Moderate deviation principles for sums of i.i.d. random compact sets

Yu Miao

Abstract. We prove a moderate deviation principle for Minkowski sums of i.i.d. random compact sets in a Banach space.

Keywords: moderate deviation, random sets

Classification: 60D05, 60F10

1. Introduction and main result

We consider a separable Banach space F with norm $\|\cdot\|$. We denote by $\mathcal{K}(F)$ the collection of all nonempty compact subsets of F. For an element A of $\mathcal{K}(F)$, we denote by $\mathrm{co}(A)$ the closed convex hull of A. Mazur's theorem [7] implies that, for A in $\mathcal{K}(F)$, $\mathrm{co}(A)$ belongs to $\mathrm{co}(K(F))$, the collection of the nonempty compact convex subsets of F. The space $\mathcal{K}(F)$ is equipped with the Minkowski addition and the scalar multiplication: for A_1 , A_2 in $\mathcal{K}(F)$ and λ a real number,

$$A_1 + A_2 = \{a_1 + a_2 : a \in A_1, a_2 \in A_2\}, \quad \lambda A_1 = \{\lambda a_1 : a_1 \in A_1\}.$$

The Hausdorff distance

$$d(A_1, A_2) = \max \left\{ \sup_{a_1 \in A_1} \inf_{a_2 \in A_2} \|a_1 - a_2\|, \sup_{a_2 \in A_2} \inf_{a_1 \in A_1} \|a_2 - a_1\| \right\}$$

makes $(\mathcal{K}(F), d)$ a complete separable metric space (i.e., a Polish space). We endow $\mathcal{K}(F)$ with the Borel σ -field associated to the Hausdorff topology. If $A \in \mathcal{K}(F)$, then we shall write $||A|| = d(A, \{0\}) = \sup_{a \in A} ||a||$.

We denote by F^* the topological dual of F and by B^* the unit ball of F^* . The Banach-Alaoglu theorem asserts that B^* endowed with the weak* topology ω^* is compact [11]. Moreover the space $(B^*; \omega^*)$ is separable and metrizable. We denote by $M(B^*)$ the set of Borel signed measures on B^* (the σ -field generated by the weak* topology). Let $(\Omega; \mathcal{F}; \mathbb{P})$ be a probability space. A random compact set of F is a measurable function from Ω to $\mathcal{K}(F)$, i.e., a random variable with

values in $\mathcal{K}(F)$. If A is a random compact convex set (i.e., $\operatorname{co}(\mathcal{K}(F))$ -valued), then $\mathbb{E}A$ is defined as

$$\mathbb{E}A = \{ \mathbb{E}f | f \in L^1(\Omega; \mathcal{F}; \mathbb{P}), f(\omega) \in A \ a.s. \}.$$

Here $f: \Omega \to F$ is a selection of A and $\mathbb{E}f$ denotes the classical expectation (via the Bochner integral). In general $\mathbb{E}A$ may be empty, but if $\mathbb{E}||A|| < \infty$, then $\mathbb{E}A \in \operatorname{co}(\mathcal{K}(F))$. If A is a random compact set, then, by definition

$$\mathbb{E}A = \mathbb{E}(\overline{\operatorname{co}}(A))$$

and so $\mathbb{E}A \in \text{co}(\mathcal{K}(F))$. Here $\overline{\text{co}}(A)$ denotes the closed convex hull of A. We suppose that F is of type p > 1, i.e., there exists a constant c such that

$$\mathbb{E} \| \sum_{i=1}^{n} f_i \|^p \le c \sum_{i=1}^{n} \mathbb{E} \| f_i \|^p$$

for any independent random variables f_1, \ldots, f_n with values in F and mean zero. Every Hilbert space is of type 2; the spaces L^p with $1 are of type <math>\min(p, 2)$. However, the space of continuous functions on [0, 1] equipped with the supremum norm is of type 1 and not of type p for any p > 1.

Denoting by \mathbb{N}^* the set of positive integers, we state our result as follows.

Theorem 1.1. (1) Let $(A_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. random compact convex subsets of F and assume that there exists a positive constant $\delta > 0$ such that

$$\mathbb{E}\exp\left(\delta\|A_1\|\right) = \mathbb{E}\exp\left(\delta\sup_{a\in A_1}\|a\|\right) < \infty.$$

For a measure λ of $M(B^*)$ we set

(1.1)
$$\Lambda(\lambda) = \frac{1}{2} \mathbb{E} \left(\int_{B^*} \sup_{a \in A_1} x^*(a) d\lambda(x^*) \right)^2$$

and for a set $U \in co(\mathcal{K}(F))$,

(1.2)
$$\Lambda^*(U) = \sup_{\lambda \in M(B^*)} \left(\int_{B^*} \sup_{x \in U} x^*(x) d\lambda(x^*) - \Lambda(\lambda) \right).$$

For a nonconvex set U in $\mathcal{K}(F)$ we set $\Lambda^*(U) = +\infty$. Moreover, suppose that the moderate deviation scale (b_n) is a sequence of positive numbers satisfying $1 \ll b_n \ll n$, i.e., as $n \to \infty$,

$$b_n \to \infty; \quad \frac{b_n}{\sqrt{n}} \to 0,$$

and

$$S_n := \frac{\sum_{i=1}^n (A_i - \mathbb{E}A_i)}{b_n \sqrt{n}} \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

(2) Let $(A_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. random compact sets of F and assume that there exist positive constant $\delta>0$ and $\delta_0>0$ such that

$$\mathbb{E}\exp\left(\delta\sup_{a\in A_1}\|a\|^{1+\delta_0}\right)<\infty.$$

Moreover, suppose that the moderate deviation scale (b_n) satisfies,

(1.3)
$$b_n \to \infty; \quad \frac{b_n}{\sqrt{n}} \to 0, \quad \frac{\log n}{b_n^2} \to 0, \quad \frac{b_n^{2+(1-\frac{2}{1+\delta_0})p}}{n^{1-p/2}} \to \infty.$$

as $n \to \infty$.

If assumption (1) or (2) holds, then the law of the random set S_n satisfies the large deviation principle with speed b_n^2 and rate function Λ^* , i.e., for any subset \mathcal{U} of $\mathcal{K}(F)$,

$$-\inf_{U\in\mathcal{U}^o}\Lambda^*(U) \le \liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}(S_n \in \mathcal{U})$$

$$\le \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}(S_n \in \mathcal{U}) \le -\inf_{U\in\mathcal{U}}\Lambda^*(U),$$

where \mathcal{U}^o and $\overline{\mathcal{U}}$ are the interior and the closure of \mathcal{U} with respect to the Hausdorff topology.

Remarks 1.2. Here the rate function Λ^* is a "good" rate function, i.e., it is lower semi-continuous and its level sets $\{U \in \mathcal{K}(F) : \Lambda^*(U) \leq \lambda\}, \lambda \in \mathbb{R}_+$, are compact.

Remarks 1.3. From the following proof, for any sequence of i.i.d. random compact convex subsets of F, the proof of moderate deviation principle is not difficult. However, if one wants to remove the convexity assumption, it is not easy. In some sense, assumption (2) is technical.

2. Proof of Theorem 1.1

Without loss of generality, in what follows we assume that $S_n = A_1 + \cdots + A_n$, where $\mathbb{E}A_1 = 0$. We first collect several results which are the main ingredients in the proof of Theorem 1.1.

Embedding theorem (see [4]). With a compact convex subset A of F we associate its support function $s_A : B^* \to \mathbb{R}$ defined by

$$\forall x^* \in B^*, \quad s_A(x^*) = \sup\{x^*(x) : x \in A\}.$$

We denote by $C(B^*, \omega^*)$ the set of all continuous functions on B^* endowed with the weak* topology. With the uniform norm $\|\cdot\|_{\infty}$, $C(B^*, \omega^*)$ is a separable Banach space (for f in $C(B^*, \omega^*)$, $\|f\|_{\infty} = \sup_{x^* \in B^*} |f(x^*)|$). Whenever A is compact, its support function s_A belongs to $C(B^*, \omega^*)$. The map $s : \operatorname{co}(\mathcal{K}(F)) \to C(B^*, \omega^*)$ has the following properties. For any A_1 , A_2 in $\operatorname{co}(\mathcal{K}(F))$ and $t \in \mathbb{R}^+$,

$$\begin{split} s_{A_1} = s_{A_2} \Leftrightarrow A_1 = A_2, & A_1 \subset A_2 \Leftrightarrow s_{A_1} \leq s_{A_2}, \\ s_{A_1 + A_2} = s_{A_1} + s_{A_2}, & s_{tA_1} = ts_{A_1} \end{split}$$

and finally $d(A_2, A_2) = \|s_{A_1} - s_{A_2}\|_{\infty}$. Hence, $\operatorname{co}(\mathcal{K}(F))$ is algebraically and topologically isomorphic to its image under s, $s(\operatorname{co}(\mathcal{K}(F)))$, which is a subset of the separable Banach space $C(B^*, \omega^*)$. This embedding theorem was used in [1] and [8] to prove limit theorems for random sets.

Moderate deviation principle. We state here a moderate deviation principle (see Chen [2], [3]).

Let E be a separable Banach space, E_1 a closed convex subset of E and let E^* denote the topological dual of E. Given an E_1 valued random variable X, we write $X \in WM_0^2$ if $\mathbb{E}\lambda(X) = 0$ and $\mathbb{E}\lambda^2(X) < \infty$ for all $\lambda \in E^*$. Let $(X_n)_{n \in N^*}$ be a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in E_1 and $X_1 \in WM_0^2$, and set $S_n = (X_1 + \dots + X_n)/(b_n\sqrt{n})$. Suppose that there exists a constant $\delta > 0$, such that

$$\mathbb{E}\exp\left(\delta\|X_1\|\right)<\infty,$$

and

$$S_n \stackrel{\mathbb{P}}{\longrightarrow} 0$$
 (in probability).

Then, S_n satisfies the moderate deviation principle with some rate function Λ^* , namely, for any $U \subset E$,

$$-\inf_{x\in U^o} \Lambda_E^*(x) \le \liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}(S_n \in U)$$

$$\le \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}(S_n \in U) \le -\inf_{x\in \overline{U}} \Lambda_E^*(x),$$

where the rate function

$$\Lambda_E^*(x) = \sup_{\lambda \in E^*} \{\lambda(x) - \Lambda_E(\lambda)\},\,$$

and

$$\Lambda_E(\lambda) = \frac{1}{2} \mathbb{E} \lambda^2(X_1).$$

Note that Λ^* is a good rate function.

Distance to the convex hull (see [4], [9]). Let A belong to $\mathcal{K}(F)$ and denote its inner radius by

$$r(A) = \sup_{a \in co(A)} \inf\{R : \exists a_1, \dots, a_s \in A, a \in co(a_1, \dots, a_s), ||a - a_i|| \le R, 1 \le i \le s\}.$$

Obviously, r(A) is zero if and only if A is convex. For any A, $r(A) \leq 2||A|| = 2 \sup_{a \in A} ||a||$. For any A_1, \ldots, A_n in $\mathcal{K}(F)$,

$$(2.1) \ d(A_1 + \dots + A_n, \operatorname{co}(A_1) + \dots + \operatorname{co}(A_n)) \le c^{1/p} (r(A_1)^p + \dots + r(A_n)^p)^{1/p}.$$

Of course, the exponent p is related to the fact that F is a Banach space of type p and the constant c is the one appearing in the functional inequality (see the definition just before Theorem 1.1).

PROOF OF THEOREM 1.1: We will divide our proof into three steps as follows.

Step 1. We suppose first that the sets $(A_n)_{n\in\mathbb{N}^*}$ are convex. We apply the moderate deviation principle with $E=C(B^*,\omega^*),\ E_1=s(\operatorname{co}(\mathcal{K}(F)))$ and the sequence of random functions $(s_{A_n})_{n\in\mathbb{N}^*}$. By the Riesz representation theorem [10], the topological dual of E is the set $M(B^*)$ of the signed Borel measures on (B^*,ω^*) . By the hypothesis of Theorem 1.1, there exists a positive constant $\delta>0$, such that

$$\mathbb{E}\exp\left(\delta\|s_{A_1}\|_{\infty}\right) = E\exp\left(\delta\sup_{a\in A_1}\|a\|\right) < \infty$$

so that the law of $(s_{A_1} + \cdots + s_{A_n})/(b_n\sqrt{n})$ satisfies the moderate deviation principle with rate function Λ_E^* (defined on E). We push back this moderate deviation principle to the space $\operatorname{co}(\mathcal{K}(F))$ with the help of the homeomorphism s. Since for any U in $\operatorname{co}(\mathcal{K}(F))$, $\Lambda^*(U) = \Lambda_E^*(s_U)$ (where Λ^* is the rate function on $\mathcal{K}(F)$ defined in Theorem 1.1), we obtain that for any $\mathcal{U} \subseteq \operatorname{co}(\mathcal{K}(F))$,

$$-\inf_{U\in\mathcal{U}_{co}^{o}}\Lambda^{*}(U) \leq \liminf_{n\to\infty} \frac{1}{b_{n}^{2}}\log \mathbb{P}(S_{n}\in\mathcal{U})$$

$$\leq \limsup_{n\to\infty} \frac{1}{b_{n}^{2}}\log \mathbb{P}(S_{n}\in\mathcal{U}) \leq -\inf_{U\in\overline{\mathcal{U}}_{co}}\Lambda^{*}(U),$$

where \mathcal{U}_{co}^{o} and $\overline{\mathcal{U}}_{co}$ are the interior and the closure of \mathcal{U} for the topology induced by the Hausdorff metric on $co(\mathcal{K}(F))$.

Step 2. In the general case, where the sets $(A_n)_{n\in\mathbb{N}^*}$ are not necessarily convex, we set $S_n=(A_1+\cdots+A_n)/(b_n\sqrt{n})$ and $S^{\text{co}}=(\cos(A_1)+\cdots+\cos(A_n))/(b_n\sqrt{n})$. We need show the following claim.

For any r > 0,

(2.2)
$$\lim_{n \to \infty} \frac{1}{b_n^2} \log \mathbb{P}(d(S_n, S_n^{\text{co}}) \ge r) = -\infty.$$

We first apply the inequality (2.1),

$$\mathbb{P}(d(S_n, S_n^{co} \ge r)) \le \mathbb{P}(c^{1/p}(r(A_1)^p + \dots + r(A_n)^p)^{1/p} \ge rb_n\sqrt{n}).$$

Notice that this inequality requires the assumption that the space F is of type p. By (1.3), there exist a positive increasing sequence of a_n and a positive constant β , such that

$$p - \beta = 1 + \delta_0$$
, $b_n^{\frac{2p}{1+\delta_0}-2} \ll a_n \ll b_n^2 n^{p/2-1}$.

That is,

(2.3)
$$b_n^{2\beta/(p-\beta)} \ll a_n \ll b_n^2 n^{p/2-1}$$

Thus we have

$$\mathbb{P}(c^{1/p}(r(A_1)^p + \dots + r(A_n)^p)^{1/p} \ge rb_n\sqrt{n})
= \mathbb{P}\left(r(A_1)^p + \dots + r(A_n)^p \ge (rb_n\sqrt{n})^p/c\right)
\le \mathbb{P}\left(\max_{1\le i\le n} r(A_i)^\beta \left[r(A_1)^{p-\beta} + \dots + r(A_n)^{p-\beta}\right] \ge (rb_n\sqrt{n})^p a_n/(ca_n)\right)
\le \mathbb{P}\left(\max_{1\le i\le n} r(A_i)^\beta \ge a_n\alpha\right) + \mathbb{P}\left(r(A_1)^{p-\beta} + \dots + r(A_n)^{p-\beta} \ge \frac{(rb_n\sqrt{n})^p}{(\alpha ca_n)}\right)
\le n\mathbb{P}\left(r(A_1)^{p-\beta} \ge (a_n\alpha)^{(p-\beta)/\beta}\right)
+ \mathbb{P}\left(r(A_1)^{p-\beta} + \dots + r(A_n)^{p-\beta} \ge \frac{(rb_n\sqrt{n})^p}{(\alpha ca_n)}\right).$$

Since $r(A_1) \leq 2||A_1|| = 2 \sup_{a \in A_1} ||a||$, by Markov inequality, we have

$$\begin{split} &\frac{1}{b_n^2} \log \mathbb{P}\left(r(A_1)^{p-\beta} \geq (a_n \alpha)^{(p-\beta)/\beta}\right) \\ &\leq \frac{1}{b_n^2} \log \left(\exp\left(-\lambda (a_n \alpha)^{(p-\beta)/\beta}\right) \mathbb{E} \exp\left(2\lambda \sup_{a \in A_1} \|a\|\right)\right) \\ &\leq \frac{-\lambda (a_n \alpha)^{(p-\beta)/\beta}}{b_n^2} + \frac{1}{b_n^2} \log \mathbb{E} \exp\left(2\lambda \sup_{a \in A_1} \|a\|\right) \end{split}$$

where $\lambda > 0$ and $2\lambda \leq \delta$. Thus from (2.3), we have

$$\lim_{n\to\infty}\frac{1}{b_n^2}\log n\mathbb{P}\left(r(A_1)^{p-\beta}\geq (a_n\alpha)^{(p-\beta)/\beta}\right)=-\infty.$$

By the independence of A_i , i = 1, ..., n and Markov inequality, we have

(2.4)
$$\frac{1}{b_n^2} \log \mathbb{P}\left(r(A_1)^{p-\beta} + \dots + r(A_n)^{p-\beta} \ge (rb_n\sqrt{n})^p/(\alpha ca_n)\right) \\ \le \frac{n}{b_n^2} \log \mathbb{E} \exp\left(2\lambda \sup_{a \in A_1} \|a\|^{p-\beta}\right) - \frac{\lambda(rb_n\sqrt{n})^p}{b_n^2 a_n c\alpha}.$$

From (2.3) we have

$$(2.5) \quad \lim_{n \to \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(r(A_1)^{p-\beta} + \dots + r(A_n)^{p-\beta} \ge (rb_n\sqrt{n})^p/(\alpha ca_n)\right) = -\infty.$$

By (2.4) and (2.5), we obtain the claim (2.2).

(Note that, taking
$$\alpha$$
 small enough and $a_n = \frac{b^p}{n^{1-p/2}}$, (2.5) still holds.)

Step 3. (Lower bound) Let \mathcal{U} be a subset of $\mathcal{K}(F)$. Let U belong to \mathcal{U}^o (if $\mathcal{U}^o \cap \operatorname{co}(\mathcal{K}(F))$) is empty, the proof is trivial). Then there exists $\gamma > 0$ such that

$${V \in \mathcal{K}(F) : d(U, V) < \gamma} \subset \mathcal{U}.$$

Then we have

$$\mathbb{P}(S_n \in \mathcal{U}) \ge \mathbb{P}(d(S_n, U) < \gamma) \ge \mathbb{P}(d(S_n^{\text{co}}, U) < \gamma/2, d(S_n, S_n^{\text{co}}) < \gamma/2)$$
$$\ge \mathbb{P}(d(S_n^{\text{co}}, U) < \gamma/2) - \mathbb{P}(d(S_n, S_n^{\text{co}}) \ge \gamma/2).$$

Applying the claim (2.2) and the moderate deviation principle for $(S_n^{\text{co}})_{n \in \mathbb{N}^*}$, we get

$$\liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}(S_n \in \mathcal{U}) \ge -\Lambda^*(U).$$

Taking the supremum over all sets U in interior (\mathcal{U}) yields the desired lower bound.

(Upper bound) Let \mathcal{U} be a subset of $\mathcal{K}(F)$. For any $\gamma > 0$ we set $\mathcal{U}^{\gamma} = \{A \in \mathcal{K}(F) : d(A,\mathcal{U}) \leq \gamma\}$. We then write

$$\mathbb{P}(S_n \in \mathcal{U}) \leq \mathbb{P}(S_n^{\text{co}} \in \mathcal{U}^{\gamma}) + \mathbb{P}(d(S_n, S_n^{\text{co}}) > \gamma).$$

Applying the claim (2.2) and the moderate deviation principle for $(S_n^{co})_{n \in N^*}$, we get

$$\liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}(S_n \in \mathcal{U}) \le -\inf\{\Lambda^*(U) : U \in \overline{\mathcal{U}^{\gamma}}_{co}\}.$$

However $\overline{\mathcal{U}^{\gamma}}_{co} = \mathcal{U}^{\gamma} \bigcap co(\mathcal{K}(F))$ and in addition, $\bigcap_{\gamma>0} \overline{\mathcal{U}^{\gamma}}_{co} = \overline{\mathcal{U}} \bigcap co(\mathcal{K}(F))$. Since Λ^* is a good rate function we have that

$$\lim_{\gamma \to 0} \inf \{ \Lambda^*(U) : U \in \overline{\mathcal{U}^{\gamma}}_{\operatorname{co}} \} = \inf \{ \Lambda^*(U) : U \in \overline{\mathcal{U}} \cap \operatorname{co}(\mathcal{K}(F)) \}.$$

The right-hand side is clearly larger than the left-hand side; let $(U_n)_{n\in N^*}$ be a sequence such that $U_n \in \overline{\mathcal{U}^{1/n}}_{\operatorname{co}}$ for all n and $\Lambda^*(U_n)$ converges to the left-hand side. The level sets of Λ^* being compact, we can extract from $(U_n)_{n\in N^*}$ a subsequence converging to a set U which necessarily belongs to $\overline{\mathcal{U}} \cap \operatorname{co}(\mathcal{K}(F))$. By the lower semi-continuity of Λ^* , $\Lambda^*(U)$ is smaller than the left-hand side. Thus, letting γ go to zero in the previous inequality gives the desired upper bound. \square

Acknowledgments. The author wants to thank Prof. M.L. Puri for providing him the key reference paper [9].

References

- [1] Artstein Z., Vitale R.A., A strong law of large numbers for random compact sets, Ann. Probability 3 (1975), 879–882.
- [2] Chen X., Probabilities of moderate deviations for independent random vectors in a Banach space, Chinese J. Appl. Prob. Stat. 7 (1991), 24–32.
- [3] Chen X., Moderate deviations for m-dependent random variables with Banach space values, Statist. Probab. Lett. 35 (1997), 123–134.
- [4] Cerf R., Large deviations for sums of i.i.d. random compact sets, Proc. Amer. Math. Soc. 127 (1999), no. 8, 2431–2436.
- [5] Dembo A., Zeitouni O., Large Deviations Techniques and Applications, second edition, Springer, New York, 1998.
- [6] Deuschel J.D., Stroock D.W., Large Deviations, Pure and Applied Mathematics, 137, Academic Press, Boston, MA, 1989.
- [7] Dunford N., Schwartz J.T., Linear Operators. Part I: General Theory, Interscience Publishers, New York-London, 1958.
- [8] Giné E., Hahn M.G., Zinn J., Limit theorems for random sets: an application of probability in Banach space results, Probability in Banach Spaces IV (Oberwolfach, 1982), Springer, Berlin, 1983, pp. 112–135.
- [9] Puri M.L., Ralescu D.A., Limit theorems for random compact sets in Banach space, Math. Proc. Cambridge Philos. Soc. 97 (1985), 151–158.
- [10] Rudin W., Real and Complex Analysis, McGraw-Hill, New York, 1966.

- [11] Rudin W., Functional Analysis, McGraw-Hill, New York, 1973.
- [12] Wu L.M., An introduction to large deviations, in Several Topics in Stochastic Analysis (J.A. Yan, S. Peng, S. Fang and L. Wu, Eds.), pp. 225–336, Academic Press of China, Beijing, 1997 (in Chinese).

College of Mathematics and Information Science, Henan Normal University, 453007 Henan, China

E-mail: yumiao728@yahoo.com.cn

(Received June 15, 2008)