A comment on simulation of MC's

Chapter 3 in the course book treat simulation on Markov chains. Given a sequence, $(U_0, U_1, ...)$, of independent identically distributed (i.i.d.) random numbers, all having distribution $U[0, 1]$, we can simulate a Markov chain.

Suppose we want to simulate the Markov chain $(X_0, X_1, ...)$ with state space $S = \{s_1, ..., s_k\}$, transition matrix P and inital distribution $\mu^{(0)}$.

We let $\psi_\nu : [0, 1] \to S$ be a function that given a random number returns a random state in S according to the distribution ν. We can simulate X_0 as follows.

$$X_0 = \psi_{\mu^{(0)}}(U_0)$$

Next we need an update function simulating the behaviour prescribed by the transition matrix. Let $\phi : S \times [0, 1] \to S$ be a function that given a state s_i and a random number returns a state s_j accoording to distribution $\mathbb{P}(X_{n+1} = s_j | X_n = s_i)$ prescribed by P. We can then simulate X_{n+1} given X_n by the following

$$X_{n+1} = \phi_P(X_n, U_n)$$

Given a certain sequence of random numbers $(U_0, U_1, ...)$ we get a realization of the Markov chain $(X_0, X_1, ...)$ by.

$$\begin{aligned} X_0 &= \psi_{\mu^{(0)}}(U_0) \\ X_1 &= \phi_P(X_0, U_1) \\ X_2 &= \phi_P(X_1, U_2) \\ &\vdots \\ X_{n+1} &= \phi_P(X_n, U_n) \\ &\vdots \end{aligned}$$

Given two idependent sequences $(U_0, U_1, ...)$ and $(U_0', U_1', ...)$ we can simulate two independent copies of the Markov chain.

$$\begin{aligned} X_0 &= \psi_{\mu^{(0)}}(U_0) & X_0' &= \psi_{\mu^{(0)}}(U_0') \\ X_1 &= \phi_P(X_0, U_1) & X_1' &= \phi_P(X_0', U_1') \\ X_2 &= \phi_P(X_1, U_2) & X_2' &= \phi_P(X_1', U_2') \\ &\vdots & &\vdots \\ X_{n+1} &= \phi_P(X_n, U_n) & X_{n+1}' &= \phi_P(X_n', U_n') \\ &\vdots & &\vdots \end{aligned}$$

Though $(X_0, X_1, ...)$ and $(X_0', X_1', ...)$ are independent they have the same probabilistic behaviour.

Another consequence of the construction is that **if** for some n we have $X_n = X_n'$, we could use X'_n and U'_n when we update to get X_{n+1}.

$$\begin{aligned}
X_0 &= \psi_{\mu^{(0)}}(U_0) \\
&\vdots \\
X_n &= \phi_P(X_{n-1}, U_n) \\
X_{n+1} &= \phi_P(X'_n, U'_n) \quad , \text{ making } X_{n+1} = X'_{n+1} \\
X_{n+2} &= \phi_P(X_{n+1}, U'_{n+1}) \quad , X_{n+1} = X'_{n+1} \;\Rightarrow\; X_{n+2} = X'_{n+2} \\
&\vdots
\end{aligned}$$

In other words, if $X_n = X'_n$ then the chain $(X_0, X_1, ..., X_{n-1}, X_n, X'_{n+1}, X'_{n+2}, ...)$ has the same probabilistic behaviour as $(X_0, X_1, ...)$ and $(X'_0, X'_1, ...)$.

The theorem ...

Theorem 5.1 ***The Markov chain convergence theorem***
Let $(X_0, X_1, ...)$ be an irreducible aperiodic Markov chain with state space $S = \{s_1, ..., s_k\}$, transition matrix P, and arbitrary initial distribution $\mu^{(0)}$. Then, for any distribution π which is stationary for the transition matrix P, we have

$$\mu^{(n)} \overset{TV}{\to} \pi.$$

Proof :
We will use a so called coupling argument. We construct two chains on S both governed by the transition matrix P but with different initial distributions. Next step is show that they will meet, that is, $X_n = X'_n$, will occur for some finite n. The last step is to use the two chains to construct a third chain, also having P as transition matrix, and see that it will in finite time have distribution π.

Step 1: Given two independent sequences of i.i.d. random numbers all uniformly distributed over $[0, 1]$ we construct two chains, $(X_0, X_1, ...)$ and $(X'_0, X'_1, ...)$. We let $\mu^{(0)}$ be the initial distribution for $(X_0, X_1, ...)$ and π be the initial distribution for $(X'_0, X'_1, ...)$.

$$\begin{aligned}
X_0 &= \psi_{\mu^{(0)}}(U_0) & X'_0 &= \psi_\pi(U'_0) \\
X_1 &= \phi_P(X_0, U_1) & X'_1 &= \phi_P(X'_0, U'_1) \\
X_2 &= \phi_P(X_1, U_2) & X'_2 &= \phi_P(X'_1, U'_2) \\
&\vdots & &\vdots
\end{aligned}$$

Step 2: Eventually these two chains will meet, that is finially we will have $X_n = X'_n$. Let T be the first time they are in the same state.

$$T = \min\{n : X_n = X'_n\}$$

With the conventionthat $T = \infty$ if they never meet. Due to irreducibility and aperiodicity and theorem 4.1 we can find $M < \infty$ such that

$$\forall i, j \in \{1, ..., k\} : \ (P^M)_{i,j} > 0$$

Let α be the smallest of those positive jump probabilities, that is

$$\alpha = \min\{(P^M)_{i,j} : \ i, j \in \{1, ..., k\}\}.$$

If we can establish that $T < M$ occur with positive probability, then eventually the two chains will meet.

$$\begin{aligned}
\mathbb{P}(T \le M) &\ge \mathbb{P}(T = M) \\
&= \mathbb{P}(X_M = X'_M) \\
&\ge \sum_{i=1}^{k} \mathbb{P}(X_M = s_1, \ X'_M = s_1) \\
&\ge \mathbb{P}(X_M = s_1, \ X'_M = s_1) \\
&= \mathbb{P}(X_M = s_1)\mathbb{P}(X'_M = s_1) \\
&= \left(\sum_{i=1}^{k} \mathbb{P}(X_0 = s_i, \ X_M = s_1)\right)\left(\sum_{i=1}^{k} \mathbb{P}(X'_0 = s_i, \ X'_M = s_1)\right) \\
&= \left(\sum_{i=1}^{k} \mathbb{P}(X_0 = s_i)\underbrace{\mathbb{P}(X_M = s_1 | X_0 = s_i)}_{\ge \alpha}\right)\left(\sum_{i=1}^{k} \mathbb{P}(X'_0 = s_i)\underbrace{\mathbb{P}(X'_M = s_1 | X'_0 = s_i)}_{\ge \alpha}\right) \\
&= \left(\alpha \sum_{i=1}^{k} \mathbb{P}(X_0 = s_i)\right)\left(\alpha \sum_{i=1}^{k} \mathbb{P}(X'_0 = s_i)\right) \\
&= \alpha^2
\end{aligned}$$

Indeed, $\mathbb{P}(T \le M) \ge \alpha^2 > 0$, and thus $\mathbb{P}(T > M) \le 1 - \alpha^2 < 1$. Given that the two chains hasn't met at time M the probability of them meeting before time $2M$ is atleast $1 - \alpha^2$ due to the Markov property.

$$\mathbb{P}(T > 2M) = \mathbb{P}(T > M)\mathbb{P}(T > 2M | T > M) \le (1 - \alpha^2)\mathbb{P}(T > 2M | T > M) \le (1 - \alpha^2)^2$$

We can iterate this argument for any l and get

$$\mathbb{P}(T > lM) \le (1 - \alpha^2)^l$$

So the probability they never meet is zero.

$$\mathbb{P}(X_n \ne X'_n \text{ for all } n) = \lim_{l \to \infty} \mathbb{P}(T > lM) \le \lim_{l \to \infty} (1 - \alpha^2)^l = 0$$

Step 3: Construct another chain $(X''_0, X''_1, ...)$ by the following

$$X''_0 = X_0 \ , \quad \text{then iterate } X''_{n+1} = \begin{cases} \phi_P(X''_n, U_{n+1}) & \text{if } X''_n \ne X'_n \\ \phi_P(X''_n, U'_{n+1}) & \text{if } X''_n = X'_n \end{cases}$$

A few observations regarding this chain ...

1. $(X''_0, X''_1, ...)$ evolves like $(X_0, X_1, ...)$ until time T when it first meets $(X'_0, X'_1, ...)$, after that it evelves like $(X'_0, X'_1, ...)$.
2. $(X''_0, X''_1, ...)$ is a Markov chain governed by transition matrix P

Now fix an abitrary state $s_i \in S$.

$$\begin{aligned} \mu_i^{(n)} - \pi_i &= \mathbb{P}(X''_n = s_i) - \mathbb{P}(X'_n = s_i) \\ &\leq \mathbb{P}(X''_n = s_i) - \mathbb{P}(X'_n = s_i,\ X''_n = s_i) \\ &= \mathbb{P}(X'_n = s_i,\ X''_n \neq s_i) \\ &\leq \mathbb{P}(X''_n \neq X'_n) \\ &= \mathbb{P}(T > n) \end{aligned}$$

If we interchange X''_n and X'_n we get $\pi_i - \mu_i^{(n)} \leq \mathbb{P}(T > n)$, and we have proved

$$|\mu_i^{(n)} - \pi_i| \leq \mathbb{P}(T > n) \to 0 \text{ as } n \to \infty$$

This is valid for any state in S, implying . . .

$$\lim_{n\to\infty} d_{\text{TV}}(\mu^{(n)}, \pi) = \lim_{n\to\infty} \left(\frac{1}{2} \sum_{j=1}^{k} |\mu_i^{(n)} - \pi_i| \right) = 0$$

This means that

$$\lim_{n\to\infty} \mu^{(n)} = \pi$$

and we are done. □

A remark about coupling

The technique used when constructing the Markov chain $(X''_0, X''_1, ...)$ is called coupling. Coupling itself is an active research area resulting in numerous articles and books. What we have seen in the poof of theorem 5.1 is just an application, but an important one, of a more general technique.

Reversible Markov chains

A reversible chain $(X_0, X_1, ...)$ has the property that if we start it in the stationary distribution and look at a typical realization

$$(..., x_{n-2}, x_{n-1}, x_n, x_{n+1}, x_{n+2}, ...)$$

and then reverse the order . . .

$$(..., x_{n+2}, x_{n+1}, x_n, x_{n-1}, x_{n-2}, ...)$$

. . . they will have the same probabilistic behaviour, that is, from a probabilistic view they could both be realizations of $(X_0, X_1, ...)$.

Note that the reversibility refers to a property of a distribution on S, which is related to the transition matrix. We say that a distribution is reversible for a transition matrix P. When say that a distribution π is stationary, it means that π is stationary for a transition matrix P

Reversibility is a stronger statement than stationarity. First we will define what we exactly mean by reversibility and the give theorem implying that reversibility is the stronger property. A couterexample will then be given to show that stationary and reversibility is not equivalent properties, that is, stationarity does not imply reversibility.

Definition 6.1 ***Reversible Markov chains***
Let $(X_0, X_1, ...)$ be a Markov chain with state space $S = \{s_1, ..., s_k\}$ and transition matrix P. A probability distribution π on S is said to be reversible for the chain, or for the transition matrix P, if for all $i, j \in \{1, ..., k\}$ we have

$$\pi_i P_{i,j} = \pi_j P_{j,i}.$$

Theorem 6.2
Let $(X_0, X_1, ...)$ be a Markov chain with state space $S = \{s_1, ..., s_k\}$ and transition matrix P. If π is a reversible distribution for the chain, then it is also a stationary distribution for the chain.

Interpretation/Consequences:
If we cannot, or if it's hard, to show that a chain has a stationary distribution, we could try to prove that it has reversible distribution, and stationarity will follow. Though the existence of a reversible distribution is a stronger statement than just existence of a stationary distribution, this clalculation might be easier.

Idea of proof:
A simple calculation will suffice to check if the reversible distribution is stationary. Let π be stationary for P.

$$\pi_i = \pi_i \sum_{j=1}^{k} P_{i,j} = \sum_{j=1}^{k} \pi_i P_{i,j} = \sum_{j=1}^{k} \pi_j P_{j,i}$$

where the third equality is due to reversibility of π. This implies that $\pi = \pi P$.

Stationarity does not imply reversibility : a counter example

Consider the Markov chain $(X_0, X_1, ..)$ having transition graph according to fig. 6.1.

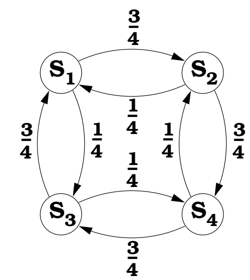

Figure 6.1: A random walk with a tendency towards clockwise movement.

We can easily check that a stationary distribution for $(X_0, X_1, ..)$ is

$$\pi = \left(\frac{1}{4}, \ \frac{1}{4}, \ \frac{1}{4}, \ \frac{1}{4} \right)$$

but it is not reversible since

$$\pi_1 P_{1,2} = \frac{1}{4} \frac{3}{4} \neq \frac{1}{4} \frac{1}{4} = \pi_2 P_{2,1}.$$