Table des matières

Première partie

Processus de Markov

1 Définitions

Définition: Processus

Un processus est un phénomène aléatoire qui se déroule au cours du temps.

Si on a un processus, son état à la date t est donné par une variable aléatoire notée par exemple X_t
$X_t(\omega) =$ l'état du processus à la date t si le hasard est $\omega \in \Omega$.

L'ensemble des temps peut être :
- discret : $\{0, ..., n\}$ ou $\mathbb{N}$
  Ce ne sont pas forcément des dates, mais par exemple des numéros d'épreuves.
- continu : $[0, T]$ ou $\mathbb{R}^+$

Dans ce cours, les processus auront leurs états dans un ensemble fini ou parfois dénombrable E, appelé l'espace d'états. Ainsi, on note :

$$X = (X_t)_{t \in T}$$

Définition: Propriété de Markov

On dit qu'un processus décrit par $X = (X_t)_{t \in T}$ a la propriété de Markov si :

$$\forall 0 \leq t_0 < ... < t_n < t,\ \mathcal{L}(X_t | X_{t_n}, ..., X_{t_0}) = \mathcal{L}(X_t | X_{t_n})$$

Définition: Homogène

On dit que le processus a la propriété de Markov homogène s'il a la propriété de Markov et, $\forall s < t$:

$$\mathcal{L}(X_t | X_s = x) = \mathcal{L}(X_{t-s} | X_0 = x)$$

1.1 Formalisme matriciel

E est ici considéré comme fini (ou dénombrable), $E = \{i, j, ..., k, ...\}$

Définition: Différents vecteurs

- Une mesure de probabilité μ sur E va être représenté par un vecteur ligne, et qu'on notera μ :
$$\mu = (\mu_j)_{j \in E} \text{ où } \mu_j = \mu(\{j\})$$
- Une fonction $f : E \to \mathbb{R}$ (ou $\mathbb{C}$) sera représenté par un vecteur colonne qui sera noté f :
$$f = (f^i)_{i \in E} \text{ où } f^i = f(i)$$

Définition: Matrices stochastiques

L'ensemble des lois $\mathcal{L}(X_t|X_0 = i)$, $i \in E$ qu'on note $\mathcal{L}(X_t|X_0)$ sera une matrice carrée, notée Π_t, appelée matrice stochastique.

— la ligne i correspondant à la mesure de probabilité

$$\mathcal{L}(X_t, X_0 = i) = (P^i_j(t))_{j\in E} = (\mathbb{P}(X_t = j|X_0 = i))_{j\in E}$$

— la colonne j représente la fonction :

$$i \in E \mapsto P^i_j(t) = \mathbb{P}(X_t = j|X_0 = i)$$

Π_t est une matrice à termes positifs dont la somme de chaque ligne vaut 1.

Exemple :

$$\begin{aligned}
\mathbb{E}(f(X_t)|X_0 = i) &= \sum_j f(j)\mathbb{P}(X_t = j|X_0 = i) \\
&= \sum_j P^i_j(t) f^j \\
&= (\Pi_t f)^i
\end{aligned}$$

Si $\mu_0 = \mathcal{L}(X_0)$, $\mu_{0j} = \mathbb{P}(X_0 = j)$

$$\begin{aligned}
\mathbb{E}_{\mu_0}[f(X_t)] &= \sum_i \mathbb{E}(f(X_t)|X_0 = i]\mathbb{P}(X_0 = i) \\
&= \sum_i \mu_{0i}(\Pi_t f)^i \\
&= \mu_0 \Pi_t f
\end{aligned}$$

et alors si $\mu_t = \mathcal{L}(X_t)$, on a $\mu_t = \mu_0 \Pi_t$ (formule de probabilité totale, μ_0 représentant la loi de X_0)

$$\begin{aligned}
\mathbb{P}(X_t = j) &= \sum_i \mathbb{P}(X_t = j|X_0 = i)\mathbb{P}(X_0 = i) \\
&= \sum_i \mu_{0,i} P^i_j(t) \\
&= (\mu_0 \Pi_t)_j
\end{aligned}$$

Dans la suite du cours, nous considérerons que nos processus ont toujours la propriété de Markov homogène.

Théorème: Relations de Kolmogorov

$$\forall s, t \geq 0, \Pi_t \Pi_s = \Pi_s \Pi_t = \Pi_{s+t}$$

$$\Pi_0 = I$$

Démonstration :

$$\begin{aligned}
\forall i,j,\ \mathbb{P}[X_{t+S}=j|X_0=i] &= \sum_{k\in E}\mathbb{P}(X_{t+s}=j|X_s=k,X_0=i)\mathbb{P}(X_s=k|X_0=i)\\
&= \sum_{k\in E}\mathbb{P}(X_{t+s}=j|X_s=k)\mathbb{P}(X_s=k|X_0=i) \text{ Par la propriété de Markov}\\
&= \sum_{k\in E}\mathbb{P}(X_t=j|X_0=k)\mathbb{P}(X_j=k|X_0=i) \text{ propriété homogène}\\
P_j^i(t+s) &= \sum_{k\in E}P_k^i(s)P_j^k(t)\\
(\Pi_{t+s})_j^i &= (\Pi_s\Pi_t)_j^i
\end{aligned}$$

Donc $\Pi_{t+s}=\Pi_s\Pi_t=\Pi_{s+t}=\Pi_t\Pi_s$.

$$(\Pi_0)_j^i=\mathbb{P}(X_0=i|X_0=j)=\delta_{ij}\Rightarrow\Pi_0=I$$

2 Processus de Markov (sous entendu homogène) en temps discret (T=$\mathbb{N}$)

D'après la relation de Kolmogorov :

$$\Pi_n=(\Pi_1)^n=\Pi^n$$

On note Π_1 par Π la matrice de transition de la chaîne de Markov.
On note $P_j^i=\mathbb{P}(X_1=j|X_0=i)$ Ainsi :

$$\Pi=[P_j^i]_{i,j\in E}$$

La ligne i de $\Pi_n=\Pi^n=\mathcal{L}(X_n|X_0=i)$
Si $\mu_0=\mathcal{L}(X_0)$, alors $\mathcal{L}(X_n)=\mu_0\Pi^n$

$$\mathcal{L}(X_0)=(\mathbb{P}(X_0=j))_{j\in E}$$
$$\mathcal{L}(X_n)=(\mathbb{P}(X_n=j))_{j\in E}=\left(\sum_{i=1}^{n}\mathbb{P}(X_n=j|X_0=i)\mathbb{P}(X_0=i)\right)_{j\in E}=\mu_0\Pi^n$$

2.1 Diagramme de la chaîne de Markov

C'est un graph orienté dont tous les sommets sont les éléments i de E, et les arêtes (orientées) sont définies ainsi :

$i \longrightarrow j$ si et seulement si $p_j^i>0$ ($\mathbb{P}(X_1=j|X_0=i)>0$), ie si et seulement si on peut passer de i à j en une étape.

2.2 Classification des états

Définition: Conduire, communiquer et classe d'équivalence

i peut conduire à j si et seulement si $i=j$ ou s'il existe un chemin allant de i à j (qu'on note $i\rightsquigarrow j$), ie :

$$\exists n\geq 0;\ p_j^i(n)>0\ (\mathbb{P}[X_n=j|X_0=i]>0)$$

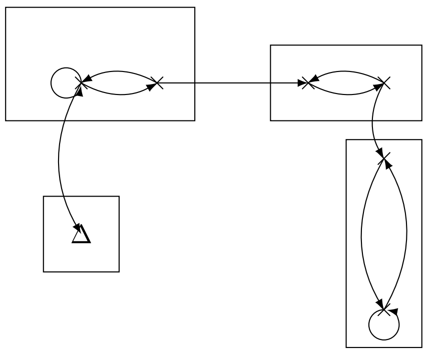
Cette relation est un préordre.
— réflexive
— transitive

A l'aide de ce préordre, on construit une relation d'équivalence.
"i et j communiquent" (noté $i \leftrightarrow j$) si et seulement si $i \rightsquigarrow j$ et $j \rightsquigarrow i$.

- Réflexive
- Symétrique
- Transitive

L'espace d'état est alors partitionné en classes d'équivalence.

Exemple :



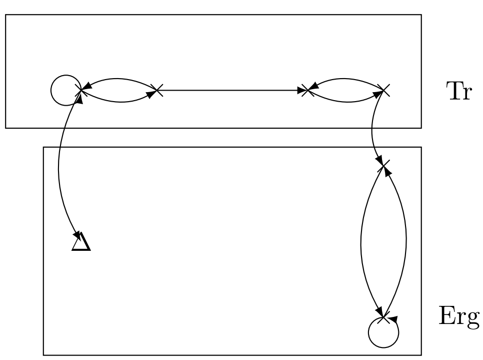
Le préordre induit une relation d'ordre sur les classes : $C \rightsquigarrow D \Leftrightarrow \exists i \in C;\ \exists j \in D;\ i \rightsquigarrow j$.
Ceci ne dépend pas des i et j choisis.

Si $C \rightsquigarrow D$ et $D \rightsquigarrow C$ alors $C = D$

Définition: Transitoire, finale, ergodique

- Une classe est dite transitive si et seulement si elle peut conduire à une autre classe. Ses éléments sont dits transitoires.
  Tr = Ensemble des états transitoires.
- Si une classe n'est pas transitoire, on dit qu'elle est finale (elle ne peut conduire qu'à elle-même). Ses éléments sont dits ergodiques.
  Erg=ensemble des états ergodiques.
- Si la classe finale n'est composée que d'un élément, on dit qu'il est absorbant ($\Leftrightarrow p_i^i = 1$)
  On le note Δ

Exemple :



Théorème:

Si E est fini, il existe toujours des classes finales, et toute classe transitoire peut conduire à au moins une classe finale (évident)

Remarque : Si E est infini, c'est faux. Prendre par exemple $E = \mathbb{N}$, où chaque $n \in \mathbb{N}$ forme une classe $\{n\}$, elles sont toutes transitoires.

Définition: Forme canonique de la matrice de transition

On regroupe les états par classe, en mettant d'abord les classes finales. Par exemple, si on considère C_1 et C_2 $\in$Erg et C_3, C_4 et C_5 $\in$Tr :

$$\begin{array}{cc} & \begin{array}{ccccc} C_1 & C_2 & C_3 & C_4 & C_5 \end{array} \\ \begin{array}{c} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{array} & \begin{pmatrix} A_1 & 0 & |0 & 0 & 0 \\ 0 & A_2 & |0 & 0 & 0 \\ — & — & |— & — & — \\ & R & | & Q & \\ & & | & & \end{pmatrix} \end{array} = \Pi$$

$$\Pi^n = \begin{pmatrix} A_1^n & 0 & |0 & 0 & 0 \\ 0 & A_2^n & |0 & 0 & 0 \\ — & — & |— & — & — \\ & R_n & | & Q^n & \\ & & | & & \end{pmatrix}$$

Q s'appelle la matrice de passage des transitoires aux transitoires.
R la matrice de passage des transitoires aux ergodiques.

Théorème:

Si E est fini, alors presque sûrement le processus finira dans une des classes finales.

Démonstration :
On suppose $X_0 = i$.
Si i est ergodique, X_n reste dans la classe finale de i.
Si i est transitoire :

$$\mathbb{P}(X_n \in Tr | X_0 = i) = \sum_{j \in Tr} \mathbb{P}(X_n = j | X_0 = i) = p_n(i) = \sum_j Q_{ji}^n$$

Il est clair que $X_{n+1} \in Tr \Rightarrow X_n \in Tr$, donc $p_{n+1}(i) \leq p_n(i)$. Donc $(p_n(i))_n$ est une suite décroissante.
E étant fini, $\exists n;\ \mathbb{P}(X_n \in Tr | X_0 = i) < 1$ (car $\exists j \in Erg$ tel que $in \to j$).

$\forall i$, soit n_i tel que $p_{n_i}(i) < 1$. Alors $p_n(i) < 1\ \forall n \geq n_i$.

Soit $N = \max_{i \in Tr} n_i < +\infty$ (Tr fini)

$$\forall i \in Tr, \forall n \geq N, p_N(i) < 1$$

Soit $p^* = \max_{i \in Tr} p_N(i) < 1$.
$\forall i,\ p_n(i)$ décroit, et $p_n(i) > 0$, donc $p_n(i)$ converge vers l_i. Montrons que $l_i = 0$, en considérant la sous-suite $(p_{kN}(i))_k$ qui converge vers l_i.

$$\begin{aligned}
p_{kN}(i) &= \mathbb{P}(X_{kN} \in Tr | X_0 = i) \\
&= \sum_{h \in Tr} \mathbb{P}(X_{kN} \in Tr | X_{(k-1)N} = h, X_0 = 1)\mathbb{P}(X_{(k-1)N} = h | X_0 = i) \\
&= \sum_{h \in Tr} \mathbb{P}(X_n \in Tr | X_0 = h)\mathbb{P}(X_{(k-1)N} = h | X_0 = i) \\
&\leq \sum_{k \in Tr} p^* \mathbb{P}(X_{(k-1)N} = h | X_0 = i) \\
&\leq p^* p_{(k-1)N}(i)
\end{aligned}$$

$$p_{kN}(i) \leq (p^*)^k \xrightarrow[n \to +\infty]{} 0$$

Donc $l_i = 0$.

De plus, $(X_n \in Tr)_n \to \cap_n (X_n \in Tr)$, donc, par propriété des mesures finies due à la σ-additivité :

$$\mathbb{P}(\cap_n (X_n \in Tr) | X_0 = i) = 0$$

Donc :

$$\mathbb{P}(\cup_n (X_n \in Tr) | X_0 = i) = 1$$

$\forall i$ point de départ, $\exists n$ presque sûrement, $X_n \in Erg$.

Remarque : Le théorème précédent équivaut à dire que :

$$Q^n \to 0$$

car :

$$\mathbb{P}(X_n \in Tr | X_0 = i) = \sum_j Q^n_{ji} \to 0$$

Ceci permettra de calculer :
- le temps moyen passé dans Tr
- la probabilité de finir dans telle ou telle classe finale

Soit une chaîne de Markov de matrice de transition Π écrite sous forme canonique :

$$\Pi = \begin{pmatrix} A & 0 \\ R & Q \end{pmatrix}$$

$$R = [p^i_j]_{i \in Tr,\ j \in Erg}\ Q = [p^i_j]_{i,j \in Tr}$$

Lemme:

$I - Q$ est inversible.

Démonstration :

$$(I - Q)(I + Q + ... + Q^n) = I - Q^{n+1} \xrightarrow[n \to +\infty]{} I$$

$$\Rightarrow \det(I - Q) \det\left(\sum_{i=0}^{n} Q^k\right) \xrightarrow[n \to +\infty]{} \det(I) = 1$$

Dont $\det(I-Q) \neq 0$ donc $I-Q$ inversible.
En multipliant par $(I-Q)^{-1}$ à gauche :

$$I + Q + ... + Q^n = (I-Q)^{-1}(I - Q^{n-1}) = (I-Q)^{-1} - (I-Q)^{-1}Q^{n+1} \xrightarrow[n\to+\infty]{} (I-Q)^{-1}$$

Donc la série des Q^k converge et :

$$\sum_{k=0}^{\infty} Q^k = (I-Q)^{-1}$$

Théorème:

Soit $N = (I-Q)^{-1} = [N_j^i]_{i,j\in Tr}$. Alors N_j^i=le nombre moyen de fois où le processus est passé par j sachant qu'il est parti de i.

Démonstration :
Soient $i \in Tr$ et $j \in Tr$. Soit Y_j le nombre de fois où le processus par par j.

$$Y_j = \sum_{n=0}^{\infty} 1_{\{j\}}(X_n)$$

Calculons $\mathbb{E}(Y_j|X_0 = i)$:

$$\begin{aligned}
\mathbb{E}(Y_j|X_0 = i) &= \mathbb{E}\left(\sum_{n=0}^{\infty} 1_{\{j\}}(X_n)|X_0 = i\right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}(1_{\{j\}}(X_n)|X_0 = i) \text{ (Corollaire de Bepo-Levi)} \\
&= \sum_{n=0}^{\infty} \mathbb{P}(X_n = j|X_0 = i) \\
&= \sum_{n=0}^{\infty} p_j^i(n) \\
&= \sum_{n=0}^{\infty} [\Pi_n]_j^i \\
&= \sum_{n=0}^{\infty} [Q^n]_j^i
\end{aligned}$$

Donc :

$$(\mathbb{E}(Y_j|X_0 = i)) = \sum_{n=0}^{\infty} Q^n = (I-Q)^{-1}$$

Corollaire:

Le nombre moyen de fois où le processus passe par les états transitoires, sachant qu'il est parti de i transitoire, vaut $N^i = \sum_{j\in Tr} N_j^i$.

Théorème:

Soit $B = NR$ $(B = [B^i_j]_i \in Tr,\ j \in Erg)$
alors B^i_j est la probabilité que le premier état ergodique j sachant que le processus est parti de i, transitoire.

Démonstration :
Soi B^i_j la probabilité que le premier état ergodique atteint soit j sachant qu'on est parti de i.
Pour que le premier état ergodique atteint soit j, deux possibilités :

- On va de i à j en 1 coup : probabilité p^i_j
- On va de i à j en au moins deux coups :
  - au premier coup, on va à $k \in Tr$, probabilité p^i_k
  - partant de k, le premier état ergodique atteint est j : probabilité B^k_j.

Donc $\forall i \in Tr$, $\forall j \in Erg$:

$$B^i_j = p^i_j + \sum_{k \in Tr} p^i_k B^k_j = [R]^i_j + [QB]^i_j$$

Donc :

$$\begin{aligned} & B = R + QB \\ \Rightarrow\ & (I - Q)B = R \\ \Rightarrow\ & B = (I - Q)^{-1} R = NR \end{aligned}$$

Corollaire:

- Si j est absorbant, $B^i_j = \mathbb{P}(\text{finir en } j | X_0 = i)$
- Si C est une classe finale :

$$B^i_C = \mathbb{P}(\text{finir en } C | X_0 = i) = \sum_{j \in C} B^i_j$$

Deuxième partie

Etudes des séjours dans une classe

1 Période de classe

1.1 Définition

Soit C une classe d'équivalence pour $i \leftrightarrow j$.

$$C = cl\{i\} = cl\{j\},\ \forall i, j \in C$$

$\forall i, j \in C$, il existe un chemin (orienté) allant de i à j. On appelera :

$$\begin{aligned} N_{ij} &= \{n > 0;\ \text{il existe un chemin allant de } i \text{ à } j \text{ de longueur } n\} \\ &= \{n > 0;\ p_j^i(n) > 0\} \end{aligned}$$

Propriété: fondamentale

$i, j, k \in C$
Si $a \in N_{ij}$ et $b \in N_{jk}$ alors $a + b \in N_{ik}$, ie $N_{ij} + N_{jk} \subset N_{ik}$.

Théorème:

Soit $d_i = PGCD(N_{ii})$ pour $i \in C$.

$$\forall i, j \in C, d_i = d_j$$

Démonstration :
Soient $a \in N_{jj}$, $b \in N_{ji}$ et $c \in N_{ij}$.

$$a + b + c = c + a + b \in N_{ii}$$

Or, $b + c = c + b \in N_{ii}$, donc :

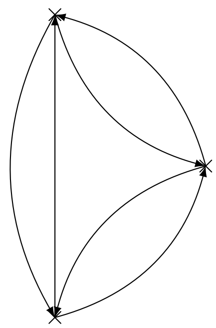
$$a + b + c \equiv 0[d_i] \text{ et } b + c \equiv 0[d_i] \Rightarrow a \equiv 0[d_i]\ \forall a \in N_{jj}$$

donc $d_j \equiv 0[d_i]$ et de même, $d_i \equiv 0[d_j]$, donc $d_i = d_j$.

Définition: Période d'une classe

Notons d cet entier commun à tous les $i \in C$. d est appelé la *période de la classe C*.
Si $d = 1$, la classe est dite *apériodique*.

Exemple :



Ici, on revient au même point en 2 ou 3 coups. On peut faire ça avec n'importe quel point, comme on l'a vu avec le théorème précédent.
$d = PGCD(2, 3) = 1$

1.2 Caractérisation

Théorème:

$\forall i \in C,\ N_{ii} = \{kd,\ k \in \mathbb{N}\backslash A_i\},\ A_i$ fini.

Démonstration :
Quitte à diviser tous les éléments de N_{ii} par d, on peut supposer que $PGCD(N_{ii}) = 1$ (pour simplifier).
N_{ii} est un semi-groupe pour $+$:

$$a, b \in N_{ii} \Rightarrow a + b \in N_{ii}$$

$N_{ii} \subset \mathbb{N}^*$
Donc $N_{ii} = \mathbb{N}\backslash A_i$, A_i ensemble fini, ie $\exists k_i$ tel que $\forall n \geq k_i,\ n \in N_{ii}$. (rien compris)

$N_{ii} \subset \mathbb{Z}$ et $PGCD(N_{ii}) = 1$, ce qui signifie que le module sur $\mathbb{Z}$ engendré par N_{ii} est celui engendré par 1 qui est $\mathbb{Z}$.
(Après quelques recherches, il ne s'agirait pas d'un module mais plutôt d'un idéal)o

$$(N_{ii}) = \left\{\sum_{k=1}^{r} \alpha_k n_k,\ \alpha_k \in \mathbb{Z}, n_k \in N_{ii}\right\}$$

$1 \in (N_{ii})$
$\Rightarrow$ Identité de Bézout, i.e. $\exists n_1, ..., n_r \in N_{ii}$, $a_1, ..., a_r \in \mathbb{Z}$ tels que $a_1 n_1 + ... + a_r n_r = 1$.
On peut soppose que $a_1, ..., a_p > 0$ et $a_{p+1}, ..., a_r \leq 0$.

$$a_1 n_1 + ... + a_p n_p = n \in N_{ii}$$

$$m = -(a_{p+1} n_{p+1} + ... + a_r n_r) \in N_{ii} \cup \{0\}$$

N_{ii} étant un semi-groupe pour $+$:

$$\exists n \in N_{ii}, m \in N_{ii} \cup \{0\};\ n - m = 1$$

Si $m \neq 0$:
Soit $k_i = m^2$. Si $k \geq m^2$

$$k = m\alpha + \beta,\ \alpha \geq m,\ 0 \leq \beta < m$$

$$\begin{aligned} k &= m\alpha + \beta(n - m) \\ &= \beta n + \underbrace{(\alpha - \beta)}_{>0} m \end{aligned}$$

Si $\beta = 0$, $k = \alpha m \in N_{ii}$
Si $\beta > 0$, $\beta n \in N_{ii}$ et $(\alpha - \beta)n \in N_{ii}$, donc $k \in N_{ii}$ ie

$$N_{ii} = \{\mathbb{N} \backslash \{0, ..., m^2\}\}$$

Théorème:

$\forall i, j \in C$, $\exists r_{i,j}$, $0 \leq r_{i,j} < d$ et

$$N_{ij} = \{kd + r_{ij},\ k \in \mathbb{N} \backslash A_{ij}$$

avec A_{ij} un ensemble fini dépendant de i et j.

Démonstration :
Montrons d'abord que deux éléments $a, b \in N_{ij}$ ont même reste dans la division par d.
Soient $a, b \in N_{ij}$, $c \in N_{ji}$.
$a + c$ et $b + c \in N_{ii}$ donc $a + c \equiv 0[d_i]$ et $c + b \equiv 0[d_i]$, donc $a + c - (b + c) = a - b \equiv 0[d_i]$. Donc a et b ont même reste dans la division par d.

Notons r_{ij} ce reste commun à tous les éléments de N_{ij}. Tout élément a de N_{ij} s'écrit $a = kd + r_{ij}$.
Soit k_i tel que $\forall l \geq k_i$, $ld \in N_{jj}$
Soit $a_0 = k_0 d + r_{ij}$, alors $\forall k \geq k_i$, $a_0 + ld \in N_{ij}$

$$\forall l \geq k_i, (k_0 + l)d + r_{ij} \in N_{ij}$$

1.3 Relation d'équivalence dans la classe C et sous-classes périodiques

Définition:

Si $i, j \in C$, on dit que $i \sim j$ si et seulement s'il existe un chemin de longueur multiple de la période d joignant i à j
Cela se traduit par $i \sim j \Leftrightarrow r_{ij} = 0$

Cette relation est évidemment (hum) réflexive et transitive.
Pour la symétrie : si $r_{ij} = 0$, soient $a \in N_{ji}$ et $b \in N_{ij}$. $a + b \in N_{jj}$ donc $a + b \equiv 0[d]$.
Or $b \equiv 0[d]$ car $r_{ij} = 0$ donc $a \equiv 0[d]$: $r_{ji} = 0$.
Donc C se partitionne en classes d'équivalences pour $\sim$. On les appelle les sous-classes cycliques.

Théorème:

Si d est la période de C alors C possède exactement d sous-classes cycliques.
Celles-ci sont atteintes successivement toujours dans le même ordre tant que l'état du processus reste dans C.

2 Chaînes régulières

Définition:

On dit que la chaîne de Markov est régulière si et seulement si :

1. Elle ne possède qu'une seule classe
2. Elle est apériodique (d=1)

⚯ *Théorème:*

Les trois points suivants sont équivalents.

1. La classe est régulière
2. $\exists n_0,\ \forall n \geq n_0,\ \forall i, j \in E,\ \mathbb{P}(X_n = j | X_0 = i) > 0$
3. $\exists n_0,\ \forall i, j \in E, \mathbb{P}(X_{n_0} = j | X_0 = i) > 0$

Démonstration :
Il est clair que $2 \Rightarrow 3$.
$3 \Rightarrow 1$ car $\forall i, j \in N_{ij} = \mathbb{N} \backslash$ ens. fini. Donc $i \rightsquigarrow j$ et $r_{ij} = 0$
$1 \Rightarrow 2$:

$$\forall i, j \in N_{ij} \supset \{n > 0,\ n \geq n_i\}$$

$(r_{ij} = 0$ et $N_{ij} \neq \emptyset)$
Soit $n_0 = \max_{i \in E}(n_i) < +\infty$ (car E fini).
$\forall i, j \in E, \exists n_0;\ \forall n \geq n_0, n \in N_{ij}$, ie $\mathbb{P}(X_n = i | X_0 = i) > 0$.

⚯ *Lemme:*

Soit $A = (a^i_j)_{i,j}$ une matrice "stochastique".
Soit $a^* = \min_{i,j} a^i_j > 0$
Soit $X_0 = (x^i_0)_i$ un vecteur colonne.
On pose $X_n = AX_{n-1} = A^n X_0$. Soit $M_n = \max_i x^i_n$ et $m_n = \min_i x^i_n$.
On a m_n croissante, M_n décroissante et $(M_n - m_n) \leq (1 - 2a^*)(M_0 - m_0)$.

Démonstration :

$$x^i_{n+1} = \sum_j a^i_j x^j_n \geq \underbrace{\sum_j a^i_j}_{=1} m_n = m_n$$

Donc $m_{n+1} \geq m_n$, donc m_n est croissante.
De même, M_n est décroissante.

A présent, soit j_0 tel que $x^{j_0}_n = m_n$.

$$\begin{aligned} x_n n + 1^i &= a^i_{j_0} m_n + \sum_{j \neq j_0} a^i_j x^j_n \\ &\leq a^i_{j_0} m_n + \underbrace{\left(\sum_{j \neq j_0} a^i_j \right)}_{=1 - a^i_{j_0}} M_n \\ &\leq M_n - a^i_{j_0}(M_n - m_n) \end{aligned}$$

Donc

$$M_{n+1} \leq M_n - a^*(M_n - m_n) \tag{1}$$

En appliquant le résultat à $-X_n$ et à $-X_{n+1}$:

$$-m_{n+1} \leq -m_n - a^*(-m_n - (-M_n))$$

$$-m_{n+1} \leq -m_n - a^*(M_n - m_n) \tag{2}$$

$$\begin{array}{rcl} (1)+(2) : M_{n+1} - m_{n+1} & \leq & M_n - m_n - 2a^*(M_n - m_n) \\ M_{n+1} - m_{n+1} & \leq & (1-2a^*)(M_n - m_n) \end{array}$$

Par récurrence, on obtient le résultat.

Théorème: fondamental

On considère une chaîne de Markov régulière et E fini.
Dans ce cas, il existe une et une seule loi de probabilité invariante sur E. Notons la μ. Celle-ci vérifie :

1. $\forall i \in E$, $\mu_i > 0$
2. $\forall \mathcal{L}(X_0), \mathcal{L}(X_n) \to \mu$ exponentiellement vite
3. μ est l'unique solution de l'équation :

$$\left\{ \begin{array}{ccc} \nu\Pi & = & \nu \\ \sum_{i\in E} \nu_i & = & 1 \end{array} \right.$$

Démonstration :
Considérons la colonne j de Π^n : C_j^n.
Que devient-elle lorsque $n \to +\infty$?

$$C_j^0 = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{pmatrix} \text{jième ligne} \qquad \begin{array}{rcl} C_j^n & = & \Pi^n C_j^0 \\ & & \\ & = & \Pi C_j^{n-1} \end{array}$$

Soit $M_j^n = \max_i (C_j^n)_i$ et $m_j^n = \min_i (C_j^n)_i$.
On a M_j^n décroissante, m_j^n croissante et $m_j^n \leq M_j^n$.
Comme $M_j^n \geq m_j^0$, M_j^n converge. De même, m_j^n converge également.

Montrons que $M_j^n - m_j^n \to l_j = 0$.
Soit n_0 tel que $A = \Pi^{n_0}$ soit à termes positifs (cela signifie que la chaîne est régulière, d'après le premier théorème de cette section).
Soit $p^* = \min_{i,j}(p_j^i(n_0)) > 0$. On a :

$$M_j^{kn_0} - m_j^{kn_0} \leq (1-2p^*)^k \underbrace{(M_j^0 - m_j^0)}_{=1} \xrightarrow[k\to+\infty]{} 0$$

Donc la sous-suite $(M_j^{kn_0} - m_j^{kn_0})_k$ de $(M_j^n - m_j^n)_n)$ converge vers 0. Or, la suite est convergente vers l_j, donc $l_j = 0$.
Donc M_j^n et m_j^n ont même limitE. Notons la μ_j.

$$C_j^n = (p_j^i(n))_i \to \begin{pmatrix} \mu_j \\ \vdots \\ \mu_j \end{pmatrix}$$

$$\Pi^n \to \begin{pmatrix} \mu_1 & \cdots & \mu_n \\ \vdots & \ddots & \vdots \\ \mu_1 & \cdots & \mu_n \end{pmatrix}$$

$$\mu_j \geq m_j^{n_0} \geq p^* > 0$$

$$\forall j,\ \mu_j > 0$$

$$m_j^n \leq \mu_j \leq M_j^n$$

$$\begin{aligned} |\mu_j - p_j^i(n)| &\leq M_j^n - m_j^n \\ &\leq M_j^{kn_0+r} - m_j^{kn_0+r} \\ &\leq (1-2p^*)^k (M_j^r - m_j^r) \\ &\leq \left((1-2p^*)^{\frac{1}{n_0}}\right)^{kn_0} (1-2p^*)^{\frac{r}{n}} \underbrace{\max_{r<n_0, j} \frac{M_j^r - m_j^r}{(1-2p^*)^{\frac{r}{n_0}}}}_{=a} \\ &\leq a\left[(1-2p^*)^{\frac{1}{n_0}}\right]^n \end{aligned}$$

Si $b = (1-2p^*)^{\frac{1}{n_0}}$, $0 \leq b < 1$

$$\forall j,\ |\mu_j - p_j^i(n)| \leq ab^n$$

On a donc une convergence exponentielle de $\mathbb{P}(X_n = j | X_0 = i)$ vers μ_j

Soit ν^0 loi de X_0.

$$\begin{aligned} \mathbb{P}(X_n = j) &= \sum_i \mathbb{P}(X_n = j | X_0 = i)\mathbb{P}(X_0 = i) \\ &= \sum_i \nu_i^0 p_j^i(n) \end{aligned}$$

Comme $\sum_i \nu_i^0 = 1$:

$$\begin{aligned} |\mathbb{P}(X_n = j) - \mu_j| &= \left| \sum_i \nu_i^0 (p_j^i(n) - \mu_j) \right| \\ &\leq \sum_i \nu_i^0 |p_j^i(n) - \mu_j| \\ &\leq \sum_i \nu_i^0 ab^n \\ &\leq ab^n \end{aligned}$$

$\forall \mathcal{L}(X_0)$, $\mathcal{L}(X_n) \to \mu$ exponentiellement vite, ie

$$|\mathbb{P}(X_n = j) - \mu_j| \leq ab^n$$

μ est une loi de probabilité :

$$1 = \sum_{j\in E} \mathbb{P}(X_n = j | X_0 = i) \to \sum_{j \in E} \mu_j = 1$$

Donc $\forall j,\ \mu_j \geq p^* > 0$ et $\sum_i \mu_j = 1$.

Soit ν un vecteur ligne.

$$\nu\Pi^n \to \nu \begin{pmatrix} \mu_1 & \cdots & \mu_n \\ \vdots & \ddots & \vdots \\ \mu_1 & \cdots & \mu_n \end{pmatrix} = \left(\sum_i \mu_i \nu_1 \ \cdots \ \sum_i \nu_i \mu_n \right) = \left(\sum_i \nu_i \right) \mu$$

$$\forall \nu,\ \nu\Pi^n \to \left(\sum_i \nu_i\right)\mu$$

Donc $\mu\Pi^n \to \mu$, donc $(\mu\Pi^n)\,\Pi \to \mu\Pi$, donc

$$\mu\Pi = \mu$$

μ est donc une loi e probabilité invariante. C'est la seule loi de probabilité invariante :
Si ν est une loi de probabilité invariante, alors $\nu\Pi = \nu$. Donc $\forall n\ \nu\Pi^n = \nu$. Or, $\nu\Pi^n \to \mu$ donc $\nu = \mu$

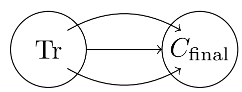
μ vérifie $\mu\Pi = \mu$, $\sum_i \mu_i = 1$. Soit ν solution de $\nu\Pi = \nu$, $\sum_i \nu_i = 1$. Aors :

$$\nu\Pi^n = \nu \to \left(\sum_i \nu_i\right)\mu = \mu$$

Donc $\mu = \nu$.

2.1 Conséquences

1. Supposons que sur E, il n'y ait qu'une seule classe finale et que celle-ci soit apérioique (d=1). Que devient $\mathcal{L}(X_n)$ lorsque $n \to +\infty$?



$\mathcal{L}(X_n) \to \mu_C$, où $\mu_C(j) = 0\ \forall j \in Tr$.
$\mu_{C_{|C}}$ = loi de probabilité invariante de la chaîne quand elle début dans C.
Sur C, X_n est une chaîne régulière.
$\mu = \mu_{C_{|C}}$ vérifie $\mu A = \mu$, $\sum_i \mu_i = 1$, dont elle est l'unique solution.

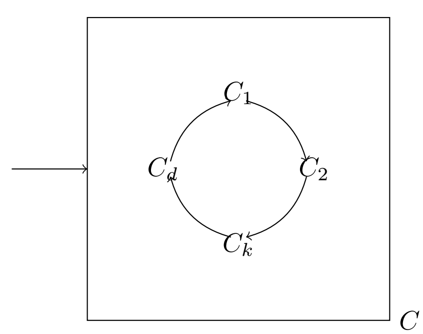
Théorème: ergodique (admis)

Supposons que la chaîne n'ait qu'une seule classe finale C et que celle-ci est apériodique.
Soit μ la mesure invariante sur C décrite précédemment. Alors :

$$\forall f : E \to \mathbb{R},\ \frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow[\forall \mathcal{L}(X_0)]{p.s.} \int_C f d\mu = \sum_{j\in C} f(j)\mu(j)$$

2.2 Etude d'une classe finale périodique

Soit d la période d'une classe C, finale. Elle possède d sous-classes cycliques, parcourues successivement, toujours dans le même ordre. Numérotons les de A à d de façon qu'on les parcourt ainsi :



Π s'écrit alors :

$$\Pi = \begin{pmatrix} 0 & A_1 & 0 & \cdots & 0 \\ 0 & 0 & A_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_{d-1} \\ A_d & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\Pi^d = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_d \end{pmatrix}$$

Avec :

$$\begin{array}{rcl} B_1 & = & A_1 \cdots A_d \\ B_2 & = & A_2 \cdots A_d A_1 \\ \vdots & & \\ B_k & = & A_k \cdots A_d A_1 \cdots A_{k-1} \\ \vdots & & \\ B_d & = & A_d A_1 \cdots A_{d-1} \end{array}$$

Si on considère C_k, B_k est la matrice de transition sur C_k. C'est une chaîne de Markov régulière. Elle admet une unique loi de probabilité invariante μ_k concentrée sur C_k.

$$\begin{cases} \mu_k B_k & = & \mu_k \\ \sum_{i \in C_k} \mu_k(i) & = & 1 \end{cases}$$

Théorème:

Il existe une unique mesure de probabilité invariante sur C. Notons la μ. On a :

$$\mu = \frac{1}{d}(\mu_1, ..., \mu_d)$$

μ_k : vecteur ligne indéxée par C_k.

Démonstration :
Montrons que $\mu = \frac{1}{d}(\mu_1, ..., \mu_d)$ est une mesure de probabilité invariante.

$$\mu\Pi = \frac{1}{d}(\mu_d A_d, \mu_1 A_1, ..., \mu_{d-1} A_{d-1}$$

$\mu_k A_k$ est une mesure de probabilité sur C_{k+1}.

$$\begin{array}{rcl} \mu_k A_k B_{k+1} & = & \mu_k A_k A_{k+1} ... A_d A_1 ... A_k \\ & = & \mu_k B_k A_k \\ & = & \mu_k A_k \end{array}$$

Donc $\mu_k A_k = \mu_{k+1}$ par unicité de la loi de probabilité invariante concentrée sur C_{k+1}. Donc $\mu\Pi = \mu$. μ est donc une loi de probabilité invariante.

Prouvons à présent son unicité.
Soit ν une mesure de probabilité invariante.

$$\nu = (\nu_1, ..., \nu_k, ..., \nu_d), \nu_k = (\nu_i)_{i \in C_k}$$

$$\sum_{i\in C_k} \nu_k(i) = \nu(C_k)$$

On a $\nu\Pi = \nu$ donc $\nu\Pi^d = \nu$. donc $\forall k, \nu_k B_k = \nu_k$. Donc :

$$\frac{\nu_k}{\nu(C_k)} = \mu_k \text{ (par unicité)}$$

Donc $\nu_k = \nu(C_k)\mu_k$.

Il reste à montrer que $\forall k$, $\nu(C_k) = \frac{1}{d}$, ie

$$\forall k, l, \nu(C_k) = \nu(C_l)$$

car

$$\sum_{k=1}^{d} \nu(C_k) = 1$$

On a $\nu\Pi = \nu$, donc : :

$$\begin{aligned} \nu_d A_d &= \nu_1 \\ \nu_1 A_1 &= \nu_2 \\ &\vdots \\ \nu_k A_k &= \nu_{k+1} \end{aligned}$$

$$(\nu_k A_k)_j = \sum_{i\in C_k} \nu_k(i) a_j^i$$

$$\sum_{j\in C_{k+1}} (\nu_k A_k)_j = \sum_{i\in C_k} \nu_k(i) \sum_{j\in C_{k+1}} a_j^i = \sum_{i\in C_k} \nu_k(i) = \nu(C_k)$$

Or, $\nu_k A_k = \nu_{k+1}$ et $\sum_{i\in C_{k+1}} \nu_{k+1}(i) = \nu(C_{k+1})$, donc

$$\forall k, \nu(C_k) = \nu(C_{k+1})$$

Ils sont donc tous égaux à $\frac{1}{d}$.

ℹ *Remarque:*

- Pour cette loi invariante μ, chaque classe C_k a même probabilité $\mu(C_k) = \frac{1}{d}$
- Pour trouver μ, 2 méthodes :
  - μ solution de $\mu\Pi = \mu$ et $\sum_{i\in C} \mu_i = 1$
  - On calcule $\Pi^d = diag(B_1, ..., B_d)$ et on résout $\mu_k B_k = \mu_k$, $\sum_{i\in C_k} \mu_k(i) = 1$. Puis $\mu = \frac{1}{d}(\mu_1, ..., \mu_d)$.
- Dans la démonstration, au lieu de supposer ν probabilité invariante, on aurait pu supposer

$$\begin{cases} \nu \text{ solution de } \nu\Pi = \nu \\ \sum \nu_i = 1 \end{cases}$$

Le reste est inchangé.

Théorème: ergodique

Supposons que la chaîne ne possède qu'une seule classe finale C. Soit μ_C la mesure invariante associée à cette classe. $\forall f : E \to \mathbb{R}$ (E fini) :

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow[\forall \mathcal{L}(X_0)]{p.s.} \int_C f d\mu_C = \sum_{i\in C} f(i)\mu_C(i)$$