MATH 507b Midterm

Qilin Ye

March 9, 2024

Problem 1

(1) Let X_n, Y_n be martingales w.r.t. $\mathcal{F}_n$ and assume $\mathbb{E}[X_n^2], \mathbb{E}[Y_n^2] < \infty$. Prove that

$$\mathbb{E}[X_n Y_n] - \mathbb{E}[X_0 Y_0] = \sum_{m=1}^{n} \mathbb{E}[(X_m - X_{m-1})(Y_m - Y_{m-1})].$$

(2) Let X_n be a martingale and set $\xi_n = X_n - X_{n-1}$. Suppose $\mathbb{E}[X_0^2] < \infty$ and $\sum_{m \geqslant 1} \mathbb{E}[\xi_m^2] < \infty$. Prove that X_n converges almost surely and in L^2 to some X_∞.

(3) Let X_n be a martingale with $X_0 = 0$ and set $\xi_n = X_n - X_{n-1}$. Suppose that $b_m > o$ is an increasing sequence with $\lim b_m = \infty$. Suppose that $\sum_{m \geqslant 1} b_m^{-2} \mathbb{E}[\xi_m^2] < \infty$. Prove that X_n / b_n converges to 0 almost surely and in L^2 as $n \to \infty$.

Proof. (1) Let $m \geqslant 1$. Note that

$$\begin{aligned} X_m Y_m - X_{m-1} Y_{m-1} &= X_m Y_m + X_{m-1} Y_{m-1} - 2 X_{m-1} Y_{m-1} \\ &\quad - X_{m-1} Y_m - X_m Y_{m-1} + X_{m-1} Y_m + X_m Y_{m-1} \\ &= (X_m - X_{m-1})(Y_m - Y_{m-1}) + X_{m-1}(Y_m - Y_{m-1}) + (X_m - X_{m-1}) Y_{m-1}. \end{aligned}$$

Manipulating the definitions we have

$$\begin{aligned} \mathbb{E}[X_{m-1}(Y_m - Y_{m-1})] &= \mathbb{E}[\mathbb{E}[X_{m-1}(Y_m - Y_{m-1}) \mid \mathcal{F}_{m-1}]] && \text{(tower property)} \\ &= \mathbb{E}[X_{m-1} \cdot \mathbb{E}[(Y_m - Y_{m-1}) \mid \mathcal{F}_{m-1}]] && \text{(Thm 4.1.14, } X_{m-1} \in \mathcal{F}_{m-1}) \\ &= \mathbb{E}[X_{m-1} \cdot 0] = 0 && \text{(martingale def)} \end{aligned}$$

and analogously $\mathbb{E}[(X_m - X_{m-1}) Y_{m-1}] = 0$. Therefore

$$\mathbb{E}[X_m Y_m] - \mathbb{E}[X_{m-1} Y_{m-1}] = \mathbb{E}[(X_m - X_{m-1})(Y_m - Y_{m-1})],$$

and it remains to rewrite the original LHS as $\mathbb{E}[X_n Y_n] - \mathbb{E}[X_0 Y_0] = \sum_{m=1}^{n} (\mathbb{E}[X_m Y_m] - \mathbb{E}[X_{m-1} Y_{m-1}])$.

(2) By letting $Y_n = X_n$ as in part (1) we see that for each n,

$$\mathbb{E}[X_n^2] = \mathbb{E}[X_0]^2 + \sum_{m=1}^{n} \mathbb{E}[(X_m - X_{m-1})^2] \mathbb{E}[X_0^2] + \sum_{m=1}^{n} \mathbb{E}[\xi_m^2] \leqslant \mathbb{E}[X_0]^2 + \sum_{m \geqslant 1} \mathbb{E}[\xi_m]^2 < \infty. \tag{*}$$

The result therefore follows from L^p convergence theorem, Thm 4.4.6, as $\sup_n \mathbb{E}[X_n^2] < \infty$.

(3) Note that $Y_n = \sum_{m=1}^n \xi_m/b_m$, with $Y_0 = 0$, is a martingale: integrability and measurability are clear, and

$$\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[Y_n \mid \mathcal{F}_n] + \mathbb{E}[\xi_{n+1} \mid \mathcal{F}] = Y_n + b_{n+1}^{-1}\mathbb{E}[\xi_{n+1} \mid \mathcal{F}_n].$$

But $\mathbb{E}[\xi_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - \mathbb{E}[X_n \mid \mathcal{F}_n] = X_n - X_n = 0$, so the condition holds.

We simply plug Y_n into (2). The prerequisites hold because $\mathbb{E}[Y_0^2] = 0$ and

$$\sum_{m\geqslant 1} \mathbb{E}[(Y_m - Y_{m-1})^2] = \sum_{m\geqslant 1} b_m^{-2}\mathbb{E}[\xi_m^2] < \infty.$$

Thus Y_n converges almost surely and in L^2 to some Y_∞. On one hand, Kronecker's lemma (Thm 2.5.9) implies that $b_n^{-1}\sum_{m=1}^n \xi_m = (X_n - X_0)/b_n = X_n/b_n$ converges to 0 almost surely. On the other hand, now that we know the limiting distribution is 0, we apply Kronecker's lemma once again and obtain L^2 convergence:

$$\begin{aligned}\mathbb{E}(X_n/b_n - 0)^2 &= b_n^{-2}\mathbb{E}(X_n^2) = b_n^{-2} \cdot \mathbb{E}(\sum_{m=1}^n \xi_m)^2 \\ &= b_n^2 \sum_{m=1}^n \mathbb{E}\xi_m^2 \qquad \text{(since cross terms cancel, by (*) or Thm 4.4.7)} \\ &\to 0. \qquad \text{(Kronecker: since by assumption } \textstyle\sum \mathbb{E}\xi_m^2/b_m^2 < \infty)\end{aligned}$$

□

Problem 2

Let $S_n = \xi_1 + \ldots + \xi_n$ where ξ_i are independent, not necessarily identically distributed, with $\mathbb{E}[\xi_i] = 0$ and $\mathrm{var}(\xi_i) = \sigma^2 \in (0, \infty)$. Put $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$.

(1) Prove that $X_n = S_n^2 - n\sigma^2$ is an $\mathcal{F}_n$-martingale.

(2) Fix $a > 0$ and define $T = \min\{n \geqslant 0 : |S_n| > a\}$. Show that $\mathbb{E}[T] \geqslant a^2/\sigma^2$.

(3) Suppose N is a stopping time with $\mathbb{E}N < \infty$. Prove that $\mathbb{E}[S_N^2] = \sigma^2\mathbb{E}N$.

Proof. (1) By independence $\mathrm{var}(S_n) < \infty$. Since $\mathrm{var}(S_n) = \mathbb{E}(S_n^2) - (\mathbb{E}S_n)^2$ and $\mathbb{E}S_n = \sum \mathbb{E}\xi_i = 0$ we have $\mathbb{E}S_n^2 < \infty$ and therefore $\mathbb{E}|X_n| < \infty$. It is also clear that $X_n \in \mathcal{F}_n$. Finally,

$$\begin{aligned}\mathbb{E}(S_{n+1}^2 - (n+1)\sigma^2 \mid \mathcal{F}_n) &= \mathbb{E}(S_n^2 + 2S_n\xi_{n+1} + \xi_{n+1}^2 \mid \mathcal{F}_n) - (n+1)\sigma^2 \\ &= S_n^2 + 2S_n\mathbb{E}(\xi_{n+1} \mid \mathcal{F}_n) + \mathbb{E}(\xi_{n+1}^2 \mid \mathcal{F}_n) - (n+1)\sigma^2 && (S_n \in \mathcal{F}_n)\sigma^2 \\ &= S_n^2 + 0 + \mathbb{E}(\xi_{n+1}^2) - (n+1)\sigma^2 && (\mathbb{E}\xi_{n+1} = 0,\ \xi_{n+1} \perp\!\!\!\perp \mathcal{F}_n) \\ &= S_n^2 + \sigma^2 - (n+1)\sigma^2 = S_n^2 - n\sigma^2.\end{aligned}$$

(This is identical to the proof of Example 4.2.2 and no assumption of identical distribution is needed.)

(2) If $\mathbb{E}T = \infty$ the result holds trivially. Otherwise the results follows from (3) as T is a stopping time (since $\{|S_k| \leqslant a \text{ for } k \leqslant n-1\} \cap \{|S_n| > a\}$ is $\mathcal{F}_n$ measurable), and $\sigma^2\mathbb{E}T = \mathbb{E}[S_T]^2 \geqslant a^2$ by definition of T.

(3) Let $X_n = S_n^2 - \sigma^2$ as in (1). Since for each n, $\mathbb{E}(S_n)^2 = \mathrm{var}(S_n) = n\sigma^2$, we have $\mathbb{E}S_{N\wedge n}^2 = \sigma^2\mathbb{E}[N \wedge n]$. By MCT, letting $n \to \infty$ we have $\mathbb{E}S_{N\wedge n}^2 \uparrow \sigma^2\mathbb{E}N < \infty$. Therefore $\sup_n \mathbb{E}S_{N\wedge n}^2 < \infty$. By Thm 4.2.9 $\{S_{N\wedge n}\}_n$ is a martingale and by Thm 4.4.6 it converges a.s. and in L^2, where the a.s. limit must be S_N. Combining

$\mathbb{E}S^2_{N\wedge n} \to \mathbb{E}S^2_N$ (convergence in L^2), $\sigma^2\mathbb{E}[N \wedge n] \to \sigma^2\mathbb{E}N$ (MCT), and $\mathbb{E}S^2_{N\wedge n} = \sigma^2\mathbb{E}[N \wedge n]$ (our earlier observation), the result follows. □

Problem 3

Suppose that X_n is a Markov chain on a finite or countable space S with transition function p. A function $f : S \to \mathbb{R}$ is said to be superharmonic if $f(x) \geqslant \sum_{y\in S} p(x,y)f(y)$ or equivalently if $f(X_n)$ is a supermartingale. Suppose that p is irreducible. Show that p is recurrent if and only if every nonnegative superharmonic function is constant.

Proof. Adopting our usual notation, let $\rho_{x,y} = \mathbb{P}_x(T_y < \infty)$, the probability of eventually visiting y from x.
For $\Rightarrow$, assume p is recurrent and let f be any nonnegative superharmonic function. Thm 5.3.2. states that recurrence is contagious so $\rho_{x,x} = 1$ for all x. On the other hand, by Thm 4.2.12 $f(X_n)$ converges to some $Y = f(X)$ almost surely. Thm 5.2.6 therefore implies $\mathbb{P}(f(X_n) = f(x) \text{ i.o.}) = 1$ for all x. In other words $Y = f(x)$ almost surely. But x is arbitrary, so f must be almost surely constant.
Conversely, let X be irreducible but transient, fix any x_0 and define $f(x) := \rho_{x,x_0}$. It follows immediately that $0 \leqslant f \leqslant 1$, and that it is superharmonic, since

$$f(x) = p(x,x_0)\cdot 1 + \sum_{y\neq x_0} p(x,y)f(y) \geqslant p(x,x_0)f(x_0) + \sum_{y\neq x_0} p(x,y)f(y) = \sum_y p(x,y)f(y).$$

If f were constant, say $f \equiv c \in [0,1]$, then for any x,

$$c = f(x) = p(x,x_0) + \sum_{y\neq x_0} p(x,y)f(y) = p(x,x_0) + c(1 - p(x,x_0))$$

from which we conclude $c = 1$ and $p(x,x_0) = 0$. Both lead to contradictions — the former implies p is recurrent, and the latter implies p is not irreducible. Thus f cannot be constant, and the proof is complete. □

Problem 4

Suppose that S is a countable state space and p a transition probability on S. Define

$$\alpha_n := \sup_{i,j\in S} \frac{1}{2}\sum_{k\in S}|p^n(i,k) - p^n(j,k)|.$$

(1) Show that $\alpha_{m+n} \leqslant \alpha_n\alpha_m$.

(2) Show that $\lim n^{-1}\log\alpha_n = \inf m^{-1}\log\alpha_m$.

(3) Show that if if $\alpha_m < 1$ for some $m \geqslant 1$ then we can find constants $A, a > 0$ such that $\alpha_m \leqslant Ae^{-am}$.

Proof. (1) *I needed to resort to Durrett's textbook to use the extremely helpful hint regarding coupling. Otherwise I couldn't have solved this using pure algebraic manipulations.*

Following Durrett's hint, for $s \in S$ and $n \in \mathbb{N}$, let us define random variables X^n_s such that $\mathbb{P}(X^n_s = k) = p^n(s,k)$. The hint gives

$$\mathbb{P}(X^n_i \neq Y^n_j) = \frac{1}{2}\sum_k |p^n(i,k) - p^n(j,k)|.$$

In addition to the hint, we establish two more identities involving it:

$$\mathbb{P}(X_i^{m+n} = k) = p^{m+n}(i,k) = \sum_s p^m(i,s)p^n(s,k) = \sum_s \mathbb{P}(X_i^n = s)\mathbb{P}(X_s^n = k), \tag{1}$$

and

$$\mathbb{P}(X_i^n \neq Y_i^n) = \frac{1}{2}\sum_s |p^n(i,s) - p^n(i,s)| = 0. \tag{2}$$

We now have the sufficient ingrediens to cook up the desired inequality. For any $i_0, j_0 \in S$,

$$\begin{aligned}
\mathbb{P}(X_{i_0}^{m+n} \neq Y_{j_0}^{m+n}) &\overset{(1)}{=} \sum_{x,y\in S} \mathbb{P}(X_{i_0}^m = x, Y_{i_0}^m = y)\mathbb{P}(X_x^n \neq Y_y^n) \\
&= \sum_{x\neq y} \mathbb{P}(X_{i_0}^m = x, Y_{i_0}^m = y)\mathbb{P}(X_x^n \neq Y_y^n) + \sum_x \mathbb{P}(X_{i_0}^m = Y_{i_0}^m = x)\underbrace{\mathbb{P}(X_x^n \neq Y_x^n)}_{=0 \text{ by } (2)} \\
&= \sum_{x\neq y} \mathbb{P}(X_{i_0}^m = x, Y_{i_0}^m = y)\mathbb{P}(X_x^n \neq Y_y^n) \\
&= \mathbb{P}(X_x^n \neq Y_y^n)\cdot \sum_{x\neq y} \mathbb{P}(X_{i_0}^m = x, Y_{i_0}^m = y) \leqslant \alpha_n\alpha_m.
\end{aligned}$$

Taking supremum over all initial states $i_0, j_0 \in S$ we are done.

(2) Since $\alpha_{m+n} \leqslant \alpha_n\alpha_m$, taking log gives $\log\alpha_{m+n} \leqslant \log\alpha_n + \log\alpha_m$. Flipping the sign we have $-\log\alpha_{m+n} \geqslant (-\log\alpha_n) + (-\log\alpha_m)$, so Lemma 2.7.1 gives

$$-n^{-1}\log\alpha_n \to \sup_n m^{-1}\log\alpha_m \qquad \Longrightarrow \qquad n^{-1}\log\alpha_n \to \inf_n m^{-1}\log\alpha_m.$$

(3) The assumption implies that $\inf m^{-1}\log\alpha_m = \lim n^{-1}\log\alpha_n = -c < 0$. Hence there exists N sufficiently large so that $n^{-1}\log\alpha_n < -c/2 < 0$ for all $n > N$. For these terms, exponentiating both sides gives

$$\alpha_n^{1/n} < \exp(-c/2) \implies \alpha_n < \exp(n\cdot(-c/2)).$$

We claim that this $a = (c/2)$ is the exponent coefficient we are looking for. The first N terms are easy to control — for example, if we let $A > \max_{1\leqslant i\leqslant N}\alpha_i \cdot \exp(aN) + 1$[1]then

$$\alpha_m \leqslant \max_{1\leqslant i\leqslant N}\alpha_i = \max_{1\leqslant i\leqslant N}\alpha_i\cdot\exp(aN)\cdot\exp(-aN) < A\exp(-aN). \qquad \square$$

[1] An additional +1 at the end to ensure $A > 1$, so that $\alpha_n < \exp(-an) < A\exp(-an)$ for $n > N$ as well.