
IE575 Stochastic Methods

Supplemental Note Set 4

Monday, December 9, 2002[1]

---

4 MARKOV CHAINS

4.1 Basic Results

A **Markov chain** is a sequence of random variables $X_0, X_1, \ldots$. It is an example of a **stochastic process.**

The range of the $\{X_n\}$ is called the **state space** and is generally denoted by $\{0, 1, 2, \ldots\}$.

If $\{X_n(\omega) = i\}$, the process is said to be in **state** i.

$$\begin{aligned} &P(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i) = \\ &P(X_{n+1} = j \mid X_n = i) \end{aligned}$$

If this is not dependent on n, we say that the Markov Chain is time-homogeneous (stationary) and we simply write

$$P(X_{n+1} = j \mid X_n = i) = P_{ij}$$

These are called the (one-step) **transition probabilities** of the process.

Definition 4.1. *The* **initial probabilities** *are*

$$\alpha_i = P(X_0 = i)$$

---

[1] Wayne F. Bialas, Department of Industrial Engineering, University at Buffalo, 342 Bell Hall, Buffalo, NY 14260-2050 USA; *E-mail:* bialas@buffalo.edu; *Web:* http://www.acsu.buffalo.edu/~bialas. Copyright © 2002 Wayne F. Bialas. All Rights Reserved. Duplication of this work is prohibited without written permission.

Note:

$$\begin{aligned} P_{ij} &\geq 0 \quad \forall i, j \geq 0 \\ \sum_{j=0}^{\infty} P_{ij} &= 1 \quad \text{for } i = 0, 1, \ldots \end{aligned}$$

Let $\mathbf{P}$ denote the matrix $[P_{ij}]$.

Example 4.1. The water in a reservoir can at one of two levels: low and high. The level is observed at the end of each day. If the water level is low one day, the probability that he will also be low the following day is 0.4.

If the level is high one day, the probability that it will be high the next day is 0.8. The state space is

$$\begin{aligned} 0 &= \text{low} \\ 1 &= \text{high} \end{aligned}$$

and the transition probabilities are

$$P_{00} = 0.4 \quad P_{01} = 0.6$$
$$P_{10} = 0.2 \quad P_{11} = 0.8$$

The transition matrix

$$\mathbf{P} = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix}$$

is then

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{bmatrix}$$

■

Some questions we might like to ask, include:

1. Given that the water level is high today. what is the probability it will be high one week (seven days) later?
2. Given that the water level is low right now, how many days will it continue to be low, before becoming high?

3. Suppose the water level becomes low. What is the expected length of time needed for it to reach a high level again?

Example 4.2. (Spitzer) Random walk (reflecting barrier)

$$\begin{aligned} P_{0,1} &= 1 \\ P_{i,i+1} &= p > 0 \\ P_{i,i-1} &= q = 1 - p \end{aligned}$$

for $i = 1, 2, \ldots$. The transition matrix is of the form

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & & & \\ q & 0 & p & & \cdots \\ & q & 0 & p & \\ & & \vdots & & \end{bmatrix}$$

■

Example 4.3. Random walk (adsorbing barrier)

$$\begin{aligned} P_{0,0} &= 1 \\ P_{i,i+1} &= p \\ P_{i,i-1} &= q = 1 - p \end{aligned}$$

for $i = 1, 2, \ldots$. The transition matrix is of the form

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & & & \\ q & 0 & p & & \cdots \\ & q & 0 & p & \\ & & \vdots & & \end{bmatrix}$$

■

Example 4.4. Bernoulli process

$$\begin{aligned} P_{i,i+1} &= p \\ P_{i,i} &= q = 1 - p \end{aligned}$$

for $i = 0, 1, \ldots$. The transition matrix is of the form

$$\mathbf{P} = \begin{bmatrix} q & p & & & \\ & q & p & & \cdots \\ & & q & p & \\ & & \vdots & & \end{bmatrix}$$

■

Example 4.5. Gambler's ruin

$$\begin{aligned} P_{i,i+1} &= p \\ P_{i,0} &= q = 1 - p \end{aligned}$$

for $i = 0, 1, \ldots$. The transition matrix is of the form

$$\mathbf{P} = \begin{bmatrix} q & p & & & \\ q & 0 & p & & \cdots \\ q & 0 & 0 & p & \\ & & \vdots & & \end{bmatrix}$$

■

Definition 4.2. *We can define the n-step transition probabilities*

$$P(X_{m+n} = j \mid X_m = i) = P_{ij}^{(n)}$$

Note that

$$P_{ij}^{(1)} = P_{ij}$$

4.1.1 Chapman-Kolmogorov Equations

$$\begin{aligned} P_{ij}^{(n+m)} &= P(X_{m+n} = j \mid X_0 = i) \\ &= \sum_{k=0}^{\infty} P(X_{m+n} = j, X_n = k \mid X_0 = i) \end{aligned}$$

$$\begin{aligned} &= \sum_{k=0}^{\infty} P(X_{m+n} = j \,|\, X_n = k, X_0 = i) \\ &\quad \cdot P(X_n = k \,|\, X_0 = i) \\ &= \sum_{k=0}^{\infty} P_{kj}^{(m)} P_{ik}^{(n)} \end{aligned}$$

And, using matrix notation, we can write

$$\begin{aligned} \mathbf{P}^{(n+m)} &= \mathbf{P}^{(n)}\mathbf{P}^{(m)} \\ \mathbf{P}^{(2)} &= \mathbf{P}^{(1)}\mathbf{P}^{(1)} = \mathbf{PP} = \mathbf{P}^2 \end{aligned}$$

and by induction

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} = \mathbf{P}^{(n-1)}\mathbf{P} = \mathbf{P}^n$$

Note: We often write

$$P_{ij}^{(n)} = P_{ij}^n$$

Example 4.6. Our reservoir

$$\begin{aligned} \mathbf{P}^2 &= \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{bmatrix} \\ &= \begin{bmatrix} 0.28 & 0.72 \\ 0.24 & 0.76 \end{bmatrix} \end{aligned}$$

■

4.1.2 Classification of States

Definition 4.3. *State j is* **accessible** *from state i ($i \to j$) if $P_{ij}^n > 0$ for some* $n \geq 0$.

Two states, i and j **communicate** *($i \leftrightarrow j$) if j is accessible from i, and i is accessible from j.* **Note:** $i \leftrightarrow i$ since

$$P_{ii}^0 = P(X_0 = i \,|\, X_0 = i) = 1$$

Properties

1. $i \leftrightarrow i$ (reflexive)
2. If $i \leftrightarrow j$ then $j \leftrightarrow i$ (symmetric)
3. If $i \leftrightarrow j$ and $j \leftrightarrow k$ then $i \leftrightarrow k$ (transitive)

   Proof: $\exists n, m$ such that $P_{ij}^n > 0$ and $P_{jk}^m > 0$. From the Chapman-Kolmogorov equations we have

$$\begin{aligned} P_{ik}^{n+m} &= \sum_{r=0}^{\infty} P_{ir}^n P_{rk}^m \\ &\geq P_{ij}^n P_{jk}^m > 0 \end{aligned}$$

■

Given the above, the relation $\leftrightarrow$ is an *equivalence relation.*

Definition 4.4. *Two states that communicate with each other are said to be in the same* **communicating class.**

A Markov chain is **irreducible** *if there is only one communicating class.*

Example 4.7.

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

had one communicating class $\{0, 1, 2\}$ and is irreducible.

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

had three communicating classes, $\{0, 1\}$, $\{2\}$ and $\{3\}$. ■

Definition 4.5. *State i is said to have* **period** *d if $P_{ii}^n = 0$ except when $n = d, 2d, 3d, \ldots$ and d is the greatest integer with this property. We denote the period of state i by $d(i)$. A state with period $d = 1$ is said to be* **aperiodic***.*

Theorem 4.1. *Periodicity is a class property. That is, if $i \leftrightarrow j$ then $d(i) = d(j)$.*

Proof: $\exists\, m, n$ such that $P_{ij}^{m} > 0$ and $P_{ji}^{n} > 0$ Suppose $P_{ii}^{s} > 0$. Then

$$P_{jj}^{n+s+m} \geq P_{ji}^{n} P_{ii}^{s} P_{ij}^{m} > 0$$

Also, $P_{ii}^{s} > 0$ implies

$$P_{ii}^{2s} \geq P_{ii}^{s} P_{ii}^{s} > 0$$

and, hence

$$P_{jj}^{n+2s+m} > 0$$

Therefore, $d(j)$ divides $n+s+m$ and $n+2s+m$. So, $d(j)$ divides

$$n + 2s + m - (n + s + m) = s$$

whenever $P_{ii}^{s} > 0$. Therefore, $d(j)$ divides $d(i)$.

Similarly, it can be shown that $d(i)$ divides $d(j)$. Hence $d(i) = d(j)$. ■

4.1.3 Recurrence

Note: The textbook calls this **persistence**.

Let

$$\begin{aligned} f_{ii}^{(n)} &\equiv P(\text{first return to state } i \text{ is at time } n) \\ &= P(X_1 \neq i, \ldots, X_{n-1} \neq i, X_n = i \mid X_0 = i) \end{aligned}$$

Note: For $n \geq 1$

$$P_{ii}^{n} = \sum_{k=1}^{n} f_{ii}^{(k)} P_{ii}^{n-k}$$

The system "restarts" by going back to i. **Note:**

$$P(\text{ever returning to } i \mid X_0 = i) = \sum_{n=1}^{\infty} f_{ii}^{(n)} \equiv f_{ii}^{\star}$$

Example 4.8. **Gambler's ruin.** Consider random walk on the integers $\{0, 1, \ldots, N\}$ where $\{0, N\}$ are adsorbing states. That is for $p + q = 1$

$$\begin{aligned} P_{i,i+1} &= p \quad \text{for } 0 < i < N \\ P_{i,i-1} &= q \quad \text{for } 0 < i < N \\ P_{0,0} &= 1 \\ P_{N,N} &= 1 \end{aligned}$$

We get

$$\begin{aligned} u_0 &\equiv f^\star_{0N} = 0 \\ u_i &\equiv f^\star_{iN} = p f^\star_{i+1,N} + q f^\star_{i-1,N} \quad i = 1, \ldots, N-1 \\ u_N &\equiv f^\star_{NN} = 1 \end{aligned}$$

This produces

$$u_{i+1} - u_i = \frac{q}{p}(u_i - u_{i-1}) \quad i = 1, \ldots, N-1$$

Therefore

$$\begin{aligned} u_2 - u_1 &= \frac{q}{p} u_1 \\ u_3 - u_2 &= \frac{q}{p}(u_2 - u_1) = \left(\frac{q}{p}\right)^2 u_1 \\ &\vdots \\ 1 - u_{N-1} &= \left(\frac{q}{p}\right)^{N-1} u_1 \end{aligned}$$

Hence

$$u_i - u_1 = u_1 \left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{i-1} \right]$$

for $i > 1$.

In other words,

$$u_i = \begin{cases} \dfrac{1 - (q/p)^i}{1 - q/p} u_1 & \text{if } q/p \neq 1 \\ i u_1 & \text{if } q/p = 1 \end{cases}$$

Also, we have

$$u_N = 1 = \begin{cases} \dfrac{1 - (q/p)^N}{1 - q/p} u_1 & \text{if } q/p \neq 1 \\ N u_1 & \text{if } q/p = 1 \end{cases}$$

hence

$$f^{\star}_{iN} = u_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)^N} & \text{if } p \neq 1/2 \\ i/N & \text{if } p = 1/2 \end{cases}$$

If the adversary has an infinite fortune, we can let $N \to \infty$ to get

$$f^{\star}_{i\infty} = \begin{cases} 1-(q/p)^i & \text{if } p > 1/2 \\ 0 & \text{if } p \leq 1/2 \end{cases}$$

■

Definition 4.6. *The* **mean recurrence time** *is given by*

$$\mu_i \equiv \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

Definition 4.7. *State i is said to be* **recurrent** *if $f^{\star}_{ii} = 1$ and* **transient** *if $f^{\star}_{ii} < 1$.*

Note:

$$P(X_n = j \text{ i.o.} \mid X_0 = i) = \begin{cases} 0 & \text{if } f_{jj} < 1 \\ & (j \text{ transient}) \\ f_{ij} & \text{if } f_{jj} = 1 \\ & (j \text{ recurrent}) \end{cases}$$

$$P(X_n = i \text{ i.o.} \mid X_0 = i) = \begin{cases} 0 & \text{if } f_{ii} < 1 \\ & (i \text{ transient}) \\ 1 & \text{if } f_{ii} = 1 \\ & (i \text{ recurrent}) \end{cases}$$

Compare this to Kolmogorov's zero-one law, but note that the events $\{X_n = i\}$ are not independent here.

Definition 4.8. *A recurrent state i is said to be* **null recurrent** *if $\mu_i = \infty$ and* **positive recurrent** *if $\mu_i < \infty$.*

Note: A recurrent state i is positive recurrent if, starting in state i, the expected time until the process returns to state i is finite.

Note: In a finite Markov chain all recurrent states are positive recurrent.

Theorem 4.2. *Transience and recurrence are both class properties*

Example 4.9. Pólya's Theorem (text page 118)

- **One dimensional** random walk on the integers. At each transition, the process moves right one unit with probability p and left one unit with probability $q = 1 - p$. State 0 (and all states) is recurrent, if and only if $p = \frac{1}{2}$.
- **Two dimensional** random walk with probabilities of right, left, up and down each equal to $\frac{1}{4}$ (symmetric random walk). State 0 (and all states) is recurrent.
- **Three dimensional** (symmetric) random walk with probability $\frac{1}{6}$ of going in each direction. State 0 (and all states) is transient.

4.2 Limiting Behavior of Markov Chains

Consider the transition matrix for a two-state Markov Chain. . .

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

Then

$$\begin{aligned} \mathbf{P}^2 &= \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix} \\ \mathbf{P}^4 &= \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix} \\ \mathbf{P}^8 &= \begin{bmatrix} 0.572 & 0.428 \\ 0.570 & 0.430 \end{bmatrix} \end{aligned}$$

Consider a two-state Markov chain,

$$\mathbf{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \quad \text{with} \begin{cases} 0 \le a \le 1 \\ 0 \le b \le 1 \\ |1-a-b| < 1 \end{cases}$$

Then

$$\mathbf{P}^n = \begin{bmatrix} \frac{b}{a+b} + \frac{a(1-a-b)^n}{a+b} & \frac{a}{a+b} - \frac{a(1-a-b)^n}{a+b} \\ \\ \frac{b}{a+b} - \frac{b(1-a-b)^n}{a+b} & \frac{a}{a+b} + \frac{b(1-a-b)^n}{a+b} \end{bmatrix}$$

Taking the limit on n, yields

$$\lim_{n\to\infty} \mathbf{P}^n = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix} \equiv \mathbf{P}^\infty$$

Note that P_{ij}^∞ is independent of i. So we define

$$\pi_j \equiv P_{ij}^\infty$$

4.2.1 Ergodicity

Definition 4.9. *If state i is positive recurrent and aperiodic, then state i is said to be* **ergodic**

Theorem 4.3. *For an irreducible, ergodic Markov chain, $\pi_j \equiv \lim_{n\to\infty} P_{ij}^n$ exists, and is independent of i. Furthermore, π_i is the unique, nonnegative solution to*

$$\begin{aligned} \pi_j &= \sum_{i=0}^{\infty} \pi_i P_{ij} \qquad \textit{for } j \geq 0 \\ 1 &= \sum_{i=0}^{\infty} \pi_i \end{aligned}$$

Proof: See textbook.

Example 4.10. Consider the Markov chain with transition matrix

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}$$

Solve, simultaneously, for $\pi_i \geq 0$ the following

$$\begin{aligned}
\pi_0 &= 0.5\pi_0 + 0.3\pi_1 + 0.2\pi_2 \\
\pi_1 &= 0.4\pi_0 + 0.4\pi_1 + 0.3\pi_2 \\
\pi_2 &= 0.1\pi_0 + 0.3\pi_1 + 0.5\pi_2 \\
1 &= \pi_0 + \pi_1 + \pi_2
\end{aligned}$$

Note that one of the equations is redundant. Remove any equation other than the last. This yields

$$\pi_0 = \tfrac{21}{62} \quad \pi_1 = \tfrac{23}{62} \quad \pi_2 = \tfrac{18}{62}$$

■

4.2.2 Computing Limiting Probabilities

Definition 4.10. *Let* **P** *be a Markov chain. Let* **p** $= (p_0, \ldots, p_n)$ *with* $\sum_i p_i = 1$ *and* $p_i \geq 0$ *for all* i. *Then* **p** *is said to be* **stationary** *with respect to* **P** *if*

$$\mathbf{pP} = \mathbf{p}$$

Note: This implies

$$\mathbf{pP}^n = \mathbf{p} \quad \forall n$$

Theorem 4.4. *Let* **P** *be a Markov chain with limiting probabilities* π_i. *Then* $\pi = (\pi_0, \ldots, \pi_n)$ *is stationary with respect to* **P**.

Proof: (For 2-state Markov chains) We see that

$$\begin{aligned}
\frac{b(1-a)}{a+b} + \frac{ab}{a+b} &= \frac{b}{a+b} \\
\frac{aa}{a+b} + \frac{a(1-b)}{a+b} &= \frac{a}{a+b}
\end{aligned}$$

■

Note: If we set the **initial probabilities**

$$\alpha_i = P(X_0 = i) = \pi_i$$

then the unconditional probabilities

$$P(X_n = i) = \pi_i$$

for all n since

$$\mathbf{p}\mathbf{P}^n = \mathbf{p} \quad \forall n$$

To compute the limiting probabilities, note that π must solve

$$\begin{aligned} \pi &= \pi\mathbf{P} \\ 1 &= \pi\mathbf{1} \end{aligned}$$

In other words,

$$\begin{aligned} \mathbf{0} &= \pi(\mathbf{P} - \mathbf{I}) \\ 1 &= \pi\mathbf{1} \end{aligned}$$

where **I** is the identity matrix. Note that one of the above equations is redundant.

4.2.3 Expected Number of Visits

Let N_{ij}^n denote the number of visits to state j in n steps, given that the Markov chain starts at state i. We would like to find

$$\mu_{ij}^{(n)} \equiv E(N_{ij}^n)$$

Define

$$Y_{ij}^{(k)} = \begin{cases} 1 & \text{if } X_k = j \text{ and } X_0 = i \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} P(Y_{ij}^{(k)} = 0) &= 1 - P_{ij}^k \\ P(Y_{ij}^{(k)} = 1) &= P_{ij}^k \end{aligned}$$

Therefore,

$$E(Y_{ij}^{(k)}) = P_{ij}^k \quad \forall k$$

But

$$\begin{aligned} N_{ij}^{(n)} &= \sum_{k=1}^{n} Y_{ij}^{(k)} \\ E(N_{ij}^{(n)}) &= \sum_{k=1}^{n} P_{ij}^{k} \end{aligned}$$

For the two-state Markov chain, these yield

$$\left[\begin{array}{cc} \frac{nb}{a+b} + \frac{a(1-a)[1-(1-a-b)^n]}{(a+b)^2} & \frac{na}{a+b} - \frac{a(1-a)[1-(1-a-b)^n]}{(a+b)^2} \\ \frac{nb}{a+b} - \frac{a(1-a-b)[1-(1-a-b)^n]}{(a+b)^2} & \frac{na}{a+b} - \frac{b(1-a-b)[1-(1-a-b)^n]}{(a+b)^2} \end{array} \right]$$

$$= \left[\begin{array}{cc} \mu_{00}^{(n)} & \mu_{01}^{(n)} \\ \mu_{10}^{(n)} & \mu_{11}^{(n)} \end{array} \right]$$

The fraction of time that you stay in state j given that you start in state i is

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n}\mu_{00}^{(n)} &= \lim_{n\to\infty} \frac{1}{n}\mu_{10}^{(n)} = \frac{b}{a+b} = \pi_0 \\ \lim_{n\to\infty} \frac{1}{n}\mu_{01}^{(n)} &= \lim_{n\to\infty} \frac{1}{n}\mu_{11}^{(n)} = \frac{a}{a+b} = \pi_1 \end{aligned}$$

Let A_i denote the number of time periods a Markov chain stays in state i before moving. Then

$$P(A_i = k) = (1 - P_{ii})(P_{ii})^k \quad \text{for } n = 0, 1, 2, \ldots$$

Note that this is a Bernoulli process, where

$$\begin{aligned} \text{stay} &= \text{failure} \\ \text{move} &= \text{success} \end{aligned}$$

4.3 Some Applications

4.3.1 Success Runs

$$\begin{aligned}
P_{0,0} &= 1 \\
P_{N,N} &= 1 \\
P_{i,0} &= p_i \\
P_{i,i} &= r_i \\
P_{i,i+1} &= q_i = 1 - r_i - p_i
\end{aligned}$$

for $i = 0, 1, \ldots$. The transition matrix is of the form

$$\mathbf{P} = \begin{bmatrix}
1 & & & & & & \\
p_1 & r_1 & q_1 & & & & \\
p_2 & & r_2 & q_2 & & \cdots & \\
p_3 & & & r_3 & q_3 & & \\
 & & \vdots & & & & \\
 & & & & & & 1
\end{bmatrix}$$

What is the probability that we reach state 0 before state N? Define

$$T = \min\{n \geq 0; X_n = 0 \text{ or } X_n = N\}$$

$$u_k = P(X_T = 0 \,|\, X_0 = k)$$

Using the Law of Total Probability

$$\begin{aligned}
u_k &= p_k + r_k P(X_T = 0 \,|\, X_1 = k) \\
&\quad + q_k P(X_T = 0 \,|\, X_1 = k+1) \\
u_k &= p_k + r_k u_k + q_k u_{k+1} \\
&\qquad \text{for } k = 1, \ldots, N-1 \\
u_0 &= 1 \\
u_N &= 0
\end{aligned}$$

It can be shown that these equations are solved by

$$\begin{aligned} u_k &= 1 - \left(\frac{q_k}{p_k + q_k}\right) \cdots \left(\frac{q_{N-1}}{p_{N-1} + q_{N-1}}\right) \\ &\text{for } k = 1, \ldots, N-1 \end{aligned}$$

4.4 Another Success Runs Example

$$\begin{aligned} P_{i,0} &= p_i \\ P_{i,i+1} &= 1 - p_i \end{aligned}$$

for $i = 0, 1, \ldots$. The transition matrix is of the form

$$\mathbf{P} = \begin{bmatrix} p_0 & 1-p_0 & & & \\ p_1 & & 1-p_1 & & \\ p_2 & & & 1-p_2 & \cdots \\ & & \vdots & & \end{bmatrix}$$

Is state 0 recurrent or transient?

$$P(\text{never return to } 0 \,|\, X = 0) = \prod_{i=0}^{\infty} (1 - p_i)$$

State 0 is transient if

$$\prod_{i=0}^{\infty} (1 - p_i) > 0$$

which occurs when

$$\sum_{n=0}^{\infty} p_n < \infty$$

Also,

$$f_{00}^{(n)} = \prod_{i=0}^{n-2} (1 - p_i) > p_{n-1}$$

Hence the probability of ever returning to state 0 is

$$\sum_{n=0}^{\infty} f_{00}^{(n)} = \sum_{n-1}^{\infty} \left(\prod_{i=0}^{n-2} (1 - p_i) > p_{n-1} \right)$$

4.4.1 Simplex Pivots (Ross)

Consider a linear programming program

$$\begin{aligned} \min \quad & cx \\ \text{st} \quad & \mathbf{A}x = b \\ & x \geq 0 \end{aligned}$$

where $\mathbf{A}$ is an $m \times n$ matrix. Consider a simple Markov chain model for the Simplex Algorithm to describe the movement from extreme point to extreme point. Suppose that the algorithm is at the j^{th} best extreme point. Then, after the next pivot, the next extreme point is equally likely to be any of the $j-1$ best extreme points. In other words,

$$\begin{aligned} P_{11} &= 1 \\ P_{ij} &= \frac{1}{i-1} \quad j = 1, \ldots, i-1; i > 0 \end{aligned}$$

Let

$$T_i = \min\{n \geq 0; X_n = 1 \mid X_0 = i\}$$

$$\begin{aligned} E(T_i) &= 1 + \sum_{j=1}^{i-1} E(T_j \mid X_1 = j) p_{ij} \\ &= 1 + \frac{1}{i-1} \sum_{j=1}^{i-1} E(T_j) \end{aligned}$$

Therefore,

$$\begin{aligned} E(T_1) &= 0 \\ E(T_2) &= 1 \\ E(T_3) &= 1 + \frac{1}{2} \\ E(T_4) &= 1 + \frac{1}{3}\left(1 + 1 + \frac{1}{2}\right) = 1 + \frac{1}{2} + \frac{1}{3} \end{aligned}$$

By induction, this yields

$$E(T_i) = \sum_{j=1}^{i-1} \frac{1}{j}$$

Furthermore, we note that

$$N = \binom{n}{m}$$

and

$$T_N = \sum_{j=1}^{N-1} I_j$$

where

$$I_j = \begin{cases} 1 & \text{if the process ever visits } j \\ 0 & \text{otherwise} \end{cases}$$

Fact 4.1. $I_1, \ldots, I_{N-1}$ *are independent and*

$$P(I_j = 1) = \frac{1}{j} \quad 1 \le j \le N-1$$

Proof: Left to reader. Given the above fact, we have

$$\begin{aligned} E(T_N) &= \sum_{j=1}^{N-1} \frac{1}{j} \\ \text{Var}(T_N) &= \sum_{j=1}^{N-1} \left(\frac{1}{j}\right)\left(1 - \frac{1}{j}\right) \end{aligned}$$

Note that

$$\int_1^N \frac{dx}{x} < \sum_{j=1}^{N-1} \frac{1}{j} < 1 + \int_1^{N-1} \frac{dx}{x}$$

$$\ln N < \sum_{j=1}^{N-1} \frac{1}{j} < 1 + \ln(N-1)$$

$$\ln N \approx \sum_{j=1}^{N-1} \frac{1}{j}$$

Hence $T_N \approx N(\ln N, \ln N)$ and for large n, m and $n-m$ (using Sterling's approximation)

$$\begin{aligned} N &= \binom{n}{m} \\ &\approx \frac{n^{n+1/2}}{(n-m)^{n-m+1/2} m^{m+1/2}\sqrt{2\pi}} \end{aligned}$$

Let $c = n/m$ and we get

$$\begin{aligned} \ln N &\approx \left(mc + \frac{1}{2}\right)\ln(mc) \\ &\quad - \left(m(c-1) + \frac{1}{2}\right)\ln(m(c-1)) \\ &\quad - \left(m + \frac{1}{2}\right)\ln m - \frac{1}{2}\ln(2\pi) \\ &\approx m\left[c\ln\left(\frac{c}{c-1}\right) + \ln(c-1)\right] \end{aligned}$$

Since

$$\lim_{c\to\infty} c\ln\left(\frac{c}{c-1}\right) = 1$$

then for large values of c, we get

$$\ln N \approx m(1 + \log(c-1))$$

Note: For example, suppose $n = 8000$ and $m = 1000$. Using the Central Limit Theorem, the number of simplex iterations required to solve the problem has a probability distribution that is approximately Gaussian (normal) with mean $\mu = (1000(1 + \ln 7)$ and variance $\sigma^2 = 1000(1 + \ln 7)$.[1] Since the probability that a standard Gaussian distribution ($\mu = 0$, $\sigma^2 = 1$) exceeds 1.96 is 0.25, we have

$$\begin{aligned} 2945 &\pm 1.96\sqrt{2945} \\ 2945 &\pm 106 \end{aligned}$$

iterations would be required with probability 0.95.

What is truly remarkable, is the narrow width of this 95% prediction interval.

---

[1] That is, letting $a = m\left(1 + \ln((n/m) - 1)\right)$ we get $T_N \approx N(a, a)$

4.5 BIBLIOGRAPHY

[1] Karlin, S. and H. M. Taylor, *A first course in stochastic processes*, Academic Press, New York, 1975.