Statistics 251/551 spring 2013
2013: Solutions to sheet 2

If you are not able to solve a part of a problem, you can still get credit for later parts: Just assume the truth of what you were unable to prove in the earlier part.

[1] Consider an irreducible Markov chain with a finite state space $\mathcal{S}$. Show that every state i is positive recurrent (that is, $\mathbb{E}_i T_i < \infty$) by following these steps.

(i) (5 points) Explain why there is a positive integer N and a $\delta > 0$ for which $\mathbb{P}_j\{T_i \le N\} \ge \delta$ for all $j \in \mathcal{S}$. Hint: For each $j \ne i$ there is a path from j to i that is taken with strictly positive $\mathbb{P}_j$ probability. What if $j = i$?

For each j in $\mathcal{S}$ there is a path from j to i that the chain follows with strictly positive $\mathbb{P}_j$ probability. If the path is of length N_j then $\mathbb{P}_j\{T_i \le N_j\} \ge \delta_j$ for some $\delta_j > 0$. Define $N^* = \max_{j\in\mathcal{S},j\ne i} N_j$ and $\delta^* = \min_{j\in\mathcal{S},j\ne i}\delta_j$. Then

$$\mathbb{P}_j\{T_j \le N\} \ge \mathbb{P}_j\{T_i \le N_j\} \ge \delta_j \ge \delta \qquad \text{for each } j \text{ not equal to } i.$$

To get from i back to i, first choose a k for which $k \ne i$ and $P(i,k) > 0$ then argue that

$$\mathbb{P}_i\{T_i \le N^* + 1\} \ge P(i,k)\mathbb{P}_k\{T_i \le N^*\} \ge P(i,k)\delta^*.$$

so that $\mathbb{P}_j\{T_i \le N\} \ge \delta$ for all j in $\mathcal{S}$, where $N = N^* + 1$ and $\delta = \delta^* P(i,k)$. Some of you used a better lower bound:

$$\begin{aligned}\mathbb{P}_i\{T_i \le N^* + 1\} &= \mathbb{P}_i\{T_i = 1\} + \sum\nolimits_{j\in\mathcal{S},j\ne i} P(i,j)\mathbb{P}_j\{T_i \le N^*\}\\ &\ge P(i,i) + \sum\nolimits_{j\in\mathcal{S},j\ne i} P(i,j)\delta^* \ge \delta^*.\end{aligned}$$

Many of you failed to explain why the same N and δ could be used for all j.

(ii) (5 points) For each $k \in \mathbb{N}$ write $\mathcal{U}_k$ for the set of times $\{n \in \mathbb{N} : (k-1)N < n \le kN\}$. Let W denote the first k for which there is a visit to i during $\mathcal{U}_k$. Explain why $T_i \le NW$.

If $W = k$ then the chain first visits i at or before the end of block $\mathcal{U}_k$, which is at time Nk.

(iii) (5 points) Intuitively, from part (i), within each $\mathcal{U}_k$ time block the chain should visit state i with probability at least δ, no matter what has happened up to time $(k-1)N$. To formalize this intuition, define

$$V_k = \{\text{chain visits } i \text{ during block } \mathcal{U}_k\} = \cup_{n\in\mathcal{U}_k}\{X_n = i\}.$$

Show that $\mathbb{P}_i(V_\ell^c \mid V_1^c \cap V_2^c \cap \cdots \cap V_{\ell-1}^c) \le 1 - \delta$ for each $\ell \ge 2$. (Hint: Condition on $X_{(\ell-1)N}$.) Deduce that $\mathbb{P}_i\{W \ge k\} \le (1-\delta)^{k-1}$ for each k in $\mathbb{N}$.

Write M for $(\ell-1)N$ and D for the event $\cap_{m<\ell} V_m^c = \{X_n \neq i : \text{for } 1 \leq n \leq M\}$. Then

$$\mathbb{P}_i(V_\ell^c \mid D\} = \sum\nolimits_{j\in\mathcal{S}} \mathbb{P}_i(X_M = j \mid D)\mathbb{P}_i(V_\ell^c \mid X_M = j, D).$$

The ith term in the sum is zero because $\mathbb{P}_i(X_M = i \mid X_M \neq i, \dots) = 0$. And by the Markov property,

$$\mathbb{P}_i(V_\ell^c \mid X_M = j, D) = \mathbb{P}_j\{T_i > N\} \leq 1-\delta \qquad \text{for } j \neq i.$$

Thus

$$\mathbb{P}_i(V_\ell^c \mid D\} \leq (1-\delta)\sum\nolimits_{j\in\mathcal{S}\setminus i} \mathbb{P}_i(X_M = j \mid D) = 1-\delta.$$

Similarly, $\mathbb{P}_i V_1^c = \mathbb{P}_i\{T_i > N\} \leq 1-\delta$. Deduce that

$$\begin{aligned}\mathbb{P}_i\{W \geq k\} &= \mathbb{P}_i\left(V_1^c V_2^c \dots V_{k-1}^c\right)\\ &= P_i(V_1^c)\mathbb{P}_i(V_2^c \mid V_1^c)\dots\mathbb{P}_i(V_{k-1}^c \mid V_1^c \cap V_2^c \cap \dots \cap V_{k-2}^c)\\ &\leq (1-\delta)^{k-1}.\end{aligned}$$

Some of you incorrectly assumed that the behaviors of the chain during each disjoint time block are independent. More precisely, you assumed independence of the events $V_1, V_2, \dots$, perhaps by a false analogy with independence of cycles (as in the next Problem).

Many of you wrote

$$\mathbb{P}_i\{V_\ell^c \mid D\} = \mathbb{P}_i\{V_\ell^c \mid X_M \neq i, \text{ past info}\} \quad ??=?? \quad \mathbb{P}_i\{V_\ell^c \mid X_M \neq i\},$$

citing the Markov property as justification for discarding the past info. (The extra question marks are mine.) In fact the Markov property lets one discard information about earlier steps only when conditioning on the exact value taken by X_M. See the remark at the bottom of Chang page 8 and the related Exercise 1.1. My solution avoids this Markov gotcha by decomposing according to which state X_M can be in when $X_M \neq i$.

(iv) (5 points) If the last inequality were actually an equality for each k then W would have a geometric distribution, which has a finite expected value. In some sense, W is smaller than a geometric, so W should also have a finite expected value. Prove that $\mathbb{E}_i W < \infty$. Hint: First explain why $W = \sum_{k\in\mathbb{N}} \mathbb{I}\{W \geq k\}$.

In class I explained why $W = \sum_{k\in\mathbb{N}} \mathbb{I}\{W \geq k\}$, by pointing out that exactly ℓ of the indicator functions equal ℓ when W takes the value ℓ. This equality implies

$$\mathbb{E}_i T_i \leq N\mathbb{E}_i W = N\sum\nolimits_{k\in\mathbb{N}} \mathbb{P}_i\{W \geq k\} \leq N\sum\nolimits_{k\in\mathbb{N}} (1-\delta)^{k-1} < \infty.$$

The state i is positive recurrent.

Some of you declared W to have a geometric distribution, which is not correct, although in some vague sense W acts as if it were bounded above by something geometric. The difficulty is: if the chain does not visit i during time block $\mathcal{U}_{\ell-1}$ then the conditional probabilty of visiting i during block $\mathcal{U}_\ell$ depends on the state that X_M is in. We only a have a uniform lower bound for the conditional probability of V_ℓ.

[2] *Suppose $i \rightsquigarrow j$ (that is, state j is accessible from state i). Suppose also that $\tau := \mathbb{E}_i T_i < \infty$. Show that $\mathbb{E}_j T_j < \infty$ by the following steps.*

Remember that $\theta = \mathbb{P}_i\{X_1 = i_1, X_2 = i_2, \ldots, X_k = j\} > 0$ for some sequence of states $i_1, i_2, \ldots, i_k = j$. Write $T_i^{(1)}$, $T_i^{(2)}$, ... for the lengths of successive cycles that start from i. That is, if starting from state i, the first return to i occurs at time $T_i^{(1)}$, the second at time $T_i^{(1)} + T_i^{(2)}$, and so on. If starting from state j, the successive visits to i occur at times T_i, $T_i + T_i^{(1)}$, $T_i + T_i^{(1)} + T_i^{(2)}, \ldots$. Also, write F_m for the event that the mth cycle (starting from i) begins with visits to states $i_1, i_2, \ldots, i_k$ in that order.

(i) *(5 points) Explain why $\mathbb{E}_i T_i^{(m)} \geq \theta \mathbb{E}_i(T_i^{(m)} \mid F_m)$ and $\mathbb{E}_i(T_i^{(m)} \mid F_m) = k + \mathbb{E}_j T_i$. Deduce that $\mathbb{E}_j T_i < \infty$.*

$$\mathbb{E}_i T_i^{(m)} = (\mathbb{P}_i F_m^c)\,\mathbb{E}_i(T_i^{(m)} \mid F_m^c) + (\mathbb{P}_i F_m)\,\mathbb{E}_i(T_i^{(m)} \mid F_m)$$

and

$$\mathbb{E}_i(T_i^{(m)} \mid F_m) = \mathbb{E}_i(T_i^{(1)} \mid F_1) = \mathbb{E}_i\left(T_i^{(1)} \mid X_1 = i_1, \ldots, X_k = i_k\right).$$

Given that conditioning information, the first cycle consists of k steps that take the chain to state j then the rest of the cycle consists of steps, starting at j, to get back to i.

(ii) *(5 points) If the chain starts in state i, explain why*

$$T_j \leq T_i^{(1)} + \sum\nolimits_{m\geq 2}\left(T_i^{(m)}\mathbb{I}(F_1^c \cap F_2^c \ldots F_{m-1}^c)\right).$$

Hint: If $m = 3$ is the first cycle for which F_m occurs, why is T_j less than $T_i^{(1)} + T_i^{(2)} + T_i^{(3)}$?

If event F_1 occurs then the chain visits j before time $T_i^{(1)}$, when the cycle ends. If $F_1^c, \ldots, F_{\ell-1}^c, F_\ell$ occur then the first visit to j is somewhere during the ℓth cycle, which ends at step $T_i^{(1)} + \ldots T_i^{(\ell)}$, which is exactly the sum on the right-hand side.

Notice that F_1^c might occur and still the chain might visit j during the first cycle: there are other ways to get to j than along the $i_1, i_2, \ldots, i_k$ path.

(iii) *(5 points) By taking $\mathbb{E}_i$ expectations of both sides of the inequality from the previous part, and by using independence between what happens in each cycle, deduce that*

$$\mathbb{E}_i T_j \leq \tau\left(1 + (1-\theta) + (1-\theta)^2 + \ldots\right) < \infty$$

By independence between cycles,

$$\mathbb{E}_i T_j \le \mathbb{E}_i T_i^{(1)} + \sum\nolimits_{m\ge 2} \mathbb{E}_i T_i^{(m)} \mathbb{E}_i(\mathbb{I}(F_1^c) \dots \mathbb{E}_i \mathbb{I}(F_{m-1}^c)$$
$$= \tau + \sum\nolimits_{m\ge 2} \tau \mathbb{P}_i(F_1^c) \dots \mathbb{P}_i(F_{m-1}^c).$$

And $\mathbb{P}_i(F_\ell^c) = 1 - \theta$ for each ℓ.

(iv) (5 points) Write $T_i + S_j$ for the first time after T_i that the chain visits state j. Show that

$$\mathbb{E}_j T_j \le \mathbb{E}_j(T_i + S_j) \le \mathbb{E}_j T_i + \mathbb{E}_i T_j < \infty.$$

Very formally,

$$\mathbb{E}_j S_j = \sum_{n\in\mathbb{N}} \mathbb{P}_j(T_i = n)\mathbb{E}_j(S_j \mid T_i = n)$$

and $\mathbb{E}_j(S_j \mid T_i = n) = \mathbb{E}_i T_j$ for each n, by the Markov property. The random variable S_j counts the number of steps to reach state i for a chain that starts in state j.