

Catégories de Markov, théorie synthétique des probabilités

Luca Brugnerotto, Janosch Laubé-Rainer

encadré par Paul Wang

Table des matières

Introduction	2
<hr/>	
1 Préliminaires catégoriques	2
<hr/>	
1.1 Généralités	2
1.2 Catégories monoïdales symétriques	5
2 Catégories de Markov	7
<hr/>	
2.1 Définition et premiers exemples	7
2.2 Notions probabilistes dans les catégories de Markov	9
3 Produits dans les catégories de Markov et loi du 0-1 de Kolmogorov	11
<hr/>	
3.1 Produits dans les catégories de Markov	11
3.2 Loi du 0-1 de Kolmogorov, version catégorique	12
4 Théorie ergodique par les catégories de Markov	14
<hr/>	
4.1 Systèmes dynamiques dans le langage catégorique	15
4.2 Inversion bayésienne, désintégrations	16
4.3 Le théorème de décomposition ergodique	17
4.4 Un exemple de décomposition ergodique	19

Introduction

La théorie des catégories a émergé au milieu du XX^e siècle, initialement développée pour résoudre des problèmes de topologie algébrique dans la publication *General theory of natural equivalences* (1942) de Mac Lane et Eilenberg. Depuis, elle a largement dépassé ce cadre, et certains domaines des mathématiques, comme la géométrie algébrique, utilisent désormais le langage catégorique. Dans ce mémoire, on emploie le formalisme catégorique pour reformuler la théorie des probabilités, proposant une version plus abstraite que l'approche classique fondée sur la théorie de la mesure.

L'idée initiale est de constituer une classe de catégories qui généralise la catégorie *Stoch*, dont les objets sont des espaces mesurables et les flèches, appelées *noyaux de Markov*, englobent les fonctions mesurables, les mesures de probabilité et, plus généralement, les transitions entre deux états d'une chaîne de Markov. Ces catégories sont nommées, suivant Fritz [Fri20], *catégories de Markov*. Intuitivement, il s'agit de catégories dans lesquelles on peut discuter du produit tensoriel et définir les notions usuelles de la théorie des probabilités : mesures, variables aléatoires, indépendance, etc.

La théorie des catégories de Markov, comme la théorie des catégories en général, offre une interprétation plus synthétique de certains concepts de par son cadre abstrait. En effet, les résultats obtenus dans le cadre des catégories de Markov s'appliquent à des catégories sans lien direct avec la théorie des probabilités. Par exemple, la loi du 0-1 de Kolmogorov peut être appliquée en topologie générale (voir la section 6 de l'article de Fritz et Rischel sur le sujet [FR20]). Ce point de vue permet donc une plus grande généralité, mais fournit surtout un langage synthétique propice à des preuves graphiques, parfois plus faciles à manipuler plus faciles à manipuler que leur équivalent usuel.

Dans ce mémoire, on utilise le formalisme développé pour apporter de nouvelles démonstrations à deux résultats centraux, respectivement en théorie des probabilités et en théorie ergodique : la loi du 0-1 de Kolmogorov et le théorème de décomposition ergodique. Bien sûr, ces théorèmes ont d'abord été prouvés sans utiliser le formalisme catégorique, notamment la loi du 0-1, démontrée par Kolmogorov en 1933, alors que la théorie des catégories n'avait pas encore émergé. Les preuves présentées reprennent ainsi des preuves déjà connues dans le cadre classique et les adaptent dans ce nouveau langage.

On commence par introduire les outils catégoriques nécessaires pour ensuite établir le formalisme des catégories de Markov, redéfinissant toutes les notions probabilistes dans ce langage, en illustrant ces notions par des exemples. Il s'agit ensuite de démontrer la loi du 0-1 de Kolmogorov dans ce cadre et de la comparer avec la preuve usuelle ; pour ce faire, il est nécessaire, entre autres, de définir une notion de produit infini, dit *de Kolmogorov*, utilisant la structure de catégorie de Markov. Enfin, on s'intéresse à la théorie ergodique : on démontre le théorème de décomposition ergodique, en l'appliquant notamment à un exemple précis afin d'explicitier la nature de la décomposition obtenue avec celle de la théorie ergodique classique.

1 Préliminaires catégoriques

On introduit dans cette section les différentes notions catégoriques nécessaires pour la suite de ce mémoire. Au fil des définitions, on définit également les notations de diagramme de corde, une écriture de calcul graphique adaptée aux catégories monoïdales.

1.1 Généralités

On commence par rappeler les définitions générales de catégorie, de foncteur et de transformation naturelle, tout en introduisant la notation utilisée dans ce mémoire.

Définition 1.1 – Une *catégorie* C est la donnée :

- d’une classe $\text{Ob}(C)$ (dont les éléments sont les *objets* de C)
- pour tout $X, Y \in \text{Ob}(C)$, d’une classe $\text{Hom}_C(X, Y)$ (dont les éléments sont les *flèches* ou *morphismes* de X à Y , on les note $f : X \rightarrow Y$)
- pour tout objet X , d’un morphisme $\mathbf{1}_X : X \rightarrow X$ appelé *identité* de X
- d’une *composition* \circ qui à $f : X \rightarrow Y$ et $g : Y \rightarrow Z$ associe $g \circ f : X \rightarrow Z$, qui est associative et dont les éléments neutres sont les identités.

Considérons, à titre d’exemple, les catégories suivantes :

- la catégorie Set des ensembles, dont les objets sont les ensembles et les morphismes sont les applications entre ensembles;
- la catégorie Grp des groupes dont les objets sont les groupes et les morphismes sont les morphismes de groupes;
- la catégorie Vect_k des k -espaces vectoriels pour k un corps, dont les objets sont les k -espaces vectoriels et les morphismes sont les applications linéaires;
- la catégorie $\text{Mat}_{\mathbb{R}}$ des espaces vectoriels réels de dimension finie, dont les objets sont les entiers $n \in \mathbb{N}$ (représentant la dimension d’un espace \mathbb{R}^n) et les morphismes $n \rightarrow m$ sont les matrices $m \times n$ (notons que ce ne sont formellement *pas* des applications) avec leur multiplication en guise de composition.

On utilisera dans ce document la notation à base de *diagramme de corde* plutôt que de diagrammes usuels. Une flèche $f : X \rightarrow Y$ s’écrit ainsi en termes de diagramme de corde de bas en haut

$$\begin{array}{c} Y \\ | \\ \boxed{f} \\ | \\ X \end{array} \quad \text{ou} \quad \begin{array}{c} | \\ \boxed{f} \\ | \end{array}$$

quand les objets d’entrée et de sorties sont clairs ou peu importants. La flèche identité s’écrit simplement sans étiquette :

$$\begin{array}{c} X \\ | \\ \boxed{\mathbf{1}_X} \\ | \\ X \end{array} = \begin{array}{c} X \\ | \\ X \end{array} = \left| \right.$$

La composition de deux flèches $f : X \rightarrow Y$, $g : Y \rightarrow Z$ s’écrit alors comme suit,

$$\begin{array}{c} Z \\ | \\ \boxed{g} \\ | \\ Y \\ | \\ \boxed{f} \\ | \\ X \end{array} \quad \text{ou} \quad \begin{array}{c} Z \\ | \\ \boxed{g} \\ | \\ \boxed{f} \\ | \\ X \end{array},$$

toujours selon si l’objet Y est clair ou non.

De manière analogue à la notion usuelle d’isomorphisme :

Définition 1.2 – On dit qu’un morphisme $f : X \rightarrow Y$ est un *isomorphisme* s’il existe $g : Y \rightarrow X$ un morphisme tel que $f \circ g = \mathbf{1}_Y$ et $g \circ f = \mathbf{1}_X$.

Définition 1.3 – Soient C et D deux catégories. Un *foncteur* $F : C \rightarrow D$ est une application qui à tout objet X de C associe un objet $F(X)$ de D et à tout morphisme $f : X \rightarrow Y$ de C associe un morphisme $F(f) : F(X) \rightarrow F(Y)$ de D , tel que $F(\mathbf{1}_X) = \mathbf{1}_{F(X)}$ pour tout $X \in C$ et $F(g \circ f) = F(g) \circ F(f)$ pour tous $f : X \rightarrow Y$ et $g : Y \rightarrow Z$, avec $X, Y, Z \in C$.

Pour donner des exemples, on peut considérer les foncteurs suivants :

- le foncteur d’oubli de Top vers Set qui envoie un espace topologique sur l’ensemble sous-jacent, et une application continue sur elle-même vue comme une simple application entre ensembles ;
- le foncteur F de la catégorie des anneaux commutatifs CRing dans Grp qui à un anneau associe son groupe des inversibles ;
- le foncteur G de CRing dans Grp qui à un anneau R associe le groupe $\text{GL}_n(R)$, et qui envoie un morphisme sur le morphisme appliqué coefficient par coefficient sur les matrices.

Définition 1.4 – Soient C et D deux catégories et F, G deux foncteurs de C dans D . Une *transformation naturelle* η de F vers G est la donnée pour tout objet X de C d’un morphisme $\eta_X : F(X) \rightarrow G(X)$ tel que, pour tous objets X, Y de C et pour tout morphisme $f : X \rightarrow Y$, le diagramme suivant commute :

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ \eta_X \downarrow & & \downarrow \eta_Y \\ G(X) & \xrightarrow{G(f)} & G(Y) \end{array}$$

On peut vérifier qu’en notant toujours F et G les foncteurs pris en exemple précédemment, le déterminant réalise une transformation naturelle de G vers F (on peut se référer au premier chapitre de [ML78] écrit par Mac Lane pour d’autres exemples de foncteurs et transformations naturelles avec plus d’explications).

Comme pour les isomorphismes parmi les morphismes, on définit :

Définition 1.5 – Soit η une transformation naturelle de F vers G . Si pour tout objet X de C , η_X est un isomorphisme, on dit que η est un *isomorphisme naturel*.

Enfin, on a besoin des notions de *limite* et *colimite*, issues de la notion de *cône* et de *cocône*, afin de définir les notions centrales de produit de Kolmogorov ou de quotient de Markov dans les catégories de Markov et les notions d’état invariant et d’observable invariant dans le cadre de la théorie ergodique.

Définition 1.6 – Soit C une catégorie. On se donne un *diagramme* dans C , i.e. un foncteur $F : I \rightarrow C$ avec I une petite catégorie (au sens où la classe des morphismes est un ensemble). Un *cône* dans F est la donnée d’un objet $X \in C$ et d’une famille $\alpha = (\alpha_i)_{i \in I}$ de morphismes $\alpha_i : X \rightarrow F(i)$ telle que pour tous objets i, j de I et $f : i \rightarrow j$, $\alpha_j = F(f) \circ \alpha_i$.

Définition 1.7 – Une *limite* dans $F : I \rightarrow C$ est un cône (X, α) *universel*, au sens où pour tout autre cône (Y, β) , il existe un unique morphisme $\psi : Y \rightarrow X$ tel que le diagramme suivant commute :

$$\forall i, j \in I, f : i \rightarrow j, \quad \begin{array}{ccc} & Y & \\ \beta_i \swarrow & \downarrow \psi & \searrow \beta_j \\ & X & \\ \alpha_i \swarrow & & \searrow \alpha_j \\ F(i) & \xrightarrow{F(f)} & F(j) \end{array}$$

Un *cocône* (resp. une *colimite*) est simplement la notion duale du cône (de la limite), obtenue en renversant les flèches :

Définition 1.8 – Soit C une catégorie. On se donne un diagramme $F : I \rightarrow C$. Un *cocône* dans F est la donnée d'un objet $X^* \in C$ et d'une famille $\alpha^* = (\alpha_i^*)_{i \in I}$ de morphismes $\alpha_i^* : F(i) \rightarrow X^*$ telle que pour tous objets i, j de I et $f : i \rightarrow j$, $\alpha_i^* = \alpha_j^* \circ F(f)$.

Définition 1.9 – Une *colimite* dans $F : I \rightarrow C$ est un cocône (X^*, α^*) *universel*, au sens où pour tout autre cocône (Y^*, β^*) , il existe un unique morphisme $\psi^* : X^* \rightarrow Y^*$ tels que le diagramme suivant commute :

$$\forall i, j \in I, f : i \rightarrow j,$$

Pour illustrer, regardons par exemple comment le produit s'écrit en termes de limite. Ce cas servira par la suite pour les produits de Kolmogorov.

Définition 1.10 – Soient C une catégorie, $(X_i)_{i \in I}$ une famille d'objets de C . Le *produit* des X_i est la limite, quand elle existe, du diagramme sans autres flèches que les identités. C'est donc un objet X , muni de flèches $p_j : X \rightarrow X_j$ pour tout $j \in I$, tel que pour tout objet Z et toute famille de flèches $(f_i : Z \rightarrow X_i)_{i \in I}$, il existe une unique flèche $f : Z \rightarrow X$ telle que le diagramme suivant commute

$$\forall i \in I,$$

Par exemple, le produit cartésien ensembliste de deux ensembles A et B munis des projections canoniques $A \times B \rightarrow A$ et $A \times B \rightarrow B$ est leur produit cartésien au sens ci-dessus. La définition en terme de propriété universelle dit que se donner une application $Z \rightarrow A \times B$, c'est exactement se donner deux applications $Z \rightarrow A$ et $Z \rightarrow B$.

1.2 Catégories monoïdales symétriques

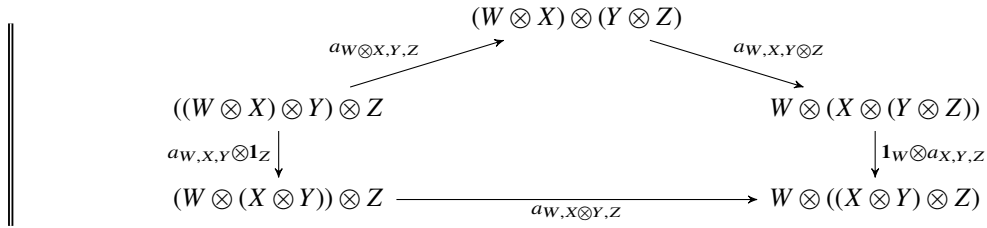
Les catégories de Markov sont des catégories particulières, appelées *catégories monoïdales symétriques*. On rappelle qu'un monoïde au sens usuel est une structure $(C, \otimes, \mathbf{1})$ avec \otimes une loi interne associative et $\mathbf{1} \in C$ un neutre pour \otimes ; on commence donc par implémenter cette notion dans le langage catégorique :

Définition 1.11 – Une *catégorie monoïdale* est la donnée d'une catégorie C , d'un foncteur $\otimes : C \times C \rightarrow C$ appelé *produit tensoriel*, d'un objet $\mathbf{1} \in C$ appelé *unité*, et de trois isomorphismes naturels

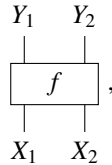
$$a = (a_{X,Y,Z} : ((X \otimes Y) \otimes Z \rightarrow X \otimes (Y \otimes Z)), \quad g = (g_X : \mathbf{1} \otimes X \rightarrow X), \quad d = (d_X : X \otimes \mathbf{1} \rightarrow X)$$

appelés respectivement *associateur*, *neutralisateur gauche* et *neutralisateur droit*. On demande de plus que les diagrammes suivants commutent, appelés respectivement *identité du triangle* et *identité du pentagone* :

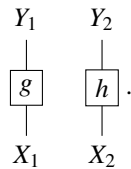
$$\begin{array}{ccc} (X \otimes \mathbf{1}) \otimes Y & \xrightarrow{a_{X,\mathbf{1},Y}} & X \otimes (\mathbf{1} \otimes Y) \\ & \searrow d_X \otimes \mathbf{1}_Y & \swarrow \mathbf{1}_X \otimes g_Y \\ & X \otimes Y & \end{array}$$



Une flèche entre deux produits tensoriels $f : X_1 \otimes X_2 \rightarrow Y_1 \otimes Y_2$ s'écrit



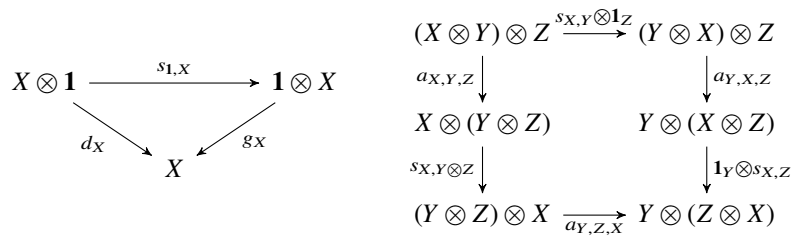
et le produit tensoriel de deux flèches $g \otimes h : X_1 \otimes X_2 \rightarrow Y_1 \otimes Y_2$ s'écrit simplement de manière parallèle



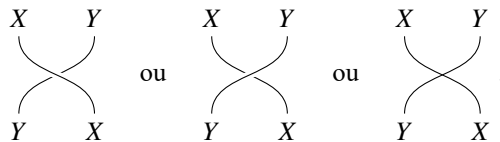
La notation utilisée suggère naturellement la catégorie $(\text{Vect}_k, \otimes, k)$ des k -espaces vectoriels munie du produit tensoriel pour un corps k , mais un exemple plus simple est celle des ensembles, munie du produit cartésien, $(\text{Set}, \times, \{*\})$. On étend sur cette notion de produit cartésien par la suite.

On enrichit ensuite la structure monoïdale par une symétrie assimilant $X \otimes Y$ et $Y \otimes X$. Plus précisément, on définit :

Définition 1.12 – Une catégorie monoïdale $(C, \otimes, \mathbf{1})$ est dite *symétrique* si elle est de plus munie d'un isomorphisme involutif naturel $s = (s_{X,Y} : X \otimes Y \rightarrow Y \otimes X)$ qui est compatible avec la structure de catégorie monoïdale au sens où les diagrammes suivant commutent :



Le morphisme $s_{X,Y}$ se notera alors



l'équivalence entre les trois étant justifiée par les isomorphismes entre $X \otimes Y$ et $Y \otimes X$.

Pour définir les catégories de Markov, nous aurons besoin de la notion de *comonoïde commutatif* dans une catégorie monoïdale.

Définition 1.13 – Soit $(C, \otimes, \mathbf{1})$ une catégorie monoïdale. Un *comonoïde* est la donnée d'un objet C dans C , d'un morphisme $\text{copy}_C : C \rightarrow C \otimes C$ appelé *coproduit* et d'un morphisme $\text{del}_C : C \rightarrow \mathbf{1}$ appelé *counité*, vérifiant les égalités (com) données ci-dessous.

On écrit les morphismes copy_C et del_C comme suit :

$$\text{del}_C = \begin{array}{c} \bullet \\ | \\ C \end{array}, \quad \text{copy}_C = \begin{array}{c} C \quad C \\ \cup \\ \bullet \\ | \\ C \end{array},$$

et, avec ces notations, on demande à ce qu'ils vérifient les égalités de commutativité

$$\begin{array}{c} \cup \\ | \\ \bullet \\ | \\ C \end{array} = \begin{array}{c} \cup \\ \bullet \\ | \\ C \end{array}, \quad \begin{array}{c} \bullet \\ | \\ C \end{array} = \begin{array}{c} \cup \\ | \\ \bullet \\ | \\ C \end{array} = \begin{array}{c} | \\ | \\ | \end{array}, \quad \begin{array}{c} \cup \\ \cup \\ \bullet \\ | \\ C \end{array} = \begin{array}{c} \cup \\ \bullet \\ | \\ C \end{array}. \quad (\text{com})$$

Enfin, les catégories de Markov formeront un exemple de *catégorie semi-cartésienne*.

Définition 1.14 – Une catégorie monoïdale $(C, \otimes, \mathbf{1})$ est dite *semi-cartésienne* si $\mathbf{1}$ est un objet terminal : pour tout objet X , il existe un unique morphisme de X vers $\mathbf{1}$.

Remarquons que le fait que $\mathbf{1}$ soit terminal permet alors de considérer $\text{Id}_X \otimes \text{del}_Y : X \otimes Y \rightarrow X \otimes \mathbf{1} = X$, avec $\text{del}_Y : Y \rightarrow \mathbf{1}$. On obtient donc le lemme suivant :

Lemme 1.15 – Si $(C, \otimes, \mathbf{1})$ est une catégorie semi-cartésienne, pour tous objets X, Y , il existe des morphismes $\alpha_{X,Y} : X \otimes Y \rightarrow X$ et $\beta_{X,Y} : X \otimes Y \rightarrow Y$ appelés *marginalisations* qui coïncident avec les isomorphismes naturels g et d quand $X = \mathbf{1}$ ou $Y = \mathbf{1}$. Ces morphismes sont *naturels*, au sens où $(\alpha_{X,Y_0})_{X \in C}$ et $(\beta_{X_0,Y})_{Y \in C}$, avec X_0 et Y_0 fixés, sont deux transformations naturelles $(\cdot \otimes Y_0) \Rightarrow \text{Id}$ et $(X_0 \otimes \cdot) \Rightarrow \text{Id}$.

En effet, les marginalisations sont données par $\alpha_{X,Y} = d_X \circ (\mathbf{1}_X \otimes \text{del}_Y)$ où $\text{del}_Y : Y \rightarrow \mathbf{1}$ est l'unique flèche de Y vers $\mathbf{1}$. Ce lemme montre une des manières dont les catégories semi-cartésiennes sont intéressantes en théorie catégorique des probabilités : elles vont nous permettre de généraliser la notion de loi marginale.

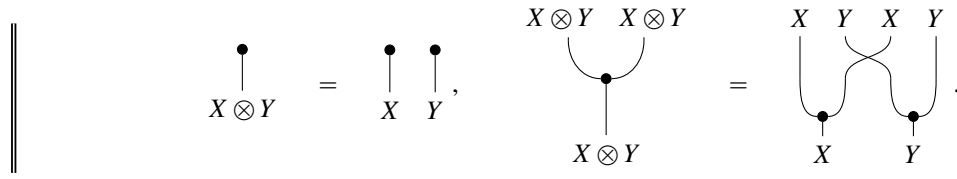
2 Catégories de Markov

Dans cette section, on définit les catégories de Markov, qui constituent le cadre formel proposé dans les articles étudiés pour généraliser les probabilités dans le langage des catégories. On présente les exemples types de catégories de Markov, ainsi que les principales notions qui seront utilisées par la suite.

2.1 Définition et premiers exemples

Avec les notions de catégories monoïdales, symétriques, semi-cartésiennes, et de comonoïde, on peut définir la notion tant attendue de catégorie de Markov :

Définition 2.1 – Une *catégorie de Markov* est une catégorie monoïdale symétrique semi-cartésienne $(C, \otimes, \mathbf{1})$ pour laquelle chaque objet est muni d'une structure de comonoïde commutatif compatible à la structure monoïdale, au sens où l'on a pour tous objets X, Y ,



Introduisons deux exemples de catégories de Markov qui serviront de prototypes, l'un sur les ensembles finis, l'autre sur les espaces mesurables.

Exemple 2.2 On définit FinStoch comme la catégorie dont les objets sont les ensembles finis, et dont les morphismes sont donnés par les matrices stochastiques. Autrement dit, un morphisme $f : X \rightarrow Y$ est une fonction $X \times Y \rightarrow [0, 1]$, $(x, y) \mapsto f(y|x)$ avec pour tout $x \in X$, $\sum_{y \in Y} f(y|x) = 1$. La quantité $f(y|x)$ représente une probabilité de transition de $x \in X$ à $y \in Y$. La composée de deux morphismes $f : X \rightarrow Y$ et $g : Y \rightarrow Z$ est donnée par la multiplication usuelle des matrices, c'est-à-dire par

$$(g \circ f)(z|x) = \sum_{y \in Y} g(z|y)f(y|x),$$

pour $x \in X, z \in Z$. La structure comonoïdale pour un objet X est donnée par $\text{del}_X(*|x) = 1$ pour tout $x \in X$, et pour $x_0, x_1, x_2 \in X$, $\text{copy}_X((x_1, x_2)|x_0) = \mathbf{1}_{x_0=x_1=x_2}$. On peut montrer qu'avec la structure monoïdale donnée par le produit cartésien usuel d'ensembles, donc notamment $(f \otimes g)((a, b)|(x, y)) = f(a|x)g(b|y)$, FinStoch présente bien une structure de catégorie de Markov.

Exemple 2.3 On définit aussi Stoch comme la catégorie dont les objets sont les espaces mesurables (X, Σ_X) et les morphismes $f : X \rightarrow Y$ sont les fonctions $X \times \Sigma_Y$, $(x, S) \mapsto f(S|x)$ telles qu'à x fixé, f est une mesure de probabilité et à S fixé, f est mesurable. On peut aussi voir f comme une famille $(f_x)_{x \in X}$ de mesures de probabilité sur l'espace mesurable Y , variant de manière mesurable en x . On dira que f est un *noyau de Markov*. La composée de deux noyaux de Markov $f : X \rightarrow Y$ et $g : Y \rightarrow Z$ est donnée par une généralisation du produit matriciel :

$$(g \circ f)(S|x) = \int_Y g(S|y) d f_x(y)$$

pour $S \in \Sigma_Z$ et $x \in X$. À x fixé, on a bien une mesure de probabilité : on envoie bien \emptyset sur 0, et on vérifie la σ -additivité en utilisant le fait que $g(\cdot|y)$ est une mesure à y fixé, puis en dominant l'intégrale par 1. Pour la mesurabilité de $(g \circ f)(S|\cdot)$ à S fixé, on peut le vérifier quand $g(S|\cdot) = \mathbf{1}_A$, A mesurable, puis pour des fonctions étagées, puis pour tout $g(S|\cdot)$ en revenant à la définition de l'intégrale de Lebesgue.

On définit, pour tout objet X , les morphismes

$$\text{del}_X(*|x) = 1 \quad \text{et} \quad \text{copy}_X(S_1 \times S_2|x) = \mathbf{1}_x(S_1 \cap S_2),$$

pour $x \in X$ et $X_1, X_2 \in \Sigma_X$. En munissant Stoch du produit tensoriel donné par les notions de tribu produit et mesure produit, comme dans FinStoch, on montre que Stoch est bien une catégorie de Markov, comme le montre Fritz dans [Fri20, section 4].

Il est bon de remarquer que dans Stoch, les morphismes sont des généralisations à la fois de la notion de mesure de probabilité et de la notion de fonction mesurable. En effet, une mesure de probabilité sur X revient à se donner un morphisme $p : \mathbf{1} \rightarrow X$. Si $f : X \rightarrow Y$ est une fonction mesurable, on peut la voir comme le morphisme $\tilde{f} : X \rightarrow Y$ défini par $\tilde{f}(B|x) = 1$ si $f(x) \in B$ et 0 sinon. On définit ainsi bien un noyau de Markov car f est supposée mesurable.

On utilisera également par la suite la catégorie BorelStoch, qui est une sous-catégorie pleine de Stoch constituée des espaces de Borel standards, c'est-à-dire des espaces polonais munis de leur tribu borélienne.

Rappelons qu'un espace polonais est un espace métrique complet séparable, c'est-à-dire qui admet une partie dénombrable dense. Cette catégorie est aussi une catégorie de Markov : le produit tensoriel est donné par le produit tensoriel de Stoch, bien défini car un produit fini d'espaces polonais est polonais. Cette catégorie se comporte en général mieux que Stoch, comme présenté dans la suite du mémoire (par exemple, les *produits de Kolmogorov* y sont bien définis, ses objets admettent des désintégrations, deux notions définies plus tard).

2.2 Notions probabilistes dans les catégories de Markov

Le formalisme des catégories de Markov ainsi posé permet de définir des notions qui encodent les propriétés probabilistes usuelles. On fixe pour les définitions une catégorie de Markov C .

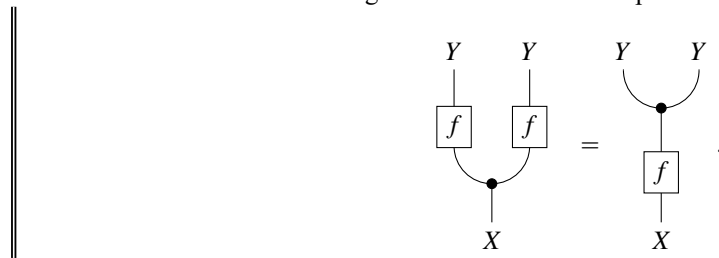
Définition 2.4 – Soit X un objet de C . On appelle *mesure* ou *loi* sur X un morphisme $p : \mathbf{1} \rightarrow X$.

En termes de diagrammes de cordes, on note un tel morphisme :



Comme il a été remarqué plus tôt, dans Stoch et FinStoch, une mesure sur X est exactement une mesure de probabilité sur X au sens usuel. Regardons donc comment se traduisent les autres objets usuels ; dès lors que les morphismes des catégories de Markov implémentent une notion de « fonction aléatoire », on s'intéressera aussi aux « fonctions » usuelles :

Définition 2.5 – Soit C une catégorie de Markov. Un morphisme $f : X \rightarrow Y$ est dit *déterministe* si



Lemme 2.6 – On peut former la sous-catégorie C_{det} ayant les mêmes objets que C et dont les morphismes sont donnés par les morphismes déterministes de C . Cette catégorie est monoïdale symétrique.

Dans FinStoch, on vérifie que les morphismes déterministes sont ceux donnés par des matrices stochastiques ne contenant que des 0 et des 1. Autrement dit, un morphisme déterministe $f : X \rightarrow Y$ est simplement une fonction au sens usuel, et $\text{FinStoch}_{\text{det}} = \text{FinSet}$.

Dans Stoch, les morphismes déterministes sont les $f : X \rightarrow Y$ pour lesquels $f(S|x) \in \{0, 1\}$ pour tout $S \in \Sigma_Y$ et tout $x \in X$. En effet, l'égalité des morphismes de la définition s'écrit $f(S \cap T|x) = f(S|x)f(T|x)$ pour tout $x \in X$ et $S, T \in \Sigma_Y$. En prenant $S = T$, on obtient le résultat voulu. Réciproquement, ce sont les seuls par additivité des mesures de probabilité. Autrement dit, les morphismes déterministes sont ceux qui n'impliquent pas d'aléatoire. Par exemple, une fonction mesurable vue comme morphisme de Stoch est un morphisme déterministe.

Définition 2.7 – Soient X et Y deux objets de C et ψ une mesure sur X . On appelle *variable aléatoire* sur (X, ψ) un morphisme $f : X \rightarrow Y$ déterministe. La *loi de la variable aléatoire* f est alors $f \circ \psi : \mathbf{1} \rightarrow Y$ qui est donc une mesure sur Y .

Définition 2.8 – Une *loi jointe* est un morphisme $\psi : \mathbf{1} \rightarrow X \otimes Y$. Les *lois marginales* sont les morphismes $\| (\mathbf{1}_X \otimes \text{del}_Y)\psi$ et $(\text{del}_X \otimes \mathbf{1}_Y)\psi$.

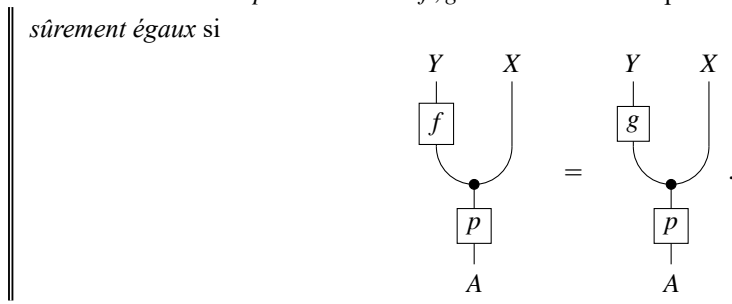
Dans Stoch, une loi jointe correspond à une mesure de probabilité sur l'espace mesurable produit muni de la tribu produit. En effet, si l'on se donne ψ une probabilité sur $X \otimes Y$, alors pour $A \in \Sigma_X$,

$$\begin{aligned} (\mathbf{1}_X \otimes \text{del}_Y)\psi(A) &= \int_{X \times X} (\mathbf{1}_X \otimes \text{del}_Y)(A|(x_1, x_2)) \, d\psi(x_1, x_2) \\ &= \int_{X \times X} \mathbf{1}_{x_1}(A) \, d\psi(x_1, x_2) \\ &= \psi(A \times X), \end{aligned}$$

ce qui est bien la mesure de A selon la première marginale de ψ au sens usuel.

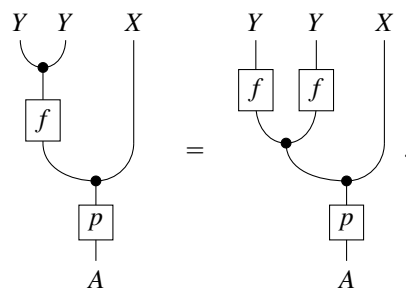
Remarquons de plus que dans Stoch toujours, avec $Y = \{0, 1\}$, un morphisme déterministe $f : X \rightarrow Y$ revient à se donner $S \in \Sigma_X$, autrement dit un évènement. Plus précisément, on peut prendre l'évènement $S = \{x \in X : f(\{1\}|x) = 1\}$. Le morphisme $f\psi$ donne alors la probabilité de l'évènement.

Définition 2.9 – Soit $p : A \rightarrow X$ et $f, g : X \rightarrow Y$ des morphismes. On dit que f et g sont *p-presque sûrement égaux* si



Dans Stoch avec des hypothèses raisonnables¹, on retrouve alors la notion usuelle d'égalité presque sûre entre variables aléatoires comme montré dans [Fri20, Exemple 13.3].

Plus généralement, si (P) est une propriété d'un morphisme f définie par une égalité de morphisme et p est une mesure, on dit que f vérifie (P) *p-presque sûrement* si les morphismes en question sont *p-presque sûrement égaux*. Par exemple, nous dirons que $f : X \rightarrow Y$ est *p-presque sûrement déterministe* si on a

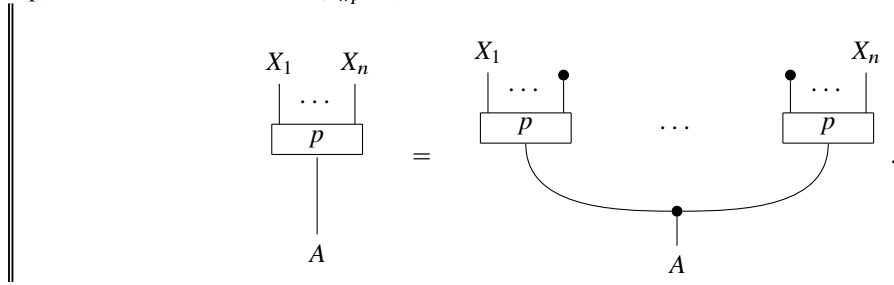


Enfin, on peut formuler dans ce cadre la notion d'indépendance entre objets, en utilisant les outils précédents.

Définition 2.10 – Soit $p : A \rightarrow X_1 \otimes \dots \otimes X_n$. On dit que X_1, \dots, X_n sont *indépendants par rapport à*

1. à savoir que l'espace probabilisé doit avoir sa tribu générée par un système dénombrable.

p , et on note $X_1 \perp \cdots \perp X_n \parallel_p A$, si



Dans Stoch, dire que deux espaces mesurables X et Y sont conditionnellement indépendants par rapport à $p : A \rightarrow X \otimes Y$ revient à dire que pour tout $T \in \Sigma_X$ et $S \in \Sigma_Y$, et pour tout $a \in A$, $p(T \times S | a) = p_X(T | a) p_Y(S | a)$, où p_X et p_Y désignent les marginales.

3 Produits dans les catégories de Markov et loi du 0-1 de Kolmogorov

Dans cette section, on introduit la notion de produit infini dans une catégorie de Markov, nécessaire pour modéliser des suites infinies de variables aléatoires, ce qui servira dans notre cas à formuler la loi du 0-1 de Kolmogorov.

3.1 Produits dans les catégories de Markov

Soit \mathcal{C} une catégorie de Markov et $(X_i)_{i \in J}$ une famille d'objets. Pour tout $F \subset J$ fini, on note $X_F = \bigotimes_{i \in F} X_i$ (qui est bien défini isomorphisme naturel près car \mathcal{C} est symétrique). Si de plus $L \subset F \subset J$ sont finis, on notera $\pi_L^F : X_F \rightarrow X_L$ les morphismes de marginalisation donnés par la structure semi-cartésienne de notre catégorie. On veut maintenant définir le produit tensoriel pour une famille infinie d'objets :

Définition 3.1 – On définit $X_J = \bigotimes_{i \in J} X_i$ comme la limite (sous condition d'existence) du diagramme $(F \mapsto X_F)_{F \subset J \text{ fini}}$, quand elle est préservée par le foncteur $X \mapsto X \otimes Y$ pour tout objet Y , au sens où

$$X_J \otimes Y = \bigotimes_{j \in J \cup \{*\}} X_j,$$

avec $* \notin J$ et $X_* = Y$.

La limite, si elle existe, vient avec les marginalisations finies $\pi_F : X_J \rightarrow X_F$.

Dans Stoch, la notion de produit tensoriel infini traduit le théorème d'extension de Kolmogorov. Ainsi, dans Stoch, les produits tensoriels infinis n'existent pas toujours et l'existence est donnée par les conditions d'application du théorème d'extension de Kolmogorov, dont on trouvera une démonstration dans [Kle14, théorème 14.35]. Ici, $T_{\#} \mu$ désigne, pour $T : X \rightarrow Y$ mesurable et μ une mesure sur X , la mesure image $\mu \circ T^{-1}$.

Théorème 3.2 (d'extension de Kolmogorov dénombrable) – Soient $(\Omega_i, \mathcal{A}_i)_{i \in J}$ un ensemble dénombrable d'espaces de Borel. On pose, pour $F \subset J$ fini, $\Omega_F = \prod_{i \in F} \Omega_i$ et $\mathcal{A}_F = \bigotimes_{i \in F} \mathcal{A}_i$, et $\Omega = \prod_{i \in J} \Omega_i$ et $\mathcal{A} = \bigotimes_{i \in J} \mathcal{A}_i$. Soit $(P_F : \mathcal{A}_F \rightarrow [0, 1] : F \subset J \text{ fini})$ une famille compatible de mesures de probabilités, au sens où pour tous $L \subset F \subset J$ finis $P_L = (\pi_L^F)_{\#} P_F$, avec $\pi_L^F : \Omega_F \rightarrow \Omega_L$ la projection canonique. Alors il existe probabilité P sur (Ω, \mathcal{A}) telle que $P_F = (\pi_F)_{\#} P$ pour tous $F \subset J$ fini, avec $\pi_F : \Omega \rightarrow \Omega_F$ la projection canonique.

Autrement dit, dans BorelStoch, les produits tensoriels dénombrables existent. Le cas pathologique est celui de FinStoch, où les produits infinis n'existent a priori pas : on se donne une famille $(X_i)_{i \in J}$, avec

$\#X_i \geq 2$ pour tout i . Alors on peut montrer que si $\bigotimes_{j \in J} X_j$ existait dans FinStoch , il ne serait pas fini, donc pas un objet de FinStoch (voir [FR20, Exemple 3.7] de Fritz).

Les lemmes suivants, dont on trouvera les preuves dans [FR20, lemmes 3.9 et 3.10], montrent que les produits tensoriels infinis se comportent bien en prenant des réunions de familles.

Lemme 3.3 – Si $J = J_1 \sqcup J_2$, alors $X_{J_1} \otimes X_{J_2}$ est un produit infini X_J avec pour morphismes de marginalisation les $(\pi_{F \cap J_1} \otimes \pi_{F \cap J_2})_{F \subset J}$ fini.

L'idée de la preuve est de d'abord fixer $F_1 \subset J_1$, regarder $X_{F_1} \otimes X_{J_2}$ la colimite des $X_{F_1} \otimes X_{F_2}$, $F_2 \subset J_2$ finis, avec les projections $\text{Id} \otimes \pi_{F_2} : X_{F_1} \otimes X_{J_2} \rightarrow X_{F_1} \otimes X_{F_2}$. Puis, en laissant varier F_1 , regarder $X_{J_1} \otimes X_{J_2}$ comme la colimite des $X_{F_1} \otimes X_{J_2}$, pour conclure en étudiant ce que la limite globale sur les F_1 et F_2 donne.

Ce fait se généralise alors :

Lemme 3.4 – Soit $(X_{j,k})_{j \in J_k, k \in K}$ une famille d'objets. Si pour tout $k \in K$, les produits tensoriels $\bigotimes_{j \in J_k} X_j$ et $\bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{j,k}$ existent, alors $\bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{j,k}$ est un produit infini $\bigotimes_{k \in K, j \in J_k}$ et les morphismes de marginalisation sont donnés par $\bigotimes_{k \in F} \pi_{F_k}$ où $F \subset K$ fini et $F_k \subset J_k$ finis.

On n'a pas utilisé pour l'instant que notre catégorie est de Markov. La bonne notion de produit dans les catégories de Markov est celle de *produit de Kolmogorov*, notamment pour l'étude de lois asymptotiques comme on le verra pour la loi du 0-1 de Kolmogorov.

Définition 3.5 – On dit que $\bigotimes_{i \in J} X_i$ est un *produit de Kolmogorov* si les marginalisations $\pi_F : X_J \rightarrow X_F$ sont déterministes.

C'est le résultat suivant qui nous permettra d'appliquer la loi du 0-1 abstraite pour obtenir la loi du 0-1 de Kolmogorov dans les espaces de Borel standard.

Propriété 3.6 – Dans BorelStoch , les produits tensoriels dénombrables sont de Kolmogorov.

En effet, les marginalisations $\pi_F : X_J \rightarrow X_F$ sont les projections canoniques ensemblistes, donc sont déterministes.

Propriété 3.7 – Un produit de Kolmogorov X_J est un produit cartésien dans C_{det}

Fritz démontre ce résultat dans [Fri20, proposition 4.3], encore une fois.

3.2 Loi du 0-1 de Kolmogorov, version catégorique

On peut maintenant formuler une version abstraite de la loi du 0-1 de Kolmogorov, redémontrer la loi usuelle dans les cas classiques. On garde les mêmes notations.

Définition 3.8 – Soit $\bigotimes_{i \in J} X_i$ un produit de Kolmogorov et soit $p : A \rightarrow \bigotimes_{i \in J} X_i$. On dit que la famille (X_j) est *indépendante par rapport à p* , et on note $\perp_{i \in J} X_i \parallel_p A$, si pour tout $F \subset J$ fini, on a $\perp_{i \in F} X_i \parallel_{\pi_F \circ p} A$

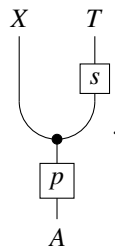
Avant d'énoncer la loi du 0-1 de Kolmogorov dans les catégories de Markov, nous avons besoin de deux lemmes : le lemme d'indépendance infinie et le lemme de déterminisme. Pour mieux comprendre ce que signifient ces lemmes, on pourra se placer dans la catégorie BorelStoch et prendre $A = \{*\}$.

Lemme 3.9 (d'indépendance infinie) – Soient $\bigotimes_{i \in J} X_i$ un produit de Kolmogorov, $p : A \rightarrow \bigotimes_{i \in J} X_i$. Si $\perp_{i \in J} X_i \parallel_p A$, alors pour tout $i \in J$, $X_i \perp X_{J-\{i\}} \parallel_p A$.

Fritz détaille la démonstration dans [FR20, lemme 5.1]. L'équivalent probabiliste de ce lemme est un cas particulier du lemme des coalitions.

Lemme 3.10 (de déterminisme) – Soit $p : A \rightarrow X$ et soit $s : X \rightarrow T$ déterministe. Soit $f : A \rightarrow X \otimes T$

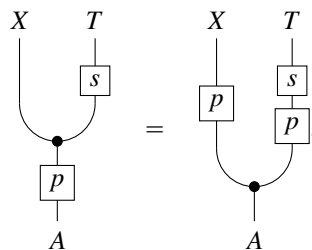
le morphisme représenté par



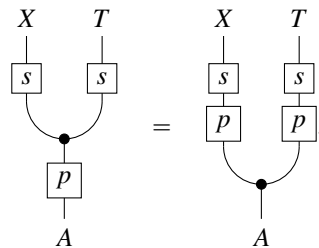
Si $X \perp T \parallel_f A$, alors sp est déterministe.

On donne la démonstration de ce lemme afin de donner une idée de comment manipuler les diagrammes de cordes.

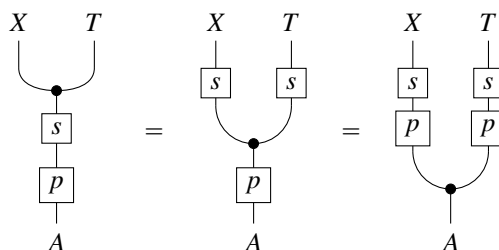
Démonstration. On se place dans les hypothèses du lemme. On écrit $X \perp T \parallel_f A$ en termes de diagrammes :



En composant par $s \otimes \mathbf{1}_T$, on obtient :



Puis on utilise que s est déterministe :



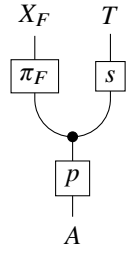
ce qui montre que sp est déterministe. □

Pour avoir une idée de la signification probabiliste de ce lemme, on se place dans Stoch. Si on prend $T = \{0, 1\}$, p représente une probabilité sur X et s représente un évènement. Le lemme dit alors que si un évènement est indépendant de la tribu support de la probabilité, alors sa probabilité est 0 ou 1 (sp donne la probabilité de l'évènement s , donc sp déterministe signifie que la probabilité de l'évènement s est 0 ou 1).

On peut maintenant énoncer et démontrer la loi du 0-1 de Kolmogorov :

Théorème 3.11 (loi du 0-1 de Kolmogorov abstraite) – Soit $p : A \rightarrow X_J$ et soit $s : X_J \rightarrow T$ déterministe. On suppose que $\perp_{i \in J} X_i \parallel_p A$ et que pour tout $F \subset J$ fini, $X_F \perp T \parallel A$ par rapport au morphisme

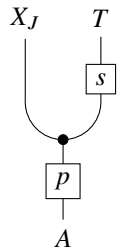
représenté par



Alors $sp : A \rightarrow T$ est déterministe.

La preuve reprend la preuve classique de la loi du 0-1 de Kolmogorov.

Démonstration. Par définition, $\perp_{i \in J \cup \{*\}} X_i \parallel_f A$ où $X_* = T$ et où f est le morphisme



Par le premier lemme, $X_J \perp T \parallel_f A$. Et par le deuxième lemme, sp est déterministe. \square

La deuxième hypothèse s'interprète comme l'indépendance de l'évènement s avec toute sous famille finie du produit de Kolmogorov, ce qui revient à dire dans le cadre classique que l'évènement est dans la tribu asymptotique. On peut déduire de ce théorème la loi du 0-1 de Kolmogorov usuelle, à ceci près qu'on se place dans les espaces de Borel (afin de pouvoir utiliser les produits de Kolmogorov).

Corollaire 3.12 – Soit (Ω, \mathcal{F}, P) un espace de Borel, $(X_i, \mathcal{F}_i)_{i \in \mathbb{N}}$ des espaces de Borel, $f_i : \Omega \rightarrow X_i$ des variables aléatoires indépendantes. On note $\mathcal{F}_\infty = \bigcap_{n \in \mathbb{N}} \sigma(f_i, i \geq n)$ la tribu asymptotique engendrée par les f_i . Alors $P(\mathcal{F}_\infty) \subset \{0, 1\}$.

Démonstration. On applique le théorème 3.11 à la catégorie BorelStoch en prenant $J = \mathbb{N}$, $A = \mathbf{1}$ et $T = \{0, 1\}$. Soit $E \in \mathcal{F}_\infty$ et notons $f_{\mathbb{N}} : \Omega \rightarrow X_{\mathbb{N}} = \prod_{i \in \mathbb{N}} X_i$ la variable aléatoire à valeurs dans $X_{\mathbb{N}}$ induite par les f_i , bien définie car les limites existent dans la catégorie BorelStoch.

Par définition de la tribu asymptotique, il existe $F \in \Sigma_{X_{\mathbb{N}}}$ tel que $E = f_{\mathbb{N}}^{-1}(F)$. On considère alors s la fonction indicatrice de F vu comme un morphisme de BorelStoch, et on prend $p = f_{\mathbb{N}} \circ P$. Comme les f_i sont indépendantes, les hypothèses du théorème précédent sont vérifiées et $sp : \mathbf{1} \rightarrow \{0, 1\}$ est déterministe, ce qui, comme on l'a vu, signifie que la probabilité de l'évènement représenté par s vaut 0 ou 1. Or, cet évènement n'est autre que $E : p(E) \in \{0, 1\}$, donc par définition de p et E , $P(F) \in \{0, 1\}$. \square

La question de l'existence d'une catégorie de Markov avec des produits de Kolmogorov dénombrables qui permette de démontrer la loi du 0-1 de Kolmogorov dans sa généralité (et pas seulement pour les espaces de Borel) reste ouverte.

4 Théorie ergodique par les catégories de Markov

Dans cette section, on s'intéresse à une autre application des catégories de Markov, qui ici vont modéliser des systèmes dynamiques. De la même manière qu'on a redéfini le formalisme probabiliste en termes catégoriques, on commence ici par définir la notion de système dynamique par les catégories.

4.1 Systèmes dynamiques dans le langage catégorique

Un système dynamique est classiquement un espace X muni d'une certaine structure et une application $T : X \rightarrow X$ préservant cette structure, induisant donc un monoïde (au sens d'un ensemble muni d'une loi associative avec un élément neutre) agissant sur X . On se donne donc un monoïde M , et on note BM la catégorie associée : un unique objet $*$ pour lequel les flèches $* \rightarrow *$ sont les éléments de M .

Définition 4.1 – Un *système dynamique* est un foncteur de BM dans une catégorie C , soit la donnée d'un objet $X \in C$ et pour tout $m \in M$, d'une flèche $X \rightarrow X$, encore notée m .

On s'intéresse alors à deux notions : les points fixes sous l'action du monoïde et les applications invariantes sous celle-ci. Dans des catégories ensemblistes, le premier est l'ensemble des $x \in X$ tels que $m(x) = x$ pour tout $m \in M$, le second est l'ensemble des flèches $r : X \rightarrow R$ telles que $r(m(x)) = r(x)$ pour tous $x \in X$ et $m : X \rightarrow X$. D'un point de vue catégorique, ces notions s'implémentent sous forme de limite et colimite :

Définition 4.2 – L'*espace des états invariants* de X , noté X^{inv} est, si elle existe, la limite du diagramme $BM \rightarrow C$.

Dans Set , X^{inv} est l'ensemble $\{x \in X \mid \forall m \in M, m(x) = x\}$ avec l'inclusion dans X canonique. C'est au sens propre l'espace des états laissés invariants sous l'action de M .

Définition 4.3 – L'*espace des orbites*, noté X_{inv} est, si elle existe, la colimite du foncteur $BM \rightarrow C$.

Dans Set , X_{inv} est l'ensemble des orbites sous l'action de M (autrement dit, le quotient X/M de X par l'action de M), avec la projection canonique.

Dans une catégorie de Markov C , nous aurons des notions plus spécifiques qui rendent compte de la structure de catégorie de Markov.

Définition 4.4 – Un *état invariant* est une mesure $p : \mathbf{1} \rightarrow X$ telle que $m \circ p = p$ pour tout $m \in M$.

Dans Stoch , si les morphismes $m \in M$ sont induits par des fonctions mesurables $\tilde{m} : X \rightarrow X$, un état invariant est exactement une mesure invariante μ , au sens de la théorie des systèmes dynamiques : pour tout $A \in \Sigma_X$ et tout $m \in M$, $\mu(\tilde{m}^{-1}(A)) = \mu(A)$. En effet, la mesure $m \circ p : \mathbf{1} \rightarrow X$ est définie par $m \circ \mu(A) = \int_X m(A|x) d\mu(x)$. Or par hypothèse, $m(A|x) = 1_{\tilde{m}(x) \in A}$, d'où $m \circ \mu(A) = \mu(\tilde{m}^{-1}(A))$.

Définition 4.5 – Une *fonction invariante* est un morphisme $r : X \rightarrow R$ tel que pour tout $m \in M$, $r \circ m = r$.

On aboutit donc à la notion duale annoncée :

Définition 4.6 – Le *quotient de Markov* de X sous l'action de M est, si elle existe, la colimite (X_{inv}, r) du foncteur $BM \rightarrow C$ à la fois dans C et dans C_{det} .

Le morphisme $r : X \rightarrow X_{\text{inv}}$ donné par la colimite est donc déterministe.

Propriété 4.7 – Dans Stoch , si (X, M) est un système dynamique déterministe (i.e. tous les $m \in M$ forment des morphismes déterministes), alors le quotient de Markov existe et est donné par l'objet (X, Σ_{inv}) où Σ_{inv} est la tribu invariante, formée de l'ensemble des $A \in \Sigma_X$ tels que pour tout $m \in M$, $m(A|x) = 1_{x \in A}$. Le morphisme de projection sur le quotient $r : X \rightarrow X_{\text{inv}}$ est donné par $r(A|x) = 1_A(x)$ pour tout $A \in \Sigma_{\text{inv}}$ et $x \in X$.

Ce résultat s'établit en vérifiant la propriété universelle à la fois dans Stoch et $\text{Stoch}_{\text{det}}$, comme Moss et Perrone le font dans [MP23, proposition 3.7].

On introduit maintenant la notion d'ergodicité dans le langage des catégories.

Définition 4.8 – Un état invariant $p : \mathbf{1} \rightarrow X$ est dit *ergodique* si pour tout $c : X \rightarrow R$ invariant et déterministe, $c \circ p$ est déterministe.

Vu la définition de quotient de Markov comme colimite, si ce quotient existe, on peut caractériser l'ergodicité en ne regardant que le morphisme $r \circ p$ où r est la projection sur le quotient de Markov, ce qui est expliqué dans le lemme suivant.

Lemme 4.9 – Supposons que le quotient de Markov existe. Alors un état invariant $p : \mathbf{1} \rightarrow X$ est ergodique si et seulement si $r \circ p$ est déterministe où (X_{inv}, r) est le quotient de Markov de X .

Démonstration. Supposons d’abord que $r \circ p$ est déterministe. Tout $c : X \rightarrow R$ invariant et déterministe se factorise en $c = r \circ \tilde{c}$, avec $\tilde{c} : X \rightarrow X_{\text{inv}}$ déterministe. Alors $c \circ p = \tilde{c} \circ (r \circ p)$ l’est lui aussi. La réciproque se démontre en appliquant l’hypothèse avec $r : X \rightarrow X_{\text{inv}}$ pour retrouver la définition de l’ergodicité. \square

Comparons avec la notion usuelle de mesure ergodique. Au sens de la théorie des systèmes dynamiques, une mesure p sur un ensemble X muni d’une tribu Σ_X , sous l’action d’un monoïde M de fonctions mesurables, est dite ergodique si pour tout $A \in \Sigma_{\text{inv}}$, c’est-à-dire $m^{-1}(A) = A$ pour tout $m \in M$, $p(A) \in \{0, 1\}$. On a déjà vu que dans Stoch, une mesure est déterministe si et seulement si son image est $\{0, 1\}$. Avec le lemme précédent et la caractérisation du quotient de Markov dans Stoch dans le cas où le monoïde est constitué de fonctions mesurables, on voit que les mesures ergodiques au sens catégorique correspondent exactement aux mesures ergodiques au sens des systèmes dynamiques.

Enfin, pour énoncer le théorème de décomposition ergodique, nous aurons besoin de la notion de morphisme presque-sûrement ergodique, qui n’est pas définie a priori, car l’ergodicité n’est pas définie par une simple égalité de morphisme.

Définition 4.10 – Soit $q : \mathbf{1} \rightarrow Y$ une mesure et $k : Y \rightarrow X$ un morphisme. On dit que k est q -presque sûrement ergodique si k est q -presque sûrement invariante et pour tout $r : X \rightarrow R$ déterministe et invariant, $r \circ k$ est q -presque sûrement déterministe.

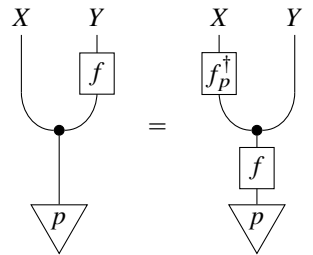
On a l’équivalent du lemme 3.10, démontré par Moss et Perrone dans [MP23, proposition 3.13] :

Lemme 4.11 – Soit (X, M) un système dynamique. On suppose que le quotient de Markov (X_{inv}, r) existe. Soit q une mesure sur X et $k : Y \rightarrow X$ un morphisme q -presque sûrement invariante. Alors k est q -presque sûrement ergodique si et seulement si $r \circ k$ est q -presque sûrement déterministe.

Dans Stoch, un morphisme $k : Y \rightarrow X$ est donc q -presque sûrement ergodique si et seulement si les mesures $k(\cdot | y)$ sont ergodiques pour q -presque tout $y \in Y$. En effet, la condition que $r \circ k$ est q -presque déterministe dit exactement que pour q -presque tout $y \in Y$, la mesure $(r \circ k)(\cdot | y)$ est à valeurs dans $\{0, 1\}$. Puisque r n’est que la projection $X \rightarrow X_{\text{inv}}$, c’est équivalent à demander à ce que pour tout $A \in \Sigma_{\text{inv}}$, $k(A|y) \in \{0, 1\}$, c’est-à-dire $k(\cdot | A)$ ergodique.

4.2 Inversion bayésienne, désintégrations

Définition 4.12 – Soit f une flèche de $X \rightarrow Y$ et p une mesure sur X . On dit que f admet une désintégration selon p s’il existe un morphisme $f_p^\dagger : Y \rightarrow X$ tel que



La source p représente le choix d’un paramètre $x \in X$, qui donne une loi pour f , et f est alors une expérience dépendant de ce paramètre. En tant qu’observateur, on n’aurait accès qu’à la donnée de la sortie de f . La désintégration f_p^\dagger permet alors de déterminer la loi du paramètre représenté par p . En effet, pour $A \subset X$ et $B \subset Y$ mesurables, l’égalité de diagramme ci-dessus s’écrit

$$\int_A f(B|x) dp(x) = \int_B f_p^\dagger(A|y) d(f \circ p)(y).$$

Avec $B = \{y_0\}$ où $(f \circ p)(\{y_0\}) > 0$, on obtient une formule de Bayes

$$f_p^\dagger(A|y_0) = \frac{\int_A f(\{y_0\})|x) dp(x)}{(f \circ p)(\{y_0\})},$$

à comparer avec la formule de Bayes usuelle

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Une désintégration est ainsi aussi appelée *inverse bayésien*.

Pour l'appliquer, on utilise notamment le théorème suivant ici admis, mais que Tao démontre dans [Tao08] (théorème 4). On rappelle qu'un *espace de Borel standard* est un espace polonais muni de sa tribu borélienne et qu'un espace mesurable est dit *régulier* s'il admet une métrique en faisant un espace de Borel standard. Par exemple, le disque dans \mathbb{R}^2 , muni de sa tribu borélienne, est régulier. De plus, on note $f_\#p$ la mesure image d'une mesure p par une fonction mesurable f . Remarquons qu'en notant \tilde{f} le noyau de Markov déterministe associé à f , la mesure $f_\#p$ s'identifie $\tilde{f} \circ p$.

Théorème 4.13 (de Rokhlin partiel) – Soient (X, p) et (Y, q) deux espaces munis de mesures de probabilité, avec (X, p) un espace régulier. Soit $f : X \rightarrow Y$ une fonction mesurable telle que $q = f_\#p$. Alors il existe un noyau de Markov $g : Y \rightarrow X$, $y \mapsto g(\cdot | y) = g_y$ tel que

$$\int_X \varphi(\psi \circ f) dp = \int_Y \left(\int_X \varphi dg_y \right) \psi(y) dq(y),$$

pour toutes fonctions $\varphi : X \rightarrow \mathbb{C}$ et $\psi : Y \rightarrow \mathbb{C}$ mesurables bornées.

Dans BorelStoch, ce théorème dit que tout espace de Borel standard *admet des désintégrations*, au sens où pour tout X espace de Borel standard et $p : \mathbf{1} \rightarrow X$, tout $f : X \rightarrow Y$ déterministe admet une désintégration selon p . En effet, si on se donne de tels X , p et f , on munit Y de la mesure image $q = f \circ p$ et il existe par le théorème de Rokhlin un noyau de Markov $g : Y \rightarrow X$ tel que, en particulier avec $\varphi = \mathbf{1}_A$ et $\psi = \mathbf{1}_B$, $A \subset X$ et $B \subset Y$ mesurables,

$$\int_A f(B|x) dp(x) = \int_X \mathbf{1}_A(\mathbf{1}_B \circ f) dp = \int_Y \left(\int_X \mathbf{1}_A dg_y \right) \mathbf{1}_B dq(y) = \int_B g(A|y) d(f \circ p)(y),$$

ce qui montre exactement que g est une désintégration de f selon p . Ce fait servira dans la sous-section 4.4, appliqué au disque.

4.3 Le théorème de décomposition ergodique

Le théorème principal dont on traite ici est le théorème de décomposition ergodique énonçant que toute mesure invariante peut s'écrire comme combinaison de mesures ergodiques.

Théorème 4.14 (synthétique de décomposition ergodique) – Soient C une catégorie de Markov et X un système dynamique déterministe dans C , de monoïde M . Supposons que :

1. l'objet $X \in C$ admet des désintégrations
2. le quotient de Markov (X_{inv}, r) du système dynamique existe.

Alors, tout état invariant $p : \mathbf{1} \rightarrow X$ s'écrit $k \circ q$ où $q : \mathbf{1} \rightarrow X_{\text{inv}}$ et $k : X_{\text{inv}} \rightarrow X$ qui est q -presque-partout ergodique.

Démonstration. L'idée est d'utiliser l'inversion bayésienne pour donner une décomposition. Soit donc $p : \mathbf{1} \rightarrow X$ un état invariant. Par hypothèse, en notant (X_{inv}, r) le quotient de Markov de X , on dispose d'un

morphisme $r_p^\dagger : X_{\text{inv}} \rightarrow X$ vérifiant :

En marginalisant à gauche, il vient que $p = r_p^\dagger \circ r \circ p$. On pose donc $q = r \circ p$ et $k = r_p^\dagger$. On a donc notre candidat pour la décomposition. Reste à montrer que k est q -presque sûrement ergodique.

On commence par montrer que k est q -presque sûrement invariant. Soit $m \in M$. Comme k est une désintégration, r est invariant, m est déterministe, p est invariante, nous avons les égalités suivantes.

Enfin, en réutilisant que k est une désintégration de r , on obtient bien que

ce qui signifie bien que k est q -presque sûrement invariant.

Reste à montrer que k est q -presque sûrement ergodique, ce qui revient à montrer par le lemme 3.11 que $r \circ k$ est q -presque sûrement déterministe. Comme r est déterministe, l'équation de désintégration se réécrit (en composant à gauche par r)

Ainsi, $r \circ r_p^\dagger$ est q -presque sûrement égal à l'identité de X qui est déterministe, donc est q -presque sûrement déterministe. □

4.4 Un exemple de décomposition ergodique

Dans BorelStoch, on retrouve le théorème de décomposition ergodique : avec les hypothèses ci-dessus, toute mesure invariante s'écrit comme mélange de mesures ergodiques. Pour conclure ce mémoire, on donne ci-dessous un exemple concret de décomposition ergodique dans BorelStoch en utilisant la construction donnée dans la preuve.

Donnons un exemple illustrant la construction de la décomposition ergodique donnée par la preuve. On se place dans BorelStoch, et on considère le système ergodique du disque unité fermé centré en 0, que l'on note X , sur lequel on agit par rotation d'un certain angle irrationnel (muni donc du monoïde \mathbb{N}); on note Σ la tribu borélienne de X . Les hypothèses du théorèmes sont vérifiées :

1. le théorème 4.13 de désintégration de Rokhlin énonce que les objets de BorelStoch admettent des désintégrations,
2. les quotients de Markov existent par la propriété 3.8.

On note Σ_{inv} la tribu invariante de X . Il va s'agir de décomposer la mesure de Lebesgue normalisée sur X , notée λ . Le théorème écrit λ comme $\lambda = r_\lambda^\dagger \circ (r \circ \lambda)$ avec $r : X \rightarrow X_{\text{inv}}$ avec le notation précédentes.

On utilise le point clé que la mesure λ_r de Lebesgue normalisée sur un cercle $\partial D(r)$ est ergodique pour une rotation d'angle irrationnel. Ce fait bien connu en théorie ergodique est démontré dans [EW11, proposition 2.16].

Calcul de la décomposition Le morphisme $r \circ \lambda$ étant simplement la mesure λ mais pour la tribu invariante, calculons r_λ^\dagger . Ce morphisme est défini à ensemble λ -négligeable près par l'égalité suivante :

$$\int_A r(B|x) d\lambda(x) = \int_B r_\lambda^\dagger(A|y) d(r \circ \lambda)(y), \quad \text{i.e.} \quad \lambda(A \cap B) = \int_B r_\lambda^\dagger(A|y) d\lambda(y),$$

où $A \in \Sigma$ et $B \in \Sigma_{\text{inv}}$. Les cercles $\partial D(r)$ de rayon $0 \leq r \leq 1$, étant invariants sous l'action du système dynamique, on décompose les deux intégrales comme suit, avec λ_r la mesure de Lebesgue normalisée sur le cercle $\partial D(r)$:

$$\int_0^1 r \cdot \left(\int_{\partial D(r) \cap B} \mathbf{1}_A(x) d\lambda_r(x) \right) dr = \int_0^1 r \cdot \left(\int_{\partial D(r) \cap B} r_\lambda^\dagger(A|x) d\lambda_r(x) \right) dr.$$

L'intuition serait de prendre $r_\lambda^\dagger(A|x) = \mathbf{1}_A(x)$, seulement cette application n'est pas un noyau de Markov $X_{\text{inv}} \rightarrow X$: elle n'est pas mesurable en y si $A \notin \Sigma_{\text{inv}}$. Plutôt, remarquons que comme $B \in \Sigma_{\text{inv}}$, $B \cap \partial D(r)$ est invariant pour la rotation induite sur le cercle, et comme λ_r est ergodique pour cette rotation,

$$\lambda_r(B \cap \partial D(r)) = 0 \text{ ou } 1.$$

Ainsi, en posant $r_p^\dagger(A|x) = \lambda_{|x|}(A \cap \partial D(|x|))$, on a, pour tout r ,

$$\begin{aligned} \int_{\partial D(r) \cap B} \mathbf{1}_A(x) d\lambda_r(x) &= \lambda_r(A \cap B \cap \partial D(r)) \\ &= \begin{cases} 0 & \text{si } \lambda_r(B \cap \partial D(r)) = 0 \\ \lambda_r(A \cap \partial D(r)) & \text{si } \lambda_r(B \cap \partial D(r)) = 1 \end{cases} \\ &= \lambda_r(B \cap \partial D(r)) \cdot \lambda_r(A \cap \partial D(r)) \\ &= \int_{\partial D(r) \cap B} r_\lambda^\dagger(A|x) d\lambda_r(x). \end{aligned}$$

La fin de la preuve consiste à vérifier qu'on a ainsi bien défini un noyau de Markov, et qu'il est bien invariant et λ -presque sûrement ergodique.

Invariance et ergodicité Ces points peuvent se vérifier à la main. Il s'agit de vérifier qu'à A fixé, la fonction $y \mapsto r_p^\dagger(A|y)$ est mesurable, on montre donc que $r \mapsto \lambda_r(A \cap \partial D(r))$ est mesurable. On fait porter r par l'argument plutôt que la mesure :

$$r_p^\dagger(A|y) = \int_{\partial D(r)} \mathbf{1}_{A \cap \partial D(r)}(x) d\lambda_r(x) = \int_{\partial D(1)} \mathbf{1}_{\frac{1}{r}A \cap \partial D(1)}(x) d\lambda_1(x), \quad (*)$$

puis comme $(r, x) \mapsto \mathbf{1}_{\frac{1}{r}A \cap \partial D(1)}(x)$ est mesurable, le théorème de Fubini montre que (*) est mesurable en r , donc en y . Enfin, cette mesure est bien invariante par rotation car c'est le cas de $y \mapsto |y|$. Pour l'ergodicité, l'argument est le même que plus haut : r_λ^\dagger hérite de l'ergodicité de λ_r sur $\partial D(r)$.

Références

- [EW11] Manfred EINSIEDLER et Thomas WARD. « Invariant measures for continuous maps ». Dans : *Ergodic theory : with a view towards number theory*. London : Springer London, 2011, p. 97-119. ISBN : 978-0-85729-021-2. DOI : 10.1007/978-0-85729-021-2_4. URL : https://doi.org/10.1007/978-0-85729-021-2_4.
- [FR20] Tobias FRITZ et Eigil Fjeldgren RISCHÉL. « Infinite products and zero-one laws in categorical probability ». Dans : *Compositionality 2* (2020), p. 3. ISSN : 2631-4444. DOI : 10.32408/compositionality-2-3. URL : <http://dx.doi.org/10.32408/compositionality-2-3>.
- [Fri20] Tobias FRITZ. « A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics ». Dans : *Advances in Mathematics* 370 (2020), p. 107239. ISSN : 0001-8708. DOI : 10.1016/j.aim.2020.107239. URL : <http://dx.doi.org/10.1016/j.aim.2020.107239>.
- [Kle14] Achim KLENKE. « Probability measures on product spaces ». Dans : *Probability theory : a comprehensive course*. London : Springer London, 2014, p. 273-293. ISBN : 978-1-4471-5361-0. DOI : 10.1007/978-1-4471-5361-0_14. URL : https://doi.org/10.1007/978-1-4471-5361-0_14.
- [ML78] Saunders MAC LANE. *Categories for the working mathematician*. New York, NY : Springer New York, 1978. ISBN : 978-1-4757-4721-8. DOI : 10.1007/978-1-4757-4721-8_8. URL : https://doi.org/10.1007/978-1-4757-4721-8_8.
- [MP23] Sean MOSS et Paolo PERRONE. « A category-theoretic proof of the ergodic decomposition theorem ». Dans : *Ergodic Theory and Dynamical Systems* 43.12 (2023), p. 4166-4192. ISSN : 1469-4417. DOI : 10.1017/etds.2023.6. URL : <http://dx.doi.org/10.1017/etds.2023.6>.
- [Per24] Paolo PERRONE. « Markov categories and entropy ». Dans : *IEEE Transactions on Information Theory* 70.3 (2024), p. 1671-1692. ISSN : 1557-9654. DOI : 10.1109/tit.2023.3328825. URL : <http://dx.doi.org/10.1109/TIT.2023.3328825>.
- [Tao08] Terence TAO. *Ergodicity, 254A lecture 9. What's New*, mathematical blog with proofs. 2008. URL : <https://terrytao.wordpress.com/2008/02/04/254a-lecture-9-ergodicity>.