# On Markov processes

Yukio NAGAHATA
Niigata univ.
nagahata@eng.niigata-u.ac.jp

2016 January 4-8

Applications of theorems.
Applying the Proposition (limit of transition matrix), which is implied by Perron-Frobenius theorem, we have followings; Suppose that $P$ is a irreducible transition matrix with period 1, then we have

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}$$

and $\mu > 0$ is a unique stationary distribution. Hence $\exists n_0$ s.t. $\forall n \geq n_0$, then $P^n_{x,x} \geq \mu_x/2 > 0$. This estimate implies

$$\sum_n P(X_n = x | X_0 = x) = \sum_n P^n_{x,x} = \infty$$

namely $\forall x$ is recurrent.

# Markov processes

Though an irreducible transition matrix $P$ has period $d \geq 2$, we can decompose $P$ into $d$ positive block matrices, namely $\exists n$ s.t.

$$P^n = \begin{pmatrix} P_1 & & & \\ & P_2 & & \\ & & \ddots & \\ & & & P_d \end{pmatrix}$$

where $P_i > 0$ for all $i$. Similarly, we have

$$\sum_m P(X_m = x | X_0 = x) = \sum_m P^m_{x,x} \geq \sum_k P^{kn}_{x,x} = \infty$$

and $\forall x$ is recurrent.

## Markov processes

### On recurrent and transient

If we consider recurrent or transient for some finite state Markov process, then we have no interesting result. If we consider some infinite state Markov process, then we have some interesting resuts. For example, if we consider mean 0 random walk on $\mathbb{Z}^d$, then we have it is recurrent if $d = 1, 2$ and transient $d \geq 3$.

# Markov processes

We are interested in a mean return time.
One of the interesting example is random knight problem a follows; Chess is one of a popular game. We play chess on $8 \times 8$ chessboard. Knight is a piece of a chess such that we can move it from black circle to white circle.

| | | | | |
|---|---|---|---|---|
| | ○ | | ○ | |
| ○ | | | | ○ |
| | | ● | | |
| ○ | | | | ○ |
| | ○ | | ○ | |

When a knight drank and move randomly, can we compute a mean return time $E[T_x|X_0 = x]$ ? Here the knight only move from black circle to white circle with equal probability.

# Markov processes

It seems hard to compute $E[T_x|X_0 = x]$ by means of $P(T_x = k|X_0 = x)$. But as a result, we know the results as

| | | | | |
|---|---|---|---|---|
| 168 | 112 | 84 | 84 | |
| 112 | 84 | 56 | 56 | |
| 84 | 56 | 42 | 42 | |
| 84 | 56 | 42 | 42 | |
| | | | | |

where we only write the mean return time only upper left part, due to symmetry.

Similarly

king

| | | |
|---|---|---|
| ○ | ○ | ○ |
| ○ | ● | ○ |
| ○ | ○ | ○ |

,

| | | | | |
|---|---|---|---|---|
| 140 | 84 | 84 | 84 | |
| 84 | 105/2 | 105/2 | 105/2 | |
| 84 | 105/2 | 105/2 | 105/2 | |
| 84 | 105/2 | 105/2 | 105/2 | |
| | | | | |

queen

| | | | | |
|---|---|---|---|---|
| ○ | | ○ | | ○ |
| | ○ | ○ | ○ | |
| ○ | ○ | ● | ○ | ○ |
| | ○ | ○ | ○ | |
| ○ | | ○ | | ○ |

,

| | | | | |
|---|---|---|---|---|
| 208/3 | 208/3 | 208/3 | 208/3 | |
| 208/3 | 1456/23 | 1456/23 | 1456/23 | |
| 208/3 | 1456/23 | 1456/25 | 1456/25 | |
| 208/3 | 1456/23 | 1456/25 | 1456/27 | |
| | | | | |

For (each) fixed $x$, we define

$$\mu_y = \mu_y^x = \sum_{n=0}^{\infty} P(X_n = y, T_x > n | X_0 = x)$$

Note that:
(1) If $x \not\to y$, then $\mu_y = 0$.
(2) If $X_n = x$, then $T_x \leq n$. This implies for $n \geq 1$, we have $P(X_n = x, T_x > n | X_0 = x) = 0$. Hence we have

$$\mu_x = P(X_0 = x, T_x > 0 | X_0 = x) = 1$$

(3) Suppose that $x$ is recurrent. If $x \to y$, then it is also $y \to x$. Hence $\mu$ is positive if $x \leftrightarrow y$.

# Markov processes

## Theorem (stationary measure)

Suppose that $x$ is recurrent. Then this $\mu$ is a stationary measure.

Proof

Suppose that $y \neq x$. By using Markov property, we have

$$\begin{aligned}
&P(X_{n+1} = z, X_n = y, T_x > n | X_0 = x) \\
&= P(X_{n+1} = z, X_n = y, X_{n-1} \neq x, \ldots, X_1 \neq x | X_0 = x) \\
&= P(X_{n+1} = z | X_n = y) P(X_n = y, X_{n-1} \neq x, \ldots, X_1 \neq x | X_0 = x) \\
&= P_{y,z} P(X_n = y, T_x > n | X_0 = x)
\end{aligned}$$

We also have

$$P(X_n = x, T_x > n | X_0 = x) = 0$$

Taking summation over all $y$, we have

$$\sum_{y \neq x} P(X_{n+1} = z, X_n = y, T_x > n | X_0 = x)$$
$$= \sum_{y} P_{y,z} P(X_n = y, T_x > n | X_0 = x)$$

Left hand side is equal to

$$P(X_{n+1} = z, T_x > n+1 | X_0 = x), \quad \text{if } z \neq x$$
$$P(T_x = n+1 | X_0 = x), \quad \text{if } z = x$$

# Markov processes

We note that $P(X_0 = z | X_0 = x) = 0$ if $z \neq x$. Taking summation over $n$, for both side. The sum of left hand side for $z \neq x$ is

$$\sum_{n=0}^{\infty} P(X_{n+1} = z, T_x > n+1 | X_0 = x)$$
$$= \sum_{n=0}^{\infty} P(X_n = z, T_x > n | X_0 = x) = \mu_z$$

and that for $z = x$ is

$$\sum_{n=0}^{\infty} P(T_x = n+1 | X_0 = x) = P(T_x < \infty | X_0 = x) = 1 = \mu_x$$

since $x$ is recurrent.

# Markov processes

By changing the order of the summation, the sum of right hand side (which does not depend on $z$) is

$$\sum_{n=0}^{\infty} \sum_{y} P_{y,z} P(X_n = y, T_x > n | X_0 = x)$$
$$= \sum_{y} \sum_{n=0}^{\infty} P(X_n = y, T_x > n | X_0 = x) P_{y,z} = \sum_{y} \mu_y P_{y,z}$$

Hence we conclude that $\mu = \mu P$. $\square$

# Markov processes

## Theorem (uniqueness)

If $P$ is an irreducible and recurrent, then stationary measure is unique up to multiple constant. Furthermore, stationary measure is positive for all $x$.

Note that in this theorem, we assumed the irreducible. Hence if there is $x$ such that $x$ is recurrent, then all $y$ is also recurrent.

# Markov processes

Note: We consider an one dimensional random walk. A stationary measure $\mu$ satisfies

$$\mu_x = q\mu_{x+1} + p\mu_{x-1}, \quad \text{for all } x$$

We can solve this system of linear equations, the basis of the solutions are given by

$$\mu^1_x = 1, \;\; \mu^2_x = \left(\frac{p}{q}\right)^x, \quad \text{if } p \neq q$$
$$\mu^1_x = 1, \;\; \mu^2_x = x, \quad \text{if } p = q$$

Combining these two basis, if $p \neq q$, we take $\nu = c_1\mu^1 + c_2\mu^2$ for $c_1, c_2 \geq 0$, $(c_1, c_2) \neq 0$, then they are stationary measures for the random walk. If $p = q$ then only $\nu = c_1\mu^1$ for $c_1 > 0$ are stationary measures. In fact, in this setting random walk is recurrent iff $p = q$.

# Markov processes

Proof

We fixed $x$ and made $\mu = \mu^x$. We have proved that this $\mu$ is stationary measure. Furthermore, by using irreducible, we have $\{n; P(X_n = y|X_0 = x) > 0\} \neq \emptyset$. Hence we set

$$n = n_y = \min\{n; P(X_n = y|X_0 = x) > 0\},$$

Then we have

$$0 < P(X_n = y|X_0 = x) = P(X_n = y, T_x > n|X_0 = x) \leq \mu_y$$

namely, this stationary measure is positive for all $y$.

# Markov processes

We show that if $\tilde{\nu}$ is a stationary measure, then we have $\exists a \in S$, $C > 0$ s.t. $\tilde{\nu} = C\mu^a$.
By using the argument of the proof of this claim, if we apply this claim for $\tilde{\nu} = \mu^y$, then we can take $a = x$. Hence we conclude that all stationary measure is the same up to multiple constant.
Since there is at least one $a \in S$ such that $\tilde{\nu}_a > 0$, we normalize $\tilde{\nu}$ as

$$\nu_x = \frac{\tilde{\nu}_x}{\tilde{\nu}_a}$$

It is obvious that this $\nu$ is also stationary measure. We show that $\nu = \mu^a$.

# Markov processes

Since $\nu$ is a stationary measure, it satisfies $\nu P = \nu$. We regard $a$ as a special point. Then we can rewrite

$$\begin{aligned}\nu_z &= \sum_y \nu_y P_{y,z} \\ &= \nu_a P_{a,z} + \sum_{y \neq a} \nu_y P_{y,z}\end{aligned}$$

Since $\nu_a = 1$, we have

$$\nu_z \geq P_{a,z}$$

Substituting this inequality into above identity, we have

$$\begin{aligned}\nu_z &\geq P_{a,z} + \sum_{y \neq a} P_{a,y} P_{y,z} \\ &= P_{a,z} + P(X_2 = z, X_1 \neq a | X_0 = a) \\ &= P_{a,z} + P(X_2 = z, T_a > 1 | X_0 = a)\end{aligned}$$

Similarly substituting this inequality into the above identity, we have

$$\begin{aligned}
\nu_z &\geq P_{a,z} + \sum_{y \neq a} P_{a,y} P_{y,z} + \sum_{y \neq a} P_{y,z} P(X_2 = y,\, T_a > 1 | X_0 = a) \\
&= P_{a,z} + P(X_2 = z, X_1 \neq a | X_0 = a) \\
&\qquad + P(X_3 = z, X_2 \neq a,\, T_a > 1 | X_0 = a) \\
&= P_{a,z} + P(X_2 = z,\, T_a > 1 | X_0 = a) \\
&\qquad + P(X_3 = z,\, T_a > 2 | X_0 = a)
\end{aligned}$$

Inductively, we conclude that

$$\nu_z \geq P_{a,z} + \sum_{n=2}^{N} P(X_n = z,\, T_a > n-1 | X_0 = a), \quad \forall N$$

If $z \neq a$, then we have

$$\nu_z \geq \sum_{n=0}^{N} P(X_n = z, T_a > n | X_0 = a) \to \mu_z^a, \quad (N \to \infty)$$

If $z = a$, then since $\nu_a = \mu_a^a = 1$. Hence we have

$$\nu_z \geq \mu_z^a, \quad \forall z$$

Since $\nu$ is a stationary measure, we have

$$1 = \nu_a = \sum_y \nu_y P_{y,a}^n \geq \sum_y \mu_y^a P_{y,a}^n = \mu_a^a = 1$$

for all $n \geq 0$.
This inequality implies that if $P_{y,a}^n > 0$ then $\nu_y = \mu_y^a$. Since we assumed that $P$ is irreducible, for any $y$ there is $n$ such that $P_{y,a}^n > 0$. Hence we conclude that $\nu = \mu^a$. $\square$

# Markov processes

**Theorem (stationary distribution and mean return time)**

Suppose that $P$ is an irreducible matrix with stationary distribution $\pi$. Then all $x$ is recurrent and it satisfies

$$\pi_x = \frac{1}{E[T_x|X_0 = x]} = \frac{1}{\sum_{n=0}^{\infty} P(T_x > n|X_0 = x)}$$

Note that if a stationary measure $\mu$ satisfiew $\sum_x \mu_x < \infty$. Then by normalizing $\mu$ as

$$\pi_x = \frac{\mu_x}{\sum_y \mu_y}$$

we have stationary distribution $\pi$.

Proof

The second identity is implied by following general result;

**Problem**

Show that if $Y$ is a nonnegative integer valued random variable, then we have

$$E[Y] = \sum_{n=0}^{\infty} P(Y > n).$$

(Here we allow $\infty = \infty$.)

On recurrent

Since $\pi$ is a stationary distribution, there is at least one $x \in S$ such that $\pi_x > 0$. Furthermore, we have

$$\pi_z = (\pi P^n)_z = \sum_y \pi_y P(X_n = z | X_0 = y) \geq \pi_x P(X_n = z | X_0 = x)$$

Since $P$ is irreducible, for each $z$, there is $n$ such that $P(X_n = z | X_0 = x) > 0$. Hence we conclude that $\pi > 0$.
Since $\pi$ is a stationary distribution, we have $\pi_z = \sum_y \pi_y P^n_{y,z}$.
Taking summation over $n$, we have

$$\infty = \sum_{n=0}^{\infty} \pi_z = \sum_{n=0}^{\infty} \sum_y \pi_y P^n_{y,z} = \sum_y \pi_y \sum_{n=0}^{\infty} P^n_{y,z}$$

# Markov processes

We decompose the probability $P^n_{y,z} = P(X_n = z|X_0 = y)$ into that for the first hitting time of the Markov process to $z$. Then we have

$$
\begin{aligned}
&P^n_{y,z} = P(X_n = z|X_0 = y) = \sum_{k=1}^{n} P(X_n = z, T_z = k|X_0 = y) \\
&= \sum_{k=1}^{n} P(X_n = z, X_k = z, T_z = k|X_0 = y) \\
&= \sum_{k=1}^{n} P(X_n = z|X_k = z)P(T_z = k|X_0 = y) \\
&= \sum_{k=1}^{n} P^{n-k}_{z,z} P(T_z = k|X_0 = y)
\end{aligned}
$$

# Markov processes

By changing the order of the summation, we have

$$\begin{aligned}\sum_{n=0}^{\infty} P_{y,z}^{n} &= \sum_{n=0}^{\infty}\sum_{k=1}^{n} P_{z,z}^{n-k} P(T_z = k | X_0 = y) \\ &= \sum_{l=0}^{\infty} P_{z,z}^{l} \sum_{k=1}^{\infty} P(T_z = k | X_0 = y) \\ &= \sum_{l=0}^{\infty} P_{z,z}^{l} P(T_z < \infty | X_0 = y) \le \sum_{l=0}^{\infty} P_{z,z}^{l}\end{aligned}$$

Since $\pi$ is a stationary distribution, we have $\sum_y \pi_y = 1$. Hence we have

$$\infty = \sum_y \pi_y \sum_{n=0}^{\infty} P_{y,z}^{n} \le \sum_y \pi_y \sum_{l=0}^{\infty} P_{z,z}^{l} \le \sum_{l=0}^{\infty} P_{z,z}^{l}$$

This inequality implies that $z$ is recurrent.

# Markov processes

Finally, we show that

$$\pi_x = \frac{1}{\sum_{n=0}^{\infty} P(T_x > n | X_0 = x)}$$

The Theorem (uniqueness) implies that $\mu^x$ defined for each $x$, which is

$$\mu_y^x = \sum_{n=0}^{\infty} P(X_n = y, T_x > n | X_0 = x),$$

is the same as $\pi$ up to multiple constant. Hence we normalize this and obtain

$$\pi = \frac{1}{\sum_y \mu_y^x} \mu^x, \quad \forall x.$$

Since we have$\mu_x^x = 1$, by changing the order of summation, we have

$$\begin{aligned}\sum_y \mu_y^x &= \sum_y \sum_{n=0}^{\infty} P(X_n = y, T_x > n | X_0 = x) \\ &= \sum_{n=0}^{\infty} \sum_y P(X_n = y, T_x > n | X_0 = x) \\ &= \sum_{n=0}^{\infty} P(T_x > n | X_0 = x)\end{aligned}$$

Hence we conclude that

$$\pi_x = \frac{1}{\sum_{n=0}^{\infty} P(T_x > n | X_0 = x)}$$

□

# Markov processes

This Theorem (stationary distribution and mean return time) implies that if we can compute the stationary distribution, then we can obtain the mean return time. Perron-Frobenius Theorem says that (if a stationary measure is finite then) if a transition matrix $P$ is irreducible, then there is an unique stationary distribution as a (normalized) nontrivial solution of $\mu P = \mu$.
Indeed, if we consider a random knight problem, then the corresponding matrix $P$ is $64 \times 64$ matrix (, which is a sparse matrix (including a lot of 0 elements)). If we solve $\mu P = \mu$, it seems hard to compute it directly.
But following results give us the stationary distribution of $P$.

# Markov processes

## Definition (random walk on a graph)

A transition matrix $P$ is called a random walk on a graph iff
$P_{x,x} = 0$
If $P_{x,y} > 0$, then $P_{y,x} > 0$
For $x, y, z \in S$, if $P_{x,y} > 0$ and $P_{x,z} > 0$ then $P_{x,y} = P_{x,z}$

Note: If a transition matrix is a random walk on graph, then we set this graph as follows;
Let $S$ be a state space. We regard $S$ as a set of vertices of the graph. We define $E$ a set of (undirected) edges as $E = \{(x, y); x, y \in S, \quad P_{x,y} > 0\}$. Then $G = (S, E)$ is a (an undirected) graph. The Markov process given by $P$ moves on this graph along the edges. If this process stands at $x$, then it chooses one of an edges $(x, y) \in E$ with the same probability.

# Markov processes

Given a random walk on a graph with transition matrix $P$, we define $\mu = (\mu_1, \dots, \mu_n)$ by $\mu_x = \#\{y; P_{x,y} > 0\}$.
Note that $\mu_x$ is given by the degree of the vertex $x$.

**Theorem (random walk on a graph)**

Suppose that $P$ is a random walk on a graph. Then $P$ is symmetric with respect to $\mu$. Namely $\mu$ is a stationary measure of $P$.

Note that by normalizing the stationary measure $\mu$, we obtain the stationary distribution as

$$\pi = \frac{1}{\sum_x \mu_x} \mu$$

The quantity $\sum_x \mu_x$ is twice as large as the total number of (undirected) edges.
Proof
It is easy to see that for any $x, y$ we have $\mu_x P_{x,y} = 1$. $\square$

# Markov processes

Transition matrices corresponding to random knight, random king and random queen are random walks on graph. Hence by counting the total number of edges, we get results.
In Japanese "Shogi" has pieces 金, 銀, whose transition matrices are not random walk on graph. But writing a computer program, we compute a stationary distribution. Result is as follows; (Due to symmetry and a lack of space, we only write the result of the left half part of $9 \times 9$ board (Shogi-board is $9 \times 9$))

# Markov processes

金

| | | |
|---|---|---|
| ○ | ○ | ○ |
| ○ | ● | ○ |
| | ○ | |

| | | | | |
|---|---|---|---|---|
| 30.791328 | 17.815554 | 17.111911 | 16.881724 | 16.8248 |
| 35.324925 | 24.915491 | 25.076442 | 25.071163 | 25.0627 |
| 85.940838 | 67.42103 | 71.810424 | 73.470653 | 73.8932 |
| 215.529198 | 179.831729 | 201.712943 | 212.137365 | 215.112 |
| 550.125457 | 476.83005 | 558.232434 | 603.699589 | 617.821 |
| 1419.270985 | 1261.152818 | 1527.8363 | 1696.096853 | 1752.161 |
| 3681.111795 | 3328.221102 | 4149.011752 | 4722.384145 | 4924.695 |
| 9499.054514 | 8719.120537 | 11269.92047 | 13163.67467 | 13854.67 |
| 29920.38808 | 28154.66878 | 38005.94435 | 45254.73054 | 47919.13 |

# Markov processes

銀

| | | |
|---|---|---|
| ○ | ○ | ○ |
| | ● | |
| ○ | | ○ |

| | | | | |
|---|---|---|---|---|
| 67.813349 | 45.504115 | 44.201301 | 45.805548 | 45.524978 |
| 49.629111 | 24.907000 | 27.663759 | 27.158334 | 27.633778 |
| 72.368415 | 43.904072 | 40.570712 | 42.020332 | 41.348125 |
| 121.449120 | 65.709962 | 65.736386 | 63.129401 | 64.097269 |
| 185.413674 | 106.341529 | 99.792808 | 99.383662 | 97.215698 |
| 296.189600 | 163.363762 | 157.864619 | 151.766057 | 152.140169 |
| 461.498716 | 258.665768 | 242.750465 | 236.878209 | 232.707452 |
| 742.552272 | 405.113035 | 376.585647 | 361.925390 | 359.450153 |
| 2025.565173 | 1020.418715 | 955.758462 | 919.546091 | 904.813474 |