Section 4: Steady-State Theory

Contents

4.1 The Concept of Stochastic Equilibrium

In the setting of a deterministic dynamical system $(x_n : n \geq 0)$ governed by a recursion

$$x_{n+1} = f(x_n),$$

a "stable" dynamical system ought typically to converge to an equilibrium x_∞, so that

$$x_n \to x_\infty$$

as $n \to \infty$. Provided that f is continuous, the equilibrium x_∞ will then necessarily satisfy the deterministic fixed point equation

$$x_\infty = f(x_\infty). \tag{4.1.1}$$

Note that if $x_0 = x_\infty$, then

$$x_n = x_\infty \tag{4.1.2}$$

for $n \geq 0$, so that the system is in equilibrium or steady-state when started in x_∞.

For a stochastic system $(X_n : n \geq 0)$, it is rarely the case that

$$X_n \to X_\infty \quad \text{a.s.} \tag{4.1.3}$$

as $n \to \infty$. This is too strong a notion of convergence to steady-state. In particular, note that if $(X_n : n \geq 0)$ satisfies the stochastic recursion

$$X_{n+1} = f(X_n, Z_{n+1}),$$

new randomness (as determined by Z_{n+1}) is being "injected" into the system, no matter how large the size of n. As a consequence, one can not usually expect that $f(X_n, Z_{n+1}) - X_n \to 0$ a.s., as would be required by (4.1.3).

Rather, in the stochastic context, one might instead hope that

$$X_n \Rightarrow X_\infty$$

as $n \to \infty$, so that we are demanding only that the distribution of X_n converges to an equilibrium distribution as $n \to \infty$. The analog to (4.1.1) is then a "stochastic fixed point equation", namely the equilibrium rv X_∞ must satisfy

$$X_\infty \stackrel{\mathcal{D}}{=} f(X_\infty, Z_\infty),$$

where $\stackrel{\mathcal{D}}{=}$ means "equality in distribution". In particular, when the stochastic recursion is such that Z_{n+1} is independent of X_n and identically distributed (as for Markov chain), the distribution π of X_∞ must satisfy

$$\pi(\cdot) = \mathrm{P}(X_\infty \in \cdot) = \int_S \mathrm{P}(X_\infty \in \mathrm{d}x)\mathrm{P}_x(X_1 \in \cdot) = \int_S \pi(x)\mathrm{P}_x(X_1 \in \cdot) \tag{4.1.4}$$

(since Z_∞ is then independent of X_∞). Equation (4.1.4) is the equation that characterizes *equilibrium distribution* (or *stead-state distribution*) of a Markov chain. Note that when S is discrete, (4.1.4) asserts that any equilibrium distribution $\pi = (\pi(x) : x \in S)$ should satisfy

$$\pi(y) = \sum_x \pi(x)p(x,y)$$

for $y \in S$, or, equivalently, the linear system

$$\pi P = \pi. \tag{4.1.5}$$

It should further be noted that if $X_0 \stackrel{\mathcal{D}}{=} X_\infty$ (so that X_0 is initialized with distribution π), then

$$X_n \stackrel{\mathcal{D}}{=} X_0 \tag{4.1.6}$$

for $n \geq 0$, in direct analog with (4.1.2).

A good deal of the theory of Markov chains is concerned with the question of existence and uniqueness of equilibrium distribution (or, in discrete state space, the question of existence and uniqueness of probability solutions of the linear system (4.1.5)). A closely related set of mathematical questions deals with the connection between the equilibrium distribution of a Markov chain and its long-run behavior when the chain is initialized from a non-equilibrium state. The typical form of such long-run (asymptotic) behavior is either a law of large numbers

$$\frac{1}{n}\sum_{j=0}^{n-1} \mathrm{I}(X_j = y) \to \pi(y) \quad \text{a.s.}$$

as $n \to \infty$ or a "pointwise limit theorem" such as

$$\mathrm{P}_x(X_n = y) \to \pi(y)$$

as $n \to \infty$.

Remark 4.1.1 The relation (4.1.6) is a manifestation of the fact that when the Markov chain is initialized with distribution π, $X = (X_n : n \geq 0)$ is a *stationary process*, so that

$$(X_{n+m} : n \geq 0) \stackrel{\mathcal{D}}{=} (X_n : n \geq 0)$$

for $m \geq 0$. For this reason, π is often called a *stationary distribution* of X.

4.2 Existence of Stationary Distributions for Finite-State Markov Chains

In this section, we use analytic (non-probabilistic) methods to show that every finite-state Markov chain has an equilibrium distribution. We first use a well-known "fixed point" theorem to assert existence of such equilibrium distributions.

Brower Fixed Point Theorem: Let $T : C \to C$ be a continuous mapping defined on a closed convex subset of $\mathbb{R}^d$. Then, there exists a fixed point z^* such that $z^* = T(z^*)$.

To apply this to finite state chains, let $\mathcal{P}$ be the set of all stochastic vectors $\mu = (\mu(x) : x \in S)$. Note that $\mathcal{P}$ can be viewed as a subset of $\mathbb{R}^d_+$, where $d = |S|$. Furthermore, $\mathcal{P}$ is closed and convex. Now, define $T(\mu) = \mu P$, where $P = (P(x,y) : x, y \in S)$ is the (one-step) transition matrix of the Markov chain. Since T maps $\mathcal{P}$ into $\mathcal{P}$ and is continuous, Brower's Theorem guarantees existence of $\pi \in \mathcal{P}$ such that $\pi = \pi P$.

But we can go further than this.

Theorem 4.2.1 Let $P = (P(x,y) : x, y \in S)$ be a stochastic matrix for which $|S| < \infty$. Then, there exists a matrix Λ such that

$$P\Lambda = \Lambda P = \Lambda^2 = \Lambda. \tag{4.2.1}$$

Furthermore,

$$\frac{1}{n} \sum_{j=0}^{n-1} P^j \to \Lambda \tag{4.2.2}$$

as $n \to \infty$.

Proof: We first show that there exists a subsequence $(n_k : k \geq 1)$ for which

$$\frac{1}{n_k} \sum_{j=0}^{n_k - 1} P^j \tag{4.2.3}$$

converges to a limit; we will call the limit matrix Λ.

Note that the sequence

$$\overline{P}_n \triangleq \frac{1}{n} \sum_{j=0}^{n-1} P^j$$

can be identified with a sequence in $\mathbb{R}^{d^2}$. Because $\overline{P}_n$ lies in the unit hypercube in $\mathbb{R}^{d^2}$ and the unit hypercube is compact, it follows that there exists a subsequence $(n_k : k \geq 1)$ and a limit Λ such that

$$\overline{P}_{n_k} \to \Lambda \tag{4.2.4}$$

as $n \to \infty$. Since $\overline{P}_n$ is stochastic for each $n \geq 1$, it follows that the limit Λ is necessarily stochastic. Furthermore,

$$\overline{P}_{n_k+1} = \frac{1}{n_k+1} \left(I + \sum_{j=1}^{n_k} P^j \right) = \frac{I}{n_k+1} + P \cdot \overline{P}_{n_k} \left(\frac{n_k}{n_k+1} \right) \to 0 + P\Lambda = P\Lambda.$$

On the other hand,

$$\overline{P}_{n_k+1} = \left(\frac{n_k}{n_k+1}\right)\overline{P}_{n_k} + \frac{1}{n_k+1}P^{n_k} \to \Lambda + 0 = \Lambda,$$

so that we conclude that

$$P\Lambda = \Lambda.$$

An essentially identical argument proves that $\Lambda P = \Lambda$ and $\Lambda^2 = \Lambda$.

It remains to establish (4.2.2), we need to show that for each $\epsilon > 0$, there exists $N = N(\epsilon)$ such that for $n \geq N$,

$$|||\overline{P}_n - \Lambda|||_e < \epsilon. \tag{4.2.5}$$

Choose n_k so that

$$|||\overline{P}_{n_k} - \Lambda|||_e < \frac{\epsilon}{2};$$

this can be done on account of (4.2.4). Then, for $m \geq 1$,

$$\overline{P}_{mn_k} - \Lambda = \frac{1}{mn_k}\sum_{j=0}^{mn_k-1} P^j - \Lambda = \frac{1}{n_k}\sum_{j=0}^{n_k-1} P^j \left(\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}\right) - \Lambda \tag{4.2.6}$$

But (4.2.1) ensures that $P^j\Lambda = \Lambda Pj = \Lambda$ for $j \geq 1$, so

$$\Lambda\left(\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}\right) = \left(\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}\right)\Lambda = \Lambda,$$

so that we can re-write (4.2.6) as

$$\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} P^j - \Lambda\right)\left(\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}\right).$$

If A is stochastic, it is easily verified that $|||A|||_e = 1$. Since

$$\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}$$

is stochastic,

$$\begin{aligned} &\left|\left|\left|\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} P^j - \Lambda\right)\left(\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}\right)\right|\right|\right|_e \\ &\leq \left|\left|\left|\frac{1}{n_k}\sum_{j=0}^{n_k-1} P^j - \Lambda\right|\right|\right|_e \cdot \left|\left|\left|\frac{I + P^{n_k} + \cdots + P^{(m-1)n_k}}{m}\right|\right|\right|_e < \frac{\epsilon}{2}. \end{aligned}$$

It follows that if we choose $N = \lceil 2n_k/\epsilon \rceil$, we have verified (4.2.5) for $n \geq N$, a multiple of n_k. To deal with non-integer multiples of n_k, say n of the form $n = mn_k + l \geq N$ (with $0 \leq l < n_k$, note that

$$\begin{aligned}
\left|\left|\left|\frac{1}{n}\sum_{j=0}^{n-1}P^j-\Lambda\right|\right|\right|_e &= \left|\left|\left|\left(\frac{mn_k}{n}\right)(\overline{P}_{mn_k}-\Lambda)+\frac{1}{n}\sum_{j=mn_k}^{n-1}(P^j-\Lambda)\right|\right|\right|_e \\
&\le \frac{mn_k}{mn_k+l}\,|||\,\overline{P}_{mn_k}-\Lambda\,|||_e+\frac{1}{n}\left|\left|\left|(P^{mn_k}-\Lambda)\sum_{j=0}^{l-1}P^j\right|\right|\right|_e \\
&\le \frac{mn_k}{mn_k+l}\,|||\,\overline{P}_{mn_k}-\Lambda\,|||_e+\frac{1}{n}\,|||P^{mn_k}-\Lambda|||_e\cdot\left|\left|\left|\sum_{j=0}^{l-1}P^j\right|\right|\right|_e \\
&\le \frac{\epsilon}{2}\cdot\frac{mn_k}{mn_k+l}+\frac{l}{n}\le\frac{\epsilon}{2}+\frac{1}{m}\le\epsilon,
\end{aligned}$$

proving the theorem. □

Remark 4.2.1 Every row of Λ is a stationary distribution of P.

4.3 Definitions and Simple Consequences

The definitions of *irreducibility*, *transient state*, and *recurrent state* are referred to the text and *Applied Probability and Queues* by S. Asmussen (2003).

Proposition 4.3.1 Suppose $X=(X_n:n\ge0)$ is an irreducible Markov chain. Let $\tau(x)=\inf\{n\ge 1:X_n=x\}$ for each $x\in S$. Then,

- If one state x is recurrent, then all states are recurrent.
- If one state x is transient, then all states are transient.
- If X is recurrent, then $P_x(\tau(y)<\infty)=1$ for all $x,y\in S$.

4.4 A Test for Recurrence

Theorem 4.4.1 Suppose $X=(X_n:n\ge0)$ irreducible. Then:

i.) X is transient if and only if there exists $x,y\in S$ such that

$$\sum_{n=0}^{\infty}P^n(x,y)<\infty.$$

ii.) X is recurrent if and only if there exists $x,y\in S$ such that

$$\sum_{n=0}^{\infty}P^n(x,y)=\infty.$$

Example 4.4.1

- Simple symmetric random walk is recurrent in $d=1,2$.
- Simple symmetric random walk is transient in $d\ge3$.

Remark 4.4.1

- A random walk $(S_n : n \geq 0)$ in $d = 1$ with $S_n = S_0 + Z_1 + \cdots + Z_n$ with the Z_i's iid and with $\mathrm{E}Z_1 = 0$ is recurrent.
- A random walk $(S_n : n \geq 0)$ in $d = 1$ with $S_n = S_0 + Z_1 + \cdots + Z_n$ with the Z_i's iid and with
$$P(Z_1 \in dx) = \frac{dx}{\pi(1+x^2)}$$
(i.e. with a Cauchy distribution) is recurrent.
- A random walk $(S_n : n \geq 0)$ in $d = 2$ with $S_n = S_0 + Z_1 + \cdots + Z_n$ with $\mathrm{E}\|Z_1\|^2 < \infty$ and with $\mathrm{E}Z_1 = 0$ is recurrent.
- A genuinely d-dimensional random walk in $\mathbb{R}^d$ is transient in $d \geq 3$.

  See R. Durrett, *Probability: Theory and Examples* (1991), p.159-170, for details.

4.5 Proving Positive Recurrence

Proposition 4.5.1 Suppose X is irreducible. Then, X is positive recurrent if and only if there exists a probability solution π to
$$\pi = \pi P. \tag{4.5.1}$$

Remark 4.5.1 If one can solve (4.5.1) explicitly, one has established positive recurrence. If one can not solve (4.5.1) explicitly, we need to look for alternatives; see below.

To prove positive recurrence, we need to prove that there exists $z \in S$ such that $\mathrm{E}_z\tau(z) < \infty$. Note that
$$\mathrm{E}_z\tau(z) = 1 + \sum_{y \neq z} P(z, y)\mathrm{E}_y\tau(z).$$
Put $C^c = \{z\}$ and $C = S - C^c$. Note that for $y \in C$,
$$\mathrm{E}_y\tau(z) \leq \mathrm{E}_y \sum_{j=0}^{T-1} w(X_j),$$
where $T = \inf\{n \geq 0 : X_n \in C^c\}$ and $w = (w(x) : x \in C)$ is a weight function as defined earlier in the quarter.

Suppose that there exists $c < 1$ such that
$$\mathrm{E}_x w(X_1) \leq c w(x)$$
for $x \in C$. Then, it follows that
$$\mathrm{E}_x w(X_1) I(X_1 \in C) \leq c w(x)$$
so
$$|||B|||_w \leq c,$$
where $B = (B(x, y) : x, y \in C)$ has entries defined by $B(x, y) = P(x, y)$. Set $f(x) = w(x)$ for $x \in C$ and note that $\|f\|_w = 1$. So,
$$\mathrm{E}_y \sum_{j=0}^{T-1} w(X_j) = \mathrm{E}_y \sum_{j=0}^{T-1} f(X_j) \leq (1 - |||B|||_w)^{-1}\|f\|_w \cdot w(x) \leq (1-c)^{-1} w(x).$$

Proposition 4.5.2 If there exists $w : S \to [1, \infty)$ such that

$$\mathrm{E}_x w(X_1) \le c w(x) \tag{4.5.2}$$

for $x \ne z$ and $c < 1$, then

$$\mathrm{E}_x \sum_{j=0}^{\tau(z)-1} w(X_j) \le (1-c)^{-1} w(x)$$

for $x \ne z$.

Example 4.5.1 Consider the embedded DTMC $(X_n : n \ge 0)$ (with $X_n = X(D_n+)$ for $n \ge 0$) for the M/G/1 queue with $\mathrm{E}\exp(\theta V_1) < \infty$ for θ in a neighborhood of the origin and with $\lambda \mathrm{E} V_1 < \infty$. Set $w(x) = \exp(\theta x)$ for $\theta > 0$ and sufficiently small. Then, there exists $c < 1$ and $\theta > 0$ small so that

$$\mathrm{E}_x e^{\theta X_1} \le c e^{\theta x}$$

for $x \ge 1$. Furthermore,

$$\sum_{y=1}^{\infty} P(0, y) e^{\theta y} < \infty.$$

Hence, $\mathrm{E}_0 \tau(0) < \infty$, so $(X_n : n \ge 0)$ is positive recurrent. Therefore, there exists a probability solution π to the linear system

$$\pi = \pi P.$$

Exercise 4.5.1 For some problems, it is hard to prove (4.5.2) for $x \ne z$, but (much) easier to prove (4.5.2) for $x \in K^c$, where K is a finite set containing z. To deal with such situations, suppose now that X is irreducible and let $T = \inf\{n \ge 1 : X_n \in K\}$.

a.) Prove that if (4.5.2) holds for $x \in K^c$, then $P_x(T < \infty) = 1$ for $x \in K^c$.

b.) (continuation of a.)) Prove that for $x \in K$,

$$\mathrm{E}_x \tau(z) = \mathrm{E}_x T + \sum_{\substack{y \in K \\ y \ne z}} P_x(X_T = y) \mathrm{E}_y \tau(z).$$

c.) (continuation of b.)) Prove that for $x \in K$,

$$\mathrm{E}_x T \le 1 + (1-c)^{-1} \sum_{y \in K^c} P(x, y) w(y).$$

d.) (continuation of c.)) Prove that if (4.5.2) holds and

$$\max_{x \in K} \mathrm{E}_x w(X_1) < \infty, \tag{4.5.3}$$

then $\mathrm{E}_z \tau(z) < \infty$.

e.) (continuation of d.)) Use a similar argument to that used above in a.)-d.) to prove that if $|f| \le w$, then

$$\mathrm{E}_x \sum_{j=0}^{\tau(z)-1} |f(X_j)| < \infty$$

if (4.5.2) and (4.5.3) hold, so that

$$\mathrm{E}_\pi |f(X_0)| = \sum_x \pi(x) |f(x)| < \infty.$$

Example 4.5.1 (continued) Under the stated conditions on V_1,

$$\mathrm{E}_0 \sum_{j=0}^{\tau(0)-1} e^{\theta X_j} < \infty$$

so

$$\widetilde{\pi}(\theta) = \sum_{x=0}^{\infty} \pi(x) e^{\theta x} < \infty$$

for θ in a neighborhood of the origin. The transition matrix P for $(X_n : n \geq 0)$ is

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & p_0 & p_1 & p_2 & \cdots \\ 0 & 0 & p_0 & p_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where

$$p_i = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^i}{i!} P(V_1 \in dt).$$

Then the steady-state equations can be written as

$$\pi(i) = \pi(0) p_i + \sum_{r=1}^{i+1} \pi(r) p_{i-r+1}, \quad i = 0, 1, 2, \ldots$$

We solve for the moment generating function $\widetilde{\pi}(\theta)$ in terms of

$$K(\theta) \triangleq \sum_{k=0}^{\infty} e^{\theta k} p_k.$$

Multiplying through by $e^{\theta i}$, we get

$$e^{\theta i} \pi(i) = \pi(0) p_i e^{\theta i} + e^{-\theta} \sum_{r=0}^{i+1} \pi(r) p_{i-r+1} e^{\theta(i+1)} - e^{-\theta} \pi(0) p_{i+1} e^{\theta(i+1)}, \quad i = 0, 1, 2, \ldots$$

Summing over i, and recognizing $\sum_{r=0}^{i+1} \pi(r) p_{i-r+1}$ as a convolution, we have

$$\sum_{i=0}^{\infty} e^{\theta i} \pi(i) = \widetilde{\pi}(\theta) = \pi(0) K(\theta) + e^{-\theta} [K(\theta) \widetilde{\pi}(\theta) - \pi(0) p_0] - e^{-\theta} \pi(0) [K(\theta) - p_0].$$

Solving for $\widetilde{\pi}(\theta)$ gives

$$\widetilde{\pi}(\theta) = \frac{\pi(0) K(\theta)(e^\theta - 1)}{e^\theta - K(\theta)}.$$

Note that

$$K(\theta) = \int_0^\infty e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} P(V_1 \in dt) = \int_0^\infty \exp\Big(\lambda t (e^\theta - 1)\Big) P(V_1 \in dt) = \varphi_V(\lambda(e^\theta - 1)),$$

where $\varphi_V(\theta) = \mathrm{E}\exp(\theta V_1)$. So,

$$\widetilde{\pi}(\theta) = \frac{\pi(0)\varphi_V(\lambda(e^\theta - 1))(e^\theta - 1)}{e^\theta - \varphi_V(\lambda(e^\theta - 1))}, \tag{4.5.4}$$

To identify $\pi(0)$, let $\theta \downarrow 0$ and use l'Hopital's rule on the right-hand side of (9.5.4), yielding

$$\pi(0) = 1 - \lambda \mathrm{E}V_1.$$

Hence, the moment-generating function of the stationary distribution for the M/G/1 queue is given by

$$\widetilde{\pi}(\theta) = \frac{(1-\rho)\varphi_V(\lambda(e^\theta - 1))(e^\theta - 1)}{e^\theta - \varphi_V(\lambda(e^\theta - 1))},$$

where $\rho \triangleq \lambda \mathrm{E}V_1$. Since $\widetilde{\pi}(\cdot)$ converges in a neighborhood of the origin, the Dominated Convergence Theorem guarantees that

$$\widetilde{\pi}'(0) = \sum_{x=0}^{\infty} x\pi(x) (\triangleq L),$$

yielding

$$L = \rho + \frac{\rho^2(c^2+1)}{2(1-\rho)}, \tag{4.5.5}$$

where c^2 is the *squared coefficient of the variation* of V_1 given by $c^2 = \mathrm{var}V_1/(\mathrm{E}V_1^2)$. Equation (4.5.5) is the celebrated Pollaczek-Khintchine formula for the mean number-in-system for the M/G/1 queue. It demonstrates the cost paid for variability in the service times.

Exercise 4.5.2 Compute the second moment $\sum_{x=0}^{\infty} x^2\pi(x)$ for Example 4.5.1.

Exercise 4.5.3 Suppose that $X = (X(t) : t \geq 0)$ is the number-in-system process for a G/M/1 queue (i.e. a G/G/1 queue with exponential service times having rate parameter μ). Suppose that $\mathrm{E}\chi_1 > 1/\mu$, where χ_1 is the first inter-arrival time.

a.) Argue that if $X_n = X(A_n-)$ (where A_n is the arrival time of customer n), then $(X_n : n \geq 0)$ is a Markov chain on state space $S = \mathbb{Z}_+$.

b.) Compute the transition matrix of P.

c.) Compute the stationary distribution of $(X_n : n \geq 0)$. (Hint: Try a geometric distribution.)

4.6 Convergence of the n-step Transition Probabilities

Definition 4.6.1 Let $X = (X_n : n \geq 0)$ be an irreducible Markov chain with transition matrix P. The state $x \in S$ is said to be *periodic of period p* if

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

(where gcd = "greatest common divisor"). If the period is 1, the state x is said to be *aperiodic.*

Proposition 4.6.1 Suppose X is irreducible. Then $x \in S$ is periodic of period p if and only if y is periodic of period p for all $y \in S$.

An irreducible Markov chain with transition matrix P is periodic of period p if and only if S can be partitioned into $P_1, P_2, \ldots, P_p$ for which P has the block structure

$$P = \begin{pmatrix} 0 & P_{12} & 0 & \cdots & 0 \\ 0 & & P_{23} & & \\ & & & \ddots & \\ & & & & P_{p-1,p} \\ P_{p,1} & 0 & & & 0 \end{pmatrix}$$

Note that if $P(x,x) > 0$ for some $x \in S$, X must be aperiodic. (The converse is false, however.)

Theorem 4.6.1 If $X = (X_n : n \geq 0)$ is an irreducible positive recurrent aperiodic Markov chain, then

$$P^n(x,y) \to \pi(y)$$

as $n \to \infty$, where $\pi = (\pi(y) : y \in S)$ is the stationary distribution of X.

For a proof via coupling, see p.273-275 of *Probability: Theory and Examples* by R. Durrett (1991).