36-771	Fall 2018
Martingales 1 : Concentration inequalities	

Improving continuous time martingale concentration

Lecturer : Aaditya Ramdas

We start with a probability space $(\Omega, \mathcal{F}, P)$. In addition to that, consider a filtration $(\mathcal{F}_t)$ meaning a sequence of sigma-algebras such that $s \leq t$ implies $\mathcal{F}_s \subseteq \mathcal{F}_t$. One typically assumes that the resulting filtered complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ satisfies the "usual" hypotheses (from Protter):

- $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$.
- The filtration is right-continuous, meaning $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$.

A random variable T is called a stopping time if $\{T \leq t\} \in \mathcal{F}_t$ for all $0 \leq t \leq \infty$. Because the filtration is right continuous, we also have that the event $\{T < t\} \in \mathcal{F}_t$ for every $0 \leq t \leq \infty$ if and only if T is a stopping time.

A stochastic process (X_t) is a collection of $\mathbb{R}$-valued (or $\mathbb{R}^d$-valued) random variables. (X_t) is called adapted if $X_t \in \mathcal{F}_t$ for every t. There are two different concepts of equality of two stochastic processes:

- (X_t) and (Y_t) are modifications if

  for each t, we have $X_t = Y_t$ almost surely.

- (X_t) and (Y_t) are indistinguishable if

  almost surely, we have $X_t = Y_t$ for each t.

The second is much stronger.

A stochastic process is called cadlag, if it almost surely has sample paths that are right-continuous with left limits. If (X_t) and (Y_t) are modifications that have right continuous paths almost surely, then they are indistinguishable. Hence for cadlag processes, the above two concepts are identical.

A continuous time process (M_t) is a martingale with respect to filtration $(\mathcal{F}_t)$ if for any $0 \leq s \leq t$, we have $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$. Every martingale has a right continuous modification that is cadlag, and without further mention, we will always assume we are dealing with this version of the martingale.

1 Brownian Motion, Poisson process, Levy process

An $(\mathcal{F}_t)$ adapted process (W_t) is a Brownian motion if

- $W_0 = 0$ a.s.
- Independent increments: for all $0 \le s \le t$, $W_t - W_s$ is independent of $\mathcal{F}_s$.
- Stationary increments: $W_t - W_s \sim N(0, t-s)$.
- The sample paths are continous a.s.

Fact 1 *W_t and $W_t^2 - t$ are martingales. Also for any $\lambda > 0$, $\exp(\lambda W_t - \frac{\lambda^2}{2}t)$ is a martingale.*

An $(\mathcal{F}_t)$ adapted process (N_t) is a Poisson process if the last two properties are

- Stationary increments: $N_t - N_s \sim Poi(\mu(t-s))$ for some $\mu > 0$.
- The sample paths are continuous in probability:

$$\forall \epsilon > 0, t \ge 0, \text{ we have } \lim_{h \to 0} P(|N_{t+h} - N_t| > \epsilon) = 0.$$

Fact 2 *$N_t - \mu t$ and $(N_t - \mu t)^2 - \mu t$ are both martingales. Also for any $\lambda > 0$, we have $\exp(\lambda N_t - \mu t(e^\lambda - 1)) = \exp(\lambda(N_t - \mu t) - \mu t(e^\lambda - \lambda - 1))$ is a martingale.*

These are both examples of Levy processes, which are all characterized by $X_0 = 0$, independent increments, stationary increments, and continuity in probability.

Fact 3 *If a Levy process (Y_t) satisfies $\mathbb{E}\exp(\lambda Y_1) < \infty$ then Assumption 1 is satisfied with $S_t = Y_t - \mathbb{E}[Y_t]$ and $\psi(\lambda) = \log \mathbb{E}\exp(\lambda Y_1)$ and $V(t) = t$. (the above two facts are special cases of this one.)*

Levy processes are "infinitely divisible" meaning that for any integer n and time t, the law of X_t matches the law of the sum of n iid random variables $X_{t/n}, X_{t2/n} - X_{t/n}, \ldots$.

Some other examples include the Gamma process, the compound poisson process and "stable" processes like the Cauchy process. In fact by the Levy-Khinchtine representation theorem, a Levy process can be uniquely defined by three components (a, σ, Π) with correspond respectively to a drift term, a brownian motion variance term, and a measure defining a compound Poisson process.

2 Local properties

For a continous stochastic process (X_t), a property π is said to hold locally if there exists a sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ increasing to infinity a.s. such that for every n, the stopped process $(X_{T_n \wedge t} 1_{T_n>0})$ has property π.

This allows us to differentiate between a square-integrable "local martingale", and a "locally square integrable" martingale. Every martingale is a local martingale. Every bounded local martingale is a martingale. Every local martingale bounded from below is a supermartingale. In general a local martingale is not a martingale, because its expectation can be distorted by large values of small probability.

Local martingales form a very important class of processes in the theory of stochastic calculus. This is because the local martingale property is preserved by the stochastic integral, but the martingale property is not. Further, all continuous local martingales are time-changes of brownian motions.

3 Variation of a function/process

For $T > 0$ let π be a partition of $[0, T]$: $0 = t_0^\pi < t_1^\pi < \cdots < t_k^\pi = T$, and define $\|\pi\| := \max_{1\le i\le k} t_i^\pi - t_{i-1}^\pi$. For a function $f : [0, T] \to \mathbb{R}$, and $p \ge 1$, define the notion of p-variation of f on $[0, T]$ as

$$V^{(p)}(f) := \lim_{\|\pi\|\to 0, k\to\infty} \sum_{i=1}^{k} |f(t_i^\pi) - f(t_{i-1}^\pi)|^p$$

provided the limit exists. $p = 1$ is called the total variation on $[0, t]$, and $p = 2$ is called the quadratic variation of f on $[0, T]$.

Analogously replacing the function f with a stochastic process (X_t), one defines the total variation process and quadratic variation process $([X]_t)$, provided the limit exists in the sense of convergence in probability. The total variation process of any nonzero continuous (M_t) is a.s. infinite on any interval.

Confusingly, there is another notion of p-variation defined in functional analysis, defined for functions $f : \mathbb{R} \to (M, d)$ where the latter is a metric space. This notion takes a supremum over finite partitions π (though they are finite, they can be arbitrarily fine).

$$\|f\|_{p-var} = \sup_\pi \left(\sum_{t_k\in\pi} [d(f(t_k) - f(t_{k-1}))]^p \right)^{1/p}$$

When $p = 1$, this is also called total variation, and the class of functions where this is finite, is called the class of bounded variation.

The 2-variation could be much larger than the quadratic variation. For eg: for a brownian motion, its quadratic variation is t, while its p-variation is infinite for $p \le 2$.

4 Doob-Meyer decomposition

In discrete time, $\langle M \rangle_t$ can be defined as the unique predictable and increasing process such that $M_t^2 - \langle M \rangle_t$ is a martingale.

Consider a locally square-integrable martingale (M_t) wrt filtration $(\mathcal{F}_t)$, meaning that $\mathbb{E}M_t^2 < \infty$ for all $t \geq 0$ and M_t. (Recall that if M_t is a martingale then M_t^2 is a submartingale.) According to the Doob-Meyer decomposition theorem, there exists a unique nondecreasing stochastic process (A_t) adapted to $(\mathcal{F}_t)$, starting at 0 with right-continuous paths, such that $(M_t^2 - A_t)$ is a martingale wrt $(\mathcal{F}_t)$. Then A_t is the quadratic variation $[M]_t$ of M_t.

There is also a unique nondecreasing right-continuous predictable process (A_t) such that $(M_t^2 - A_t)$ is a martingale, and this A_t is denoted by $\langle M \rangle_t$.

Fact 4 *For continuous local martingales we also have* $[M]_t = \langle M \rangle_t$.

For a Brownian motion or Wiener process W_t, we have $\langle W \rangle_t = [W]_t = t$ almost surely, so we say that the BM accumulates quadratic variation at rate one per unit time. (in stochastic calculus, we write $dW_t dW_t = dt$, and also $dW_t dt = 0$ and $dtdt = 0$.) In fact, this property is characteristic of a BM:

Fact 5 *If* (M_t) *is a martingale with continuous paths and* $(M_t^2 - t)$ *is a martingale, then* (M_t) *is a BM.*

5 Concentration inequalities

While many of the earlier results already generalize results known in discrete time to new results for continuous-time martingales [C], here we summarize a few more useful bounds explicitly for continuous-time processes which are corollaries of the mother theorem.

Corollary 6 *Let* $(S_t)_{t \in (0,\infty)}$ *be a real-valued process.*

(a) If (S_t) *is a locally square-integrable martingale with a.s. continuous paths, then*

$$\mathcal{P}\left(\exists t \in (0, \infty) : S_t \geq a + b \left\langle S \right\rangle_t\right) \leq \exp\{-2ab\}.$$

If $\langle S \rangle_t \uparrow \infty$ *as* $t \uparrow \infty$, *then the probability inequality may be replaced with an equality. This recovers as a special case the standard line-crossing probability for Brownian motion (e.g., Durrett 2017, Exercise 7.5.2).*

(b) *If* (S_t) *is a local martingale with* $\Delta S_t \le c$ *for all* t*, then*

$$\mathcal{P}\left(\exists t \in (0,\infty) : S_t \ge x + \mathfrak{s}_P\left(\frac{x}{m}\right) \cdot (\langle S \rangle_t - m)\right) \le d \exp\{-m\psi_P^\star\left(\frac{x}{m}\right)\} \le d \exp\{-\frac{x^2}{2(m + cx/3)}\}. \tag{1}$$

This strengthens Appendix B, Inequality 1 of Shorack & Wellner (1986) [B].

(c) *If* (S_t) *is any locally square-integrable martingale satisfying the Bernstein condition of (see the Big Table lecture) for some predictable process* (W_t)*, then*

$$\mathcal{P}\left(\exists t \in (0,\infty) : S_t \ge x + \mathfrak{s}_G\left(\frac{x}{m}\right) \cdot (V_t - m)\right) \le d \exp\{-m\psi_G^\star\left(\frac{x}{m}\right)\} \le d \exp\{-\frac{x^2}{2(m + cx)}\}.$$

This strengthens Lemma 2.2 of van de Geer (1995) [B,E].

Clearly, statement (b) applies to centered Poisson processes with $c = 1$, but this can be inferred directly from the fact that it is a Levy process. The point of statement (b) is that any local martingale with bounded jumps obeys this inequality, and so concentrates like a centered Poisson process in this sense.

References

Durrett, R. (2017), *Probability: Theory and Examples*, 5a edn.

Shorack, G. R. & Wellner, J. A. (1986), *Empirical processes with applications to statistics*, Wiley, New York.

van de Geer, S. (1995), 'Exponential Inequalities for Martingales, with Application to Maximum Likelihood Estimation for Counting Processes', *The Annals of Statistics* **23**(5), 1779–1801.