Igor Kortchemski
Probality theory,
Autumn 2023
ETH

Chapter 4: Part 2, Martingales and their a.s. convergence

Outline:

1) Definitions and first properties
2) The (sub/super) martingale a.s. convergence theorem
3) Example: the Bienaymé – Galton – Watson branching process

1) Definitions and first properties

We work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition A filtration $(\mathcal{F}_n)_{n\geq 0}$ is an increasing sequence of σ-fields in $\mathcal{F}$: $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}$

Interpretation: n is time and $\mathcal{F}_n$ represents the information accessible at time n

Definition Let $(M_n)_{n\geq 0}$ be a sequence of real-valued r.v. such that $M_n \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$ (we say "(M_n) is adapted and integrable"). It is called:

- a $(\mathcal{F}_n)$ martingale if $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n$ for every $n \geq 0$
- a $(\mathcal{F}_n)$ supermartingale if $\mathbb{E}[M_{n+1} | \mathcal{F}_n] \leq M_n$ for every $n \geq 0$
- a $(\mathcal{F}_n)$ submartingale if $\mathbb{E}[M_{n+1} | \mathcal{F}_n] \geq M_n$ for every $n \geq 0$

interpretation: imagine a player betting at a casino; M_n corresponds to her wealth at time n and $\mathcal{F}_n$ the information the player has at time n to place a bet and "win" an amount $M_{n+1} - M_n$

- (M_n) martingale corresponds to a "fair" game
- (M_n) supermartingale corresponds to a "defavorable" game ("supermartingales tend to decrease")
- (M_n) submartingale corresponds to a "favorable" game ("submartingales tend to increase")

Remarks

- The definition is always with respect to some filtration: it is a property of $((M_n),(\mathcal{F}_n))$ and not (M_n) alone. However, if (M_n) is a $(\mathcal{F}_n)$ martingale, set $\mathcal{A}_n = \sigma(M_0,\dots,M_n)$ called the canonical filtration of (M_n), then (M_n) is a $(\mathcal{A}_n)$-martingale. Indeed, $M_n \in L^1(\Omega,\mathcal{A}_n,\mathbb{P})$ and $\mathbb{E}[M_{n+1}\mid\mathcal{A}_n] = \mathbb{E}[\,\mathbb{E}[M_{n+1}\mid\mathcal{F}_n]\mid\mathcal{A}_n] = \mathbb{E}[M_n\mid\mathcal{A}_n] = M_n$ where we have used $\mathcal{A}_n\subset\mathcal{F}_n$ (because M_n is $\mathcal{F}_n$ mesurable) and the tower property.
- If (M_n) is a $(\mathcal{F}_n)$-martingale, then $\mathbb{E}[M_n\mid\mathcal{F}_m] = M_m$ for $0\le m\le n$.
  This follows by induction: for $n=m$ it's clear, for $n=m+1$ it's the definition, if $\mathbb{E}[M_n\mid\mathcal{F}_m]=M_m$, then $\mathbb{E}[M_{n+1}\mid\mathcal{F}_m] = \mathbb{E}[\,\mathbb{E}[M_{n+1}\mid\mathcal{F}_n]\mid\mathcal{F}_m]$ since $\mathcal{F}_m\subset\mathcal{F}_n$
  which is $\mathbb{E}[M_n\mid\mathcal{F}_m] = M_m$ (tower property)
  By taking expectations, we get $\mathbb{E}[M_n]=\mathbb{E}[M_m]$ so $\mathbb{E}[M_n]=\mathbb{E}[M_0]$ $\forall n\geq 1$
  Similarly, for a submartingale, $\mathbb{E}[M_n\mid\mathcal{F}_m]\geq M_m$ and $(\mathbb{E}[M_n])_{n\geq 0}$ is increasing
  and for a supermartingale, $\mathbb{E}[M_n\mid\mathcal{F}_m]\leq M_m$ and $(\mathbb{E}[M_n])_{n\geq 0}$ is decreasing.
- (M_n) is a $(\mathcal{F}_n)$ supermartingale iff $(-M_n)$ is a $(\mathcal{F}_n)$-submartingale. For this reason results are often written using either supermartingales, or submartingales.

Examples

① If $M\in L^1(\Omega,\mathcal{F},\mathbb{P})$, $M_n = \mathbb{E}[M\mid\mathcal{F}_n]$ is a $(\mathcal{F}_n)$ martingale, called closed martingale. Indeed:
- M_n is $\mathcal{F}_n$ measurable and $\mathbb{E}[\,|M_n|\,] = \mathbb{E}[\,|\mathbb{E}[M\mid\mathcal{F}_n]|\,] \leq \mathbb{E}[\,\mathbb{E}[\,|M|\mid\mathcal{F}_n]] = \mathbb{E}[\,|M|\,]<\infty$
- $\mathbb{E}[M_{n+1}\mid\mathcal{F}_n] = \mathbb{E}[\,\mathbb{E}[M\mid\mathcal{F}_{n+1}]\mid\mathcal{F}_n] = \mathbb{E}[M\mid\mathcal{F}_n]$ (tower property)
  $= M_n$

② Random walk in $\mathbb{R}$: Fix $x\in\mathbb{R}$ and let $(X_n)_{n\geq 1}$ be iid integrable random variables. Set $M_0 = x$, $\mathcal{F}_0=\{\emptyset,\Omega\}$ and $M_n = x+X_1+\dots+X_n$, $\mathcal{F}_n=\sigma(X_1,\dots,X_n)$ for $n\geq 1$.
Then
- $M_n\in L^1(\Omega,\mathcal{F}_n,\mathbb{P})$
- $\mathbb{E}[M_{n+1}\mid\mathcal{F}_n] = x+\dots+X_n+\mathbb{E}[X_{n+1}\mid\mathcal{F}_n] = M_n+\mathbb{E}[X_1]$

Thus (M_n) is a $(\mathcal{F}_n)$ martingale if $\mathbb{E}[X_1]=0$
$(\mathcal{F}_n)$ submartingale if $\mathbb{E}[X_1]\geq 0$
$(\mathcal{F}_n)$ supermartingale if $\mathbb{E}[X_1]\leq 0$

③ If $M_n \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$ and $M_{n+1} \leq M_n$ for every $n\geq 0$, then (M_n) is a $(\mathcal{F}_n)$ supermartingale: $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \leq \mathbb{E}[M_n \mid \mathcal{F}_n] = M_n$.

We now perform some operations:

Proposition Assume that $M_n \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$ and let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be a convex function such that $\mathbb{E}[|\varphi(M_n)|] < \infty$ $\forall n \geq 0$.

ⓐ If (M_n) is a $(\mathcal{F}_n)$ martingale, then $(\varphi(M_n))$ is a $(\mathcal{F}_n)$ submartingale

ⓑ If (M_n) is a $(\mathcal{F}_n)$ submartingale and φ is weakly increasing, then $(\varphi(M_n))$ is a $(\mathcal{F}_n)$ submartingale

Proof: First, in both cases $\varphi(M_n) \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$.

ⓐ By Jensen's inequality $\mathbb{E}[\varphi(M_{n+1}) \mid \mathcal{F}_n] \geq \varphi(\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]) = \varphi(M_n)$

ⓑ Similarly, $\mathbb{E}[\varphi(M_{n+1}) \mid \mathcal{F}_n] \geq \varphi(\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]) \geq \varphi(M_n)$

$\sim$

Useful corollary: If (M_n) is a martingale:

- $(|M_n|)$ is a sub-martingale
- (M_n^+) is a sub-martingale, where $M_n^+ = \max(0, M_n)$
- If $\forall n \geq 1$, $\mathbb{E}[M_n^2] < \infty$, (M_n^2) is a sub martingale

If (M_n) is a submartingale, (M_n^+) is a submartingale

End of lecture 17

Proposition (discrete stochastic calculus) A sequence $(H_n)_{n\geq 1}$ of real-valued random variables is called predictable if $\forall n\geq 1$ H_n is bounded and $\mathcal{F}_{n-1}$ measurable.

For a sequence $(M_n)_{n\geq 0}$ we define $(H\cdot M)_n = \sum_{k=1}^{n} H_k (M_k - M_{k-1})$

(1) If (M_n) is a martingale, then $((H\cdot M)_n)$ is a martingale

(2) If $\forall n\geq 1$, $H_n \geq 0$ and (M_n) is a sub/supermartingale, then $((H\cdot M)_n)$ is a sub/super martingale

Interpretation: if M_n represents the wealth of a player at time n, $M_{n+1} - M_n$ represents the amount "won" at time n. $H_{n+1}(M_{n+1} - M_n)$ the amount won if the player had multiplied by H_{n+1} the bet at time n (H_{n+1} has to be $\mathcal{F}_n$ measurable: the player bets knowing information at time n)

Proof (1) We have $(H\cdot M)_n \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$

We check that $\mathbb{E}[(H\cdot M)_{n+1} - (H\cdot M)_n \mid \mathcal{F}_n] = 0$ (indeed, $\mathbb{E}[(H\cdot M)_n \mid \mathcal{F}_n] = (H\cdot M)_n$)

We have $(H\cdot M)_{n+1} - (H\cdot M)_n = H_{n+1}(M_{n+1} - M_n)$, so

$\mathbb{E}[(H\cdot M)_{n+1} - (H\cdot M)_n \mid \mathcal{F}_n] = H_{n+1}\, \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n]$ since H_{n+1} is $\mathcal{F}_n$ measurable

$= 0.$

(2) Analog proof $\sim$

2) The (sub/super) martingale a.s. convergence theorem

Recall that a family $(X_i)_{i\in I}$ of real-valued random variables is bounded in L^1 if $\sup_{i\in I} \mathbb{E}[|X_i|] < \infty$

Theorem Let (M_n) be a submartingale / supermartingale or martingale bounded in L^1. Then (M_n) converges almost surely to some real-valued random variable M_∞ with $\mathbb{E}[|M_\infty|] < \infty$.

Corollary A non-negative supermartingale or martingale converges a.s

(This comes from the fact that then $\mathbb{E}[|M_n|] = \mathbb{E}[M_n]$ which is $\searrow$ for a supermartingale (and a martingale is a supermartingale).

It is enough to show the theorem for a supermartingale (M_n) (then apply it to $(-M_n)$ when M_n is a a submartingale.

The key idea is to introduce the notion of upcrossing.

First, for $a<b$ define $S_1 = \inf\{n\geq 0 : M_n \leq a\}$, $T_1 = \inf\{n \geq S_1 : M_n \geq b\}$ and by induction

$S_{k+1} = \inf\{n \geq T_k : M_n \leq a\}$, $T_{k+1} = \inf\{n \geq S_{k+1} : M_n \geq b\}$ with the convention $\inf \phi = \infty$.

Then, for $n \geq 1$ we define $N_n([a,b]) = \sum_{k=1}^{\infty} \mathbb{1}_{\{T_k \leq n\}}$ and $N_\infty([a,b]) = \sum_{k=1}^{\infty} \mathbb{1}_{\{T_k < \infty\}}$, which represents the number of "up-crossings" of $[a,b]$ by (M_n)

Lemma $(M_n)_{n\geq 1}$ converges in $[-\infty,\infty]$ iff $\forall a<b$ $a,b\in\mathbb{Q}$ $N_\infty([a,b])<\infty$

This comes from the fact that if (M_n) has several cluster points, it will upcross infinitely many times some interval.

The connection with supermartingales is the following

Lemma (Doob upcrossing lemma) Let (M_n) be a supermartingale. Then for every $a<b$ and $n\geq 1$,

$$\mathbb{E}[N_n([a,b])] \leq \frac{1}{b-a}\mathbb{E}[(a-M_n)^+] = \frac{1}{b-a}\mathbb{E}[(a-M_n)\mathbb{1}_{a-M_n \geq 0}]$$

Proof of the theorem using the lemmas Let (M_n) be a supermartingale with $\sup_{n\geq 1}\mathbb{E}[|M_n|] = K < \infty$.

It is enough to show that $\forall a<b$, $a,b\in\mathbb{Q}$, a.s $N_\infty([a,b])<\infty$.

We have $\mathbb{E}[N_n([a,b])] \leq \frac{1}{b-a}(a+K)$.

By monotone convergence $\mathbb{E}[N_\infty([a,b])] = \lim_{n\to\infty}\mathbb{E}[N_n([a,b])] \leq \frac{1}{b-a}(a+K) < \infty$

Thus a.s. $N_\infty([a,b])<\infty$

This shows that a.s. $M_n \longrightarrow M_\infty$ with M_∞ having values in $[-\infty,\infty]$.

But by Fatou's lemma, $\mathbb{E}[|M_\infty|] = \mathbb{E}[\liminf_{n\to\infty}|M_n|] \leq \liminf_{n\to\infty}\mathbb{E}[|M_n|] \leq K < \infty$

Thus a.s. $M_\infty < \infty$

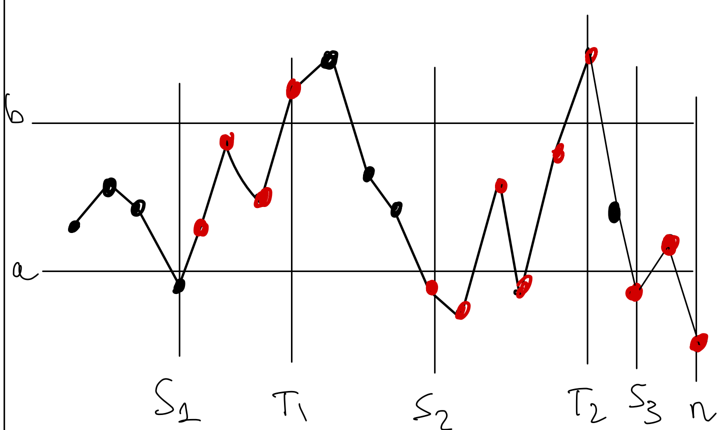
Proof of the upcrossing lemma

First observe that for every $k, n \geq 1$, $\{T_k \leq n\} \in \mathcal{F}_n$ and $\{S_k \leq n\} \in \mathcal{F}_n$.

Indeed, for example $\{T_k \leq n\} = \bigcup_{0 \leq m_1 < n_1 < \dots < m_k \leq n_k \leq n} \{M_{m_1} \leq a, X_{n_1} \geq b, \dots, X_{m_k} \leq a, X_{n_k} \geq b\}$

The idea is to define $H_n = \sum_{k=1}^{\infty} \mathbb{1}_{\{S_k < n \leq T_k\}} = \mathbb{1}_{\{M \text{ is in the process of doing an upcrossing at time } n\}}$, which is predictable since $\{S_k < n \leq T_k\} = \{S_k \leq n-1\} \setminus \{T_k \leq n-1\} \in \mathcal{F}_{n-1}$, and consider $(H \cdot M)_n$

Interpretation



Imagine that (M_n) is the price of an asset on the stock market. As soon as it is low $(\leq a)$ by a unit, so that your fortune then evolves as (M_n). Sell as soon as the price is $\geq b$, and repeat the strategy. When you sell, you win at least $b-a$.

$(H \cdot M)_n$ is your fortune at time n.

In red points such that $H_n = 1$ (the points are (n, M_n)). In this example $N_n([a,b]) = 2$.

Write $(H \cdot M)_n = \sum_{\ell=1}^{n} H_\ell (M_\ell - M_{\ell-1}) = \sum_{\ell=1}^{n} \sum_{k=1}^{\infty} \mathbb{1}_{\{S_k < \ell \leq T_k\}} (M_\ell - M_{\ell-1})$

$$= \sum_{k=1}^{\infty} \sum_{\ell=S_k+1}^{\min(T_k, n)} (M_\ell - M_{\ell-1}) = \sum_{k=1}^{N_n([a,b])} (M_{T_k} - M_{S_k}) + \mathbb{1}_{S_{N_n([a,b])+1} \leq n} \Big(M_n - \underbrace{M_{S_{N_n([a,b])+1}}}_{\leq a}\Big)$$

$$\geq (b-a) N_n([a,b]) - (a - M_n)^+$$

(the term $(a-M_n)^+$ comes from forgetting the temporary positive gain, but not the loss, during the last incomplete upcrossing)

Since (M_n) is a supermartingale, so is $(H \cdot M)_n$ thus

$\mathbb{E}[(H \cdot M)_n] \leq \mathbb{E}[(H \cdot M)_0] = 0$.

thus $\mathbb{E}[N_n([a,b])] \leq \dfrac{\mathbb{E}[(a-M_n)^+]}{b-a}$

∞

Remark A (sub/super)martingale (M_n) bounded in L^p $(p>1)$ is also bounded in L^1 and thus converges a.s. Indeed, since $x \mapsto |x|^p$ is convex, $\mathbb{E}[|X|]^p \leq \mathbb{E}[|X|^p]$, so $\mathbb{E}[|X|] \leq \mathbb{E}[|X|^p]^{1/p}$ so $\sup_{n\geq 1} \mathbb{E}[|M_n|^p] < \infty$ implies $\sup_{n \geq 1} \mathbb{E}[|M_n|] < \infty$.

We have also seen that (M_n) bounded in L^p implies uniform integrability, so (M_n) also converges in L^1. We will study in more detail UI martingales later in chapter 5.

3) Example: the Bienaymé-Galton-Watson branching process

Goal: introduce a simple model for the evolution of a population

Let μ be a probability distribution on $\mathbb{N} = \{0, 1, 2, \dots\}$ and let $(K_{n,j})_{n\geq 0, j \geq 1}$ be a family of iid random variables with law μ. Define by induction $X_0 = 1$ and for $n \geq 0$:

$$X_{n+1}(\omega) = \sum_{j=1}^{X_n(\omega)} K_{n,j}(\omega)$$

(if $X_n = 0$, then $X_{n+1} = 0$)

Interpretation: X_n represents the number of individuals at generation n, which have iid number of children with law μ

Studied by Bienaymé and Galton & Watson in the 19th century (motivation: extinction of noble names)

What is the behavior of X_n as $n \to \infty$? **End of lecture 18**

To avoid degenerate cases we assume $\mu(0) \neq 1$, $\mu(1) \neq 1$. Our main assumption is $R = \sum_{i=0}^{\infty} i \mu(i) < \infty$.

Set $\mathcal{F}_0 = \{\phi, \Omega\}$ and for $n \geq 1$ $\mathcal{F}_n = \sigma(K_{i,j} : i < n, j \geq 1)$

Claim Set $M_n = \frac{X_n}{R^n}$. Then (M_n) is a $(\mathcal{F}_n)$ martingale.

- X_n is $\mathcal{F}_n$ measurable since the definition of X_n only involves $K_{i,j}$ for $i<n$ and $j \geq 1$.

- $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = \mathbb{E}[\sum_{j=1}^{\infty} \mathbb{1}_{j \leq X_n} K_{n,j} | \mathcal{F}_n]$

$$= \sum_{j=1}^{\infty} \mathbb{E}[\mathbb{1}_{j\leq X_n} K_{n,j} | \mathcal{F}_n] \quad \text{(monotone convergence)}$$

$$= \sum_{j=1}^{\infty} \mathbb{1}_{j \leq X_n} \mathbb{E}[K_{n,j} | \mathcal{F}_n] \quad (\mathbb{1}_{j\leq X_n} \text{ is } \mathcal{F}_n \text{ measurable})$$

$$= X_n \cdot R \quad (\mathbb{E}[K_{n,j} | \mathcal{F}_n] = \mathbb{E}[K_{n,j}] \text{ because } K_{n,j} \perp\!\!\!\perp \mathcal{F}_n)$$

Thus $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n$.

In particular $\mathbb{E}[M_{n+1}] = \mathbb{E}[M_n]$, so $\mathbb{E}[M_n] = \mathbb{E}[M_0] = 1$.

Thus $M_n \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$ and $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n$, so (M_n) is a $(\mathcal{F}_n)$ martingale.

Since $M_n \geq 0$, it converges a.s. to a r.v. denoted by M_∞.

We distinguish 3 cases:

① $\underline{R<1}$ From $\frac{X_n}{R^n} \xrightarrow[n\to\infty]{} M_\infty$, it follows that a.s $X_n = 0$ for n sufficiently large (if $X_n > 0$ then $\frac{X_n}{R^n} \geq \frac{1}{R^n} \xrightarrow[n\to\infty]{} \infty$). Thus we have extinction and $M_\infty = 0$ a.s

② $\underline{R=1}$ Then $X_n \xrightarrow{a.s} M_\infty$

For every $k \geq 1$ the events $\{\sum_{j=1}^{k} K_{n,j} \neq k\}$ are $\perp\!\!\!\perp$ and have >0 probability (because $\mu(1) \neq 1$), so by Borel–Cantelli 2 they happen a.s infinitely often. This shows that a.s $M_\infty \neq k$

Thus a.s $M_\infty = 0$: we have extinction.

Observe that here $M_n = X_n$ does not converge in L^1 since $\mathbb{E}[X_n] = 1 \not\to 0$

③ $\underline{R>1}$ $\frac{X_n}{R^n} \xrightarrow{a.s} M_\infty$. This raises the question of whether $M_\infty > 0$ or not: when $M_\infty > 0$, X_n is of order $R^n \cdot M_\infty$. This question is rather delicate!

Let us show that $\mathbb{P}(M_\infty > 0) > 0$ under the assumption $\sum_{k=1}^{\infty} k^2 \mu(k) < \infty$ (i.e $\mathbb{E}[K^2] < \infty$)

<u>Claim</u> $(M_n)_{n\geq 1}$ is bounded in L^2

Once the claim is proved this implies than $M_n \longrightarrow M_\infty$ in L^1, so $\mathbb{E}[M_\infty] = 1$ and

$\mathbb{P}(M_\infty > 0) > 0.$

Let us show the claim. Write

$$\mathbb{E}[X_{n+1}^2 \mid \mathcal{F}_n] = \mathbb{E}\Big[\sum_{j,j'\geq 1} \mathbb{1}_{j\leq X_n, j'\leq X_n} K_{n,j} K_{n,j'} \mid \mathcal{F}_n\Big] = \sum_{j,j'\geq 1} \mathbb{E}[\mathbb{1}_{j\leq X_n, j'\leq X_n} K_{n,j} K_{n,j'} \mid \mathcal{F}_n]$$

$$= \sum_{j,j'\geq 1} \mathbb{1}_{j\leq X_n, j'\leq X_n} \mathbb{E}[K_{n,j} K_{n,j'} \mid \mathcal{F}_n] \quad \text{because } X_n \text{ is } \mathcal{F}_n\text{-measurable}$$

$$= \sum_{\substack{j,j'\geq 1\\ j\neq j'}} \mathbb{1}_{\substack{j\leq X_n\\ j'\leq X_n}} R^2 + \sum_{j=1}^{\infty} \mathbb{1}_{j\leq X_n} \mathbb{E}[K^2] \quad \text{where } K \text{ has law } \mu$$

$$= X_n(X_n - 1) R^2 + X_n \mathbb{E}[K^2]$$

$$= R^2 X_n^2 + X_n \mathrm{Var}(K)$$

Thus $\mathbb{E}[X_{n+1}^2] = R^2\, \mathbb{E}[X_n^2] + R\, \mathrm{Var}(K)$, so by induction $X_n \in L^2$ $\forall n \geq 1$ and since $M_n = \frac{X_n}{R^n}$, $\mathbb{E}[M_{n+1}^2] = \mathbb{E}[M_n^2] + \frac{\mathrm{Var}(K)}{R^{n+2}}$,

so $\mathbb{E}[M_n^2] = 1 + \mathrm{Var}(K) \sum_{j=1}^{n} \frac{1}{R^{j+1}}$.

Since $R > 1$, this indeed shows that (M_n) is bounded in L^2

∾

We next study more specifically when martingales convergence in L^1!