* Invariant distributions

Let $X = (X_n)_{n \geq 0}$ be a (P, μ) Markov chain on S.

As for deterministic dynamics, we are interested in the existence (and uniqueness) of "equilibriums" or "stationary states", that possibly describe the long time behavior of X_n.

[These are in general non-deterministic but rather described by a probability measure $\mu \in \mathcal{P}(S)$ s.t. if $X_0 \sim \mu \Rightarrow X_n \sim \mu \quad \forall n$.]

Def: A distribution $\pi \in \mathcal{P}(S)$ is stationary (or invariant) for a MC with transition matrix P if

$$\boxed{\pi P = \pi}$$

, i.e. π is the left eigenvector of P with eigenvalue 1

Remark: In that case: $\pi = \pi P^n \quad \forall n \in \mathbb{N}$

$$\Rightarrow X_n \sim \pi \quad \forall n \in \mathbb{N}_0 \qquad \underbrace{\pi P^n}_{=\, \mathrm{Law}_\pi(X_n)}$$

▸ Existence of invariant distribution

We first consider the case S finite, and then move to S infinite (countable), where more tools are required.

Proposition. Let X be a MC on S, finite space. Then an invariant distribution exists.

Proof: Let $\mu \in \mathcal{P}(S)$, and $\forall n \in \mathbb{N}$, set $\mu_n := \mu P^n \in \mathcal{P}(S)$

We then define a sequence $(\bar{\mu}^{(N)})_{N \in \mathbb{N}}$, where $\bar{\mu}^{(N)} \in \mathcal{P}(S)$ s.t.

$$\bar{\mu}^{(N)}(x) = \frac{1}{N} \sum_{n=0}^{N-1} \mu_n(x) = \frac{1}{N} \sum_{n=0}^{N-1} \mu P^n(x) \qquad \forall x \in S$$

We then define a sequence $(\bar\mu^{(N)})_{N\in\mathbb{N}}$, with $\bar\mu^{(N)} \in \ldots$, by:

$$\bar\mu^{(N)}(x) = \frac{1}{N}\sum_{m=0}^{N-1} \mu_m(x) = \frac{1}{N}\sum_{m=0}^{N-1} \mu P^m(x) \quad , \forall x\in S$$

Since $\mathcal{P}(S) = \{\nu \in \mathbb{R}^S : \nu(x)\geq 0 \ , \ \sum_{x\in S}\nu(x) = 1\}$ is compact,

we can extract from $(\bar\mu^{(N)})_{N\in\mathbb{N}}$ a subsequence $(\bar\mu^{(N_k)})_{k\in\mathbb{N}}$

convergent in $\mathcal{P}(S)$, and let $\tilde\pi$ denote its limit.

We claim that $\tilde\pi$ is stationary for P. Indeed:

$$\bar\mu^{(N)}P = \frac{1}{N}\sum_{n=0}^{N-1}\mu P^{n+1} = \frac{1}{N}\sum_{m=0}^{N-1}\mu P^m + \frac{(\mu_N - \mu_0)}{N} = \bar\mu^{(N)} + \frac{(\mu_N-\mu_0)}{N}$$

limit along subsequence $(N_k)_{k\geq 0}$, $k\to\infty$

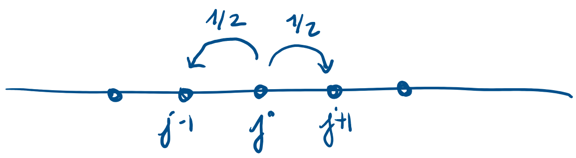
$$\tilde\pi P \qquad\qquad \tilde\pi \qquad 0$$

$$\boxed{\tilde\pi P = \tilde\pi}$$

#

- Counterexample (for S infinite)

Let $(X_n)_{n\geq 0}$ be a simple symmetric RW on $\mathbb{Z}$:



$$\left(P_{jk} = \frac{1}{2}\mathbb{1}_{\{j-1;\,j+1\}}(k)\right)$$

By definition, an invariant distribution π is s.t.

$$\pi = \pi P \iff \pi(j) = \tfrac{1}{2}\pi(j-1) + \tfrac{1}{2}\pi(j+1) \ , \ \forall j\in\mathbb{Z}$$

$$\iff \pi(j) = c \quad \forall j\in\mathbb{Z}$$

But then $\displaystyle\sum_{j\in\mathbb{Z}}\pi(j) = \sum_{j\in\mathbb{Z}} c \begin{cases} +\infty & (\text{if } c>0) \\ =0 & (\text{if } c=0)\end{cases} \implies \pi\notin\mathcal{P}(S)$

Note: $\pi = c$ is however an invariant measure for P

- Uniqueness of invariant distribution

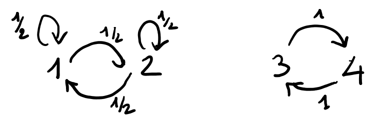
The problem of uniqueness is related to the possibility of the MC

The problem of uniqueness is related to the possibility of the MC of visiting any state, starting from every initial state.

<u>Examples</u>

$\boxed{1.}$ Let P a transition matrix on $S = \{1,2,3,4\}$

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & \\ 1/2 & 1/2 & & \\ & & 0 & 1 \\ 0 & & 1 & 0 \end{pmatrix} \xrightarrow{\text{graphical repr.}}$$



One can find algebraically that the eigenvalue 1 has multiplicity 2.
Intuitively, the MC can be splitted in two non-interacting MC's: one evolving in $\{1,2\} \to P^{I}$, with a unique invariant measure μ_1 (verify that $\mu_1 = \mathcal{U}\{1,2\}$), and the other evolving in $\{3,4\} \to P^{II}$ with unique invariant measure μ_2 (verify that $\mu_2 = \mathcal{U}\{3,4\}$).

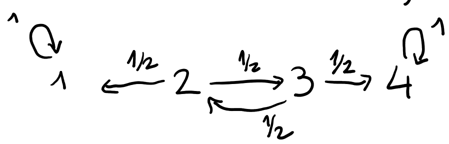
Then, any convex combination of μ_1 and μ_2 is also invariant:

$\alpha \in [0,1]$: $\mu = \alpha \mu_1 + (1-\alpha)\mu_2 \Rightarrow \mu P = \alpha \mu_1 P^{I} + (1-\alpha)\mu_2 P^{II} = \alpha \mu_1 + (1-\alpha)\mu_2$,

and we are far from uniqueness.

$\boxed{2.}$ Let P a transition matrix on $S = \{1,2,3,4\}$ s.t.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{graphical repr.}}$$



From definition, it is obvious that δ_1 and δ_4 are invariant measures, and hence every convex combination of them:

$\alpha \in [0,1]$: $\mu = \alpha \delta_1 + (1-\alpha)\delta_4$

In that case, the MC can not be splitted in disjoint subsets, but still we can infer by a Borel-Cantelli argument that the MC will be absorbed a.s. in $\{1,4\}$:

$$\mathbb{P}_x\left(\liminf_m \{X_m \in \{1,4\}\}\right) = 1 \quad \forall x \in S$$

$\mathbb{P}_x(\text{visiting}_m \{X_m \in \{1,4\}\}) = 1 \quad \forall x \in S$

Under hypotheses which garantee $\exists!$ of π, invariant distribution, we would like to show that $Law_\mu(X_n)$ "converges" to π, $\forall \mu \in \mathcal{P}(S)$

▶ Convergence to the stationary measure

Under hypotheses which garantee $\exists!$ of π, invariant distribution, we would like to show that $Law_\mu(X_n)$ "converges" to π, $\forall \mu \in \mathcal{P}(S)$

Def: Let X be a MC on S, with transition matrix P.

1. X is irreducible if $\forall x,y \in S$, $\exists\, n = n(x,y) \in \mathbb{N}$ s.t. $P^{(n)}_{x,y} > 0$ (all states communicate among each other)
2. X is strongly irreducible if $\exists\, m \in \mathbb{N}$ and $\varepsilon > 0$ s.t. $P^{(m)}_{x,y} \geq \varepsilon \quad \forall x,y \in S$

Remark

1. Irreducibility excludes the two counterexamples

Ex: Prove that if S is finite and X irreducible $\Rightarrow \exists!\, \pi$ inv. distr.

2. Strong irreducibility (which implies irreducibility) corresponds to the assumptions that P^m has all positive components, and exclude irreducible but periodic behavior as the following:

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \longrightarrow a \underset{1}{\overset{1}{\rightleftarrows}} b$$

This leads to the following definition.

Let X be an irreducible MC on S $\left[\forall x,y \in S, \exists\, n = n(x,y) : P^{(n)}_{xy} > 0\right]$

Def [period of a state]: The period of the MC on a state $x \in S$ is

$$d_x = \gcd\{n \geq 1 : P^{(n)}_{x,x} > 0\}$$

Properties:

Properties:

1. Under irreducibility, all the states $x \in S$ have the same period, that is then called period of the MC, and denoted by d.
   If $d=1 \Rightarrow X$ is called aperiodic MC.
2. If S finite: strong irred. $\Leftrightarrow$ irredu. + aperiodicity

To state the main result, we introduce a distance on $\mathcal{P}(S)$.

Def [Total variation distance]

later on, $\Omega = S^m$, for $m \in \mathbb{N}$

Let $\mu, \nu \in \mathcal{P}(\Omega)$, with $(\Omega, \mathcal{F})$ a countable measurable space.
The total variation distance of μ and ν is

$$\boxed{\|\mu - \nu\|_{TV} = \max_{A \in \mathcal{F}} |\mu(A) - \nu(A)|}$$

Remark: It is an exercise to show that, when $\mathcal{F} = \{A : A \subseteq \Omega\}$. (power set)

$$(1) \quad \|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|$$

(Hint: Consider $A = \{x \in \Omega : \mu(x) \geq \nu(x)\}$. Then ...)

Theorem [Convergence to stationarity - S finite]

Let $X = (X_n)_{n \in \mathbb{N}_0}$ be a strongly irreducible MC on S finite. Then

* it exists a unique invariant distribution $\pi \in \mathcal{P}(S)$
* π is s.t. $\pi(x) > 0 \;\; \forall x \in S$
* it exists $\gamma = \gamma(m, \varepsilon) \in (0,1)$ s.t., $\forall \mu \in \mathcal{P}(S)$:

related to condition of strong irreducibility

$$(2) \qquad \boxed{\|\mu P^m - \pi\|_{TV} \leq \gamma^m}$$

Remark:

Inequality (2.) hence implies an exponentially fast convergence of μP^n (law of $X_n \mid X_0 \sim \mu$) to π. In particular, from the def. of total variation distance given (1), it also implies

$$\forall y \in S: \quad \mu P^n(y) \xrightarrow[n\to\infty]{} \pi(y) \quad \text{exp. fast}$$

and in particular, for $\mu = \delta_x$

$$\forall x, y \in S: \quad \delta_x P^n(y) = P^{(n)}_{x,y} \xrightarrow[n\to\infty]{} \pi(y) \quad \text{exp. fast}$$

which best expresses the loss of memory from the initial state.

- Tools for the proof: Coupling of measures

Def: (Coupling) Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ countable measurable spaces and let $\mu \in \mathcal{P}(\Omega_1)$ and $\nu \in \mathcal{P}(\Omega_2)$.
A coupling of μ and ν is a probability measure $q \in \mathcal{P}(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ with marginals μ and ν:

$$q(A \times \Omega_2) = \mu(A), \quad q(\Omega_1 \times B) = \nu(B), \quad \forall A \in \mathcal{F}_1,\ B \in \mathcal{F}_2$$

Remark:
- The product measure $\mu \times \nu$ is an obvious coupling, but there are many others possibility.
- We will mainly deal with the situation: $X \sim \mu$ and $Y \sim \nu$, so that we speak equivalently of coupling among random variables. In fact, X and Y will be rather MC's (whole trajectories).

Proposition [Coupling and TV-distance]
Let $\mu, \nu \in \mathcal{P}(\Omega, \mathcal{F})$ and q a coupling of μ and ν.
Consider a r.v. $(X, Y) \sim q$ (e.g. using canonical representation) Then

Let $\mu, \nu \in \mathcal{P}(\Omega, \mathcal{F})$ and q a coupling of μ and ν.

Consider a r.v. $(X, Y) \sim q$ (e.g, using canonical representation). Then

i. $$\|\mu - \nu\|_{TV} \leq \mathbb{P}_q(X \neq Y)$$

perfect coupling

ii. Moreover, it exists a coupling q^* s.t $\|\mu - \nu\|_{TV} = \mathbb{P}_{q^*}(X \neq Y)$

Proof (only of i. - complete proof on "Markov Chains and Mixing time" by Levin, Peres)

$\forall$ coupling q and $\forall A \in \mathcal{F}$:

$$\mu(A) - \nu(A) = \mathbb{P}_q(X \in A) - \mathbb{P}_q(Y \in A) \leq \mathbb{P}_q(X \in A, Y \notin A) \leq \mathbb{P}_q(X \neq Y)$$

$$\implies \|\mu - \nu\|_{TV} = \min_{A \in \mathcal{F}} |\mu(A) - \nu(A)) \leq \mathbb{P}_q(X \neq Y) \qquad \#$$