On the general transformation of Markov chains

Soumia Benbakreti [a], **Abderrahmane Yousfate** [a]

[a] Laboratoire de Mathématiques (LDM), Université Djillali Liabès, BP 89, Sidi bel Abbès, 22000, Algeria.
souben2223@gmail.com, yousfate@univ-sba.dz.

Abstract: In this work, we consider a class of homogeneous irreducible discrete states Markov process $(X_t)_{t\in\mathbb{T}}, \mathbb{T}\subset\mathbb{Z}$, taking values in a partition of some topological space E. Using a notion of fractional state, we introduce a general transformation φ such that $(\varphi(X_t))_{t\in\mathbb{T}}$ remains a Markov process. This general transformation includes also *lumpability*. Infinite kinds of this expanded Markov chain $(\varphi(X_t))_{t\in\mathbb{T}}$ can be constructed, however lumped Markov chains are finite or may not exist .
Using general transformation φ and fractional states, we give a necessary and sufficient condition under which the transformed process $(\varphi(X_t))_{t\in\mathbb{T}}$ remains still a Markov chain.

Keywords: Expanded Markov chain; fractional state; Markov chain; multivalued transformation.

1. Introduction

Let $(U_t)_{t\in\mathbb{T}\subset\mathbb{Z}}$, denoted $\mathbb{U}$, an irreducible homogeneous Markov chain verifying, for all $t\in\mathbb{T}$, $U_t:(\Omega,\mathcal{A},\mathbb{P})\longrightarrow(E,\mathcal{B})$, and let $X_t:(\Omega,\mathcal{A},\mathbb{P})\longrightarrow(E,\mathcal{E})$.

We consider $X_t=E^{\mathcal{E}}(U_t)=E(U_t/\ \mathcal{E})$ where $\mathcal{E}$ is the σ-algebra generated by a partition $\{E_1,\ E_2,\ \ldots,\ E_r\}$ of the topological space (E,τ_E). Parts $E_i,\ i=1,\ \ldots,\ r$ are assumed to be states of the Markov chain $\mathbb{X}$.

Denoting $\mathcal{F}_t$ the σ-algebra generated by $\{X_s;\ s\leq t\}$, $\mathbb{X}$ is $\mathcal{F}_t$-adapted verifying for all $\tau,\tau_1,\tau_2,\ldots\in\mathbb{N}^*$:

$$\mathbb{P}(X_{t+\tau}\in A/\ X_t=x_t,X_{t-\tau_1}=x_{t-\tau_1},X_{t-\tau_2}=x_{t-\tau_2},\ldots)=\mathbb{P}(X_{t+\tau}\in A/\ X_t=x_t)\quad(1)$$

This characterization of Markov chain remains valid, even if $-\tau,-\tau_1,-\tau_2,\ldots\in\mathbb{N}^*$. It means that $\mathbb{X}$ is a Markov chain for any direction of the time. For the descending t, $(X_t)_t$ is adapted to another filtration $(\mathcal{G}_t)_t$.

That allows us to present the other equivalent definition of a Markov process : $\mathbb{X}$ is a Markov process iff, knowing X_t (the present), σ-algebras generated by $\mathcal{G}_t$ (the future) and $\mathcal{F}_t$ (the past) are independent.

So, if $\widetilde{\mathbb{X}}$ is a measurable transformation of $\mathbb{X}$, taking values in $(\widetilde{E},\widetilde{\mathcal{E}})$, then $\widetilde{\mathbb{X}}$ is a Markov process conditionally to the present of $\mathbb{X}$, but conditionally to present of $\widetilde{\mathbb{X}}$, the past and the future of $\widetilde{\mathbb{X}}$ aren't necessary independent. In this study, we are interesting essentially by general transformations which conserve Markov nature of Markov chains conditioned by the present of the transformed process $\widetilde{\mathbb{X}}$.

We consider that $\mathbb{X}$ is a homogeneous Markov chain, *i. e.*

$$\mathbb{P}(X_{t+\tau}\in A/\ X_t=E_i)=\mathbb{P}(X_\tau\in A/\ X_0=E_i)=\sum_{j;E_j\in A}\mathbb{P}(X_\tau=E_j/\ X_0=E_i).$$

For $\tau = 1$ and varying $i, j \in \{1, \dots, r\}$, one obtains a square matrix P where conditional distributions lay out on matrix rows. P is a fundamental matrix, called transition matrix, which gives general probability term $\mathbb{P}(X_\tau = E_j / X_0 = E_i)$ by extracting the term (i, j) of the matrix P^τ where $\{P^\tau;\ \tau \in \mathbb{N}\}$ constitutes a semigroup for matrix product.

2. Problematic

Problem of univalued transformation conserving the property of Markov Process $\mathbb{X}$ is formulated by many authors since the middle of the last century. We cite Burke and Rosenblatt **(Burke & Rosenblatt, 1958)** which consider essentially aggregation on recurrent discrete states of a Markov processes with an invariant measure on states. That aggregated states were called "collapsed states". Some detailed aspects of this problem known as problem of Markov chain lumpability or Markov states aggregation were developed by others authors; we cite works of Hachigian **(Hachigian, 1963)**, Dynkin **(Dynkin, 1965)**, Kemeny and Snell **(Kemeny & Snell, 1976)**, Rogers-Pitman's **(Rogers & Pitman's, 1981)**, Ledoux, Rubino and Sericola **(Rubino & Sericola, 1991), (Rubino & Sericola, 1993), (Ledoux, Rubino & Sericola, 1994)** , Peng **(Peng , 1996)**, Tian and Kannan **(Tian & Kannan, 2006)**, Geiger and Temmel (**Geiger &Temmel, 2014)**, Pang **(Pang, 2019)**, ...

The focus of all these works concerns essentially how is it possible to reduce the number of states of E , by a subpartition (less fine) which conserves the Markovian nature of the process.

Their main results are based on some aggregation of states also known as *lumpability* which consists to aggregate states which generate identical rows of P after aggregation. More details are presented below.

In this approach, knowing a Markov chain $\mathbb{X}$, we build the set of Markov chains which give $\mathbb{X}$ by aggregation (lumpability) or by disaggregation (state expansion).

Using multivalued transformations **(Berge, 1959)**, the current work gives a general method to study transformed Markov chain preserving the Markov nature of the process. As univalued transformations are particular cases of multivalued ones, the lumpability becomes a particular case in this study.

Many other authors use “aggregation” and “disggregation” terms to investigate decompositional methods in order to obtain a partition of state space. This partition allows process to be decomposed into some subprocesses which must be quasi-independant allowing the user to study numerically the subprocesses separately. So solutions of aggregation/disaggregation in this context carry out to some approximated novel Markov chains which have the same asymptotic stationary measure using some algorithm named Iterative Aggregation/Disaggregation (IAD). Stewart **(Stewart, 1994)** gives a lot of methods to investigate this approach. The technique uses essentially perturbed transition matrices which conserve the positive recurrent property and the stationary measure. Even if we use sames terms “aggregation” and “disaggregation”, our work is related to a general transformation (aggregation/disaggregation) of Markov chain which gives us exactly another markovian representation of the original Markov chain.

Let $(X_t)_{t\in\mathbb{T}}$, denoted $\mathbb{X}$, an irreducible homogeneous Markov chain. We have to transform $(X_t)_t$, $t \in \mathbb{T},\ \mathbb{T} \subset \mathbb{Z}$ by a multivalued function φ, to a new process $(Y_t)_t$ such that this last must be a homogeneous Markov chain. In the following we denote $\mathbb{X} = (X_t)_{t\in\mathbb{Z}}$ and $\mathbb{Y} = (Y_t)_{t\in\mathbb{Z}}$.

Using multivalued transformations $\varphi : X_t \longrightarrow Y_t$, we may to characterise φ such that $(Y_t)_t$ remains a Markov process.

Let $\mathbb{P}_X$ and $\mathbb{P}_Y$ the image distributions of $\mathbb{P}$ via $\mathbb{X}$ and $\mathbb{Y}$ respectively and let build a deterministic coupling $(\mathbb{P}_{X_t},\ \mathbb{P}_{Y_t})\,;\ t \in \mathbb{Z}$ such that the problem becomes to construct some measurable function

$\hat{\varphi} : (\Omega,\ \mathcal{A},\ \mathbb{P}) \longrightarrow (\Omega,\ \hat{\mathcal{A}},\ \mathbb{P})$ verifying for all t in $\mathbb{T}$, $\varphi \circ X_t = Y_t \circ \hat{\varphi}$ which is noted $\varphi \circ \mathbb{X} = \mathbb{Y} \circ \hat{\varphi}$. Multivalued φ is presented as follows :

$(E, \mathcal{E}, \mathbb{P}_{X_t}) \overset{\varphi}{\multimap} (\hat{E}, \hat{\mathcal{E}}, \mathbb{P}_{Y_t})$ where $\hat{E}$ is a partition of E extracted from the set of all non-empty parts of E generating the σ-algebra $\hat{\mathcal{E}}$. This new partition will be the new states of E and $\mathbb{P}_{X_t}$ (resp. $\mathbb{P}_{Y_t}$) is the image measure associated to X_t (resp. Y_t).

When φ becomes univalued (which is a particular multivalued case) or is a selection of a multivalued transformation, then φ is denoted f which verifies $f\ :\ (E, \mathcal{E}, \mathbb{P}_{X_t}) \longrightarrow (\hat{E}, \hat{\mathcal{E}}, \mathbb{P}_{Y_t})$.

The problem is illustrated by the next diagrams

$$
\begin{array}{ccc}
(\Omega,\, \mathcal{A},\, \mathbb{P}) & \xrightarrow{\hat{f}} & (\Omega,\, \hat{\mathcal{A}},\, \mathbb{P}) \\
X_t \downarrow \;\uparrow X_t^{-1} & & Y_t^{-1} \uparrow \;\downarrow Y_t \\
(E, \mathcal{E}, \mathbb{P}_{X_t}) & \overset{\varphi}{\multimap} & (\hat{E}, \hat{\mathcal{E}}, \mathbb{P}_{Y_t})
\end{array}
\qquad
\begin{array}{ccc}
(\Omega,\, \mathcal{A},\, \mathbb{P}) & \xrightarrow{\hat{f}} & (\Omega,\, \hat{\mathcal{A}},\, \mathbb{P}) \\
X_t \downarrow \;\uparrow X_t^{-1} & & Y_t^{-1} \uparrow \;\downarrow Y_t \\
(E, \mathcal{E}, \mathbb{P}_{X_t}) & \xrightarrow{f} & (\hat{E}, \hat{\mathcal{E}}, \mathbb{P}_{Y_t})
\end{array}
$$

where X_t^{-1} and Y_t^{-1} are set inverse functions and $\hat{f}$ is a measurable function which verifies $Y_t \circ \hat{f} = f \circ X_t$.

According to compare $\hat{\mathcal{E}}$ and $\mathcal{E}$, three types of situations can arise :

1. If $\hat{\mathcal{E}} \subsetneq \mathcal{E}$, then f is an aggregation (case where f is surjective and non injective).
2. If $\mathcal{E} \subsetneq \hat{\mathcal{E}}$, then f is a selection of the multivalued transformation φ which is considered as a state expansion.
3. If $\mathcal{E} = \hat{\mathcal{E}}$, then f is a bijection which can be explained by a permutation of the states.

3. Aggregation

Let $\mathbb{X}$ a homogeneous Markov chain and f a measurable function from $(E, \mathcal{E})$ to $(\widetilde{E},\, \widetilde{\mathcal{E}})$, we denote for all $t \in \mathbb{T},\; f \circ X_t = \widetilde{X}_t$ and $f \circ \mathbb{X} = \widetilde{\mathbb{X}}$. The states of the transformed process are obtained from a new partition of E . This new partition induces on E a new structure denoted $\widetilde{E}$. The problem becomes to find a univalued function f such that $(f(X_t))_t$ conserves Markov property by reducing states number. We define aggregation by identical profiles *a posteriori* as a Markov transformation provided by the relation $\mathcal{R}$ verifying $\forall t \in \mathbb{T}$ and $\tau > 0$,

$$
\Big((E_i, E_j) \in \mathcal{R}\Big) \Longleftrightarrow \Big(\exists A \in \mathcal{E}\,;\; E_i, E_j \in A, \forall B \in \mathcal{E}, B \subset E - A,
$$
$$
\mathbb{P}(X_{t+\tau} \in B / X_t = E_i) = \mathbb{P}(X_{t+\tau} \in B / X_t = E_j)\Big), \quad (2)
$$

It is obvious that $\mathcal{R}$ is an equivalence relation and A is the class of E_i (or E_j). So we set up a new measurable function f from $(E, \mathcal{E})$ to $(\widetilde{E},\, \widetilde{\mathcal{E}})$ such for all $t \in \mathbb{T}$, $(\Omega, \mathcal{A}, \mathbb{P}) \xrightarrow{X_t} (E, \mathcal{E}) \xrightarrow{f} (\widetilde{E}, \widetilde{\mathcal{E}})$ where $\widetilde{E} = E/\mathcal{R}$. In this construction, it is obvious that $\widetilde{\mathcal{E}} \subset \mathcal{E}$.

Lemma 1.

Let $\mathbb{X}$ a discrete states Markov chain with values in $(E,\, \mathcal{E})$, a measurable surjective application $f : (E, \mathcal{E}) \longrightarrow (\widetilde{E}, \widetilde{\mathcal{E}})$. Let also $A = \{E_i,\, E_j\}$ with non-zero weight verifying $f(E_i) = f(E_j) = E_{i,j} \in \widetilde{\mathcal{E}}$ and the restriction $f_{|E-A}$ is a bijection.

If $\forall (t, \tau) \in \mathbb{T} \times \mathbb{N}^*,\; \forall B \in \mathcal{E}\,; B \subset E - A$,

$$
\begin{aligned}
\mathrm{P}(\mathrm{X}_{t+\tau} \in B / X_t = E_i) &= \mathbb{P}(X_{t+\tau} \in B / X_t = E_j) \\
&= \mathrm{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = E_{i,j})
\end{aligned} \quad (3)
$$

is fulfilled, then $\widetilde{\mathbb{X}}$ is a Markov chain.

Proof.

Let's denote for all $t \in \mathbb{T}, \widetilde{X}_t = f \circ X_t$. To proof the proposition, it is enough to show that

$$\forall\, t, t_1, \cdots, t_k \in \mathbb{T};\ t > t_1 > \cdots > t_k$$

$$\mathbb{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = f(E_t), \widetilde{X}_{t_1} = f(E_{t_1}), \cdots, \widetilde{X}_{t_k} = f(E_{t_k})) = \mathbb{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = f(E_t)) \quad (4)$$

for all $E_t \in E$.

It is obvious that for all $E_t \in E$, $f(E_t) \neq E_{i,j}$, equation (4) is satisfied.

If $f(E_t) = E_{i,j}$, the first term of equation (4) yields to

$$\mathbb{P}(X_{t+\tau} \in B / X_t \in \{E_i, E_j\}, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k})) = \frac{\mathbb{P}(X_{t+\tau} \in B, X_t \in \{E_i, E_j\}, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k}))}{\mathbb{P}(X_t \in \{E_i, E_j\}, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k}))} =$$

$$\Big(\mathbb{P}(X_{t+\tau} \in B, X_t = E_i, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k})) + \mathbb{P}(X_{t+\tau} \in B, X_t = E_j, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k}))\Big) \Big/ \Big(\mathbb{P}(X_t = E_i, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k})) + \mathbb{P}(X_t = E_j, X_{t_1} \in f^{-1} \circ f(E_{t_1}), \cdots, X_{t_k} \in f^{-1} \circ f(E_{t_k}))\Big) =$$

$$\Big(\mathbb{P}(X_{t+\tau} \in B / X_t = E_i).\mathbb{P}(X_t = E_i, X_{t_1} \in f^{-1} \circ f(E_{t_1}), .., X_{t_k} \in f^{-1} \circ f(E_{t_k})) + \mathbb{P}(X_{t+\tau} \in B / X_t = E_j).\mathbb{P}(X_t = E_j, X_{t_1} \in f^{-1} \circ f(E_{t_1}), .., X_{t_k} \in f^{-1} \circ f(E_{t_k}))\Big) \Big/ \Big(\mathbb{P}(X_t = E_i, X_{t_1} \in f^{-1} \circ f(E_{t_1}), .., X_{t_k} \in f^{-1} \circ f(E_{t_k})) + \mathbb{P}(X_t = E_j, X_{t_1} \in f^{-1} \circ f(E_{t_1}), .., X_{t_k} \in f^{-1} \circ f(E_{t_k}))\Big)$$

We require to have $\mathbb{P}(X_{t+\tau} \in B / X_t = E_i) = \mathbb{P}(X_{t+\tau} \in B / X_t = E_j)$ in order to have last term equal to

$$\mathrm{P}(\mathrm{X}_{t+\tau} \in B / X_t = E_i) = \mathbb{P}(f \circ X_{t+\tau} \in f(B) / f \circ X_t = f(E_i)) = \mathrm{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = E_{i,j}) \quad (5)$$

which verifies (4). □

This result can be extended to more than two states.

Lemma 2.

Let $\mathbb{X}$ homogeneous Markov chain and $\widetilde{\mathbb{X}} = f \circ \mathbb{X}$ where f is a measurable function from $(E, \mathcal{E})$ to $(\widetilde{E}, \widetilde{\mathcal{E}})$ such that : $\forall (t,\tau) \in \mathbb{T} \times \mathbb{N}^*$, $\forall E_i,\, E_j \in E$, $\forall B \in \mathcal{E}$ where E_i and E_j verify $(E_i,\, E_j \in B)$ or $(E_i,\, E_j \in E - B)$

$$\mathrm{P}(\mathrm{X}_{t+\tau} \in B / X_t = E_i) = \mathbb{P}(X_{t+\tau} \in B / X_t = E_j) \iff f(E_i) = f(E_j) \qquad (6)$$

then $\widetilde{\mathbb{X}}$ is a Markov process.

Proof.

- Let E_i, E_j in $B \in \mathcal{E}, \tau > 0$ and $\{t,\ t_1,\, t_2, ...\}$ such that $t > t_i, i = 1,\, 2,\, \ldots$,
  $\mathbb{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = \widetilde{E}_t, \widetilde{X}_{t_1} = \widetilde{E}_{t_1}, \cdots, \widetilde{X}_{t_k} = \widetilde{E}_{t_k}) =$

$$\mathbb{P}(X_{t+\tau} \in B /\, X_t \in f^{-1}(\{\widetilde{E}_t\}), X_{t_1} \in f^{-1}(\{\widetilde{E}_{t_1}\}), \cdots, X_{t_k} \in f^{-1}(\{\widetilde{E}_{t_k}\})).$$

  1. If $\widetilde{E}_t \neq f(E_i)$ Markov property is fulfilled.
  2. If $\widetilde{E}_t = f(E_i)$, using lemma 1, one finds

$$\mathbb{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = \widetilde{E}_t, \widetilde{X}_{t_1} = \widetilde{E}_{t_1}, \cdots, \widetilde{X}_{t_k} = \widetilde{E}_{t_k}) = \mathbb{P}(\widetilde{X}_{t+\tau} \in B / \widetilde{X}_t = \widetilde{E}_t)$$

which establishes that $(\widetilde{X}_t)_t$ is a Markov process.

- If E_i and E_j are not in B, let $\tau > 0$, $\{t,\ t_1,\, t_2, ...\}$ such that $t > t_1 > t_2 > \ldots$, then
  1. If $f^{-1}(\{\widetilde{E}_t\})$ is a singleton (i.e $\# f^{-1}(\{\widetilde{E}_t\}) = 1$), then $\widetilde{X}$ verifies Markov property.
  2. Else if $f^{-1}(\{\widetilde{E}_t\})$ is not a singleton (i.e $\# f^{-1}(\{\widetilde{E}_t\}) = k > 1$), let

$$\mathbb{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = \widetilde{E}_t, \widetilde{X}_{t_1} = \widetilde{E}_{t_1},\, \widetilde{X}_{t_2} = \widetilde{E}_{t_2}, \cdots) =$$
$$\mathbb{P}(X_{t+\tau} \in B / X_t \in f^{-1}(\{\widetilde{E}_t\}), X_{t_1} \in f^{-1}(\{\widetilde{E}_{t_1}\}), X_{t_2} \in f^{-1}(\{\widetilde{E}_{t_2}\}), \ldots) =$$
$$\frac{\sum\limits_{E_i \in f^{-1}(\{\widetilde{E}_t\})} \mathbb{P}(X_{t+\tau} \in B / X_t = E_i).\mathbb{P}(X_t = E_i, X_{t_1} \in f^{-1}(\{\widetilde{E}_{t_1}\}), X_{t_2} \in f^{-1}(\{\widetilde{E}_{t_2}\}), \ldots)}{\sum\limits_{E_i \in f^{-1}(\{\widetilde{E}_t\})} \mathbb{P}(X_t = E_i, X_{t_1} \in f^{-1}(\{\widetilde{E}_{t_1}\}), X_{t_2} \in f^{-1}(\{\widetilde{E}_{t_2}\}), \ldots)} =$$
$$\mathbb{P}(X_{t+\tau} \in B / X_t \in f^{-1}(\{\widetilde{E}_t\})) = \mathbb{P}(\widetilde{X}_{t+\tau} \in f(B) / \widetilde{X}_t = \widetilde{E}_t)$$

which yields that $f(\mathbb{X})$ is a Markov chain. □

Under Dynkin's criterion, $\mathcal{R}$ characterises the strong lumpability defined by Kemeny and Snell **(Kemeny & Snell, 1976)**; we call it *strong Markov aggregation*. So the transition matrix of Markov chain transformed is presented in the following proposition.

Proposition 1. (Aggregation by identical profiles a posteriori)

Let $(X_t)_{t\in\mathbb{T}}$ be a Markov chain on a state space E with transition matrix $P_{t,t+\tau}$, such that $P_{t,t+\tau}(x,y)=\mathbb{P}(X_{t+\tau}=y/X_t=x)$ and let $f:E\to E/\mathcal{R}$. Then $(X_t)_t$ has a strong Markov aggregation via f if for every $E_i,E_j\in E$ with $f(E_i)=f(E_j)$, and every $\widetilde{y}\in E/\mathcal{R}$, the transition probability sums satisfy

$$\sum_{\{y;f(y)=\widetilde{y}\}} P_{t,t+\tau}(E_i,\,y)=\sum_{\{y;f(y)=\widetilde{y}\}} P_{t,t+\tau}(E_j,\,y). \tag{7}$$

The aggregated Markov chain has transition matrix $\widetilde{P}_{t,t+\tau}$ which verifies

$$\widetilde{P}_{t,t+\tau}(\widetilde{X},\,\widetilde{y})=\sum_{\{y;f(y)=\widetilde{y}\}} P_{t,t+\tau}(x,\,y) \tag{8}$$

for any x with $f(x)=\widetilde{X}$.

Remarks

1) Using the reverse chain, one can do an analogous approach using Rogers-Pitman criterion **(Rogers & Pitman's, 1981)**. Weak Markov aggregation depends on the initial distribution; otherwise Markov aggregation is said strong.
2) A process is Markov aggregated if it allows a Markov aggregation under Dynkin criterion or under that Rogers-Pitman.
3) A process is Markov aggregated weakly if it is Markov aggregated and depends on the initial distribution ; otherwise it is strong Markov aggregated process.

Proposition 2.

Let $\mathbb{X}$ an homogeneous Markov chain and $\widetilde{\mathbb{X}}$ its transformed process such that for all $t\in\mathbb{T}$, $\widetilde{X}_t=f\circ X_t$ with f a measurable application from $(E,\mathcal{E})$ to $(\widetilde{E},\widetilde{\mathcal{E}})$.

Process $\widetilde{X}$ is a Markov chain if and only if for any measurable set $A\in\widetilde{\mathcal{E}}$ and any $\widetilde{E}_k\in\widetilde{E}$, we have

$$\mathbb{P}(X_{t+\tau}\in f^{-1}(A)/\,X_t=E_j)=c(A,\widetilde{E}_k,\tau),\ \forall\,E_j\in f^{-1}(\{\widetilde{E}_k\}) \tag{9}$$

Proof.

- If (9) is fulfilled, then, using lemma 2, one finds that $\widetilde{\mathbb{X}}$ is a homogeneous Markov chain.
- Let $A\in\mathcal{E}$, $\tau>0$ and $\{t,\ t_1,\,t_2,\ldots\}$ such that $t>t_1>t_2>\ldots$ where $\widetilde{E}_t=\widetilde{E}_k$.

  If $\widetilde{\mathbb{X}}$ is a homogeneous Markov chain, then

$$\mathbb{P}(\widetilde{X}_{t+\tau}\in A/\,\widetilde{X}_t=\widetilde{E}_t,\widetilde{X}_{t_1}=\widetilde{E}_{t_1},\widetilde{X}_{t_2}=\widetilde{E}_{t_2},\cdots)=$$

$$\mathbb{P}(\widetilde{X}_{t+\tau}\in A/\,\widetilde{X}_t=\widetilde{E}_t)=\mathbb{P}(X_{t+\tau}\in f^{-1}(A)/\,X_t\in f^{-1}(\{\widetilde{E}_k\})=c(A,\widetilde{E}_k,\tau) \tag{10}$$

As $\widetilde{\mathbb{X}}$ is markovian by aggregation, from (10), we end up with the result $\mathbb{P}(X_{t+\tau}\in f^{-1}(A)/\,X_t=E_j)=c(A,\widetilde{E}_k,\tau),\ \forall\,E_j\in f^{-1}(\{\widetilde{E}_k\})$. □

Under assumptions of proposition 2, let us consider $\widetilde{\mathbb{X}}$ as a transformation of an homogeneous Markov chain $\mathbb{X}$ by the function f. Without loss of generality, let's denote $\{E_{k_1},E_{k_2},\ldots,E_{k_s}\}=f^{-1}(\{\widetilde{E}_k\})$, then for all $A\in\widetilde{\mathcal{E}}$ and $\widetilde{E}_k$ a state of $\widetilde{E}$

$$\mathbb{P}(\widetilde{X}_{t+\tau} \in A/ \widetilde{X}_t = \widetilde{E}_k) = \mathbb{P}(X_{t+\tau} \in f^{-1}(A)/ X_t \in f^{-1}(\{\widetilde{E}_k\})$$

More precisely, one has

$$\mathbb{P}(X_{t+\tau} = E_j/X_t \in f^{-1}(\{\widetilde{E}_k\}) = \frac{\sum_{i=1}^{s} \mathbb{P}(X_{t+\tau} = E_j/X_t = E_{k_i}).\mathbb{P}(X_t = E_{k_i})}{\sum_{i=1}^{s} \mathbb{P}(X_t = E_{k_i})}$$

So $\mathbb{P}(\widetilde{X}_{t+\tau} \in A/ \widetilde{X}_t = \widetilde{E}_k)$ can be written

$$\sum_{\{E_j \in f^{-1}(A)\}} \sum_{i=1}^{s} \alpha_i \mathbb{P}(X_{t+\tau} = E_j/X_t = E_{k_i}) \quad \text{with } \alpha_i \geq 0 \text{ and } \sum_{i=1}^{s} \alpha_i = 1$$

this means that aggregative state is built from a linear convex combination of initial conditional distributions as vectors on unitary sphere of $\mathbb{R}^s$ for the 1-norm. One remarks that we have built a novel state which is a combination of other states or a sum of substates which we name “fractional state”.

Definition 1.

Let $E_i,\ i = 1,\ \ldots,\ r$, a state of the Markov chain $\mathbb{X}$ and $E_{ik}\,;\ k = 1, ..., k_i$ a partition of E_i.

Sub-state E_{ik} $(i = 1,\ \ldots,\ r$ and $k = 1, ..., k_i)$ is said α_k-fractional state *a posteriori* of E_i $(\alpha_k \in [0,\ 1])$ relatively to $\mathbb{X}$,

1. if $\mathbb{P}(\hat{X}_t = E_{ik}) = \alpha_k \mathbb{P}(X_t = E_i)$
2. and if $\exists A \in \mathcal{E}\,;\ \forall E_i, E_j \in A, \quad \text{and} \quad \forall B \in \mathcal{E}, B \subset E - A$ verifying

$$\sum_{k=1}^{k_i} \mathbb{P}(\hat{X}_t = E_{ik}) = \mathbb{P}(X_t = E_i) \quad \text{and} \qquad \sum_{k'=1}^{k_j} \mathbb{P}(\hat{X}_t = E_{jk'}) = \mathbb{P}(X_t = E_j)$$

$$\begin{aligned} \mathbb{P}(X_{t+\tau} \in B/X_t = E_i) \quad &= \mathbb{P}(X_{t+\tau} \in B/X_t = E_j) \\ &= \mathbb{P}(\hat{X}_{t+\tau} \in B/\hat{X}_t = E_{ik}) \\ &= \mathbb{P}(\hat{X}_{t+\tau} \in B/\hat{X}_t = E_{jk'}) \end{aligned}$$

$$\forall k = 1, ..., k_i \quad \forall k' = 1, ..., k_j.$$

The previous definition means that row of transition matrix $\hat{P}_{t,t+\tau}$ associated to α_k-fractional state E_{ik} obtained from E_i is the same as that row of transition matrix $P_{t,t+\tau}$ outside columns associated to E_i where columns of the novel transition matrix verify for transition matrix $\hat{P}_{t,t+\tau}$

$$\sum_{k=1}^{k_i} \hat{P}_{t,t+\tau}(E_{i1},\ E_{ik}) = P_{t,t+\tau}(E_i,\ E_i)$$

such that

$$\hat{P}_{t,t+\tau}(E_{i1},\, E_{ik}) = \alpha_k\, P_{t,t+\tau}(E_i,\, E_i)$$

So conditional distributions $\{\mathbb{P}(\ .\ /X_t = E_{k_i}); i = 1, \ldots, s\}$ generate a vectorial subspace with dimension $d; 1 \leq d \leq s$. If $d < s$, there are endless ways to build an initial Markov chain which gives $\widetilde{\mathbb{X}}$. If generated subspace dimension is 1, then the solution consists to aggregate only states which have the same conditional distributions. So contribution of each E_{k_i} is a fraction of that $f^{-1}(\widetilde{E}_k)$.

This algebraic perception of the problem gives broad perspectives to track the problem linearly. Generally, there exists a lot of Markov chains giving $\widetilde{\mathbb{X}}$. The reverse approach which is an expanded approach, named "disaggregation", introduces naturally fractional states and multivalued functions.

4. Disaggregation

Let $\mathbb{X}$ a homogeneous Markov Chain and $\widetilde{\mathbb{X}}$ a Markov chain which is an aggregating transformation of $\mathbb{X}$ such that for all $t \in \mathbb{T}, \widetilde{X}_t = f \circ X_t$ with f a measurable application from $(E, \mathcal{E})$ to $(\widetilde{E}, \widetilde{\mathcal{E}})$.

Using aggregation, it is obvious that f is surjective and its inverse function, say φ, is multivalued.

In the sequel, one considers $\hat{X}_t = \varphi(X_t)$

$$(\Omega, \mathcal{A}, \mathbb{P}) \quad \xrightarrow{X_t} \quad (E, \mathcal{E}) \quad \stackrel{\varphi}{\multimap} (\hat{E}, \hat{\mathcal{E}})$$

where φ is multivalued and $\mathcal{E} \subset \hat{\mathcal{E}}$. So $\{\hat{E}_1,\, \hat{E}_2,\, \ldots,\, \hat{E}_s\}$ is a new partition of E constituting the set of $\hat{X}_t$ states $(s \geq r)$ where $\hat{\mathcal{E}} \subset 2^E$ which is set of all the subsets of the topological space E ($\mathbb{X}$ can be one case of $\hat{\mathbb{X}}$).

We are interested, in the following, to insure conditions that a derived chain $\Big(\varphi(X_t)\Big)_t$, denoted also $\varphi \circ \mathbb{X}$, obtained from Markov chain $\mathbb{X}$ with initial measure ν and transition matrix P be Markovian whatever the initial distribution ν may be. One remarks that there exist two types of expansion : Time expansion and State expansion.

4.1 Time expansion

Time expansion for an irreducible homogeneous Markov Chain $\mathbb{X}$ consists to consider different finite trajectories as states of the novel process. So, instead studying $\mathbb{X}$, one considers process $\mathbb{S}$ whose states are obtained from ($k+1$)-uples (trajectories) of $\mathbb{X}$. It is obvious that $(S_t)_t$, $S_t = (X_t,\, X_{t-1},\, \ldots,\, X_{t-k})_t$ is a Markov process verifying $\forall\, t \in \mathbb{T}$;

$$\mathbb{P}\Big(S_{t+1} \in B_0 \times B_1 \cdots \times B_k / S_t = (x_t, x_{t-1}, \cdots, x_{t-k}), S_{t-1} = (x_{t-1}, x_{t-2}, \cdots, x_{t-k}), \cdots\Big)$$

$$= \begin{cases} \mathbb{P}\Big(S_{t+1} \in B_0 \times B_1 \cdots \times B_k \quad / \quad S_t = (x_t, x_{t-1}, \cdots, x_{t-k})\Big) \\ \qquad\qquad\qquad \text{if } x_t \in B_1 \text{ and } x_{t-1} \in B_2 \text{ and } \cdots \text{ and } x_{t-k} \in B_k \\ 0 \\ \qquad\qquad\qquad \text{if } x_t \notin B_1 \text{ or } x_{t-1} \notin B_2 \text{ or } \cdots \text{ or } x_{t-k} \notin B_k. \end{cases}$$

For $k = 1$, one finds Kemeny and Snell approach for expanding Markov chain **(Kemeny & Snell, 1976) [in Chap VI, § 6.5]**. On the new process obtained by time expansion, one can study as well as Markov aggregation than Markov disaggregation.

Time expansion can detect order expansion which transforms a non markovian process to a markovian one where states of the Markov chain may be finite trajectories instead original states. It's another problem which is not developped in this work.

4.2 State expansion

Proposition 3.

Let $\mathbb{X}$ a Markov chain taking values in a partition of a topological space E and $\varphi : (E,\mathcal{E}) \multimap (\hat{E},\hat{\mathcal{E}})$ a multivalued transformation which expands $\mathbb{X}$ by fractioning its states *a posteriori*, then $\varphi \circ \mathbb{X} = \hat{\mathbb{X}}$ is a Markov chain.

Proof.

Using definition 1, let $E_{ik}; k = 1, ..., k_i$ a partition of $E_i;\ i = 1, \ldots, r.$ α_k-fractional state *a posteriori* of $E_i(\alpha_k \in [0,\ 1])$ relatively to $\mathbb{X}$. One finds that these disaggregation of $\mathbb{X}$ leaves the process Markovian. Just check that for all strictly decreasing sequence $(t_i)_{i=1,\ldots,n},\ n \in \mathbb{T}$ and $\tau \in \mathbb{N}^*$ such that $t > t_1$,

$$\mathbb{P}(\hat{X}_{t+\tau} = E_{ij}/\hat{X}_t = E_{l_0 m_0}, \hat{X}_{t_1} = E_{l_1 m_1}, \ldots, \hat{X}_{t_n} = E_{l_n m_n}) = \mathbb{P}(\hat{X}_{t+\tau} = E_{ij}/\hat{X}_t = E_{l_0 m_0})$$

which is always verified. □

Proposition 4.

Let $(\Omega,\mathcal{A},\mathbb{P}) \xrightarrow{X_t} (E,\mathcal{E}) \overset{\varphi}{\multimap} (\hat{E},\hat{\mathcal{E}})$ where φ is a multivalued function fractioning E_i. If

$$\mathbb{P}(X_{t+\tau} = ./\,X_t = E_i) = \sum_{j=1;\ j\neq i}^{r} \alpha_j(t,\tau).\mathbb{P}(X_{t+\tau} = ./\,X_t = E_j),$$

E_i is decomposed into $E_{i1}, \ldots, E_{i(i-1)}, E_{i(i+1)}, \ldots, E_{ir}$ which are respectively α_j-fractional state of E_i; $\alpha_j(t,\tau) \geq 0,\ \ j \neq i;\ 1 \leq j \leq r$ checking $\sum_{j=1;\ j\neq i}^{r} \alpha_j(t,\tau) = 1$ then the new process $\widetilde{\mathbb{X}}$ taking values in $\widetilde{E} = E - \{E_i\}$ and verifying $\forall\, E_j,\ E_k \in \widetilde{E}$,

$$\mathbb{P}(\widetilde{X}_{t+\tau} = E_k/\,\widetilde{X}_t = E_j) = \mathbb{P}(X_{t+\tau} = E_k/\,X_t = E_j) + \alpha_k(t,\tau).\mathbb{P}(X_{t+\tau} = E_i/\,X_t = E_j)$$

is a Markov process.

Proof.

Without loss of generality, let $i = 1$ and $\{E_{1j}\ ;\ j = 2\ldots, r\}$ a set of fractional states of E_1. Let $\hat{\mathbb{X}} = \varphi \circ \mathbb{X}$ which takes values in $\widetilde{E} \bigcup \{E_{i2}, \ldots, E_{ir}\}$. One can verify easily that $\hat{\mathbb{X}}$ is markovian (see proposition 3).

It is obvious that $\widetilde{\mathbb{X}}$ is obtained from $\hat{\mathbb{X}}$ by aggregating E_j and E_{1j}. So $\widetilde{\mathbb{X}}$ is a Markov process which verifies

$$\mathbb{P}(\widetilde{X}_{t+\tau} = E_k/\,\widetilde{X}_t = E_j) = \mathbb{P}(X_{t+\tau} = E_k/\,X_t = E_j) + \alpha_k(t,\tau).\mathbb{P}(X_{t+\tau} = E_i/\,X_t = E_j)$$

□

The transformation of $\mathbb{X}$ to $\hat{\mathbb{X}}$ (disaggregation by a multivalued transformation) preserving the Markovian property of the process can be extended even to an infinity of states.

Proposition 2 can be extended to $\varphi \circ \mathbb{X}$ where φ is multivalued. Thus allows us to state the

Proposition 5.

Let $\mathbb{X}$ homogeneous Markov chain and $\hat{\mathbb{X}}$ its transformed process such that for all $t \in \mathbb{T},\ \hat{X}_t = \varphi \circ X_t$ with φ a multivalued measurable application from $(E, \mathcal{E})$ to $(\hat{E}, \hat{\mathcal{E}})$.

Process $\hat{X}$ is a Markov chain if and only if for any measurable set $B \in \hat{\mathcal{E}}$ and any $\hat{E}_k \in \hat{E}$, we have

$$\mathbb{P}(X_{t+\tau} \in \varphi^{-1}(B)/\, X_t = E_j) = c(B, \hat{E}_k, \tau),\ \forall\, E_j \in \varphi^{-1}(\{\hat{E}_k\}) \tag{11}$$

Proof.

Here φ is strong measurable, because space of states is discrete. Furthermore, considering all the selections f of multivalued transformation φ, the proof becomes similar to that of proposition 2. □

4.3 A method to construct a markovian state expansion

Let $(\Omega, \mathcal{A}, \mathbb{P}) \quad \overset{X_t}{\longrightarrow} \quad (E, \mathcal{E}) \quad \overset{\varphi}{\multimap}\ (\hat{E}, \hat{\mathcal{E}})$ where φ is a multivalued function.

Let $E_i \in E$ a state of the Markov chain which will be fractioned in k_i substates $E_{il},\ l = 1, \cdots, k_i$; $(k_i > 1)$.

If $k_i = 1$ we presume that E_i is not fractioned. The topological space E is transformed into $\hat{E} = \varphi(E)$ where the novel state space is presented as follows $\hat{E} = (E - \{E_i\}) \bigcup \{E_{i1}, \cdots, E_{ik_i}\}$ which is denoted $E^{[1,\ldots,1,k_i,1\ldots,1]}$ where k_i is in i^{th} position.

Generally $\hat{E}$ with pure fractional states must be presented $\hat{E} = E^{[k_1,\ldots,k_j,\ldots,k_r]}$

where k_j represents the number of fractional states obtained from E_j; $j = 1, \ldots, r$.

We can obtain other fractional states by convex combinations of pure fractional states. These states obtained by linear combination satisfy an aggregation. When $\hat{E}$ is constituted only by pure fractional states, $\hat{X}$, defined previously, is a Markov chain with $\sum_{j=1}^{r} k_j$ states. Generally $\hat{E}$ is constituted by pure fractional states et by convex combination of pure fractional states.

To illustrate a case of the transformation of a transition matrix where $\hat{E} = E^{[1,\ldots,1,k_j,1\ldots,1]}$, we present, in the following, a construction of the transition matrix $\hat{P}$ knowing P:

$$\left(\begin{array}{ccccccccc}
p_{11} & \cdots & p_{1\,(j-1)} & \hat{p}_{1j} & \cdots & \hat{p}_{1(j+k_j-1)} & p_{1\,(j+1)} & \cdots & p_{1\,r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
p_{(j-1)\,1} & \cdots & p_{(j-1)\,(j-1)} & \hat{p}_{(j-1)j} & \cdots & \hat{p}_{(j-1)(j+k_j-1)} & p_{(j-1)\,(j+1)} & \cdots & p_{(j-1)\,r} \\
\hline\hline
p_{j1}^{(0)} & \cdots & p_{j\,(j-1)}^{(0)} & \hat{p}_{jj}^{(0)} & \cdots & \hat{p}_{j(j+k_j-1)}^{(0)} & p_{j\,(j+1)}^{(0)} & \cdots & p_{j\,r}^{(0)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
p_{j1}^{(k_j-1)} & \cdots & p_{j\,(j-1)}^{(k_j-1)} & \hat{p}_{jj}^{(k_j-1)} & \cdots & \hat{p}_{j(j+k_j-1)}^{(k_j-1)} & p_{j\,(j+1)}^{(k_j-1)} & \cdots & p_{j\,r}^{(k_j-1)} \\
\hline\hline
p_{(j+1)\,1} & \cdots & p_{(j+1)\,(j-1)} & \hat{p}_{(j+1)j} & \cdots & \hat{p}_{(j+1)(j+k_j-1)} & p_{(j+1)\,(j+1)} & \cdots & p_{(i+1)\,r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
p_{r\,1} & \cdots & p_{r\,(j-1)} & \hat{p}_{rj} & \cdots & \hat{p}_{r(j+k_j-1)} & p_{r\,(j+1)} & \cdots & p_{r\,r}
\end{array}\right)$$

with :

$$\sum_{m=j}^{j+k_j-1} \hat{p}_{im} = p_{ij}\ ;\ i = 1, \cdots, r;\ i \neq j.$$

$$\sum_{m=j}^{j+k_j-1} \hat{p}_{jm}^{(n)} = p_{jj}\ ;\ n = 0, \cdots, k_j - 1.$$

$$p_{jm}^{(n)} = p_{jm};\ n = 0, \cdots, k_j - 1.$$

To construct a transition matrix $\hat{P}^{(1)}$ knowing $\hat{P}$ with $\hat{E} = E^{[1,\ldots,1,k_j,1\ldots,1,k_{j'},1\ldots,1]}$, we use the same approach; and so on.

On the reverse dynamic of the Markov chain, the same technique can be done.

5. Application on graph

Let $G = (V, E)$ be a directed graph with $|V| = r$ and $|E| = m$; where V and E represent respectively vertices and edges. We can consider it as a representation of a Markov chain with r states. The weight of the edge $i \to j$ is equals to conditional probability

$$p_{ij} = P(X_{t+1} = E_j /\ X_t = E_i)$$

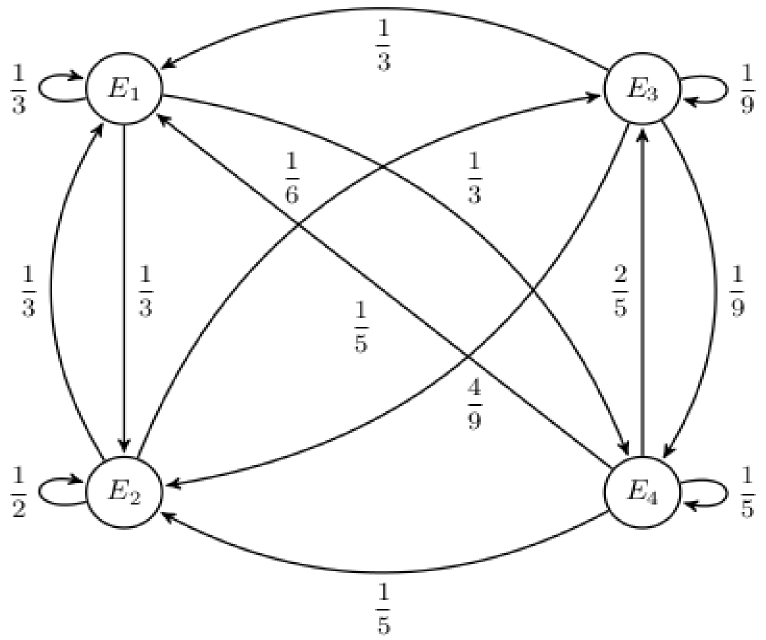
5.1 Exemple 1

Let $\mathbb{X}$ an homogeneous continuous time Markov process with values in state space $E = \{E_1, E_2, E_3, E_4\}$ and transition matrix $P_{t,t+1}$ denoted P:

$$P = \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/2 & 1/6 & 0 \\ 1/3 & 4/9 & 1/9 & 1/9 \\ 1/5 & 1/5 & 2/5 & 1/5 \end{pmatrix}$$

$\mathbb{X}$ is represented by the following graph:

Figure.1 Graph of $\mathbb{X}$



Since

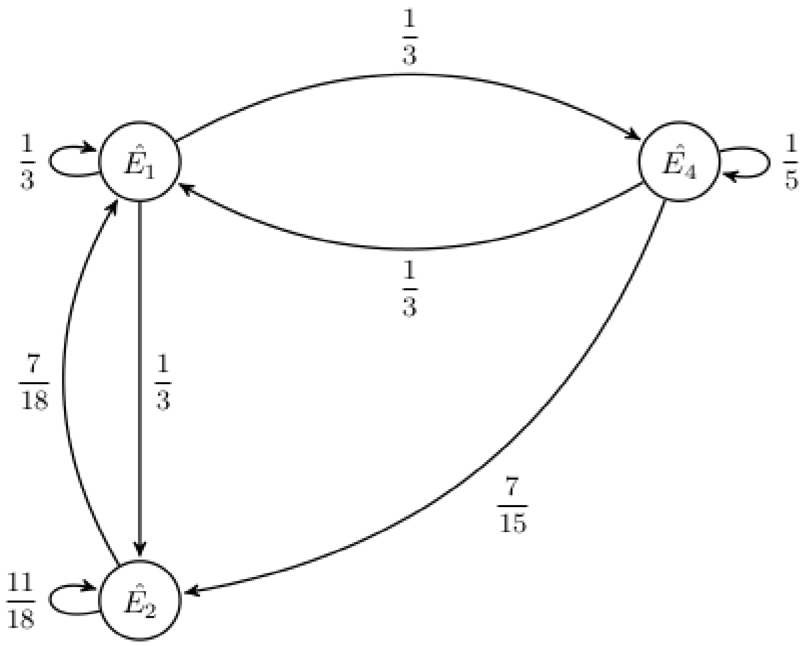
$\mathbb{P}(X_{t+1} = ./X_t = E_3) = \mathbb{P}(X_{t+1} = ./X_t = E_1)/3 + 2.\mathbb{P}(X_{t+1} = ./X_t = E_2)/3$ and using proposition 4, E is transformed to $\hat{E} = \{\hat{E}_1,\ \hat{E}_2,\ \hat{E}_4\}$ which is composed from fractional states.

Thus, one gets a Markov process $\widetilde{\hat{\mathbb{X}}}$ having transition matrix from t to $t+1$ noted $\widetilde{\hat{P}}$:

$$\widetilde{\hat{P}} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 7/18 & 11/18 & 0 \\ 1/3 & 7/15 & 1/5 \end{pmatrix}$$

and following graph:

Figure. 2 Graph aggregated



which is non lumpable neither under Dynkin criterion nor Rogers-Pitman's one. The stationary distribution of $\mathbb{X}$ (resp. $\widetilde{\hat{\mathbb{X}}}$) is

$$(p_1 \quad p_2 \quad p_3 \quad p_4) = (0.3133047 \quad 0.3948498 \quad 0.1416309 \quad 0.1502146)$$
$$\left(\text{resp. } (\widetilde{p_1} \quad \widetilde{p_2} \quad \widetilde{p_3}) = (0.360515 \quad 0.4892704 \quad 0.1502146) \right)$$

as expected, one gets

$\widetilde{p_1} = p_1 + p_3/3$, $\widetilde{p_2} = p_2 + 2p_3/3$.

One verifies also that P and $\widetilde{\hat{P}}$

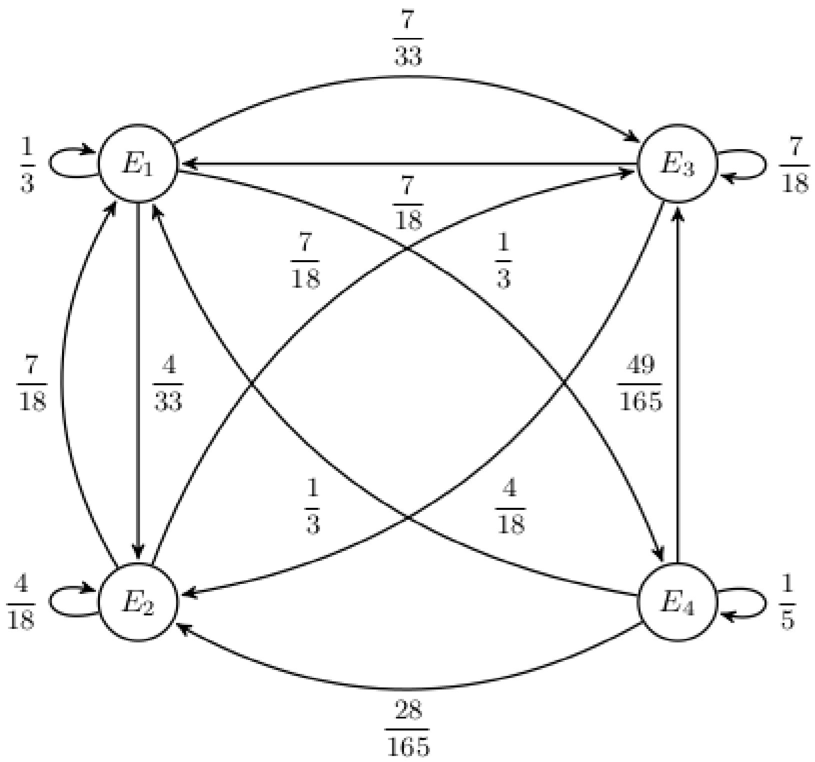
have same non null eigenvalues :
$1,\ 0.07222222 + 0.04681194\,i,\ 0.07222222 - 0.04681194\,i$.

Disaggregating the state E_2 into two fractional states E_{21} and E_{22}, and using the previous construction, we can construct an infinity of Markov processes which can be lumped to the Markov chain $\mathbb{X}$.

In the following, we present two cases : first, we endow the fractional states with relatively identical conditional weights; let 4 for E_{21} and 7 for E_{22}. The transition matrix and graph from t to $t+1$ of Markov chain obtained by this first type of expansion noted $\hat{P}^{(1)}$ is presented below :

$$\hat{P}^{(1)} = \begin{pmatrix} 1/3 & 4/33 & 7/33 & 1/3 \\ 7/18 & 4/18 & 7/18 & 0 \\ 7/18 & 4/18 & 7/18 & 0 \\ 1/3 & 28/165 & 49/165 & 1/5 \end{pmatrix}$$

Figure. 3 Graph expanded by identical conditional weights



Stationary distribution of $\widetilde{\hat{\mathbb{X}}}$ (resp. $\hat{\mathbb{X}}^{(1)} \equiv \hat{\widetilde{\hat{\mathbb{X}}}}$) is

$$\begin{pmatrix} \widetilde{\hat{p}_1} & \widetilde{\hat{p}_2} & \widetilde{\hat{p}_3} \end{pmatrix} = \begin{pmatrix} 0.360515 & 0.4892704 & 0.1502146 \end{pmatrix}$$

$$\left(\text{resp. } \begin{pmatrix} \hat{p_1}^{(1)} & \hat{p_{21}}^{(1)} & \hat{p_{22}}^{(1)} & \hat{p_3}^{(1)} \end{pmatrix} = \begin{pmatrix} 0.3605150 & 0.1779165 = \frac{4}{11}p_2 & 0.3113539 = \frac{7}{11}p_2 & 0.1502146 \end{pmatrix} \right)$$

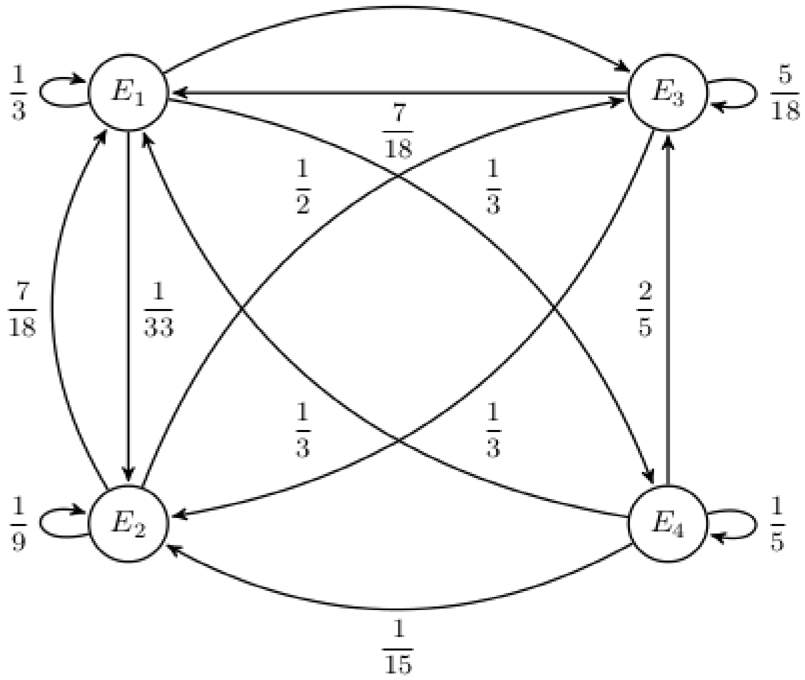
One verifies also that $\widetilde{\hat{P}}$ and $\hat{P}^{(1)}$ have same non null eigenvalues :

$$1,\ 0.07222222 + 0.04681194\,i,\ 0.07222222 - 0.04681194\,i.$$

For the second case of expansion, we endow the fractional states with relatively random conditional weights. The transition matrix and graph from t to $t+1$ of Markov chain obtained by second type of expansion noted $\hat{P}^{(2)}$ is presented below

$$\hat{P}^{(2)} = \begin{pmatrix} 1/3 & 1/33 & 10/33 & 1/3 \\ 7/18 & 1/9 & 1/2 & 0 \\ 7/18 & 1/3 & 5/18 & 0 \\ 1/3 & 1/15 & 2/5 & 1/5 \end{pmatrix}$$

Figure. 4 Graph expanded by random conditional weights



Stationary distribution of $\hat{\mathbb{X}}^{(2)}$ is

$$(\hat{p_1}^{(2)} \quad \hat{p_{21}}^{(2)} \quad \hat{p_{22}}^{(2)} \quad \hat{p_3}^{(2)}) = (0.3605150 \quad 0.1505693 \quad 0.3387011 \quad 0.1502146)$$

One verifies that all non null eigenvalues of $\widetilde{\hat{P}}$ are included in the set of non null eigenvalues of $\hat{P}^{(2)}$ which are :

$$1,\ -0.22222222,\ 0.07222222 + 0.04681194\, i,\ 0.07222222 - 0.04681194\, i.$$

Another expansion based on convex combination of fractional states extracted from states of original process can be envisaged.

Through theses examples, using spectral analysis of the transition matrices of aggregated or disaggregated processes, one can remark that this study can be extended easily to a topological analysis approach of aggregation or disaggregation.

5.2 Exemple 2

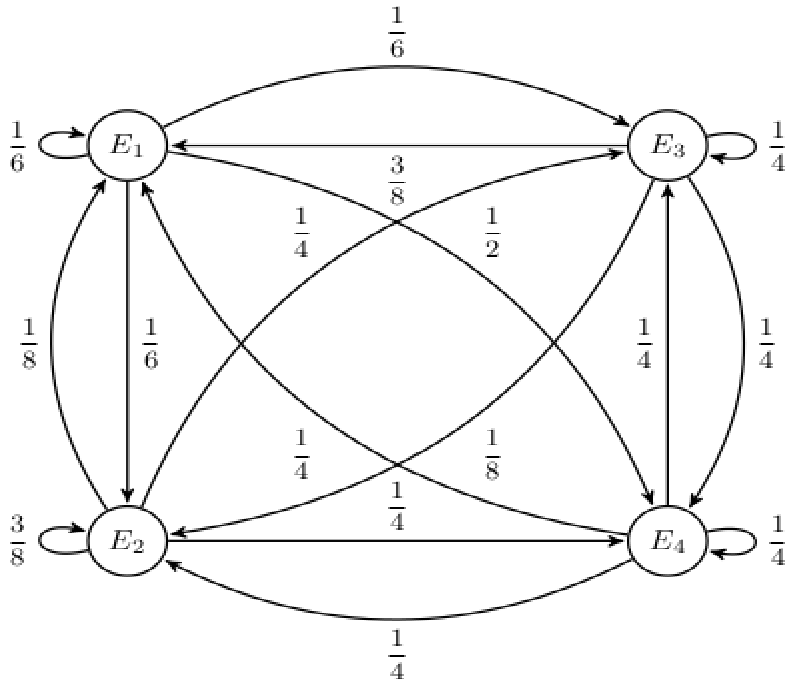
Rubino and Sericola **(Rubino & Sericola, 1989),** give the following example :

Let $\mathbb{X}$ an homogeneous continuous time Markov process with values in state space $E = \{E_1, E_2, E_3, E_4\}$ and transition matrix $P_{t,t+1}$ denoted P :

$$P = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/2 \\ 1/8 & 3/8 & 1/4 & 1/4 \\ 3/8 & 1/8 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

represented by the following graph :

Figure. 5 Graph exemple given by Rubino and Sericola **(Rubino & Sericola, 1989)**

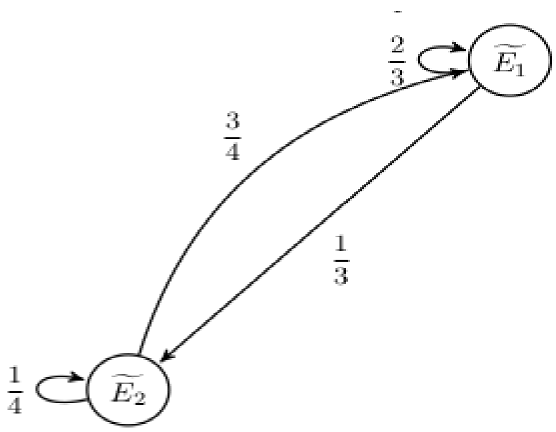


Under Roger-Pitman's criterion, the authors lump $\mathbb{X}$ to $\widetilde{\mathbb{X}}$ which has two states $\widetilde{E}_1 = \{E_1,\ E_2,\ E_3\}$ and $\widetilde{E}_2 = \{E_4\}$. So $\widetilde{\mathbb{X}}$ has following transition matrix (reversed time) from t to $t-1$ noted $\widetilde{P}$:

$$\widetilde{P} = \begin{pmatrix} 2/3 & 1/3 \\ 3/4 & 1/4 \end{pmatrix}$$

and graph bellow:

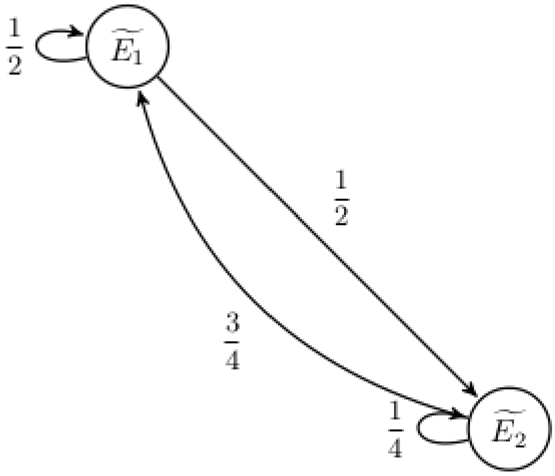
Figure. 6 Graph reduced (reversed time) **(Rubino & Sericola, 1989)**



Under Dynkin criterion, lumping $\mathbb{X}$ to $\widetilde{\mathbb{X}}$ by $\mathcal{R}$ aggregation gives another partition of E denoted also $\widetilde{E}_1,\ \widetilde{E}_2$ where $\widetilde{E}_1 = \{E_1\}$ and $\widetilde{E}_2 = \{E_2,\ E_3,\ E_4\}$. Associated transition matrix and graph of from t to $t+1$ noted $\widetilde{P}$ is:

$$\widetilde{P} = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

Figure. 7 Graph aggregated **(Rubino & Sericola, 1989)**



Using spectral analysis, one extracts firstly all linear combinations before aggregating states (or fractional states) to finally uses lumping under Dynkin criterion. Here, one has

$$\mathbb{P}(X_{t+1} = ./X_t = E_4) = \mathbb{P}(X_{t+1} = ./X_t = E_2)/2 + \mathbb{P}(X_{t+1} = ./X_t = E_3)/2$$

and using proposition 4, E is transformed to $\hat{E} = \{\hat{E}_1,\ \hat{E}_2,\ \hat{E}_3\}$ where $\hat{E}_1 = E_1$ and $\hat{E}_2$ (resp. $\hat{E}_3$) contains E_2 (resp. $\hat{E}_3$) with a fractional state from E_4.

$\widetilde{\hat{\mathbb{X}}}$ having transition matrix from t to $t+1$ noted $\widetilde{\hat{P}}$:

$$\widetilde{\hat{P}} = \begin{pmatrix} 1/6 & 1/3 & 1/2 \\ 1/8 & 5/8 & 1/4 \\ 3/8 & 3/8 & 1/4 \end{pmatrix}$$

Then we use Dynkin criterion to obtain an aggregate Markov process $\widetilde{\widetilde{\hat{\mathbb{X}}}}$ having transition matrix from t to t+1 noted $\widetilde{\widetilde{\hat{P}}}$:

$$\widetilde{\widetilde{\hat{P}}} = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

The same technique can be used to study aggregation conserving Markov nature of $\mathbb{X}$ with Roger-Pitman criterion.

5. Conclusion

This approach allows Markov chain us to build many other Markov chain which describe the same process. Reducing will be equivalent to lumping and expanding to disaggregation using fractional states.

In the literarture, one finds some necessary and sufficient condition preserving Markov chain naure. But its the first time where a theorem gives a necessary ans sufficient condition to aggregate or disaggregate state of Markov chain whitout lossing Markovian dynamic of the process.

Acknowledgement

This research was partially funded with support from a grant of DGRSDT and ATRST (Algiers) to the Mathematics Laboratory (LDM) as well as a grant from MESRS (PRFU : C00L03UN220120200001 and PRFU : C00L03UN220120180012).

References

Burke, C. J., & Rosenblatt, M. (1958). A Markovian function of a Markov chain. The Annals of Mathematical Statistics, 29(4), 1112-1122.

Berge, C. (1959). Espaces Topologiques: Fonctions Multivoques, 2ème éd., Topological Spaces.

Dynkin, E.B. (1965). Markov Processes (Volume 1), Springer-Verlag Berlin Heidelberg.

Geiger, B. C., & Temmel, C. (2014). Lumpings of Markov chains, entropy rate preservation, and higher-order lumpability. Journal of Applied Probability, 51(4), 1114-1132.

Hachigian, J. (1963). Collapsed Markov chains and the Chapman-Kolmogorov equation. The Annals of Mathematical Statistics, 34(1), 233-237.

Kemeny, J. G., Snell, J.L. (1976). Finite Markov chains, Springer-Verlag.

Ledoux, J., Rubino, G., & Sericola, B. (1994). Exact aggregation of absorbing Markov processes using the

quasi-stationary distribution. Journal of Applied Probability, 31(3), 626-634.

Pang, C. Y. (2019). Lumpings of algebraic Markov chains arise from subquotients. Journal of Theoretical Probability, 32(4), 1804-1844.

Peng, N. F. (1996). On weak lumpability of a finite Markov chain. Statistics & probability letters, 27(4), 313-318.

Rogers, L. C., & Pitman, J. W. (1981). Markov functions. The Annals of Probability, 573-582.

Rubino, G., & Sericola, B. (1989). On Weak Lumpability in Markov Chains. Journal of Applied Probability, 26(3), 446–457.

Rubino, G., & Sericola, B. (1991). A finite characterization of weak lumpable Markov processes. Part I: The discrete time case. Stochastic Processes and their Applications, 38(2), 195-204.

Rubino, G., & Sericola, B. (1993). A finite characterization of weak lumpable Markov processes. Part II: The continuous time case. Stochastic Processes and their Applications, 45(1), 115-125.

Stewart, W.J. (1994). Introduction to the numerical solution of Markov Chains.

Tian, P., Kannan, D. (2006). Lumpability and Commutativity of Markov Processes, Stochastic Analysis and Applications, 24(3), 685-702.