Introduction to Ito calculus.

Content.

1. Spaces $L_2, M_2, M_{2,c}$.
2. Quadratic variation property of continuous martingales.

1 Doob-Kolmogorov inequality. Continuous time version

Let us establish the following continuous time version of the Doob-Kolmogorov inequality. We use RCLL as abbreviation for right-continuous function with left limits.

Proposition 1. Suppose $X_t \geq 0$ is a RCLL sub-martingale. Then for every $T, x \geq 0$

$$
\mathbb{P}(\sup_{0 \leq t \leq T} X_t \geq x) \leq \frac{\mathbb{E}[X_T^2]}{x^2}.
$$

Proof. Consider any sequence of partitions $\Pi_n = \{0 = t^n_0 < t^n_1 < \ldots < t^n_n = T\}$ such that $\Delta(\Pi_n) = \max_j |t^n_{j+1} - t^n_j| \to 0$. Additionally, suppose that the sequence $\Pi_n$ is nested, in the sense that for every $n_1 \leq n_2$, every point in $\Pi_{n_1}$ is also a point in $\Pi_{n_2}$. Let $X^n_t = X_{t^n}$ where $j = \max\{i : t_i \leq t\}$. Then $X^n_t$ is a sub-martingale adopted to the same filtration (notice that this would not be the case if we instead chose right ends of the intervals). By the discrete version of the D-K inequality (see previous lectures), we have

$$
\mathbb{P}(\max_{j \leq N_n} X^n_{t^j} \geq x) = \mathbb{P}(\sup_{t \leq T} X^n_t \geq x) \leq \frac{\mathbb{E}[X_T^2]}{x^2}.
$$

By RCLL, we have $\sup_{t \leq T} X^n_t \to \sup_{t \leq T} X_t$ a.s. Indeed, fix $\epsilon > 0$ and find $t_0 = t_0(\omega)$ such that $X_{t_0} \geq \sup_{t \leq T} X_t - \epsilon$. Find $n$ large enough and
\( j = j(n) \) such that \( t^{j(n)}_{j(n)-1} \leq t_0 \leq t^{j(n)}_{j(n)} \). Then \( t^{j(n)}_{j(n)} \to t_0 \) as \( n \to \infty \).

By right-continuity of \( X \), \( X_{t^{j(n)}} \to X_{t_0} \). This implies that for sufficiently large \( n \), \( \sup_{t \leq T} X^n_t \geq X_{t^{j(n)}} \geq X_{t_0} - 2\epsilon \), and the a.s. convergence is established.

On the other hand, since the sequence \( \Pi_n \) is nested, then the sequence \( \sup_{t \leq T} X^n_t \) is non-decreasing.

By continuity of probabilities, we obtain
\[
\mathbb{P}(\sup_{t \leq T} X^n_t \geq x) \to \mathbb{P}(\sup_{t \leq T} X_t \geq x).
\]

\( \square \)

## 2 Stochastic processes and martingales

Consider a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a filtration \((\mathcal{F}_t, t \in \mathbb{R}_+)\). We assume that all zero-measure events are "added" to \( \mathcal{F}_0 \). Namely, for every \( A \subset \Omega \), such that for some \( A' \in \mathcal{F} \) with \( \mathbb{P}(A') = 0 \) we have \( A \subset A' \in \mathcal{F} \), then \( A \) also belongs to \( \mathcal{F}_0 \). A filtration is called right-continuous if \( \mathcal{F}_t = \cap_{\epsilon > 0} \mathcal{F}_{t+\epsilon} \).

From now on we consider exclusively right-continuous filtrations. A stochastic process \( X_t \) adopted to this filtration is a measurable function \( X : \Omega \times [0, \infty) \to \mathbb{R} \), such that \( X_t \in \mathcal{F}_t \) for every \( t \). Denote by \( \mathcal{L}_2 \) the space of processes s.t. the Riemann integral \( \int_0^T X_t(\omega)dt \) exists a.s. and moreover \( \mathbb{E}[\int_0^T X^2_t dt] < \infty \) for every \( T > 0 \). This implies \( \mathbb{P}(\omega : \int_0^T |X_t(\omega)|dt < \infty, \forall T) = 1 \).

Let \( \mathcal{M}_2 \) consist of square integrable right-continuous martingales with left limits (RCLL). Namely \( \mathbb{E}[X^2_t] < \infty \) for every \( X \in \mathcal{M}_2 \) and \( t \geq 0 \). Finally \( \mathcal{M}_{2,c} \subset \mathcal{M}_2 \) is a further subset of processes consisting of a.s. continuous processes. For each \( T > 0 \) we define a norm on \( \mathcal{M}_2 \) by \( \|X\| = \|X\|_T = (\mathbb{E}[X^2_T])^{1/2} \). Applying sub-martingale property of \( X^2_t \) we have \( \mathbb{E}[X^2_{T_1}] \leq \mathbb{E}[X^2_{T_2}] \) for every \( T_1 \leq T_2 \).

A stochastic process \( Y_t \) is called a version of \( X_t \) if for every \( t \in \mathbb{R}_+ \), \( \mathbb{P}(X_t = Y_t) = 1 \). Notice, this is weaker than saying \( \mathbb{P}(X_t = Y_t, \forall t) = 1 \).

**Proposition 2.** Suppose \((X_t, \mathcal{F}_t)\) is a submartingale and \( t \to \mathbb{E}[X_t] \) is a continuous function. Then there exists a version \( Y_t \) of \( X_t \) which is RCLL.

We skip the proof of this fact.

**Proposition 3.** \( \mathcal{M}_2 \) is a complete metric space and (w.r.t. \( \| \cdot \| \)) \( \mathcal{M}_{2,c} \) is a closed subspace of \( \mathcal{M}_2 \).

**Proof.** We need to show that if \( X^{(n)} \in \mathcal{M}_2 \) is Cauchy, then there exists \( X \in \mathcal{M}_2 \) with \( \|X^{(n)} - X\| \to 0 \).

Assume \( X^{(n)} \) is Cauchy. Fix \( t \leq T \) Since \( X^{(n)} - X^{(m)} \) is a martingale as well, \( \mathbb{E}[(X^{(n)}_t - X^{(m)}_t)^2] \leq \mathbb{E}[(X^{(n)}_T - X^{(m)}_T)^2] \). Thus \( X^{(n)}_t \) is Cauchy as well. We know that the space \( \mathcal{L}_2 \) of random variables with finite second moment
is closed. Thus for each \( t \) there exists a r.v. \( X_t \) s.t. \( \mathbb{E}[(X_t^{(n)} - X_t)^2] \to 0 \) as \( n \to \infty \). We claim that since \( X_t^{(n)} \in \mathcal{F}_t \) and \( X_t^{(n)} \) is RCLL, then \( (X_t, t \geq 0) \) is adopted to \( \mathcal{F}_t \) as well (exercise). Let us show it is a martingale. First \( \mathbb{E}[|X_{t}] < \infty \) since in fact \( \mathbb{E}[X_t^2] < \infty \). Fix \( s < t \) and \( A \in \mathcal{F}_s \). Since each \( X_t^{(n)} \) is a martingale, then \( \mathbb{E}[X_t^{(n)}1(A)] = \mathbb{E}[X_s^{(n)}1(A)] \). We have

\[
\mathbb{E}[X_t1(A)] - \mathbb{E}[X_s1(A)] = \mathbb{E}[(X_t - X_t^{(n)})1(A)] - \mathbb{E}[(X_s - X_s^{(n)})1(A)]
\]

We have \( \mathbb{E}[|X_t - X_t^{(n)}|1(A)] \leq \mathbb{E}[|X_t - X_t^{(n)}|] \leq (\mathbb{E}[(X_t - X_t^{(n)})^2])^{1/2} \to 0 \) as \( n \to \infty \). A similar statement holds for \( s \). Since the left-hand side does not depend on \( n \), we conclude \( \mathbb{E}[X_t1(A)] = \mathbb{E}[X_s1(A)] \) implying \( \mathbb{E}[X_t|\mathcal{F}_s] = X_s \), namely \( X_t \) is a martingale. Since \( \mathbb{E}[X_t] = \mathbb{E}[X_0] \) is constant and therefore continuous as a function of \( t \), then there exists version of \( X_t \) which is RCLL. For simplicity we denote it by \( X_t \) as well. We constructed a process \( X_t \in \mathcal{M}_2 \) s.t. \( \mathbb{E}[(X_t^{(n)} - X_t)^2] \to 0 \) for all \( t \leq T \). This proves completeness of \( \mathcal{M}_2 \).

Now we deal with closeness of \( \mathcal{M}_{2,c} \). Since \( X_t^{(n)} - X_t \) is a martingale, \( (X_t^{(n)} - X_t)^2 \) is a submartingale. Since \( X_t \in \mathcal{M}_2 \), then \( (X_t^{(n)} - X_t)^2 \) is RCLL. Then submartingale inequality applies. Fix \( \epsilon > 0 \). By submartingale inequality we have

\[
\mathbb{P}(\sup_{t \leq T}|X_t^{(n)} - X_t| > \epsilon) \leq \frac{1}{\epsilon^2} \mathbb{E}[(X_T^{(n)} - X_T)^2] \to 0,
\]

as \( n \to \infty \). Then we can choose subsequence \( n_k \) such that

\[
\mathbb{P}(\sup_{t \leq T}|X_t^{(n_k)} - X_t| > 1/k) \leq \frac{1}{2^k}.
\]

Since \( 1/2^k \) is summable, by Borel-Cantelli Lemma we have \( \sup_{t \leq T}|X_t^{(n_k)} - X_t| \to 0 \) almost surely: \( \mathbb{P}(\{\omega \in \Omega : \sup_{t \leq T}|X_t^{(n_k)}(\omega) - X_t(\omega)| \to 0\}) = 1 \). Recall that a uniform limit of continuous functions is continuous as well (first lecture). Thus \( X_t \) is continuous a.s. As a result \( X_t \in \mathcal{M}_{2,c} \) and \( \mathcal{M}_{2,c} \) is closed.

### 3 Doob-Meyer decomposition and quadratic variation of processes in \( \mathcal{M}_{2,c} \)

Consider a Brownian motion \( B_t \) adopted to a filtration \( \mathcal{F}_t \). Suppose this filtration makes \( B_t \) a strong Markov process (for example \( \mathcal{F}_t \) is generated by \( B \) itself). Recall that both \( B_t \) and \( B_t^2 - t \) are martingales and also \( B \in \mathcal{M}_{2,c} \). Finally recall that the quadratic variation of \( B \) over any interval \([0, t]\) is \( t \). There is a
generalization of these observations to processes in $\mathcal{M}_{2,c}$. For this we need to recall the following result.

**Theorem 1 (Doob-Meyer decomposition).** Suppose $(X_t, \mathcal{F}_t)$ is a continuous non-negative sub-martingale. Then there exist a continuous martingale $M_t$ and a.s. non-decreasing continuous process $A_t$ with $A_0 = 0$, both adapted to $\mathcal{F}_t$ such that $X_t = A_t + M_t$. The decomposition is unique in the almost sure sense.

The proof of this theorem is skipped. It is obtained by appropriate discretization and passing to limits. The discrete version of this result we did earlier. See [1] for details.

Now suppose $X_t \in \mathcal{M}_{2,c}$. Then $X_t^2$ is a continuous non-negative submartingale and thus DM theorem applies. The part $A_t$ in the unique decomposition of $X_t^2$ is called quadratic variation of $X_t$ (we will shortly justify this) and denoted $\langle X_t \rangle$.

**Theorem 2.** Suppose $X_t \in \mathcal{M}_{2,c}$. Then for every $t > 0$ the following convergence in probability takes place

$$
\lim_{\Pi_n: \Delta(\Pi_n) \to 0} \sum_{0 \leq j \leq n-1} (X_{t_{j+1}} - X_{t_j})^2 \to \langle X_t \rangle,
$$

where the limit is over all partitions $\Pi_n = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ and $\Delta(\Pi_n) = \max_j |t_j - t_{j-1}|$.

**Proof.** Fix $s < t$. Let $X \in \mathcal{M}_{2,c}$. We have

$$
\mathbb{E}[(X_t - X_s)^2 - (\langle X_t \rangle - \langle X_s \rangle)] = \mathbb{E}[X_t^2 - 2X_tX_s + X_s^2 - (\langle X_t \rangle - \langle X_s \rangle)]
$$

$$
= \mathbb{E}[X_t^2|\mathcal{F}_s] - 2X_s \mathbb{E}[X_t|\mathcal{F}_s] + X_s^2 - \mathbb{E}[\langle X_t \rangle|\mathcal{F}_s] + \langle X_s \rangle
$$

$$
= \mathbb{E}[X_t^2 - (\langle X_t \rangle)|\mathcal{F}_s] - X_s^2 + \langle X_s \rangle
$$

$$
= 0.
$$

Thus for every $s < t \leq u < v$ by conditioning first on $\mathcal{F}_u$ and using tower property we obtain

$$
\mathbb{E}\left((X_t - X_s)^2 - (\langle X_t \rangle - \langle X_s \rangle)\right) \left((X_u - X_v)^2 - (\langle X_u \rangle - \langle X_v \rangle)\right) = 0 \tag{1}
$$

The proof of the following lemma is application of various "carefully placed" tower properties and is omitted. See [1] Lemma 1.5.9 for details.

**Lemma 1.** Suppose $X \in \mathcal{M}_2$ satisfies $|X_s| \leq M$ a.s. for all $s \leq t$. Then for every partition $0 = t_0 \leq \cdots \leq t_n = t$

$$
\mathbb{E}\left(\sum_j (X_{t_{j+1}} - X_{t_j})^2\right) \leq 6M^4.
$$
Lemma 2. Suppose $X \in \mathcal{M}_2$ satisfies $|X_s| \leq M$ a.s. for all $s \leq t$. Then

$$\lim_{\Delta(\Pi_n) \to 0} E[\sum_j (X_{t_{j+1}} - X_{t_j})^4] = 0,$$

where $\Pi_n = \{0 = t_0 < \cdots < t_n = t\}$, $\Delta(\Pi_n) = \max_j |t_{j+1} - t_j|$.

**Proof.** We have

$$\sum_j (X_{t_{j+1}} - X_{t_j})^4 \leq \sum_j (X_{t_{j+1}} - X_{t_j})^2 \sup\{|X_r - X_s|^2 : |r - s| \leq \Delta(\Pi_n)\}.$$

Applying Cauchy-Schwartz inequality and Lemma 1 we obtain

$$\left( E[\sum_j (X_{t_{j+1}} - X_{t_j})^4] \right)^2 \leq E\left( \sum_j (X_{t_{j+1}} - X_{t_j})^2 \right)^2 E[\sup\{|X_r - X_s|^4 : |r - s| \leq \Delta(\Pi_n)\}] \leq 6M^4 E[\sup\{|X_r - X_s|^4 : |r - s| \leq \Delta(\Pi_n)\}].$$

Now $X(\omega)$ is a.s. continuous and therefore uniformly continuous on $[0, t]$. Therefore, a.s. $\sup\{|X_r - X_s|^2 : |r - s| \leq \Delta(\Pi_n)\} \to 0$ as $\Delta(\Pi_n) \to 0$. Also $|X_r - X_s| \leq 2M$ a.s. Applying Bounded Convergence Theorem, we obtain that $E[\sup\{|X_r - X_s|^4 : |r - s| \leq \Delta(\Pi_n)\}]$ converges to zero as well and the result is obtained.

We now return to the proof of the proposition. We first assume $|X_s| \leq M$ and $\langle X_s \rangle \leq M$ a.s. for $s \in [0, t]$.

We have (using a telescoping sum)

$$E\left( \sum_j (X_{t_{j+1}} - X_{t_j})^2 - \langle X_t \rangle \right)^2 = E\left( \sum_j ((X_{t_{j+1}} - X_{t_j})^2 - (\langle X_{t_{j+1}} \rangle - \langle X_{t_j} \rangle)) \right)^2.$$

When we expand the square the terms corresponding to cross products with $j_1 \neq j_2$ disappear due to (1). Thus the expression is equal to

$$E\sum_j ((X_{t_{j+1}} - X_{t_j})^2 - (\langle X_{t_{j+1}} \rangle - \langle X_{t_j} \rangle))^2 \leq 2E\left[ \sum_j (X_{t_{j+1}} - X_{t_j})^4 \right] + 2E[\sum_j ((\langle X_{t_{j+1}} \rangle - \langle X_{t_j} \rangle)^2].$$

The first term converges to zero as $\Delta(\Pi_n) \to 0$ by Lemma 2.
We now analyze the second term. Since \( \langle X_t \rangle \) is a.s. non-decreasing, then
\[
\sum_j (\langle X_{t_j+1} \rangle - \langle X_{t_j} \rangle)^2 \leq \sum_j (\langle X_{t_j+1} \rangle - \langle X_{t_j} \rangle) \sup_{0 \leq s \leq r \leq t} \{ \langle X_r \rangle - \langle X_s \rangle : |r - s| \leq \Delta(P_n) \}
\]
Thus the expectation is upper bounded by
\[
\mathbb{E}[\langle X_t \rangle \sup_{0 \leq s \leq r \leq t} \{ \langle X_r \rangle - \langle X_s \rangle : |r - s| \leq \Delta(P_n) \}] \tag{2}
\]
Now \( \langle X_t \rangle \) is a.s. continuous and thus the supremum term converges to zero a.s. as \( n \to \infty \). On the other hand a.s. \( \langle X_t \rangle(\langle X_r \rangle - \langle X_s \rangle) \leq 2M^2 \). Thus using Bounded Convergence Theorem, we obtain that the expectation in (2) converges to zero as well. We conclude that in the bounded case Bounded Convergence Theorem, we obtain that the expectation in (2) converges to zero as well. We conclude that in the bounded case, \( \langle X_s \rangle \leq M \) on \([0, t]\), the quadratic variation of \( X_s \) over \([0, t]\) converges to \( \langle X_t \rangle \) in \( \mathbb{L}_2 \) sense. This implies convergence in probability as well.

It remains to analyze the general (unbounded) case. Introduce stopping times \( T_M \) for every \( M \in \mathbb{R}_+ \) as follows
\[
T_M = \min\{t : |X_t| \geq M \text{ or } \langle X_t \rangle \geq M \}
\]
Consider \( X_t^M \triangleq X_{t \wedge T_M} \). Then \( X_t^M \in \mathcal{M}_{2,c} \) and is a.s. bounded. Further since \( X_t^2 - \langle X_t \rangle \) is a martingale, then \( X_t^2 - \langle X_t \wedge T_M \rangle \) is a bounded martingale. Since Doob-Meyer decomposition is unique, we that \( \langle X_t \wedge T_M \rangle \) is indeed the unique non-decreasing component of the stopped martingale \( X_{t \wedge T_M} \). There is a subtlety here: \( X_t^M \) is a continuous martingale and therefore it has its own quadratic variation \( \langle X_t^M \rangle \) - the unique non-decreasing a.s. process such that \( (X_t^M)^2 - \langle X_t^M \rangle \) is a martingale. It is a priori non obvious that \( \langle X_t^M \rangle \) is the same as \( \langle X_{t \wedge T_M} \rangle \) - quadratic variation of \( X_t \) stopped at \( T_M \). But due to uniqueness of the D-M decomposition, it is.

Fix \( \epsilon > 0, t \geq 0 \) and find \( M \) large enough so that \( \mathbb{P}(T_M < t) < \epsilon/2 \). This is possible since \( X_t \) and \( \langle X_t \rangle \) are continuous processes. Now we have
\[
\mathbb{P}\left( \left| \sum_j (X_{t_{j+1}} - X_{t_j})^2 - \langle X_t \rangle \right| > \epsilon \right)
\]
\[
\leq \mathbb{P}\left( \left| \sum_j (X_{t_{j+1}} - X_{t_j})^2 - \langle X_t \rangle \right| > \epsilon, t \leq T_M \right) + \mathbb{P}(T_M < t)
\]
\[
= \mathbb{P}\left( \left| \sum_j (X_{t_{j+1} \wedge T_M} - X_{t_j \wedge T_M})^2 - \langle X_{t \wedge T_M} \rangle \right| > \epsilon, t \leq T_M \right) + \mathbb{P}(T_M < t)
\]
\[
\leq \mathbb{P}\left( \left| \sum_j (X_{t_{j+1} \wedge T_M} - X_{t_j \wedge T_M})^2 - \langle X_{t \wedge T_M} \rangle \right| > \epsilon \right) + \mathbb{P}(T_M < t).
\]
We already established the result for bounded martingales and quadratic variations. Thus, there exists $\delta = \delta(\epsilon) > 0$ such that, provided $\Delta(\Pi) < \delta$, we have

$$\mathbb{P}\left(\left|\sum_j (X_{t_{j+1} \wedge T_M} - X_{t_j \wedge T_M})^2 - \langle X_{t \wedge T_M} \rangle \right| > \epsilon\right) < \epsilon/2.$$

We conclude that for $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ with $\Delta(\Pi) < \delta$, we have

$$\mathbb{P}\left(\left|\sum_j (X_{t_{j+1}} - X_{t_j})^2 - \langle X_t \rangle \right| > \epsilon\right) < \epsilon.$$

4 Additional reading materials

- Chapter I. Karatzas and Shreve [1]

References

15.070J / 6.265J Advanced Stochastic Processes
Fall 2013

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.