Probability and Computation: Problem sheet 4 Solutions

You are encouraged to submit your solutions by emailing them to luca.zanetti@cl.cam.ac.uk by Wednesday 20th of February.

Question 1. *Let $(X_i)_{i\geq 1}$ be independent random variables with $\mathbf{P}[\,X_i = 1\,] = \mathbf{P}[\,X_i = -1\,] = 1/2$. Let $S_n = \sum_{i=0}^{n} X_n$, with $X_0 = K > 0$. For $N > K$ define*

$$T = T_{0,N} = \min\{n \geq 0 : S_n = 0 \text{ or } S_n = N\}.$$

1. *Prove that $\mathbf{E}[\,T\,] < \infty$ (you cannot use the OST).*
2. *Find a deterministic sequence of values $a_n \in \mathbb{R}$ such that $Z_n = S_n^3 + a_n S_n$ is a martingale w.r.t. $X_0, X_1, \ldots$*
3. *Find deterministic sequences $b_n, c_n \in \mathbb{R}$ such that $W_n = S_n^4 + b_n S_n^2 + c_n$ is a martingale w.r.t. $X_0, X_1, \ldots$*

Solution: 1. Let T' be the first time we win N times in a row. When T' happens we either have that $S_{T'} \geq N$ or not. If $S_{T'} \geq N$ it means that at some time $t < T'$ we had $S_t = N$. If $S_{T'} < N$ then $S_{T'-N} < 0$ then at some time $t < T' - N$ we had $S_t = 0$. In both cases it exist some $t < T'$ such that $S_t = 0$ or $S_t = N$, therefore $T < T'$. The expectation of T' was computed in Question 6 of the previous problem sheet.

2. Clearly Z_n is a function of $X_0, \ldots, X_n$ and $|Z_n| \leq n^3 + |a_n|n$, which check the first two parts of the definition of a martingale. We need to find a_n such that

$$\mathbf{E}[\,Z_{n+1}|X_0, \ldots, X_n\,] = Z_n.$$

note that

$$S_{n+1}^3 = (S_n + X_{n+1})^3 + a_n(S_n + X_{n+1}) = S_n^3 + 3S_n^2 X_{n+1} + 3S_n X_{n+1}^2 + X_{n+1}^3$$

Note that $X_{n+1}^2 = 1$ and $X_{n+1}^3 = X_{n+1}$, then

$$\mathbf{E}[\,Z_{n+1}|X_0, \ldots, X_n\,] = S_n^3 + 3S_n + a_{n+1}S_n \tag{1}$$

we want to force $Z_n = S_n^3 + 3S_n + a_{n+1}S_n$, which implies that $a_n = 3 + a_{n+1}$. The solutions of the recursion for a_n is $a_n = -3n + a_0$. We just set the initial value $a_0 = 0$. Therefore $Z_n = S_n^3 - 3nS_n$ is a martingale

3. W_n is a function of $X_0, \ldots, X_n$ and $|W_n| \leq n^4 + |b_n|n^2 + |c_n|$, which check the first two parts of the definition of a martingale.

We proceed to check that $\mathbf{E}[\,W_{n+1}|X_0, \ldots, X_n\,] = W_n$ for some values of b_n and c_n.

Using that $X_n^{2k} = 1$ and $X_n^{2k+1} = X_n$ for $k \geq 0$, then

$$S_{n+1}^4 = (S_n + X_{n+1})^4 = S_n^4 + 4S_n^3 X_{n+1} + 6S_n^2 + 4S_n X_{n+1} + 1,$$

and

$$S_{n+1}^2 = S_n^2 + 2S_n X_{n+1} + 1.$$

Hence

$$\mathbf{E}[\,W_{n+1}|X_0, \ldots, X_n\,] = S_n^4 + (6 + b_{n+1})S_n^2 + (1 + b_{n+1} + c_{n+1}) = S_n^4 + b_n S_n^2 + c_n$$

This suggest that $b_n = 6 + b_{n+1}$ and $c_n = 1 + c_{n+1} + b_{n+1}$. Clearly $b_n = -6n + b_0$, so we choose $b_0 = 0$. Then $c_n = 1 - 6n + c_{n+1}$, hence

$$c_n - c_0 = \sum_{i=1}^{n} c_i - c_{i-1} = \sum_{i=1}^{n}[6(i-1) - 1] = 3n^2 - 4n$$

Again, we choose $c_0 = 0$ and we conclude that

$$W_n = S_n^4 - 6nS_n^2 + (3n^2 - 4n)$$

is a martingale

Question 2.

1. *Consider out path on vertices $\{0, \dots, N\}$, and suppose $X_0 = K$. Compute $h_{K,N}$*
2. *Compute the cover time of a path on $\{0, \dots, N\}$ when N is even. What about when N is odd?*
3. *Consider a cycle on vertices $\{0, 1, \dots, N\}$ where vertex i is adjacent to $i + 1$, and 0 is adjacent to N. Compute the cover time.*
4. *Consider a cycle on vertices $\{0, 1, \dots, N\}$. Define T by*

$$T = \min\{m : \cup_{i=0}^{m} X_i = \{0, \dots, N\}\}$$

T is the first time all vertices has been covered. Compute $\mathbf{P}[\,X_T = i | X_0 = 0\,]$ for $i \in \{1, \dots, N\}$.

Solution: 1. A random walk on the path can be seen as a folded version of a random walk on $\mathbb{Z}$ where we associate point $-k$ with point k. Let $S_n = \sum_{i=1}^{n} X_i$ be a random walk on $\mathbb{Z}$ starting from K (recall that X_i are independent, $\mathbf{P}[\,X_i = 1\,] = \mathbf{P}[\,X_i = -1\,] = 1/2$, and $X_0 = 0$.)

Therefore, the problem of computing $h_{K,N}$ is equivalent to the problem to finding $\mathbf{E}[\,T_{N-,N} | S_0 = K\,]$. It was shown in class that this is equivalent to $\mathbf{E}[\,T_{0,2N} | S_n = N + K\,] = (N + K)(2N - (N + K)) = (N + K)(N - K)$

2. Starting from K we need to hit either 0 or N which takes $K(N - K)$ times in expectation. Later from one of the extremes we need to hit the other extreme which takes N^2 times. Then, the total time to cover the graph is $N^2 + K(N - K)$. Recall the cover time is considered over the worst starting point, so maximizing K we get that the cover time is $N^2 + \lfloor N/2 \rfloor \lceil N/2 \rceil$

3. First of all, note that the set of covered vertices is a path in the cycle. Second, note when we just discover a new vertex this is one of the extreme points of the path. Third, we can always relabel the vertices of the cycle, so the set of cover vertices is $\{0, 1, \dots, K\}$ and the last discovered vertex is K. The expected time to discover a new vertex is equivalent the expected time to move outside this path. As a random walk problem this is equivalent to $\mathbf{E}[\,T_{-1,K+1} | S_0 = K\,] = K + 1$. Now we just need to sum from $K = 0$ to $n - 1$ which equals $n(n - 1)/2$.

4. We will prove that $\mathbf{P}[\,X_T = i | X_0 = 0\,] = 1/N$. For that, note that before hitting vertex i we are either in vertex $i - 1$ or vertex $i + 1$ ($\mod N$). Without lost of generality, suppose we hit vertex $i - 1$ before vertex $i + 1$ (otherwise just rename all the vertices). Then if i is the last vertex we hit, we have to hit $i + 1$ before vertex i starting from vertex $i - 1$. Note this is equivalent to start a path in 1, and we want to hit $N - 1$ before vertex 0. In lectures we analysed that problem and deduced that such a probability is $1/(N - 1)$. Therefore $\mathbf{P}[\,X_T = i | X_0 = 0\,] \geq 1/(N - 1)$. Since there are $N - 1$ other vertices besides 0 and for all of them $\mathbf{P}[\,X_T = i | X_0 = 0\,] \geq 1/(N - 1)$, then the only possibility is that$\mathbf{P}[\,X_T = i | X_0 = 0\,] = 1/(N - 1)$

Question 3. ***Wald's Equation:*** *Let $X_1, \dots,$ i.i.d. non-negative random variables with finite expectation. Let T be a stopping time with respect to this sequence and suppose that $\mathbf{E}[\,T\,] < \infty$ and that $\mathbf{E}[\,|X_1|\,] < \infty$. Prove that*

$$\mathbf{E}\left[\sum_{i=1}^{T} X_i\right] = \mathbf{E}[\,T\,]\,\mathbf{E}[\,X_1\,].$$

Solution: Define $Z_n = \sum_{i=1}^n X_i - n\mathbf{E}[\, X_1 \,]$ and $Z_0 = 0$. Then Z_n is a martingale w.r.t $X_1, \dots, X_n$. Clearly, Z_n is a function of $X_1, \dots, X_n$. Also

$$\mathbf{E}[\, |Z_n| \,] \le 2n\mathbf{E}[\, |X_1| \,] < \infty.$$

Finally,

$$\mathbf{E}[\, Z_{n+1} | X_1, \dots, X_n \,] = Z_n + \mathbf{E}[\, X_{n+1} - \mathbf{E}[\, X_1 \,] \,|X_1, \dots, X_n \,] = Z_n,$$

where in the last equality we use that the X_i's are independent and they have the same distribution and thus $\mathbf{E}[\, X_1 \,] = \mathbf{E}[\, X_{n+1} \,]$.

Finally, as T is a stopping time, we can use condition iii) of the OST and then

$$\mathbf{E}[\, Z_T \,] = \mathbf{E}[\, Z_0 \,] = 0,$$

but

$$Z_T = \sum_{i=1}^T X_i - T\mathbf{E}[\, X_1 \,],$$

from which we conclude the result.

Question 4. *A weighted undirected graph $G = (V, E, w)$ is defined by a set vertices V, a collection of edges $E \subseteq V \times V$, and a weight function $w \colon V \times V \to \mathbb{R}_{\ge 0}$ such that, for any $u, v \in V$, $w(u, v) = w(v, u)$ and $w(u, v) > 0$ if and only if $(u, v) \in E$. Self-loops of the kind (u, u) are allowed. A random walk on $G = (V, E, w)$ is a Markov chain with transition matrix P such that, for any $u, v \in V$, $P(u, v) = w(u, v)/d(u)$, where $d(u) = \sum_{z \in V} w(u, z)$.*

1. *What is the stationary distribution of this Markov chain?*
2. *What does being aperiodic amounts to?*
3. *Prove that a Markov chain is reversible if and only if it can be represented by a random walk on a weighted undirected graph.*
4. *Prove that if P is reversible, then P^t is also reversible for any $t \in \mathbb{N}$.*

Solution:

1. Let $\pi \colon V \to \mathbb{R}$ such that $\pi(u) = \frac{d(u)}{\sum_{z \in V} d(z)}$. Then, π is stationary for P. To prove this, we just need to check that $\pi P = \pi$. Let u be an arbitrary vertex. Then,

$$(\pi P)(u) = \sum_{v \in V} \pi(v) P(v, u) = \sum_{v \in V} \frac{d(v) P(v, u)}{\sum_{z \in V} d(z)} \sum_{v \in V} \frac{w(v, u)}{\sum_{z \in V} d(z)} = \frac{d(v)}{\sum_{z \in V} d(z)} = \pi(v).$$

   where the third equality follows from $P(v, u) = w(v, u)/d(v)$.

2. Being aperiodic for a random walk on an undirected graphs simply means the graph does not contain bipartite connected components.

3. We start showing that a random walk on an undirected graph is always reversible, i.e., it satisfies the detailed balance condition: $\pi(u)P(u, v) = \pi(v)P(v, u)$ for any $u, v \in V$. Let u, v be arbitrary vertices. Then,

$$\pi(u)P(u, v) = \frac{d(u)}{\sum_{z \in V} d(z)} \cdot \frac{w(u, v)}{d(u)} = \frac{d(v)}{\sum_{z \in V} d(z)} \cdot \frac{w(v, u)}{d(v)} = \pi(v)P(v, u)$$

   where the second equation follows from the fact that, since G is undirected, $w(u, v) = w(v, u)$.
   We now show the reverse implication: we are given a transition matrix P on Ω with stationary distribution π such that $\pi(u)P(u, v) = \pi(v)P(v, u)$ for any $u, v \in \Omega$, and we want to show we can construct an undirected weighted graph $G = (V, E, w)$ such that P is the transition matrix

of a random walk on G. First of all, we choose $\Omega = V$. Then, we construct the weight function $w\colon V \to \mathbb{R}_{\geq 0}$ as $w(u,v) = \pi(u)P(u,v)$, and we set $E = \{\{u,v\} : w(u,v) > 0\}$. We need to show that w is a proper weight function. Clearly, w is nonnegative and strictly positive exactly on E. Moreover, $w(u,v) = \pi(u)P(u,v) = \pi(v)P(v,u) = w(v,u)$ since P is reversible. Finally,

$$\frac{w(u)}{\sum_{z\in V} w(u,z)} = \frac{\pi(u)P(u,v)}{\sum_{z\in V} \pi(u)P(u,z)} = \frac{P(u,v)}{\sum_{z\in V} P(u,z)} = P(u,v),$$

where the last equality follows from the fact that each row of P sum up to 1 (by definition of transition matrix). Hence, we have shown that a random walk on G has transition matrix P.

4. There are several ways to show this fact. The easiest way is probably to use the fact that P is reversible if and only if $\langle Pf, g\rangle_\pi = \langle f, Pg\rangle_\pi$ for any $f, g \in V \to \mathbb{R}$. Hence, we need to show that, for arbitrary $f, g \in V \to \mathbb{R}$, $\langle P^t f, g\rangle_\pi = \langle f, P^t g\rangle_\pi$:

$$\langle P^t f, g\rangle_\pi = \langle P^{t-1} f, Pg\rangle_\pi = \langle P^{t-2} f, P^2 g\rangle_\pi = \cdots = \langle f, P^t g\rangle_\pi$$

where at each step we have applied the reversibility of P.

Question 5. *Let P be the transition matrix of a (simple) random walk on an undirected graph $G = (V,E)$. Let $\lambda_1 \geq \cdots \geq \lambda_n$. Prove the following.*

1. $\lambda_1 = 1$.
2. $\lambda_2 = 1$ *if and only if the graph is disconnected.*
3. $\lambda_n = -1$ *if and only if there exists a bipartite connected component.*
4. *Suppose now that the random walk is lazy (i.e., $P(u,u) \geq 1/2$ for any $u \in V$). Prove that all the eigenvalues of P are non-negative.*

Solution:

1. Since the all-one vector is an eigenvector of P with eigenvalue 1, $\lambda_1 \geq 1$. We just have to show that we cannot have an eigenvalue strictly greater than 1. Actually, we will show something stronger: P doesn't have eigenvalues strictly greater than 1 *in absolute value.*
   Let f be an eigenvector of eigenvalue λ for P, i.e., $Pf = \lambda f$. Take x maximising $|f(x)|$. Since f is an eigenvector of eigenvalue λ,

$$|\lambda||f(x)| = |(Pf)(x)| = \left|\sum_y P(x,y)f(y)\right| \leq |f(x)| \left|\sum_y P(x,y)\right| = |f(x)|.$$

   Clearly, $f(x) \neq 0$. Therefore, $|\lambda| \leq 1$.

2. We will first show that if G is disconnected, then $\lambda_2 = 1$. Recall the variational characterisation of λ_2:

$$1 - \lambda_2 = \min_{0 \neq f \perp 1} \frac{\sum_{x,y}(f(x) - f(y))^2 P(x,y)\pi(x)}{\|f\|_{2,\pi}^2} \tag{2}$$

   This expression is clearly nonnegative. Hence, we just need to find a nonzero $f \perp 1$ such that the numerator is 0. Since G is disconnected, we can partition V in two sets, S and $V \setminus S$, such that there are no edges between the two . Then, as long as f is constant on S and $V \setminus S$, the numerator is 0 ($P(x,y)$ is nonnegative only if there exists an edge between x and y). We just need to be careful to construct f so that $\langle f, 1\rangle_\pi = 0$. The following is a good choice (check!):

$$f(u) = \begin{cases} 1/\operatorname{vol}(S) & \text{if } u \in S \\ -1/\operatorname{vol}(V \setminus S) & \text{if } u \notin S \end{cases}$$

where $\mathrm{vol}(S) = \sum_{u \in S} d(u) = 2|E| \sum_{u \in S} \pi(u)$.
We now show that, if G is connected, $\lambda_2 < 1$. Suppose by contradiction $0 \neq f \perp 1$ is an eigenvector of eigenvalue $\lambda_2 = 1$. Again, by (2),

$$\sum_{x,y} (f(x) - f(y))^2 P(x,y)\pi(x) = \sum_{\{x,y\} \in E} (f(x) - f(y))^2 P(x,y)\pi(x) = 0.$$

Therefore, for any $\{x, y\} \in E$, $f(x) = f(y)$. But notice that $0 = \langle f, 1 \rangle_\pi = \sum_z f(z)\pi(z)$ implies that f must have strictly positive and strictly negative entries. Take $u, v \in V$ such that $f(u) > 0$ and $f(v) < 0$. Since G is connected, there exists a path $u = x_0, x_1, \ldots, x_\ell = v$ from u to v. But then, since for any $\{x, y\} \in E$ $f(x) = f(y)$, we have that $0 > f(u) = f(x_0) = f(x_1) = \cdots f(x_\ell) = f(v) < 0$, reaching a contradiction.

3. Without loss of generality, we assume G is connected and prove that P has eigenvalue -1 if and only if G is bipartite. This is without loss of generality because if G were disconnected, the eigenvalues of P would simply be the union of the eigenvalues of the transition matrices of the simple random walks in each one of the connected component of G. First we assume that G is bipartite and prove it has eigenvalue -1. Recall that a graph is bipartite if it can be partitioned in two sets S and $V \setminus S$ such that no edge connect two vertices in S or two vertices in $V \setminus S$ (i.e., it has a cut containing all the edges in the graph). Assuming such S and $V \setminus S$ exist, we construct a function $f \colon V \to \mathbb{R}$ as follows:

$$f(u) = \begin{cases} 1 & \text{if } u \in S \\ -1 & \text{if } u \notin S \end{cases}$$

We claim this function is an eigenvector of eigenvalue -1 for P. Let $u \in S$. then,

$$(Pf)(u) = \sum_{v \in V} P(u,v)f(v) = \sum_{v \in S} P(u,v) - \sum_{v \notin S} P(u,v) = -\sum_{v \colon \{u,v\} \in E} \frac{1}{d(u)} = -1 = -f(u),$$

where the third equality follows from the fact that all the neighbours of u are in $V \setminus S$ and the fourth follows from u having $d(u)$ neighbours. Analogously, we can prove that $(Pf)(u) = -f(u)$ for any $u \in V \setminus S$. This proves that $Pf = -f$, which completes one direction of the proof.
For the reverse direction, we assume that P has eigenvalue -1 and prove that G must be bipartite. Let f (nonzero) such that $Pf = -f$. We define $S = \{u \in V \colon f(u) \geq 0\}$ and we claim that $(S, V \setminus S)$ is a bipartition of the graph such that no edge connects two vertices in S or in $V \setminus S$. To this end, pick any $u \in V$ such that $|f(u)| = \max_{v \in V} |f(v)|$. Then,

$$-f(u) = (Pf)(u) = \sum_{v \in V} P(u,v)f(v) = \sum_{v \colon \{u,v\} \in E} \frac{f(v)}{d(u)}. \tag{3}$$

But since u has only $d(u)$ neighbours and $|f(u)|$ achieves the maximum of f in absolute value, $f(v) = -f(u)$ for any v such that $\{u, v\} \in E$. Therefore, since G is connected, we have proved that $|f(u)| = |f(v)| \neq 0$ for any $v \in V$. Hence, (3) must hold for any $u \in V$. This implies that, for any $\{u, v\} \in E$, $f(u) = -f(v)$. Therefore, $u \in S$ and $v \notin S$, or vice versa. In other words, we have proved we can partition V in two sets such that no edge connects two vertices in the same set, i.e., G is bipartite.

4. Since P is lazy, $P = \frac{1}{2}(I + P')$ where I is the identity matrix and P' is the transition matrix of a lazy random walk on G. We proved above that the eigenvalues of P' are between 1 and -1. Hence, $\lambda_n(P) = \frac{1}{2} + \frac{1}{2}\lambda_n(P') \geq 0$.
