Frontières et données au bord, visions mathématiques de la marge.

Thomas Leblé (CNRS - Université de Paris (-Descartes) - MAP5)

17 janvier 2020

Le but de cet exposé est de présenter certains concepts ou problèmes que les "marges" et le "paradigme des marges" peuvent évoquer en mathématiques. Le terme de "marge" lui-même n'apparaît pas dans la terminologie usuelle ; on emploie l'adjectif "marginal" en statistiques, où l'on parle notamment de "loi marginale" (*marginal law*), mais il s'agit d'un concept peu profond, quoique très utile. Je parlerai ici plutôt de frontière et de bord (qui correspondent tous deux à l'anglais *boundary*), ou d'événements rares et de déviations, compris comme événements marginaux. Je voudrais notamment :

- Donner un exemple de la façon dont les mathématiques pensent la marge comme un concept géométrique.
- Illustrer le rôle des marges, ou "conditions au bord", dans la modélisation mathématique des phénomènes naturels.
- Enfin, évoquer une situation, la théorie des grandes déviations, où le paradigme des marges pourrait être aperçu au sein même des mathématiques.

1 Frontière et intérieur

Fixons une partie, notée Ω, du plan Euclidien (un espace plat, à deux dimensions, tel celui d'une feuille de papier ou d'un tableau noir). Pour parler de Ω comme d'un objet spécifique, le plus simple est de faire une distinction *ensembliste* entre Ω et son complémentaire Ω^c : il y a les points du plan qui appartiennent à Ω et ceux qui n'appartiennent pas à Ω (et appartiennent donc à son complémentaire). C'est un point de vue ensembliste au sens où le seul critère est celui de l'appartenance, on ne manipule pas d'information *géométrique*.

Pour affiner notre point de vue, on peut faire usage des concepts fournis par la topologie générale. La topologie est l'étude mathématique des formes en un sens assez lâche : deux formes sont considérées comme équivalentes si l'on peut passer de l'une à l'autre par une déformation continue, sans les déchirer ni les trouer. Pour la topologie, un disque et deux demi-disques sont des objets différents, en revanche on ne peut pas distinguer un carré d'un disque, ni un petit disque d'un grand disque : ils ont exactement les mêmes propriétés topologiques.

Dans le langage topologique, on dispose des termes *intérieur* et *frontière.*

— Un point est à l'intérieur de Ω s'il existe autour de lui une petite zone toute entière contenue dans Ω.
— À l'inverse, un point est à la frontière de Ω si, dès que l'on regarde autour de lui, on trouve à la fois des points de Ω et des points de son complémentaire.

Au lieu de dire "une petite zone autour", on emploie le terme de "voisinage" (en anglais : *neighborhood*). La description *topologique* d'un espace suppose de dresser la liste des "voisinages" de chaque point, ces voisinages doivent vérifier certaines règles structurelles, au prix de quoi on bénéficie de la force de la topologie générale, qui fournit des notions et des résultats valables quelle que soit la notion de voisinage que l'on a choisie. Formellement, les mathématiques définissent la frontière de Ω comme l'ensemble des points dont tous les voisinages rencontrent à la fois Ω et Ω^c. Il est intéressant de noter qu'on ne sait pas *a priori* si un point de la frontière appartient à Ω ou non, les deux cas de figure sont possibles.

Lorsque qu'on passe du langage ensembliste au langage topologique, on peut distinguer un point qui est à l'intérieur de Ω d'un point qui appartient à Ω mais vit sur sa frontière. Ces propriétés seront préservées lors d'une déformation continue : si l'on transforme un disque en un carré, l'intérieur du disque devient l'intérieur du carré et le bord est envoyé sur le bord. Le langage topologique ne permet cependant pas une description *métrique*, on ne dit par exemple rien sur les distances entre les points. Le fait d'être plus ou moins "près du bord", en particulier, n'est pas capturé par ce langage. Pour cela, on pourrait encore enrichir la description de notre espace en le munissant d'une distance. Structurellement, cela signifie passer de la catégorie des ensembles à celle des espaces topologiques, puis à celle des espaces métriques.

2 L'épaisseur du trait

Pour parler des "voisinages" en langage courant, au lieu de dire "une petite zone autour", on serait tenté d'utiliser "marge de sécurité". En effet, dire qu'un point est à l'intérieur signifie intuitivement qu'il n'est pas au bord, qu'il ne risque pas de tomber dehors, qu'il existe autour de lui une petite marge de sécurité contenue dans Ω. L'expression "marge de sécurité" rappelle qu'une "marge" a, dans le sens usuel, une certaine épaisseur. La frontière mathématique, elle, n'a intuitivement pas d'épaisseur. Cet énoncé est à nuancer, ne serait-ce que parce que l' "épaisseur" est une notion délicate à définir. La frontière d'un disque, c'est le cercle : on passe d'un objet de dimension 2 à un objet de dimension 1, qui n'a pas d'épaisseur dans le plan. De même, dans l'espace usuel à trois dimensions, la frontière d'une boule est une sphère :

on perd une dimension (de 3 à 2) et l'objet "frontière" n'a pas d'épaisseur. Cependant certains domaines plus tortueux ont des frontières plus torturées, qui peuvent avoir une certaine épaisseur, c'est le cas notamment des objets fractals. Il existe plusieurs manières de mesurer la "taille" ou "l'épaisseur" de ces objets, l'une des plus utilisées est la "dimension de Hausdorff", qui peut prendre des valeurs non-entières, par exemple comprises entre 1 et 2 pour des objets dans le plan. Se rapprochant davantage de l'idée usuelle de marge comme étant "une certaine épaisseur près du bord", il arrive qu'on munisse l'espace d'une distance, et qu'on introduise des "voisinages tubulaires" : c'est l'ensemble des points qui sont au plus à une certaine distance de la frontière. Le choix de cette distance est cependant arbitraire, il n'y a pas de choix naturel, ce qui, en mathématiques, est souvent mauvais signe.

3 Points extrémaux

On peut encore mentionner la notion de faces et de points extrémaux. Beaucoup d'objets (géométriques ou non) sont *convexes* : lorsque deux points A et B appartiennent à l'objet, tous les points situés entre A et B appartiennent aussi à l'objet : penser aux polygones réguliers, aux solides platoniciens, ou à une situation dans laquelle un point est repéré par un certain nombre de paramètres qui peuvent prendre n'importe quelle valeur dans un intervalle donné. Cela crée un "espace des possibles" qui est un objet convexe avec, parfois, beaucoup de dimensions. On peut alors chercher à décrire cet objet en termes d'éléments plus simples : les points extrémaux.

Prenons le cas d'un polyèdre convexe en dimension 3, par exemple un cube. Sa frontière est composée de six *faces*. Chacune de ces faces est un carré, dont la frontière est composée de quatre côtés. Chacun de ces côtés, enfin, est un segment dont la frontière est composée de deux points, appelés couramment *sommets*. On peut observer que tout point du cube s'exprime comme un barycentre des sommets. Une telle procédure s'applique dans le cas général : on peut construire les faces, puis les faces des faces, jusqu'à arriver à des points extrémaux. Le théorème de Minkowski énonce que tout point du convexe s'obtient comme une combinaison géométrique des points extrémaux. Ce résultat stipule donc qu'on peut reconstruire entièrement les objets d'un certain type à partir de la donnée de leurs points extrémaux : c'est la structure convexe qui permet une telle réduction. En sus, il est intéressant d'observer qu'on a ici décomposé la frontière selon une hiérarchie de nature géométrique, en descendant d'une dimension à chaque fois.

4 Conditions au bord

La modélisation mathématique des phénomènes naturels utilise de façon centrale le langage des "équations différentielles", qui permet d'exprimer les relations gouvernant l'évolution des différentes grandeurs physiques. Une telle équation est dite "différentielle" parce qu'elle parle toujours de *variations* (en espace, en temps), qu'elle lie entre elles. L'équation de la chaleur, par exemple, obtenue à partir de la loi de propagation, phénoménologique, obtenue par Fourier, dit que "la variation de la température par rapport au temps est proportionnelle à la seconde variation de la température par rapport à la position". Mathématiquement, on utilise des variations infinitésimales, c'est-à-dire des *dérivées*, et le domaine est celui des *équations aux dérivées partielles.*

Comme elles ne régissent que les *variations* à l'*intérieur* des systèmes, ces équations ne sont pas suffisantes pour décrire pratiquement l'état à un instant donné. Pour cela, il faut leur adjoindre une liste de "données" ou "conditions au bord" (*boundary data*), qui spécifient d'une part l'état du système à l'instant initial, et d'autre part donnent une information sur ce qu'il se passe, à tout instant, au bord du domaine, c'est à dire le long de sa frontière. Selon les situations, il pourra s'agir de dire que la température au bord du système est connue (par exemple parce ce que celui-ci est au contact d'un thermostat), de demander à ce qu'aucune particule ne puisse sortir du domaine, ou bien encore d'observer que le champ électrique doit être nul. Ces conditions peuvent revêtir de nombreuses formes, mais sans elles le problème de prédire l'évolution du système est sous-déterminé : il existe une infinité de solutions à l'équation de la chaleur, mais une seule qui satisfasse à une certain jeu de données au bord. Si l'on impose un certain comportement au bord, il se répercutera dans tout le système via le jeu des relations encodée par l'équation différentielle, qui gouverne uniquement l'évolution des grandeurs.

Notons que là encore, la marge mathématique n'a aucune épaisseur. Cependant, par exemple dans l'étude mathématique de la mécanique des fluides, la présence d'un bord donne naissance au phénomène crucial, et particulièrement difficile à analyser, des *couches limites.* Pour schématiser, une couche limite est une bande le long du bord du domaine dans laquelle le comportement "libre" du fluide et les "conditions au bord" imposées à la frontière sont en conflit. Il ne s'agit cette fois pas d'une marge d'épaisseur arbitraire : les couches limites possèdent une taille caractéristique déterminée par les grandeurs physiques du problème considéré.

Dans la pratique de l'analyse mathématique, les "conditions au bord" apparaissent parfois comme une contrainte technique ennuyeuse. Elles ont toutefois un caractère essentiel, qui ne fait que traduire mathématiquement leur importance "dans la vraie vie", c'est-à-dire dans l'évolution réelle des systèmes physiques, biologiques, économiques, etc. L'interaction subtile entre l'intérieur et le bord, qui apparaît dans le cadre des couches limites en mécanique des fluides, recèle des difficultés d'analyse considérables qui requièrent une compréhension fine des équations en jeu.

5 Cygnes noirs et tableaux blancs

Nous avons évoqué deux situations où la marge apparaissait en mathématiques : d'abord comme notion géométrique, puis comme élément nécessaire à la modélisation. Pour finir, voici un cas où l'on peut deviner le paradigme des marges en action au sein même du raisonnement mathématique.

Lorsqu'on on étudie un événement aléatoire, on cherche souvent à comprendre son comportement *typique*, c'est-à-dire le plus probable, par opposition à des événements que l'on pourrait qualifier de *marginaux*, qui n'ont qu'une petite masse dans l'espace probabilisé. Notons qu'il est parfois important de bien connaître ces événements rares, ces "cygnes noirs", car bien que très rares ils peuvent, par exemple, être très coûteux (penser à une catastrophe écologique ou financière). Dans les situations les plus simples, le comportement typique est facile à déterminer, et l'exercice délicat consiste à quantifier la probabilité de certains événements rares. Si on lance un dé à six faces un très grand nombre de fois, on s'attend à relever une proportion à peu près équitable de chaque chiffre entre 1 et 6, et un résultat fondamental, le théorème central limite, nous enseigne quelles sont les fluctuations raisonnables autour d'un équilibre parfait. Si l'on s'interroge sur la probabilité d'un comportement vraiment étrange, par exemple de n'avoir lancé que des 2 et des 4, on fait appel à la théorie dite "des grandes déviations".

Cette théorie s'applique à des événements extrêmement marginaux, dont la probabilité décroît exponentiellement vite. Son objet principal est *la fonction de taux*, c'est une fonction qui permet de comparer la rareté de différentes situations. Plus la valeur de cette fonction est grande, plus l'événement est rare. À l'inverse, pour des événements raisonnables, la valeur de cette fonction est presque zéro. Savoir calculer la *fonction de taux* est donc une étape cruciale pour l'étude quantitative des événements marginaux. Notons que cette "marge dans l'aléa" possède, cette fois, une certaine épaisseur, certes très petite, et qu'il s'agit précisément de quantifier : ces grandes déviations sont très peu probables, mais pas impossibles, c'est à dire de mesure nulle.

Dans le cas du lancer de dés, la fonction de taux est un objet bien connu en statistique : c'est ce qu'on appelle l'entropie relative. On comprend alors à la fois le comportement typique (une répartition à peu près équitable entre toutes les faces du dé) et les déviations vis-à-vis de ce comportement. L'idée que je voudrais soulever ici, pour répondre au sujet de cette journée du point de vue de mon expérience de recherche, est que l'étude des événements marginaux est parfois la clef non seulement d'une compréhension des "grandes déviations" mais aussi des événements typiques. Dans des situations complexes, où plusieurs phénomènes de natures différentes entrent en ligne de compte pour déterminer la probabilité d'un événement, il n'est pas toujours possible de deviner *a priori* quel est le comportement "typique" attendu, et quel est le comportement "déviant". Il arrive cependant que l'on puisse calculer ladite fonction de taux et quantifier la rareté de chaque événement. Selon le principe originel, les déviations se lisent comme les zones où cette fonction de taux prend des grandes valeurs, mais, dans un dédoublement du paradigme, le comportement "normal" lui aussi se lit, en relief, comme l'ensemble des événements où la fonction de taux est minimale. Dans cette méthode, on détermine les propriétés typiques comme étant "celles qui ne sont pas rares", celles qui ne sont pas trop loin pour la distance mesurée par la fonction de taux. À mon sens, il y a là une interprétation possible de l'expression employée par Robert Castel "la distance qui les sépare du centre permettrait de mettre au jour les fondements", puisque les propriétés caractéristiques du système ne sont mises au jour que lorsqu'on se met à étudier et mesurer, au contraire, ce qui pourrait conduire un événement à être peu probable, non-typique, marginal.

6 Conclusion

Que peut-on rapporter de cette promenade au long des marges mathématiques ? D'une part, que les objets "marginaux" ont, en mathématiques, un intérêt propre : les caractériser, déterminer leur régularité, leur dimension, leur taille, et relier leurs propriétés aux propriétés de l'objet "global" sont autant de questions intéressantes, et parfois très difficiles. Il n'est d'ailleurs pas rare que les phénomènes les plus délicats et les plus riches se nichent près du bord : on a mentionné le cas des couches limites, mais il y a bien d'autres exemples, dans des domaines divers des mathématiques. Pour aller plus particulièrement dans le sens du thème abordé aujourd'hui, il est remarquable d'observer que le paradigme des marges trouve sa place dans la tactique mathématique elle-même, et qu'il peut être fructueux, voire crucial, d'aborder l'analyse d'un objet par l'étude de ses limites, de ce qui en lui se trouve au bord, ou déviant, de se demander ce que la frontière nous dit du tout.

Pour conclure, jouons au jeu suivant : traçons une courbe sur un tableau. Cette courbe est-elle une *frontière* ? Délimite-t-elle un intérieur et un extérieur ? L'intuition visuelle nous dit que oui, et la réponse est, en effet, positive, mais assez difficile : c'est le théorème de Jordan [1]. Imaginons que nous vivions quelque part dans le plan, comment ferions-nous pour savoir de quel côté (intérieur ou extérieur) nous nous trouvons ? Une façon de procéder est de se déplacer tout droit. Si on est à l'intérieur, on passera la frontière un nombre *impair* de fois, contre un nombre *pair* de fois si l'on démarre à l'extérieur. Toutefois, plus la courbe est tordue, plus ce nombre (pair ou impair)

1. Camille Jordan (1838-1922), mathématicien français qui fit d'importantes contributions à l'algèbre, et fut l'un des premiers à reconnaître l'importance des travaux d'Évariste Galois. La démonstration qu'il donne de son énoncé (le fait que toute courbe fermée, continue, simple, ou "courbe de Jordan", sépare le plan en deux composantes connexes) a longtemps été jugée non rigoureuse et incomplète, mais a été récemment ré-habilitée.

peut être grand, et plus il devient difficile de tenir les comptes, or on peut imaginer des frontières arbitrairement torturées. D'ailleurs, le théorème dit qu'il existe un extérieur et un intérieur, mais ne prétend pas fournir une règle pour déterminer qui est qui en général ! Devant la délicate de tâche de savoir qui est à l'intérieur, qui à l'extérieur, je me risque, en guise de conclusion, à évoquer ces vers du *Roman inachevé* :

Il existe près des écluses / Un bas quartier de bohémiens / Où la belle jeunesse s'use / À démêler le tien du mien.