Exercise Sheet 1

Exercise 1

(1) We need to prove that for any two points $x, y \in A+B$ and $\forall t \in [0,1]$ then $tx + (1-t)y \in A+B$.

Let $a, c \in A$, $b, d \in B$. Set $x = a+b$ and $y = c+d$. Then

$$tx + (1-t)y = t(a+b) + (1-t)(c+d) = ta + tb + (1-t)c + (1-t)d$$

$ta + (1-t)c \in A$ since A convex

$tb + (1-t)b \in B$ since B convex

and therefore $tx + (1-t)y \in A+B$. $\square$

(2) „$\Rightarrow$" Let $x, y \in K$ and $\lambda \in [0,1]$ $\overset{\text{def}}{\Longrightarrow}$ $\lambda x + (1-\lambda)y \in K$

We can write $\lambda x + (1-\lambda)y = \lambda\underbrace{(\lambda x + (1-\lambda x))}_{=1} + (1-\lambda)\underbrace{(\lambda y + (1-\lambda)y)}_{=1}$

Thus we have expressed $\lambda x + (1-\lambda)y$ as a convex combination of points in K

$\Rightarrow$ $\lambda x + (1-\lambda)y \in \lambda K + (1-\lambda)K$.

„$\Leftarrow$" Let $x, y \in K$, $\lambda \in [0,1]$. Then $\lambda x + (1-\lambda)y \in \lambda K + (1-\lambda)K$ and hence $\lambda x + (1-\lambda)y \in K$ by assumption $\Rightarrow$ K convex. $\square$

(3) f is convex on $[a,b]$ $\Rightarrow$ f is continuous on $[a,b]$ $\Rightarrow$ f is Riemann integrable on $[a,b]$

④ Let $x, y \in S_t$, $\lambda \in [0,1]$.

Since φ is convex, $\varphi(\lambda x + (1-\lambda)y) \leq \lambda \varphi(x) + (1-\lambda)\varphi(y) \leq \lambda t + (1-\lambda)t = t$

$\Rightarrow \varphi(\lambda x + (1-\lambda)y) \in S_t$ $\square$

⑤ Suppose that f is not lower-semicontinuous.

Then $\exists x_0 \in \mathbb{R}$ and $\exists (x_n)_{n \geq 1}$ with $x_n \to x_0$ st $f(x_n) < f(x_0)$ $\forall n$.

Since $x_n \to x_0$, we have $f(x_n) \to f(x_0)$ as $n \to \infty$ $\not{\downarrow}$ to f is closed. $\square$

⑥ "$\supseteq$" Since f is differentiable at x_0, $\nabla f(x_0)$ exists and is unique.

Also, $f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0)$ $\forall x \in \mathbb{R}^n$ $\Rightarrow \nabla f(x_0) \in \partial f(x_0)$

"$\subseteq$" Let $g \in \partial f(x_0)$. Then $f(x) \geq f(x_0) + g^T(x - x_0)$ $\forall x \in \mathbb{R}^n$

f is convex $\Rightarrow g = \nabla f(x_0)$. $\square$

Exercise 2

1. $\mathbb{E}[M_t] = \mathbb{E}[\mathbb{E}[M_t \mid \mathcal{F}_{t-1}]] = \mathbb{E}[M_{t-1}] \qquad \forall t \geq 1$

Therefore, all M_t have same expectation.

In particular, $\mathbb{E}[M_{t-1}] = \mathbb{E}[M_0]$. $\square$

2. We apply Jensen's inequality

Recall: for a convex g it holds $g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)]$

Let $g(x) := \max\{x, 0\} \to g$ is convex

In fact, let $x, y \in \mathbb{R}$

1st case: $\alpha x + (1-\alpha)y \geq 0 \Rightarrow g(\alpha x + (1-\alpha)y) = \max\{(\alpha x + (1-\alpha)y\}$

$$\leq \max\{\alpha x, 0\} + \max\{(1-\alpha)y, 0\}$$

$$= \alpha g(x) + (1-\alpha) g(y)$$

2nd case: $\alpha x + (1-\alpha)y < 0 \Rightarrow g(\alpha x + (1-\alpha)y) = \max\{\alpha x + (1-\alpha)y, 0\} = 0$

$$\leq \alpha \max\{x, 0\} + (1-\alpha)\max\{x, 0\}$$

$$= \alpha g(x) + (1-\alpha) g(y)$$

Let $s \leq t$

$$\max\{M_t, 0\} = g(M_t)$$

$$= g(\mathbb{E}[M_t \mid \mathcal{F}_s] + M_t - \mathbb{E}[M_t \mid \mathcal{F}_s])$$

$$\overset{g \text{ convex}}{\leq} g(\mathbb{E}[M_t \mid \mathcal{F}_s]) + g(\mathbb{E}[M_t \mid \mathcal{F}_s] - \mathbb{E}[M_t \mid \mathcal{F}_s])$$

$$= \mathbb{E}[\max\{M_t, 0\} \mid \mathcal{F}_s] \qquad \square$$

Exercise 3

Note that

- (X_i) bounded $\Rightarrow$ $S_t + a_t$, $S_t^2 + a_t S_t b_t$ and $(a_t^\lambda)^{-1} \exp(\lambda S_t)$ are integrable
- $\mathcal{F}_t$ - measurability is clear.

$\Rightarrow$ we just need to check the martingale property

① $$\mathbb{E}[S_{t+1} + a_{t+1} \mid \mathcal{F}_t] = \mathbb{E}\Big[\sum_{i=1}^{t+1} X_i + a_{t+1} \mid \mathcal{F}_t\Big]$$

$$= \mathbb{E}[S_t + X_{t+1} + a_{t+1} \mid \mathcal{F}_t]$$

S_t is $\mathcal{F}_t$- meas.
a_t is constant

$$= S_t + \mathbb{E}[X_{t+1} \mid \mathcal{F}_t] + a_{t+1}$$

$X_{t+1} \perp \mathcal{F}_t$

$$\overset{\perp}{=} S_t + \mathbb{E}[X_{t+1}] + a_{t+1}$$

$$= S_t + m_t + a_{t+1}$$

$\Rightarrow$ $a_t = m_t + a_{t+1}$ $\Rightarrow$ $a_t = -\sum_{k=1}^{t} m_t$.

② $$\mathbb{E}[S_{t+1}^2 + a_{t+1} S_{t+1} + b_{t+1} \mid \mathcal{F}_t] = \mathbb{E}[(S_t + X_{t+1})^2 + a_{t+1}(S_t + X_{t+1}) + b_{t+1} \mid \mathcal{F}_t]$$

$$= \mathbb{E}[S_t^2 + X_{t+1}^2 + 2 S_t X_{t+1} + a_{t+1} S_t + a_{t+1} X_{t+1} + b_{t+1} \mid \mathcal{F}_t]$$

$$= S_t^2 + \mathbb{E}[X_{t+1}^2 \mid \mathcal{F}_t] + 2 S_t \mathbb{E}[X_{t+1} \mid \mathcal{F}_t] + a_{t+1} S_t + a_{t+1} \mathbb{E}[X_{t+1} \mid \mathcal{F}_t] + b_{t+1}$$

$$= S_t^2 + \mathbb{E}[X_{t+1}^2] + 2 S_t \mathbb{E}[X_{t+1}] + a_{t+1} S_t + a_{t+1} \mathbb{E}[X_{t+1}] + b_{t+1}$$

$$= S_t^2 + \sigma_{t+1}^2 + m_{t+1}^2 + 2 S_t m_{t+1} + a_{t+1} S_t + a_{t+1} m_{t+1} + b_{t+1}$$

$$= S_t^2 + S_t (2 m_{t+1} + a_{t+1}) + \sigma_{t+1}^2 + a_{t+1} m_{t+1} + b_{t+1}$$

Then it must hold

$$a_t = 2 m_{t+1} + a_{t+1} \quad \text{and} \quad b_t = \sigma_{t+1}^2 + a_{t+1} m_{t+1} + b_{t+1}$$

$$\Rightarrow a_t = -\sum_{k=1}^{t} m_k \quad \text{and} \quad b_t = -\sum_{k=1}^{t} (\sigma_k^2 + m_k^2 + 2 a_k m_k)$$

③ $$\mathbb{E}\left[\frac{1}{a_{t+1}} e^{\lambda S_{t+1}} \,\middle|\, \mathcal{F}_t \right] = \frac{1}{a_{t+1}} \mathbb{E}\left[e^{\lambda S_t} e^{\lambda X_{t+1}} \,\middle|\, \mathcal{F}_t \right]$$

$$= \frac{1}{a_{t+1}} e^{\lambda S_t} \mathbb{E}\left[e^{\lambda X_{t+1}} \,\middle|\, \mathcal{F}_t \right]$$

$$= \frac{1}{a_{t+1}} e^{\lambda S_t} \mathbb{E}\left[e^{\lambda X_{t+1}} \right]$$

$$\Rightarrow \quad \frac{a_{t+1}}{a_t} = \mathbb{E}\left[e^{\lambda X_{t+1}} \right]$$

$$\Rightarrow \quad a_t = \prod_{k=1}^{t} \mathbb{E}\left[e^{\lambda X_k} \right]$$

Exercise 4

① By Markov's ineq.: $\mathbb{P}(X \geq \alpha) = \mathbb{P}(e^{tX} \geq e^{t\alpha}) \leq e^{-t\alpha} M_X(t)$

take the infimum.

② $M_X(t) = \mathbb{E}[e^{tX}] = e^{t\cdot 0}\, \mathbb{P}(X=0) + e^{t\cdot 1}\, \mathbb{P}(X=1) = (1-p) + pe^t$

③ Note $\forall x \in \mathbb{R}$: $1 + x \leq e^x$

$$M_{S_N}(\lambda) = \mathbb{E}[e^{\lambda \sum_{i=1}^n X_i}] = \prod_{i=1}^n \mathbb{E}[e^{\lambda X_i}] = \prod_{i=1}^n (1 + p_i (e^\lambda - 1))$$
$$\leq \prod_{i=1}^n \exp\left(p_i (e^\lambda - 1)\right) = \exp\left((e^\lambda - 1) \sum_{i=1}^n p_i\right) = \exp\left(\mu (e^\lambda - 1)\right)$$

Therefore,

$$\mathbb{P}(S_n \geq t) \leq e^{-\lambda t} \exp\left((e^\lambda - 1)\mu\right)$$

Substitute $\lambda = \ln\left(\frac{t}{\mu}\right)$. $\square$

$\boxed{\text{Exercise 5}}$

1. $\mathbb{E}[X_{n+1} \mid X_n = 0] = 0$

   $\mathbb{E}[X_{n+1} \mid X_n = 2^n] = 2^n$

2. Since $X_n \geqslant 0 \quad \forall n$

   $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] = \mathbb{E}[X_0] = 1 \quad \forall n$

3. We use the following result

$$(X_n)_n \text{ bounded martingale in } L^1 \Rightarrow X_n \xrightarrow[n\to\infty]{a.s.} X_\infty \text{ and } X_\infty \in L^1$$

(This claim actually holds for sub-martingales)

$$\mathbb{P}\left(\liminf_k X_k > 0\right) \leq \mathbb{P}(X_n = 0) = 2^{-n} \longrightarrow 0$$

Note that $\mathbb{E}[X_\infty] \neq \lim_{n\to\infty} \mathbb{E}[X_n] \Rightarrow$ it doesn't converge in L^1

Exercise 6

① Recall $\operatorname{diam}(A) = \sup_{x,y \in A} \|x-y\|_2$

For all $x, y \in Q^n$: $\|x-y\|_2^2 = \sum_{i=1}^n |x_i - y_i|^2 \le \sum_{i=1}^n (|x_i|+|y_i|)^2 \le \sum_{i=1}^n 1 = n$

On the other hand, consider

$$x = \frac{1}{2}(1,\dots,1) \qquad y = \frac{1}{2}(-1,\dots,-1)$$

Clearly, $x, y \in Q^n$ and

$$\|x-y\|_2^2 = \sum_{i=1}^n \left(\frac{1}{2}+\frac{1}{2}\right)^2 = n$$

Thus, $\sqrt{n} \le \operatorname{diam}(Q^n) \le \sqrt{n} \Rightarrow \operatorname{diam}(Q^n) = \sqrt{n}$. $\square$

② $X_1, \dots, X_n, Y_1, \dots, Y_n$ are independent r.v. distributed on $[-\frac{1}{2}, \frac{1}{2}]$

$$\Rightarrow \mathbb{E}[\|X-Y\|_2^2] = \sum_{i=1}^n \mathbb{E}[|X_i - Y_i|^2] = n\,\mathbb{E}[|X_1 - Y_1|^2] = \frac{n}{6}$$

Conclusion: the L^2-norm of the distance btw two random points X and Y has order of magnitude $\sqrt{n}$.

③ Apply CLT: $\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \longrightarrow \mathcal{N}(0,1)$

④ In ③ the weights were $\theta = \frac{1}{\sqrt{n}}$ but it is not essential.

$\mathbb{E}[\theta_i Y_i] = \theta_i \mathbb{E}[Y_i] = 0$

$\operatorname{var}(\theta_i Y_i) = \theta_i^2 \operatorname{var}(Y_i) = \theta_i^2$

By Berry-Esseen thm we get

$$\forall t \in \mathbb{R} \quad \left| \mathbb{P}\left(\sum_{i=1}^n \theta_i Y_i \le t\right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2}\, ds \right| \le C \sum_{i=1}^n \theta_i^4 \qquad (*)$$

where $C>0$ is an absolute constant.

<u>Remark</u> $\theta_i = \frac{1}{\sqrt{n}} \Rightarrow$ RHS of $(*)$ is of order $\frac{1}{n}$

Exercise 7

① $$\frac{1}{2}\int_K\int_K |f(x)-f(y)|^2\,dx\,dy = \frac{1}{2}\iint_{K\,K} f^2(x)\,dx\,dy + \frac{1}{2}\iint_{K\,K} f^2(x)\,dx\,dy + \frac{1}{2}\cdot 2\underbrace{\int_K f(x)\,dx}_{=0}\underbrace{\int_K f(y)\,dy}_{=0}$$
$$= \int_K f\,d\lambda$$

② By fundamental thm of calculus

$$f(x)-f(y) = \int_0^1 \frac{d}{dt} f\big((1-t)x+ty\big)\,dt = \int_0^1 \nabla f\big((1-t)x+ty\big)(y-x)\,dt$$

③ $$|f(x)-f(y)| \leq \operatorname{diam}(K)\int_0^1 |\nabla f((1-t)x+ty)|\,dt$$

since $|x-y| \leq \operatorname{diam}(K) \quad \forall x, y \in K$

We take the squares and apply Cauchy-Schwarz

$$|f(x)-f(y)|^2 \leq \operatorname{diam}^2(K)\int_0^1 |\nabla f((1-t)x+ty)|^2\,dt$$

We integrate over K:

$$\int_K f^2(x)\,d\lambda(x) = \frac{1}{2}\operatorname{diam}^2(K)\int_K\int_K\int_0^1 |\nabla f((1-t)x+ty)|^2\,dx\,dy\,dt$$
$$= \operatorname{diam}^2(K)\int_K\int_K\int_{1/2}^1 |\nabla f((1-t)x+ty)|^2\,dx\,dy\,dt$$

④ Fix x, change variables $z(y) = (1-t)x+ty \qquad \forall t \geq \frac{1}{2}$

$$\Rightarrow dz = t^n\,dy$$

$$\int_K |\nabla f((1-t)x+ty)|\,dy = \int_{(1-t)x+tK} |\nabla f(z)|\frac{dt}{t^n} \leq 2^n\int_K |\nabla f(z)|^2\,dz$$

⑤ $$\int_K f^2\,d\lambda \leq \left\{\int_{1/2}^1 dt\int_K dx\int_K |\nabla f(z)|^2\,dz\right\} 2^n\operatorname{diam}^2(K)$$
$$= \underbrace{2^{n-1}\operatorname{diam}^2(K)}_{=:\,C_P(K)}\int_K |\nabla f(z)|^2\,dz$$

This is the first Poincaré inequality

Poincaré constant

Exercise 8

1 X_t is $\mathcal{F}_t$ - measurable $\Rightarrow$ $f(X_t)$ is $\mathcal{F}_t$ measurable

convex functions are continuous on the interior of their domain and therefore f is Borel-measurable

$f(X_t) \in L^1$ by assumption

Jensen: $\mathbb{E}[f(X_t) | \mathcal{F}_s] \geq f(\mathbb{E}[X_t | \mathcal{F}_s]) \geq f(X_s)$

$\Rightarrow$ $(f(X_t))_{t \geq 0}$ is a submartingale

2 Adaptness and $f(X_t) \in L^1$ as in 1

Jensen: $\mathbb{E}[f(X_t) | \mathcal{F}_s] \geq f(\mathbb{E}[X_t | \mathcal{F}_s]) = f(X_s)$

$\Rightarrow$ $(f(X_t))_{t \geq 0}$ is a submartingale

3 $f(x) = x^2$

$f(X_t) - t$ is a martingale $\Rightarrow$ $f(X_t)$ cannot be a martingale

Exercise 9 This is Proposition 2.4.1 of Vershynin, High-dimensional probability

Exercise 10

Taylor: $\mathbb{E}[e^{\lambda X}] = \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{(\lambda X)^k}{k!}\right]$

$= \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k \mathbb{E}[X^k]$

$= 1 + \sum_{k=2}^{\infty} \frac{\lambda^k}{k!} \mathbb{E}[X^2 X^{k-2}]$

$\leq 1 + \sum_{k=2}^{\infty} \frac{\lambda^k}{k!} M^{k-2} \sigma^2$

$= 1 + \frac{\sigma^2}{M^2} \sum_{k=2}^{\infty} \frac{\lambda^k}{k!} M^k$

$= 1 + \frac{\sigma^2}{M^2} \left(e^{\lambda M} - 1 - \lambda M\right)$

$1 + x \leq e^x \rightarrow \leq \exp\left(\frac{\sigma^2}{M^2}\left(e^{\lambda M} - 1 - \lambda M\right)\right)$

Therefore,

$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq t\right) \leq e^{-\lambda t} \mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n} X_i\right)\right]$

$= e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}[e^{\lambda X_i}]$

$= e^{-\lambda t} \left(\mathbb{E}[e^{\lambda X_1}]\right)^n$

$\leq e^{-\lambda t} \exp\left(\frac{n\sigma^2}{M^2}\left(e^{\lambda M} - 1 - \lambda M\right)\right)$

$= \exp\left\{-\lambda t + \frac{n\sigma^2}{M^2}\left(e^{\lambda M} - 1 - \lambda M\right)\right\}$

$\frac{\partial}{\partial \lambda} \exp\left\{-\lambda t + \frac{n\sigma^2}{M^2}\left(e^{\lambda M} - 1 - \lambda M\right)\right\} = -t + \frac{n\sigma^2}{M^2}\left(M e^{\lambda M} - M\right) = 0$

$\Rightarrow \lambda = \frac{1}{M} \log\left(1 + \frac{tM}{n\sigma^2}\right)$

Plug-in and we get the claim. $\square$

Exercise 11

Let L be the Lipschitz constant of f.

Let $X \sim \mathcal{U}nif([0,1])$ and define the r.v.

$$X_n = 2^{-n} \lfloor 2^n X_n \rfloor$$

$$Z_n = 2^n \left(f(X_n + 2^{-n}) - f(X_n) \right)$$

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

Then (Z_n) is a $\mathcal{F}_n$- martingale bounded by L.

In fact, conditioned on X_n, there are two possible values (X_n and $X_n + 2^{-n-1}$)

which have same probability.

$$\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = 2^{n+1} \left\{ \frac{f(X_n + 2^{-n-1}) - f(X_n)}{2} + \frac{f(X_n + 2^{-n-1} + 2^{-n-1}) - f(X_n + 2^{-n-1})}{2} \right\}$$

$$= 2^n \left\{ f(X_n + 2^{-n}) - f(X_n) \right\}$$

$$= Z_n$$

Z_n converges a.s. and in L^1 to a r.v. Z, which is σ- measurable

Then $\exists g$ measurable function with $Z = g(X)$.

Conditionally on X_n, X is uniformly distributed on $[X_n, X_n + 2^{-n-1})$ and therefore

$$Z_n = 2^n \int_{X_n}^{X_n + 2^{-n}} g(u)\, du$$

$$\Rightarrow \quad f(X_n + 2^{-n}) - f(X_n) = \int_{X_n}^{X_n + 2^{-n}} g(u)\, du$$

Therefore, for all $k < 2^n$ we have

$$f((k+1)2^{-n}) - f(k 2^{-n}) = \int_{k2^{-n}}^{(k+1)2^{-n}} g(u)\, du$$

Summing we get

$$f(k2^{-n}) = f(0) + \int_0^{k2^{-n}} g(u)\,du$$

Let $k2^{-n} \to x$ and by continuity of f we get

$$f(x) = f(0) + \int_0^{x} g(u)\,du \qquad \square$$