5. Martingale Processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathbb{F} = \{\mathcal{F}_n; n \geq 0\}$ a filtration on $(\Omega, \mathcal{F})$. Let $\{M_n; n \geq 1\}$ be a stochastic process adapted to $\mathbb{F}$ (note: we can consider $\mathcal{F}_n = \sigma(M_1, \dots, M_n)$ so that it is automatically adapted).

Definition 1: If $\{M_n; n \geq 1\}$ is adapted to $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, and $E|M_n| < \infty$ for all n, then we say that:

- $\{M_n\}$ is a $\mathbb{F}$-martingale if $M_n = E(M_{n+1} | \mathcal{F}_n)$ w.p.1 for all $n \geq 1$,
- $\{M_n\}$ is a $\mathbb{F}$-submartingale if $M_n \leq E(M_{n+1} | \mathcal{F}_n)$ w.p.1 for all $n \geq 1$,
- $\{M_n\}$ is a $\mathbb{F}$-supermartingale if $M_n \geq E(M_{n+1} | \mathcal{F}_n)$ w.p.1 for all $n \geq 1$.

Example 11: The "gambler" problem is usually stated in terms of his fortune M_n at the time of the n-th game. If the results of consecutive games are iid (he always gambles \$1) the resulting process is a martingale when the game is fair:

$$M_n = M_{n-1} + X_m, \quad E(X_{n+1} | \mathcal{F}_n) = 0$$

$\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$.

Definition 2: The process $\{\delta M_n\}$, $\delta M_n = M_{n+1} - M_n$, called the martingale difference process. If $E|M_n|^2 < \alpha$ then the martingale diferences are uncorrelated:

$$E[\delta M_n \, \delta M_m] = 0 \quad \forall\, m \neq n$$

Example 12: If Y is a $\mathcal{F}$-mbl random variable with $E|Y|$ then the random variables

$$M_n = E(Y | \mathcal{F}_n)$$

form a martingale, since $E(M_{n+1} | \mathcal{F}_m) = E(E(Y | \mathcal{F}_m) = E(Y | \mathcal{F}_n) = M_n$. This is called Doob's martingal and it corresponds to a model where information about Y "accumulated" as the numberof observations grows.

Example 13: Let $\{Y_n\}$ be a sequence of random varic on $(\Omega, \mathcal{F}, \mathbb{P})$ and form the natural filtration $\mathbb{F}$ with $\mathcal{F}_n$ $\sigma(Y_1, \dots, Y_n)$. Write:

$$\begin{aligned} Y_n &= (Y_n - E(Y_n | \mathcal{F}_{n-1})) + E(Y_n | \mathcal{F}_{n-1}) \\ &= \delta M_n + E(Y_n | \mathcal{F}_{n-1}) \end{aligned}$$

The "unpredictable" term in Y_n is a martingale difference, where:

$$M_n = \sum_{j=0}^{n} \left(Y_j - E(Y_j \mid \mathcal{F}_{j-1})\right)$$

is a martingale. Vector-valued martingales are stochastic processes where each component is a martingale.

Proposition 1: Let $\{M_n\}$ be a $\mathbb{F}$-martingale and $q(\cdot)$ a non-decreasing non-negative convex function. Then $\forall n<N$ and $\lambda>0$,

$$\mathbb{P}\left\{ \sup_{n\le m\le N} |M_n| \ge \lambda \mid \mathcal{F}_n \right\} \le \frac{E(q(M_N \mid \mathcal{F}_n))}{q(\lambda)}$$

Commonly, $q(x) = |x|$, x^2 or $e^{\alpha x}$, $\alpha>0$.

Proposition 2: Let $\{M_n\}$ be a $\mathbb{F}$-martingale. Then

$$E\left\{ \sup_{n\le m\le N} |M_m|^2 \mid \mathcal{F}_n \right\} \le 4\,E\left(|M_N|^2 \mid \mathcal{F}_n\right)$$

$$\mathbb{P}\left\{ \sup_{n\le m\le N} M_m \ge \lambda \mid \mathcal{F}_n \right\} \le \frac{M_n}{\lambda}$$

Proposition 3: If $\{M_n\}$ is a martingale and τ is a stopping time, then $E\, M_n \mathbb{1}_{(\tau\le n)} = E\, X_\tau \mathbb{1}_{(\tau\le n)}$ $\forall n\ge 1$.

Definition 3: Let $\{M_n\}$ be a $\mathbb{F}$-martingale on $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\mathcal{F}_n; n\ge 1\}$ and τ an adapted stopping time. The stopped martingale $\{M_{\tau\wedge n}\}$ is defined by:

$$M_{\tau\wedge n} = \begin{cases} M_n & \text{if } n<\tau \\ M_\tau & \text{if } n\le\tau \end{cases}$$

and $\tau\wedge n = \min(n,\tau)$.

Proposition 4: If τ is an $\mathbb{F}$-adapted stopping time and it is a.s. bounded (uniformly in ϖ), then $\{M_{n\wedge\tau}\}$ is a martingale (sub- or supermartingale) if $\{M_n\}$ is a martingale (respectively, a sub- or supermartingale).

Theorem 1: Let $\{M_n\}$ be a real-valued ~~submartingale~~ supermartingale with $\sup_n E|M_n|<\infty$, on a space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Then $\{M_n\}$ converges w.p.1, as $n\to\infty$. If $\{M_n\}$ is a supermartingale, $\{M_n\}$ converges w.p.1 as $n\to\infty$ if $E[M_n^-]<\infty$, where $x^- = \max(0,-x)$.

Definition 4: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $\mathbb{F} = \{\mathcal{F}_t \,;\, t\in\mathbb{R}^+\}$. A process $\{M(t); t\ge 0\}$ adapted to the filtration $\mathbb{F}$ is a:

- martingale if $M(s) = E(M(s+t) \mid \mathcal{F}_s)$ $\forall t\ge 0$
- submartingale if $M(s) \le E[M(s+t) \mid \mathcal{F}_s]$ $\forall t\ge 0$
- supermartingale if $M(s) \ge E[M(s+t) \mid \mathcal{F}_s]$ $\forall t\ge 0$

with probability 1.

Definition: Let $X(t)$ be a continuous time stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ with natural filtration $\mathbb{F} = \{\mathcal{F}_t ; t \geq 0\}$. We say that X is locally Lipschitz continuous w.p.1 if for each $T>0$ there is a random variable $K(T)>0$ $\mathcal{F}_T$-mbl, such that

$$|X(t+s) - X(t)| \leq K(T)\, s, \text{ for all } t \leq t+s \leq T.$$

Theorem 2: A continuous time martingale $\{M(t), t \geq 0\}$ whose paths are locally Lipschitz continuous w.p.1 on each bounded interval is a constant w.p.1.

Characterization of a Martingale

In many of the proofs of convergence of stochastic approximations one uses Theorem 2 to establish continuity of some (limit) processes. In order to do this, it is first necessary to verify that the process in question is a martingale.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and a "state" process (or "control" process) $\{\vartheta(t) ; t \geq 0\}$ adapted to the filtration.

Theorem 3: Let $\{Y(t) ; t \geq 0\}$ be a stochastic proce adapted to the filtration $\mathbb{F}$, $\sigma(Y(s); s \leq t) \subset \mathcal{F}_t$ $\forall$ Suppose that for each $t, r \geq 0$ and for each $p \in \mathbb{N}$ $s_i \leq t$ and any bounded and continuous real-valued f $h(\cdot)$:

$$E\, h(\vartheta(s_i); i \leq p)[Y(t+r) - Y(t)] = 0$$

then $\{Y(t)\}$ is a $\mathbb{F}$-martingale.

Proof: To give the idea of the proof, let Y be a ranc variable on $(\Omega, \mathcal{F})$ with $E|Y| < \infty$ and suppose that

$$E\, h(\vartheta(s_i); i \leq p)\, Y = 0$$

for all p; $s_i \leq t$ and any continuous and bounded h

Since

$$0 = E\, h(\vartheta(s_i); i \leq p)\, Y = E\{h(\vartheta(s_i); i \leq p) \times E(Y \mid \vartheta(s_i); i \leq p)\}$$

and we can use $h \equiv 1$, then $E(Y \mid \vartheta(s_i); i \leq p) = 0$, but from the arbitrariness of p and $s_i \leq t$, this in turn implies $E(Y \mid \vartheta(s); s \leq t) = 0$. Using $Y(t+\tau) - Y(t)$ instead of Y for each t, τ, yields the desire result.