Definition of Martingale.

Def: A martingale is a discrete time stochastic process $\{X_n\}$, $n \geq 0$ that satisfies:

(1) $E[|X_n|] < \infty$ for all n

(2). $E[X_n | X_{n-1}, X_{n-2}, \cdots, X_0] = X_{n-1}$, for all $n \geq 2$ #

Example: (Random walk) Let $Z_1, Z_2, \cdots, Z_n$ be i.i.d. with $E[Z_i] = 0$

let $X_n = \sum_{i=1}^{n} Z_i$, then X_n is a martingale

$$E[X_n | X_{n-1}, \cdots, X_0] = E[Z_n + X_{n-1} | X_{n-1}, \cdots, X_0]$$

$$= E[Z_n | Z_{n-1}, \cdots, Z_0] + E[X_{n-1} | X_{n-1}, \cdots, X_0] = X_n$$ #

Question: • what if $\{Z_n\}$ is not iid? Can $\{X_n\}$ still be a martingale?

• Keep the iid assumption, let $E[Z_n] = 1$, is $X_n = \prod_{i=1}^{n} Z_i$ a martingale?

a particular example: $Z_i = 0$ w.p. $\frac{1}{2}$ and 2 w.p. $\frac{1}{2}$

in this case, $P(X_n = 2^n) = 2^{-n}$, however $E(X_n) = 1$ $\forall n$

$X_n \to 0$ a.s.

Fact: Let $\{X_n\}$ be a martingale. Then $\forall\, n > i \geq 0$, we have

$$E[X_n | X_i, X_{i-1}, \cdots, X_0] = X_i$$ #

Question: Proof? It implies $E[X_n] = E[X_0]$ $\forall n$ for a martingale.

Submartingales and Super-martingales

Def: A submartingale (supermartingale) is a discrete time stochastic process $\{X_n\}$, $n \geq 0$, that satisfies

(1) $E[|X_n|] < \infty$

(2) $E[X_n | X_{n-1}, \cdots, X_0] \geq X_{n-1}$

$(\leq)$ #

For a submartingale (supermartingale) $\{X_n\}$, $\forall n > i \geq 1$, we have

$$E[X_n \mid X_i, \cdots, X_0] \geq X_i \qquad (\leq X_i)$$

which implies

$$E[X_n] \geq X_i \qquad (\leq X_i)$$

Examples: Convex functions of martingales

(convex function: $f(x)$ is a convex function if $\forall x_1, x_2$, $\forall \alpha \in [0,1]$,
$f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$)

(Jensen's Inequality: $f(x)$ is a convex function, X is a r.v., then
$E[f(x)] \geq f(E[X])$)

- If $\{X_n\}$ is a martingale (or a submartingale), h is a convex function, and $E[|h(X_n)|] < \infty$, then $\{h(X_n)\}$ is a submartingale.

Proof: (Question: how? Use Jensen's Inequality.)

$\forall x_0, \cdots, x_{n-1}$, we have

$E[h(X_n) \mid X_{n-1} = x_{n-1}, \cdots, X_0 = x_0] \geq h(E[X_n \mid X_{n-1} = x_{n-1}, \cdots, X_0 = x_0])$

$= h(x_{n-1})$

$\forall h_0, \cdots, h_{n-1}$ in the range of h, let $x_0, \cdots, x_{n-1}$ be such that
$h(x_0) = h_0, \cdots, h(x_{n-1}) = h_{n-1}$

Then

$E[h(X_n) \mid h(X_{n-1}) = h_{n-1}, \cdots, h(X_0) = h_0]$

$\geq h(E[X_n \mid h(X_{n-1}) = h_{n-1}, \cdots, h(X_0) = h_0])$

$= h(x_{n-1}) = h_{n-1}$ #

- Similarly, if $\{X_n\}$ is a martingale (or a supermartingale), h is a concave function, and $E[|h(X_n)|] < \infty$, then $\{h(X_n)\}$ is a supermartingale.

The Martingale Convergence Theorem.

Theorem 1: Let $\{X_n\}$ be either a non-negative supermartingale, or a bounded submartingale, then $\lim_{n\to\infty} X_n$ exists and is finite almost surely. #

Martingale and the Recurrence of the Markov Chain

Let $\{X_n\}$ be a HMC on the countable state space E with transition matrix P.

Def: (Harmonic, Subharmonic, Superharmonic functions)
A function $h: E \mapsto \mathbb{R}$ is called harmonic (subharmonic, superharmonic) iff

$$\sum_{j\in E} P_{ij}\, h(j) = h(i) \qquad (\geq h(i),\ \leq h(i))$$

or in other words,

$$Ph = h \qquad (\geq h,\ \leq h)$$ #

Def: (Martingales with respect to a stochastic process)
A discrete time stochastic process $\{Y_n\}$, $n \geq 0$ is a martingale (submartingale, supermartingale) w.r.t to $\{X_n\}$, $n\geq 0$ if
(i) Y_n is a function of $X_0, \cdots, X_n$
(ii) $E[|Y_n|] < \infty$ (or $Y_n \geq 0$)
(iii) $E[Y_n \mid X_{n-1}, \cdots, X_0] = Y_{n-1}$ ($\geq Y_{n-1}$, $\leq Y_{n-1}$) #

Note: In the above definition, $\{X_n\}$ is not necessarily a Markov Chain.

Example: (Harmonic Functions Produce Martingales)
Let $\{X_n\}_{n\geq 0}$ be a HMC with state space E. If $h: E\mapsto \mathbb{R}$ is a harmonic (subharmonic, superharmonic) function, and $E[h(X_n)|] < \infty$ for all $n \geq 0$ (or h is non-negative), then $\{h(X_n)\}_{n\geq 0}$ is a martingale (submartingale, supermartingale) #

Question: Why?

Remark: The above mentioned properties for martingales still hold for martingales w.r.t. a stochastic process.

Theorem 2: An irreducible recurrent HMC has no nonnegative superharmonic or bounded subharmonic functions besides the constant functions. #

Proof: (Question: why? Hint: use the martingale convergence theorem.)
If h is nonnegative superharmonic (or bounded subharmonic), then $\{h(X_n)\}_{n\geq 0}$ is a nonnegative supermartingale (or bounded submartingale). By Theorem 1 it converges a.s. to a finite

limit Y. Since $\{X_n\}_{n\geq 0}$ visit any state $i \in E$ infinitely often, we must have $Y = h(i)$ a.s. for all $i \in E$ (Question: why?), in particular, h is a constant. #

Theorem 3. (A Transience Criterion)

A necessary and sufficient condition for an irreducible HMC to be transient is that there exists some state (that for convenience, we call it state 0) and a bounded function $h: E \to \mathbb{R}$, not identically null and satisfies

$$h(j) = \sum_{k \neq 0} P_{jk} h(k) \quad , \forall j \neq 0 \qquad (*)$$

#

Proof: Necessity. the HMC is transient $\Rightarrow$ there exists such a function (Question: find such a function?)

Let T_0 be the return time to state 0. Let

$$h(j) = P(T_0 = \infty \mid X_0 = j)$$

If $\{X_n\}_{n\geq 0}$ is transient, then $h(j)$ is nontrivial, and it satisfies $(*)$.

Sufficiency:

Suppose $(*)$ holds for a not identically null bounded function

$$\tilde{h}(j) = \begin{cases} h(j) & \text{if } j \neq 0 \\ 0 & \text{if } j = 0 \end{cases}$$

and let $\alpha = \sum_{k \in E} P_{0k} \tilde{h}(k)$. Changing signs if necessary, we can assume $\alpha \geq 0$. then $\tilde{h}$ is subharmonic.

If the chain is recurrent, then by Theorem 2, $\tilde{h}$ would be a constant, and it would be equal to $\tilde{h}(0) = 0$. This contradicts the assumption of nontriviality of h. #

Theorem 4. (A Recurrence Criterion)

Let the HMC with transition matrix P be irreducible, and suppose there exists a function $h: E \to \mathbb{R}$ such that the set (called level set) $\{i : h(i) \leq K\}$ is finite for all finite K, and such that

$$\sum_{k\in E} P_{ik}\, h(k) \le h(i), \text{ for all } i \notin F,$$

for some finite $F \subset E$. Then the chain is recurrent. #

Note: This condition is necessary (which we are not going to prove here) and sufficient.

Proof: (of sufficiency)

Since $\{i : h(i) < 0\}$ is finite, $\inf h(i) > -\infty$. Thus without loss of generality, we assume $h \ge 0$.

Let $\tau = \tau(F)$ be the return time to F, and define

$$Y_n = h(X_n)\, 1_{\{n<\tau\}}.$$

[We have shown last time, that $(X_0^n = X_0 \cdots X_n)$

$$E_i[Y_{n+1} \mid X_0^n] = E_i[Y_{n+1} 1_{\{n<\tau\}} \mid X_0^n] + E_i[Y_{n+1} 1_{\{n\ge\tau\}} \mid X_0^n]$$

$$= E_i[Y_{n+1} 1_{\{n<\tau\}} \mid X_0^n] \le E_i[h(X_{n+1}) 1_{\{n<\tau\}} \mid X_0^n]$$

$$= 1_{\{n<\tau\}} E_i[h(X_{n+1}) \mid X_0^n] = 1_{\{n<\tau\}} E_i[h(X_{n+1}) \mid X_n]$$

$$\le 1_{\{n<\tau\}} h(X_n) = Y_n$$]

Thus, $\forall\, i \notin F$, P_i almost surely (condition on $X_0 = i$)

$$E_i[Y_{n+1} \mid X_0^n] \le Y_n$$

Therefore, $\{Y_n\}_{n\ge 0}$, under P_i, is a nonnegative supermartingale w.r.t $\{X_n\}_{n\ge 0}$. By Theorem 1,

$$\lim_{n\to\infty} Y_n = Y_\infty$$

exists and is finite, P_i. a.s.

Suppose, the chain is transient. Then it must visit any finite subset of E only a finite times. In particular, $\forall K < \infty$, we have $h(X_n) < K$ only for a finite number of indices n. This implies that

$$\lim_{n\to\infty} h(X_n) = \infty, \quad P_i \text{ a.s. } (\forall j \in E)$$

But $\{Y_n\} = \{1_{\{n<\tau\}} h(X_n)\}$ has a P_i a.s. finite limit for $i \notin F$. So we must have $P_i(\tau<\infty) = 1$.

Hence, we have $P_i(\tau<\infty)=1$ for all $i \notin F$. Since F is finite, some states in F must be recurrent. This contradicts the assumption that the chain is transient. #

Example: A queueing system with service capacity 1, the arrival at time n is a iid R.V. A_n, with distribution

$$P(A_n = k) = a_k, \quad k \geq 0$$

Let the queue length at time n be X_n, and assume that the packet arrive at time n can not be served in the same slot, then we have

$$X_{n+1} = (X_n - 1)^+ + A_n$$

X_n is a Markov Chain, with state space $\{0, 1, \cdots\}$.
Use the above theorems to determine whether the chain is recurrent or transient when

(1) $E(A_0) > 1$ (2) $E(A_0) < 1$ (3) $E(A_0) = 1$

Solution. Note that for this chain,

$$P_{jk} = P(X_{n+1}=k \mid X_n=j) = P((j-1)^+ + A_n = k)$$

$$= P(A_n = k-(j-1)^+) = a_{k-(j-1)^+}$$

(1). Let state "0" be the state $X_n = 0$.

$\forall j \neq 0$, we need to find

$$h(j) = \sum_{k\geq 1} P_{jk}\, h(k)$$

Note that state j can only transit to the states

$j-1,\ j,\ j+1 \cdots$

Choose $h(j) = 1 - r^j$, then we need to check

$$1 - r^j = \begin{cases} \sum_{k \geq j-1} P(A_1 = k-(j-1)) \cdot (1 - r^k) \\ = 1 - \sum_{m \geq 0} P(A_1 = m) r^m \cdot r^{j-1} & (j \geq 2) \\ = 1 - \sum_{k \geq 1} P(A_1 = k) \cdot r^k & (j = 1) \end{cases}$$

$\Rightarrow \quad r = \sum_{m \geq 0} P(A_1 = m) r^m$ has a non-trivial solution

Let $g(r) = \sum_{m \geq 0} P(A_1 = m) r^m = E[r^A]$, $r \in [0,1]$

(the probability generating function of A_1, $g^{(k)}(0)/k! = P(A_1 = k)$)

Note that $g(0) = P(A_1 = 0) \geq 0$, $g(1) = 1$

$g'(r) = \sum_{m \geq 1} m P(A_1 = m) r^{m-1} \geq 0 \leftarrow g(r)$ increasing

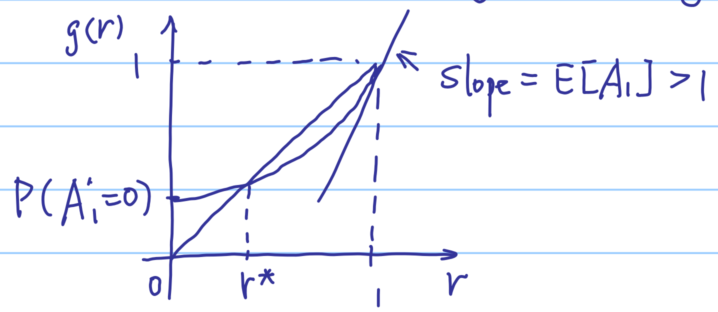
$(= 0$ if $P(A_1 = m) = 0 \ \forall m \geq 1$, i.e. $P(A_1 = 0) = 1)$

$g'(0) = P(A_1 = 1)$, $g'(1) = E[A] > 1.$

$g''(r) = \sum_{m \geq 2} m(m-1) P(A_1 = m) r^{m-2} \geq 0 \leftarrow g(r)$ concave.

$(= 0$ if $P(A_1 = m) = 0 \ \forall m \geq 2$, i.e. $P(A_1 = 0) + P(A_1 = 1) = 1)$

Since $E[A_1] > 1$, then $g'(r) > 0$, $g''(r) > 0$.



Thus $g(r) = r$ has a solution $r^* \in (0,1)$.

Then $h(j) = 1 - (r^*)^j$ is a function that is not a constant. #

(2). Use a corollary of Foster's theorem.

Corollary: (Pakes's Lemma).

Let $\{X_n\}_{n\geq 0}$ be an irreducible HMC on $E=\mathbb{N}$ s.t. $\forall n\geq 0$ and all $i\in E$, we have

$$E[X_{n+1}\mid X_n=i]<\infty$$
$$\limsup_{n\to\infty} E[X_{n+1}-X_n\mid X_n=i]<0$$

Then such a HMC is positive recurrent.

Proof of the corollary:

Let $\limsup\limits_{i\to\infty} E[X_{n+1}-X_n\mid X_n=i]=-2\varepsilon$, $\varepsilon>0$.

Then $\exists\, i_0$ sufficiently large s.t.

$$E[X_{n+1}-X_n\mid X_n=i]<-\varepsilon$$

Then choose $h(i)=i$ and $F=\{i: i<i_0\}$, we have the conditions of the Foster's theorem. #

Then we use this result to show $E[A_1]<1 \Rightarrow$ positive recurrence.

$$E[X_{n+1}-X_n\mid X_n=i]$$
$$=E[X_{n+1}-i\mid X_n=i]$$
$$=E[(i-1)^+ + A_n - i]$$
$$=E[A_1]-1\{i\geq 1\}$$

Thus $\limsup\limits_{i\to\infty} E[X_{n+1}-X_n\mid X_n=i]=E[A_1]-1<0$

Thus the chain is positive recurrent.

Above, we show that $E[A_1]<1$ is a sufficient condition for the chain to be positive recurrent.

Actually, we can also show that $E[A_1]<1$ is also necessary * for the chain to be positive recurrent. (How?)

Proof of necessity:

Assume the chain is positive recurrent, and has a stationary distribution of π_i, $i=0,1,\cdots$

$\forall r$, $|r|\leq 1$,

$$r^{X_{n+1}+1}=\left(r^{(X_n-1)^+ +1}\right)r^{A_n}$$
$$=\left(r^{X_n}1\{X_n>0\}+r\,1\{X_n=0\}\right)r^{A_n}$$
$$=\left(r^{X_n}-1\{X_n=0\}+r\,1\{X_n=0\}\right)r^{A_n}$$

Therefore, $r\cdot r^{X_{n+1}}-r^{X_n}r^{A_n}=(r-1)1\{X_n=0\}r^{A_n}$

Taking expectation on both sides, note that X_n and A_n are independent,
we have $rE[r^{X_{n+1}}] - g_A(r)E[r^{X_n}] = (r-1)\pi_0 g_A(r)$

In steady state, $E[r^{X_{n+1}}] = E[r^{X_n}] = g_X(r)$, thus

$$g_X(r)(r - g_A(r)) = \pi_0 (r-1) g_A(r) \quad (*)$$

Thus if we can get π_0, we can obtain the generating function $g_X(r)$

Take derivative on both sides, we have

$$g_X'(r)(r-g_A(r)) + g_X(r)(1-g_A'(r)) = \pi_0 g_A(r) + \pi_0(r-1)g_A'(r)$$

Note that $g_X(1) = g_A(1) = 1$, $g_A'(1) = E[A_1]$, we have

$$g_X'(1)(1-1) + 1\cdot(1-E[A_1]) = \pi_0 + \pi_0(1-1)E[A_1]$$

$$\Rightarrow \pi_0 = 1 - E[A_1]$$

Since $\pi_0 \geq 0$, we have $E[A_1] \leq 1$.

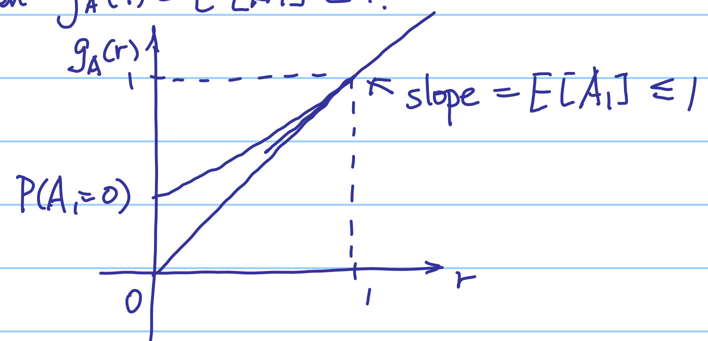
It remains to show that $E[A_1] \neq 1$.

Case 1, $A_1 \equiv 1$, in this case, $X_n = 1 \ \forall n \geq 1$ $(X_0 = 0)$, the chain is not positive recurrent (since it does not go back to state 0.

Case 2, $A_1 \not\equiv 1$, $E[A_1] = 1$, thus $\pi_0 = 1 - E[A_1] = 0$.

Then $(*)$ reduces to $g_X(r)(r - g_A(r)) = 0$, $\forall r \in [0,1]$.

However, $r - g_A(r) = 0$ has only $r = 1$ for a solution when $g_A'(1) = E[A_1] \leq 1$.



Therefore $g_X(r) \equiv 0 \ \forall r \in [0,1)$, thus $g_X(r) \equiv 0$ $\forall r \in [0,1]$ (Since its radius of convergence > 1). This leads to a contradiction, since $g_X(1) = 1$.

Thus, $E[A_1] < 1$ is a also a necessary condition for the chain to be positive recurrent. #

(3). Check Theorem holds when $E[A_1] = 1$ with

$$h(i) = i, \quad F = \{0\}$$

Therefore, the chain is recurrent. Since it is not positive recurrent (by (2.)), it is null recurrent. #

Theorem 5. (A sufficient condition of transience)

Let HMC $\{X_n\}_{n\geq 0}$ with transition matrix P be irreducible and let $h: E \to \mathbb{R}$ be a bounded function such that

$$\sum_{k\in E} P_{ik}\, h(k) \leq h(i) \quad \text{for all } i \notin F,$$

for some set $F \subset E$, not assumed finite. Suppose, moreover, that there exists $i \notin F$ such that

$$h(i) < h(j) \quad \forall j \in F \tag{**}$$

Then the chain is transient.

Proof: Let τ be the return time to F and let $i \notin F$ satisfy $(**)$. Defining $Y_n = h(X_{n\wedge\tau})$, we have

$$E_i[Y_{n+1} \mid X_0^n] = \underbrace{E_i[1_{\{n<\tau\}} h(X_{n+1}) \mid X_0^n]}_{①} + \underbrace{E_i[1_{\{n\geq\tau\}} h(X_\tau) \mid X_0^n]}_{②}$$

$$① \leq 1_{\{n<\tau\}} h(X_n) = 1_{\{n<\tau\}} Y_n$$ (same reasoning as in the proof of Foster's thm)

$$② = 1_{\{n\geq\tau\}} h(X_\tau) = 1_{\{n\geq\tau\}} Y_\tau$$ (Since $1_{\{n\geq\tau\}} h(X_\tau)$ is a function of X_0^n)

Thus $E_i[Y_{n+1} \mid X_0^n] \leq Y_n$, Y_n is a bounded supermartingale w.r.t $\{X_n\}_{n\geq 0}$. Then by the Martingale Convergence Theorem, we have

$$Y_n \to Y \quad P_i.\ a.s.$$

Thus $\quad E_i[Y] = \lim_{n\to\infty} E_i[Y_n]$.

Since $E_i[Y_n] \leq E_i[Y_0] = h(i)$ (by supermartingale property), we have $E_i[Y] \leq h(i)$.

If τ is P_i. a.s finite, then $n \geq \tau$ will eventually happen for a large enough n. Then after that n, $Y_n = h(X_\tau)$, where $X_\tau \in F$. Thus in this case, $E_i[Y] = h(j)$ for some $j \in F$, and $h(j) > h(i)$, contradicting the last inequality.

Therefore, $P_i(\tau < \infty) < 1$, which means the chain is transient. #

Additional Examples of Martingale

1. A gambling strategy: fair game. reward = 2× bet.

A gambler bets \$1 on the 1st play, and double his bet in the next play when he loses. Eventually, he will win, say, on T-th play, giving him a profit of $2\cdot 2^{T-1} - (1+2+4+\cdots+2^{T-1}) = 1$ dollar. It looks like that if he restart this process whenever he wins (i.e., bet \$1 in the next play when he wins) he will eventually end up with any desired profit.

Is this strategy really as good as it sounds?

Let Y_n be the "profit" the gambler makes after n-th play, we have

$$Y_{n+1} = \begin{cases} Y_n - 2^n & \text{w.p. } \frac{1}{2} \\ Y_n + 2^n & \text{w.p. } \frac{1}{2} \end{cases}$$

$\Rightarrow E(Y_{n+1} \mid Y_1, \cdots, Y_n) = Y_n$, thus Y_n is a martingale, which implies $E(Y_n) = 0 \ \forall n$

Let T be the time of his 1st win, $P(T=n) = 2^{-n} \ \forall n \geq 1$.

$$E(\text{money lost before 1st win}) = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n (1+\cdots+2^{n-2}) = \infty.$$

This does not sound good for a gambler starts with finite budget.

2. De Moivre's Martingale.

A gambler plays a game with $P(\text{win}) = p$. When he wins, he wins \$1, otherwise he lose \$1. The gambler starts with k dollars, and he will leave the game when he bankrupts or until he has N dollars.

Question:

This is a RW, with $S_0 = k$, $X_n = \begin{cases} 1 & \text{w.p. } p \\ -1 & \text{w.p. } 1-p \end{cases}$

Define $Y_n = (q/p)^{S_n}$, then Y_n is a martingale w.r.t. X_n (check)

Let T be the first time S_n hits 0 or N, then T is a stopping time.

$$E(Y_T) = E(Y_0) = (q/p)^k \quad \text{(the optional sampling theorem)}$$

$$E(Y_T) = (q/p)^0 p_k + (q/p)^N (1-p_k) = (q/p)^k$$

Thus $p_k = \dfrac{\rho^k - \rho^N}{1-\rho^N}$, where $\rho = q/p$.