. Remarquons que dans la définition de $\Delta(\mathcal{B},\mathcal{C})$ le premier terme est nul si et seulement si $\mathcal{B} \vee \mathcal{N}(\mathcal{A},\mathbb{P}) \subset \mathcal{C} \vee \mathcal{N}(\mathcal{A},\mathbb{P})$, et de même pour le second terme.

Évidemment, la pseudo-distance entre événements s'étend aux variables aléatoires grâce à l'identité (1) et on constate que ce n'est rien d'autre que la distance quadratique — et ce, toujours grâce à l'exposant $1/2$.

Nous noterons $A = L^2(\Omega, \mathcal{A}, \mathbb{P})$, et pour toutes sous-tribus $\mathcal{B}$, $\mathcal{C}$ de $\mathcal{A}$, B, respectivement C, le sous-espace vectoriel fermé de A formé des classes de variables

1. Rappelons que l'inégalité triangulaire est aussi satisfaite sans l'exposant $1/2$.

aléatoires $\mathcal{B}$-mesurables, respectivement $\mathcal{C}$-mesurables. La boule unité dans A sera notée $B(0,1)$ et la sphère unité $S(0,1)$, et on notera $B_B(0,1) = B(0,1) \cap B$, $B_C(0,1) = B(0,1) \cap C$, $S_B(0,1) = S(0,1) \cap B$ et $S_C(0,1) = S(0,1) \cap C$.

Les considérations précédentes nous amènent à introduire la pseudo-distance Θ entre sous-tribus définie par :

$$\Theta(\mathcal{B},\mathcal{C}) = \max\Big(\sup_{X\in B_B(0,1)} \inf_{Y\in B_C(0,1)} \|X-Y\|_{L^2}, \sup_{Y\in B_C(0,1)} \inf_{X\in B_B(0,1)} \|Y-X\|_{L^2} \Big).$$

Comme chaque infimum est réalisé par les espérances conditionnelles, qui sont les projections orthogonales P_B et P_C sur les espaces correspondants,

$$\begin{aligned}
\Theta(\mathcal{B},\mathcal{C}) &= \max\Big(\sup_{X\in B_B(0,1)} \|X - P_C X\|_{L^2}, \sup_{Y\in B_C(0,1)} \|Y - P_B Y\|_{L^2} \Big) \\
&= \max\Big(\sup_{Z\in B(0,1)} \|P_B Z - P_C \circ P_B Z\|_{L^2}, \sup_{Z\in B(0,1)} \|P_C Z - P_B \circ P_C Z\|_{L^2} \Big), \\
&= \max\big(|||(\mathrm{Id} - P_C)\circ P_B|||,\ |||(\mathrm{Id} - P_B)\circ P_C||| \big),
\end{aligned}$$

où $|||\,.\,|||$ est la norme opérateur. Si E et F sont deux sous-espaces vectoriels fermés de A, nous notons $\Theta_0(E,F) = |||(\mathrm{Id} - P_F)\circ P_E|||$, ainsi, avec un abus de notations certain, on a $\Theta(E,F) = \max(\Theta_0(E,F), \Theta_0(F,E))$. On voit facilement que

$$\Theta_0(E,F) = \Theta_0(F^\perp, E^\perp) \qquad \text{et} \qquad \Theta(E,F) = \Theta(E^\perp, F^\perp).$$

Nous tirons la proposition suivante d'un *preprint* de M. I. Ostrovskii.[2]

Proposition. — *Pour tous sous-espaces vectoriels fermés B et C de A,*

$$\Theta(B,C) = |||P_B - P_C|||.$$

Démonstration. — Montrons que $\Theta(B,C) \geqslant |||P_B - P_C|||$. On a

$$\begin{aligned}
|||P_B - P_C||| &= \sup_{x\in S(0,1)} \|P_B x - P_C x\| \\
&= \sup_{x\in S(0,1)} \|P_B x - P_B c \circ P_C x + P_B c \circ P_C x - P_C x\| \\
&= \sup_{x\in S(0,1)} \|P_B \circ (\mathrm{Id} - P_C)x + (\mathrm{Id} - P_B)\circ P_C x\| \\
&= \sup_{x\in S(0,1)} \big(\|P_B \circ (\mathrm{Id} - P_C)x\|^2 + \|(\mathrm{Id} - P_B)\circ P_C x\|^2 \big)^{1/2}
\end{aligned}$$

car le premier terme appartient à B, l'autre à $B^\perp$. Ainsi, avec la propriété d'idem-

2. « Topologies on the set of all subspaces of a Banach Space and related questions of Banach space geometry », `http://www.arts-sciences.cua.edu/math/MIO/`.

potence des projections,

$$\begin{aligned}
\|\!| P_B - P_C |\!\| &= \sup_{x\in S(0,1)} \left(\|P_B\circ(\mathrm{Id}-P_C)^2x\|^2 + \|(\mathrm{Id}-P_B)\circ P_C^2 x\|^2\right)^{1/2}\\
&\leqslant \sup_{x\in S(0,1)} \left(\Theta_0^2(C^\perp,B^\perp)\|(\mathrm{Id}-P_C)x\|^2 + \Theta_0^2(C,B)\|P_Cx\|^2\right)^{1/2}\\
&= \sup_{x\in S(0,1)} \left(\Theta_0^2(B,C)\|(\mathrm{Id}-P_C)x\|^2 + \Theta_0^2(C,B)\|P_Cx\|^2\right)^{1/2}\\
&\leqslant \Theta(B,C)\sup_{x\in S(0,1)} \left(\|(\mathrm{Id}-P_C)x\|^2 + \|P_Cx\|^2\right)^{1/2}\\
&\leqslant \Theta(B,C).
\end{aligned}$$

Démontrons l'inégalité réciproque. Revenons à

$$\|\!| P_B - P_C |\!\| = \sup_{x\in S(0,1)} \left(\|P_B\circ(\mathrm{Id}-P_C)x\|^2 + \|(\mathrm{Id}-P_B)\circ P_Cx\|^2\right)^{1/2}$$

d'où on déduit que

$$\begin{aligned}
\|\!| P_B - P_C |\!\| &\geqslant \sup_{x\in S_C(0,1)} \left(\|P_B\circ(\mathrm{Id}-P_C)x\|^2 + \|(\mathrm{Id}-P_B)\circ P_Cx\|^2\right)^{1/2}\\
&= \sup_{x\in S_C(0,1)} \|(\mathrm{Id}-P_B)\circ P_Cx\| = \Theta_0(C,B),
\end{aligned}$$

de même,

$$\begin{aligned}
\|\!| P_B - P_C |\!\| &\geqslant \sup_{x\in S_{C^\perp}(0,1)} \left(\|P_B\circ(\mathrm{Id}-P_C)x\|^2 + \|(\mathrm{Id}-P_B)\circ P_Cx\|^2\right)^{1/2}\\
&= \sup_{x\in S_C(0,1)} \|P_B\circ(\mathrm{Id}-P_C)x\| = \sup_{x\in S_C(0,1)} \|(\mathrm{Id}-P_{B^\perp})\circ P_{C^\perp}x\|\\
&= \Theta_0(C^\perp,B^\perp) = \Theta_0(B,C).
\end{aligned}$$

Ce qui permet de voir que $\|\!| P_B - P_C |\!\| \geqslant \max(\Theta_0(B,C),\Theta_0(B,C)) = \Theta(B,C)$, et finalement qu'il y a égalité. □

Remarques. — *a*) D'autres pseudo-distances peuvent être considérées, en particulier en remplaçant L^2 par L^p pour $p\in[1,\infty[$. Cependant, la proposition précédente — qui est particulière au contexte hilbertien — ne s'applique plus.

b) Nous voulions au départ définir la pseudo-distance

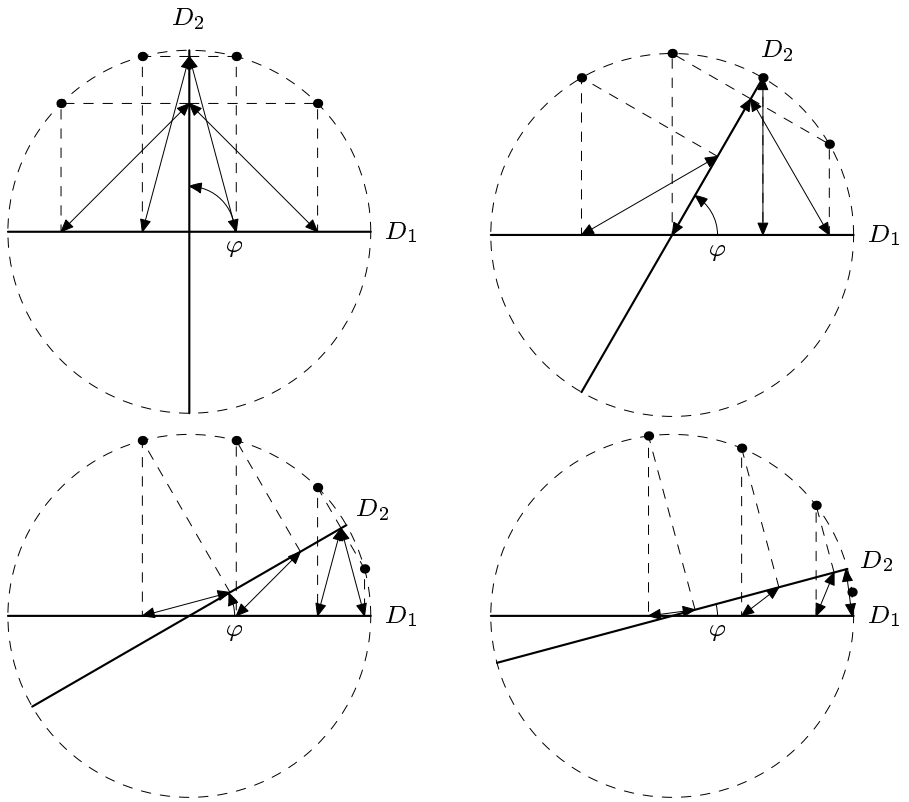
$$\begin{aligned}
\Theta'(\mathcal{B},\mathcal{C}) &= \sup_{X\in B_B(0,1)}\inf_{Y\in B_C(0,1)} \|X-Y\|_{L^2} + \sup_{Y\in B_C(0,1)}\inf_{X\in B_B(0,1)} \|Y-X\|_{L^2}\\
&= \Theta_0(B,C)+\Theta_0(C,B).
\end{aligned}$$

Mais on voit que $\Theta(\mathcal{B},\mathcal{C})\leqslant\Theta'(\mathcal{B},\mathcal{C})\leqslant 2\Theta(\mathcal{B},\mathcal{C})$, et donc qu'elle ne présente aucun intérêt particulier.

2. **Remarque géométrique**

Lemme. — *Dans le plan euclidien, soient D_1 et D_2 deux droites concourantes en un point O. Notons φ l'angle aigü entre ces deux droites. Si M est un point du*

plan à distance unité du point O et si M_i désigne la projection orthogonale de M sur D_i, alors $\|M_1 - M_2\| = \sin\varphi$.



Démonstration. — Nous abordons le problème en notation complexe. Nous identifions le point O au nombre 0, la droite D_1 à la droite réelle et $M = \mathrm{e}^{\mathrm{i}(\varphi+\vartheta)}$. Alors $M_1 = \cos(\varphi+\vartheta)$, $M_2 = \mathrm{e}^{\mathrm{i}\varphi}\cos\vartheta$ et $z = M_2 - M_1 = \mathrm{e}^{\mathrm{i}\varphi}\cos\vartheta - \cos(\varphi+\vartheta)$.

$$\begin{aligned}
z\bar{z} &= \big(\mathrm{e}^{\mathrm{i}\varphi}\cos\vartheta - \cos(\varphi+\vartheta)\big)\big(\mathrm{e}^{-\mathrm{i}\varphi}\cos\vartheta - \cos(\varphi+\vartheta)\big)\\
&= \cos^2\vartheta - 2\cos\varphi\cos\vartheta\cos(\varphi+\vartheta) + \cos^2(\varphi+\vartheta)\\
&= \cos^2\vartheta - \big(\cos(\varphi+\vartheta) + \cos(\varphi-\vartheta)\big)\cos(\varphi+\vartheta) + \cos^2(\varphi+\vartheta)\\
&= \cos^2\vartheta - \cos^2(\varphi+\vartheta) + \cos(\varphi-\vartheta)\cos(\varphi+\vartheta) + \cos^2(\varphi+\vartheta)\\
&= \cos^2\vartheta + \cos(\varphi-\vartheta)\cos(\varphi+\vartheta)\\
&= \frac{1}{2}(1+\cos 2\vartheta) - \frac{1}{2}\cos 2\vartheta - \frac{1}{2}\cos 2\varphi\\
&= \frac{1}{2}(1-\cos 2\varphi)\\
&= \sin^2\varphi,
\end{aligned}$$

d'où la conclusion. □

Remarque. — Une preuve élémentaire et plus géométrique devrait être évidente, mais nous ne l'avons pas vue. Michel Coornaert donne l'explication suivante :

On remarque tout d'abord que les points O, M, M_1 et M_2 sont sur le cercle de rayon $1/2$ centré au milieu I de OM. Cela résulte du fait que les triangles OM_1M et

OM_2M sont rectangles. L'angle M_1IM_2 vaut donc 2φ (théorème de l'angle capable). On a donc $M_1M_2 = sin\varphi$.

Les considérations qui précèdent amènent à la définition suivante :

Définition. — L'angle entre deux sous-tribus $\mathcal{B}$ et $\mathcal{C}$ est défini comme étant l'unique réel $\varphi \in [0, \pi/2]$ tel que $\Theta(\mathcal{B}, \mathcal{C}) = \sin\varphi$.

3. **Convergence de sous-tribus**

Dans la suite, les sous-tribus de $\mathcal{A}$ qu'on considérera seront complètes. On pourrait aussi convenir d'identifier les sous-tribus équivalentes modulo les négligeables de $(\Omega, \mathcal{A}, \mathbb{P})$. Ce qui importe est de considérer un espace séparé afin que la notion de limite soit définie sans ambiguïté.

Définition. — (i) Une suite de sous-tribus $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *converge* vers une sous-tribu $\mathcal{B}$ si et seulement si pour tout $\varepsilon > 0$ il existe $N \geqslant 0$ tel que, pour tout $n \geqslant N$, $\Theta(\mathcal{B}_n, \mathcal{B}) \leqslant \varepsilon$.

(ii) Une suite de sous-tribus $(\mathcal{B}_n)_{n\in\mathbb{N}}$ est une *suite de Cauchy* si et seulement si pour tout $\varepsilon > 0$ il existe $N \geqslant 0$ tel que, pour tout m, $n \geqslant N$, $\Theta(\mathcal{B}_m, \mathcal{B}_n) \leqslant \varepsilon$.

Soit $(\mathcal{B}_n)_{n\in\mathbb{N}}$ une suite de Cauchy de sous-tribus. Alors, pour toute variable aléatoire $Z \in L^2(\Omega, \mathcal{A}, \mathbb{P})$, la suite $(Z_n)_{n\in\mathbb{N}}$ définie par $Z_n = \mathbb{E}[Z \,|\, \mathcal{B}_n]$ est une suite de Cauchy. En effet, quitte à considérer $Z/\|Z\|_{L^2}$ lorsque la norme de Z est non nulle, on peut supposer que $Z \in B_{\mathcal{A}}(0, 1)$. On constate alors que $(Z_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de même « module de convergence » que $(\mathcal{B}_n)_{n\in\mathbb{N}}$. Puisque $L^2(\Omega, \mathcal{A}, \mathbb{P})$ est complet, $(Z_n)_{n\in\mathbb{N}}$ converge et notons Z_∞ sa limite — qui est une classe d'équivalence de variables aléatoires. Posons

$$\mathcal{B}_\infty = \sigma\{Z_\infty \,|\, Z \in L^2(\Omega, \mathcal{A}, \mathbb{P})\}.$$

L'espace vectoriel $L^2(\Omega, \mathcal{B}_\infty, \mathbb{P})$ est l'adhérence de l'ensemble des limites Z_∞ pour Z parcourant $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Il y a en fait égalité et ceci résulte du lemme suivant :

Lemme. — *Si $Z \in L^2(\Omega, \mathcal{A}, \mathbb{P})$, alors $Z_\infty = \mathbb{E}[Z \,|\, \mathcal{B}_\infty]$. En particulier, si $Z \in L^2(\Omega, \mathcal{B}_\infty, \mathbb{P})$, alors $Z_\infty = Z$.*

Démonstration. — Pour montrer que Z_∞ est la projection orthogonale de Z sur $L^2(\Omega, \mathcal{B}_\infty, \mathbb{P})$, il suffit de prouver que

$$\mathbb{E}[YZ] = \mathbb{E}[YZ_\infty]$$

pour toute $Y \in L^2(\Omega, \mathcal{B}_\infty, \mathbb{P})$. Comme chacun des termes est continu en Y, il suffit d'établir cette identité sur un sous-espace dense, en particulier pour les Y qui sont de la forme X_∞. On a

$$\begin{aligned}\mathbb{E}[X_\infty Z] = \mathbb{E}\Big[\lim_{n\to\infty} X_n Z\Big] &= \lim_{n\to\infty} \mathbb{E}[X_n Z] = \lim_{n\to\infty} \mathbb{E}\big[\mathbb{E}[X_n Z \,|\, \mathcal{B}_n]\big] \\ &= \lim_{n\to\infty} \mathbb{E}\big[X_n\, \mathbb{E}[Z \,|\, \mathcal{B}_n]\big] = \lim_{n\to\infty} \mathbb{E}[X_n Z_n] = \mathbb{E}[X_\infty Z_\infty].\end{aligned}$$

D'où $Z_\infty = \mathbb{E}[Z \,|\, \mathcal{B}_\infty]$. □

PROPOSITION. — *Si la suite $(\mathcal{B}_n)_{n\in\mathbb{N}}$ est une suite de Cauchy, elle est alors convergente et sa limite est $\mathcal{B}_\infty$. L'espace des sous-tribus complètes muni de la distance Θ est complet.*

Démonstration. — Ce qu'il nous faut montrer est que $(\mathcal{B}_n)_{n\in\mathbb{N}}$ tend vers $\mathcal{B}_\infty$. Soient $\varepsilon > 0$ et $N \geqslant 0$ tel que pour tous m, $n \geqslant N$, $\Theta(\mathcal{B}_m, \mathcal{B}_n) \leqslant \varepsilon$. Ainsi, pour tout $Z \in B_{\mathcal{A}}(0,1)$ et tous m, $n \geqslant N$, $\|Z_m - Z_n\| \leqslant \varepsilon$. En laissant m tendre vers l'infini, par continuité de la norme on a $\|Z_\infty - Z_n\| \leqslant \varepsilon$. Ceci montre que $\Theta(\mathcal{B}_\infty, \mathcal{B}_n) \leqslant \varepsilon$ et donc que $(\mathcal{B}_n)_{n\in\mathbb{N}}$ converge vers $\mathcal{B}_\infty$. □

Remarque. — Ce qui précède n'est pas propre au contexte des sous-tribus, ce ne sont que des propriétés hilbertiennes. En effet, dire qu'une suite de sous-tribus est de Cauchy, c'est dire que les projecteurs orthogonaux associés forme une suite de Cauchy. Comme l'espace des opérateurs linéaires continus sur un espace de Banach est un espace de Banach, ces suites de Cauchy convergent. L'identité de projecteur $P^2 - \mathrm{Id} = 0$ passant à la limite, la limite est un projecteur. La préservation de l'orthogonalité suit le même principe puisque c'est dire que la forme quadratique $\langle P \,|\, \mathrm{Id} - P\rangle$ est nulle.

4. **Cas des filtrations**

Soit T un ensemble ordonné. Définir une distance entre deux filtrations indexées par T est aisé, il suffit de prendre le supremum des distances entre tribus de même indice : si $(\mathcal{F}_t)_{t\in T}$ et $(\mathcal{G}_t)_{t\in T}$ sont deux filtrations sur $(\Omega, \mathcal{A}, \mathbb{P}$, leur distance est

$$\Theta\big((\mathcal{F}_t)_{t\in T}, (\mathcal{G}_t)_{t\in T}\big) = \sup_{t\in T} \Theta(\mathcal{F}_t, \mathcal{G}_t),$$

et il n'y a pas grand chose à en dire, sauf qu'elle est peut-être un peu forte.

Département de Mathématiques
Téléport 2 – BP 30 179
Boulevard Marie et Pierre CURIE
86 962 Futuroscope Chasseneuil cedex
phan@mathlabo.univ-poitiers.fr