Sums of independent random variables

This lecture collects a number of estimates for sums of independent random variables with values in a Banach space E. We concentrate on sums of the form $\sum_{n=1}^{N} \gamma_n x_n$, where the γ_n are real-valued Gaussian variables and the x_n are vectors in E. As we shall see later on such sums are the building blocks of general E-valued Gaussian random variables and, perhaps more importantly, stochastic integrals of E-valued step functions are of this form. Furthermore, they are used in the definition of various geometric properties of Banach spaces, such as type and cotype.

The highlights of this lecture are the Kahane contraction principle (Theorem 3.1), a covariance domination principle (Theorem 3.9) and the Kahane-Khintchine inequalities (Theorems 3.11 and 3.12).

3.1 Gaussian sums

We begin with an important inequality for sums of independent symmetric random variables, due to KAHANE.

Theorem 3.1 (Kahane contraction principle). Let $(X_n)_{n=1}^{\infty}$ be a sequence of independent symmetric *E*-valued random variables. Then for all $a_1, \ldots, a_N \in \mathbb{R}$ and $1 \leq p < \infty$,

$$\mathbb{E} \bigg\| \sum_{n=1}^{N} a_n X_n \bigg\|^p \leq \big(\max_{1 \leq n \leq N} |a_n| \big)^p \mathbb{E} \bigg\| \sum_{n=1}^{N} X_n \bigg\|^p.$$

Proof. For all $(\varepsilon_1, \ldots, \varepsilon_N) \in \{-1, +1\}^N$ the E^N -valued random variables (X_1, \ldots, X_N) and $(\varepsilon_1 X_1, \ldots, \varepsilon_N X_N)$ are identically distributed and therefore

$$\mathbb{E} \Big\| \sum_{n=1}^{N} \varepsilon_n X_n \Big\|^p = \mathbb{E} \Big\| \sum_{n=1}^{N} X_n \Big\|^p.$$

For the general case we may assume that $|a_n| \leq 1$ for all n = 1, ..., N. Then $a = (a_1, ..., a_N)$ is a convex combination of the 2^N elements of $\{-1, +1\}^N$, say $a = \sum_{j=1}^{2^N} \lambda^{(j)} \varepsilon^{(j)}$. Hence,

$$\mathbb{E} \left\| \sum_{n=1}^{N} a_n X_n \right\|^p = \mathbb{E} \left\| \sum_{j=1}^{2^N} \lambda^{(j)} \sum_{n=1}^{N} \varepsilon_n^{(j)} X_n \right\|^p$$
$$\leq \mathbb{E} \left(\sum_{j=1}^{2^N} \lambda^{(j)} \right\| \sum_{n=1}^{N} \varepsilon_n^{(j)} X_n \right\|^p$$
$$\leq \mathbb{E} \sum_{j=1}^{2^N} \lambda^{(j)} \left\| \sum_{n=1}^{N} \varepsilon_n^{(j)} X_n \right\|^p$$
$$= \sum_{j=1}^{2^N} \lambda^{(j)} \mathbb{E} \left\| \sum_{n=1}^{N} X_n \right\|^p = \mathbb{E} \left\| \sum_{n=1}^{N} X_n \right\|^p,$$

where the third step follows from the convexity of the function $t \mapsto t^p$ (or an application of Jensen's inequality).

As an application of the Kahane contraction principle we shall prove an inequality which shows that Rademacher sums have the 'smallest' L^p -norms among all random sums. Rademacher sums are easier to handle than the Gaussian sums in which we are ultimately interested, and, as we shall see, there are various techniques to pass on results for Rademacher sums to Gaussian sums.

Let us begin with a definition. An $\{-1, +1\}$ -valued random variable r is called a *Rademacher variable* if

$$\mathbb{P}\{r = -1\} = \mathbb{P}\{r = +1\} = \frac{1}{2}$$

Throughout these lectures, the notation $(r_n)_{n=1}^{\infty}$ will be used for a *Rademacher* sequence, that is, a sequence of independent Rademacher variables.

Theorem 3.2 (Comparison). Let $(\varphi_n)_{n=1}^{\infty}$ be a sequence of independent symmetric integrable real-valued random variables satisfying $\mathbb{E}|\varphi_n| \ge 1$ for all $n \ge 1$. Then for all $x_1, \ldots, x_N \in E$ and $1 \le p < \infty$ we have

$$\mathbb{E} \Big\| \sum_{n=1}^{N} r_n x_n \Big\|^p \leqslant \mathbb{E} \Big\| \sum_{n=1}^{N} \varphi_n x_n \Big\|^p.$$

The proof of this theorem relies on an auxiliary lemma, for which we need two definitions based on the following easy observation: if X_1, \ldots, X_N are random variables with values in E_1, \ldots, E_N , then (X_1, \ldots, X_N) is a random variable with values in $E_1 \times \cdots \times E_N$. **Definition 3.3.** Two families of random variables $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$, where I is some index set and X_i and Y_i take values in a Banach space E_i , are identically distributed if for all choices of $i_1, \ldots, i_N \in I$ the random variables $(X_{i_1}, \ldots, X_{i_N})$ and $(Y_{i_1}, \ldots, Y_{i_N})$ are identically distributed.

Note that by Proposition 2.13, if $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$ are families of independent random variables such that X_i and Y_i are identically distributed for all $i \in I$, then $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$ are identically distributed.

Definition 3.4. Two families of random variables $(X_i)_{i \in I}$ and $(Y_j)_{j \in J}$, where I and J are index sets, X_i takes values in E_i for all $i \in I$ and Y_j takes values in F_j for all $j \in J$, are independent of each other if for all choices $i_1, \ldots, i_M \in I$ and $j_1, \ldots, j_N \in I$ the random variables $(X_{i_1}, \ldots, X_{i_M})$ and $(Y_{j_1}, \ldots, Y_{i_N})$ are independent.

Lemma 3.5. Let $(\varphi_n)_{n=1}^{\infty}$ be a sequence of independent symmetric real-valued random variables and let $(r_n)_{n=1}^{\infty}$ be a Rademacher sequence independent of $(\varphi_n)_{n=1}^{\infty}$. The sequences $(\varphi_n)_{n=1}^{\infty}$ and $(r_n|\varphi_n|)_{n=1}^{\infty}$ are identically distributed.

Proof. By independence and symmetry we have

$$\begin{split} & \mathbb{P}\{r_{n}|\varphi_{n}| \in B\} \\ &= \mathbb{P}\{r_{n} = 1, \ \varphi_{n} \ge 0, \ \varphi_{n} \in B\} + \mathbb{P}\{r_{n} = 1, \ \varphi_{n} < 0, \ \varphi_{n} \in -B\} \\ &+ \mathbb{P}\{r_{n} = -1, \ \varphi_{n} \ge 0, \ \varphi_{n} \in -B\} + \mathbb{P}\{r_{n} = -1, \ \varphi_{n} < 0, \ \varphi_{n} \in B\} \\ &= \frac{1}{2}\mathbb{P}\{\varphi_{n} \ge 0, \ \varphi_{n} \in B\} + \frac{1}{2}\mathbb{P}\{\varphi_{n} < 0, \ \varphi_{n} \in -B\} \\ &+ \frac{1}{2}\mathbb{P}\{\varphi_{n} \ge 0, \ \varphi_{n} \in -B\} + \frac{1}{2}\mathbb{P}\{\varphi_{n} < 0, \ \varphi_{n} \in B\} \\ &= \frac{1}{2}\mathbb{P}\{\varphi_{n} \ge 0, \ \varphi_{n} \in B\} + \frac{1}{2}\mathbb{P}\{\varphi_{n} > 0, \ \varphi_{n} \in B\} \\ &+ \frac{1}{2}\mathbb{P}\{\varphi_{n} \leqslant 0, \ \varphi_{n} \in B\} + \frac{1}{2}\mathbb{P}\{\varphi_{n} < 0, \ \varphi_{n} \in B\} \\ &= \mathbb{P}\{\varphi_{n} \in B\}. \end{split}$$

Since $(\varphi_n)_{n=1}^{\infty}$ and $(r_n|\varphi_n|)_{n=1}^{\infty}$ are sequences of independent random variables, the lemma now follows from the observation preceding Definition 3.4.

Proof (Proof of Theorem 3.2). We may assume that the sequences $(\varphi_n)_{n=1}^{\infty}$ and $(r_n)_{n=1}^{\infty}$ are defined on distinct probability spaces Ω_{φ} and Ω_r . By considering the φ_n and r_n as random variables on the probability space $\Omega_{\varphi} \times \Omega_r$, we may assume that $(\varphi_n)_{n=1}^{\infty}$ and $(r_n)_{n=1}^{\infty}$ are independent of each other.

Since $\mathbb{E}_{\varphi}|\varphi_n| \ge 1$, with the Kahane contraction principle and Jensen's inequality we obtain

$$\mathbb{E}_{r} \left\| \sum_{n=1}^{N} r_{n} x_{n} \right\|^{p} \leq \mathbb{E}_{r} \left\| \mathbb{E}_{\varphi} \sum_{n=1}^{N} r_{n} |\varphi_{n}| x_{n} \right\|^{p} \\ \leq \mathbb{E}_{r} \mathbb{E}_{\varphi} \left\| \sum_{n=1}^{N} r_{n} |\varphi_{n}| x_{n} \right\|^{p} = \mathbb{E}_{\varphi} \left\| \sum_{n=1}^{N} \varphi_{n} x_{n} \right\|^{p},$$

32 3 Sums of independent random variables

where the last identity follows from Lemma 3.5.

A real-valued random variable γ is called standard Gaussian if its distribution has density

$$f_{\gamma}(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2)$$

with respect to the Lebesgue measure on \mathbb{R} . For later reference we note that γ is standard Gaussian if and only if its Fourier transform is given by

$$\mathbb{E}\exp(-i\xi\gamma) = \exp(-\frac{1}{2}\xi^2), \quad \xi \in \mathbb{R}.$$
(3.1)

The 'only if' statement follows from the identity

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-i\xi t - \frac{1}{2}t^2) \, dt = \exp(-\frac{1}{2}\xi^2)$$

which can be proved by completing the squares in the exponential and then shifting the path of integration from $i\xi + \mathbb{R}$ to \mathbb{R} by using Cauchy's formula; the 'if' part then follows from the injectivity of the Fourier transform (Theorem 2.8).

For a standard Gaussian random variable γ we have

$$\mathbb{E}|\gamma| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |t| \exp(-\frac{1}{2}t^2) dt = \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} t \exp(-\frac{1}{2}t^2) dt = \sqrt{2/\pi}.$$
 (3.2)

From this point on, $(\gamma_n)_{n=1}^{\infty}$ will always denote a *Gaussian sequence*, that is, a sequence of independent standard Gaussian variables.

From (3.2) and Theorem 3.2 we obtain the following comparison result.

Corollary 3.6. For all $x_1, \ldots, x_N \in E$ and $1 \leq p < \infty$,

$$\mathbb{E} \left\| \sum_{n=1}^{N} r_n x_n \right\|^p \leqslant (\pi/2)^{\frac{p}{2}} \mathbb{E} \left\| \sum_{n=1}^{N} \gamma_n x_n \right\|^p.$$
(3.3)

The geometric notions of *type* and *cotype* will be introduced in the exercises. Without proof we state the following important converse to Corollary 3.6 for Banach spaces with finite cotype. Examples of spaces with finite cotype are Hilbert spaces, L^p -spaces for $1 \leq p < \infty$, and the UMD spaces which will be introduced in later lectures.

Theorem 3.7. If E has finite cotype, there exists a constant $C \ge 0$ such that for all $x_1, \ldots, x_N \in E$,

$$\mathbb{E} \left\| \sum_{n=1}^{N} \gamma_n x_n \right\|^2 \leqslant C^2 \mathbb{E} \left\| \sum_{n=1}^{N} r_n x_n \right\|^2.$$

The Kahane-Khintichine inequalities (Theorems 3.11 and 3.12 below) can be used to extend this inequality to arbitrary exponents $1 \leq p < \infty$.

The proof of Theorem 3.7 is beyond the scope of these lectures; we refer to the Notes at the end of the lecture for references to the literature. When taken together, Corollary 3.6 and Theorem 3.7 show that in spaces with finite cotype, Gaussian sequences and Rademacher sums can be used interchangeably.

Without any assumptions on E, Theorem 3.7 fails. This is shown by the next example.

Example 3.8. Let $E = c_0$ and let $(u_n)_{n=1}^{\infty}$ be the standard unit basis of c_0 . Then

$$\mathbb{E} \Big\| \sum_{n=1}^{N} r_n u_n \Big\|_{c_0} = \mathbb{E} \Big(\max_{1 \le n \le N} |r_n| \Big) = 1.$$

Next we estimate $\mathbb{E} \| \sum_{n=1}^{N} \gamma_n u_n \|_{c_0}$ from below. First, if γ is standard Gaussian, the inequality $1 - x \leq e^{-x}$ implies

$$\mathbb{P}\left\{\max_{1\leqslant n\leqslant N}|\gamma_n|\leqslant r\right\} = \left[1-\mathbb{P}\{|\gamma|>r\}\right]^N \leqslant \exp(-N\mathbb{P}\{|\gamma|>r\}).$$

For $r = \frac{1}{2}\sqrt{\log N}$ we estimate

$$\mathbb{P}\{|\gamma| > \frac{1}{2}\sqrt{\log N}\} \ge \frac{2}{\sqrt{2\pi}} \int_{\frac{1}{2}\sqrt{\log N}}^{\sqrt{\log N}} e^{-\frac{1}{2}x^2} dx$$
$$\ge \frac{2}{\sqrt{2\pi}} \cdot \frac{1}{2}\sqrt{\log N} \cdot e^{-\frac{1}{2}\log N} = \sqrt{\frac{\log N}{2\pi N}}$$

Hence, using the integration by parts formula of Exercise 2.1,

$$\mathbb{E} \left\| \sum_{n=1}^{N} \gamma_n u_n \right\|_{c_0} = \mathbb{E} \left(\max_{1 \leq n \leq N} |\gamma_n| \right)$$
$$= \int_0^\infty \mathbb{P} \left\{ \max_{1 \leq n \leq N} |\gamma_n| > r \right\} dr$$
$$\geqslant \int_0^{\frac{1}{2}\sqrt{\log N}} \left[1 - \exp(-N\mathbb{P}\{|\gamma| > r\}) \right] dr$$
$$\geqslant \frac{1}{2}\sqrt{\log N} \cdot \left[1 - \exp\left(-\sqrt{\frac{N\log N}{2\pi}}\right) \right]$$
$$= \frac{1}{2}\sqrt{\log N} \quad \text{as} \quad N \to \infty.$$

Similar estimates show that the bound $\mathscr{O}(\sqrt{\log N})$ for $N \to \infty$ is of the correct order.

We conclude this section with an important comparison result for Gaussian sums.

Theorem 3.9 (Covariance domination). Let $(\gamma_m)_{m=1}^{\infty}$ and $(\gamma'_n)_{n=1}^{\infty}$ be Gaussian sequences on probability spaces Ω and Ω' , respectively, and let x_1, \ldots, x_M and y_1, \ldots, y_N be elements of E satisfying

$$\sum_{m=1}^{M} \langle x_m, x^* \rangle^2 \leqslant \sum_{n=1}^{N} \langle y_n, x^* \rangle^2 \qquad \forall x^* \in E^*.$$

Then, for all $1 \leq p < \infty$,

$$\mathbb{E} \Big\| \sum_{m=1}^{M} \gamma_m x_m \Big\|^p \leqslant \mathbb{E}' \Big\| \sum_{n=1}^{N} \gamma'_n y_n \Big\|^p.$$

Proof. Denote by F the linear span of $\{x_1, \ldots, x_M, y_1, \ldots, y_N\}$ in E. Define $Q \in \mathscr{L}(F^*, F)$ by

$$Qz^* := \sum_{n=1}^N \langle y_n, z^* \rangle y_n - \sum_{m=1}^M \langle x_m, z^* \rangle x_m, \qquad z^* \in F^*.$$

The assumption of the theorem implies that $\langle Qz^*, z^* \rangle \ge 0$ for all $z^* \in F^*$, and it is clear that $\langle Qz_1^*, z_2^* \rangle = \langle Qz_2^*, z_1^* \rangle$ for all $z_1^*, z_2^* \in F^*$. Since F is finitedimensional, by linear algebra we can find a sequence $(x_j)_{j=M+1}^{M+k}$ in F such that Q is represented as

$$Qz^* = \sum_{j=M+1}^{M+k} \langle x_j, z^* \rangle x_j, \quad z^* \in F^*.$$

We leave the verification of this statement as an exercise for the moment and shall return to this issue from a more general point of view in the next lecture. Now,

$$\sum_{m=1}^{M+k} \langle x_m, z^* \rangle^2 = \sum_{n=1}^N \langle y_n, z^* \rangle^2, \qquad z^* \in F^*.$$
(3.4)

It follows from (3.1) that the random variables $X := \sum_{m=1}^{M+k} \gamma_m x_m$ and $Y := \sum_{n=1}^{N} \gamma'_n y_n$ have Fourier transforms

$$\mathbb{E}\exp(-i\langle X, x^*\rangle) = \prod_{m=1}^{M+k} \mathbb{E}\exp(-i\gamma_m \langle x_m, x^*\rangle)$$
$$= \prod_{m=1}^{M+k} \exp\left(-\frac{1}{2}\langle x_m, x^*\rangle^2\right) = \exp\left(-\frac{1}{2}\sum_{m=1}^{M+k} \langle x_m, x^*\rangle^2\right)$$

and similarly $\mathbb{E}' \exp(-i\langle Y, x^* \rangle) = \exp(-\frac{1}{2}\sum_{n=1}^N \langle y_n, x^* \rangle^2)$. Hence by (3.4) and Theorem 2.8, X and Y are identically distributed. Thus, for all $1 \leq p < \infty$,

$$\mathbb{E}\Big\|\sum_{m=1}^{M+k}\gamma_m x_m\Big\|^p = \mathbb{E}'\Big\|\sum_{n=1}^N\gamma'_n y_n\Big\|^p.$$

By Proposition 2.16,

$$\mathbb{E} \left\| \sum_{m=1}^{M} \gamma_m x_m \right\|^p \leqslant \mathbb{E} \left\| \sum_{m=1}^{M+k} \gamma_m x_m \right\|^p,$$

and the proof is complete.

3.2 The Kahane-Khintchine inequality

The main result of this section states that all L^p -norms of an E-valued Gaussian sum are comparable, with universal constants depending only on p. First we prove the analogous result for Rademacher sums; then we use the central limit theorem to pass it on to Gaussian sums.

The starting point is the following inequality, which is a consequence of Lévy's inequality.

Lemma 3.10. For all $x_1, \ldots, x_N \in E$ and r > 0 we have

$$\mathbb{P}\Big\{\Big\|\sum_{n=1}^{N}r_nx_n\Big\|>2r\Big\}\leqslant 4\Big[\mathbb{P}\Big\{\Big\|\sum_{n=1}^{N}r_nx_n\Big\|>r\Big\}\Big]^2.$$

Proof. Let us write $S_n := \sum_{j=1}^n r_j x_j$. As in the proof of Lemma 2.18 we put

 $A_n := \{ \|S_1\| \leqslant r, \, \dots, \, \|S_{n-1}\| \leqslant r, \, \|S_n\| > r \}.$

If for an $\omega \in A_n$ we have $||S_N(\omega)|| > 2r$, then $||S_N(\omega) - S_{n-1}(\omega)|| > r$. Now the crucial observation is that (r_1, \ldots, r_N) and $(r_1, \ldots, r_n, r_n r_{n+1}, \ldots, r_n r_N)$ are identically distributed; we leave the easy proof as an exercise. From this and the fact that $|r_n| = 1$ almost surely we obtain

$$\mathbb{P}(A_n \cap \{ \|S_N - S_{n-1}\| > r \}) = \mathbb{P}\left(A_n \cap \left\{ \left\| \sum_{j=n}^N r_j x_j \right\| > r \right\} \right)$$
$$= \mathbb{P}\left(A_n \cap \left\{ \left\| r_n \sum_{j=n}^N r_j x_j \right\| > r \right\} \right)$$
$$= \mathbb{P}\left(A_n \cap \left\{ \left\| x_n + \sum_{j=n+1}^N r_n r_j x_j \right\| > r \right\} \right)$$
$$= \mathbb{P}\left(A_n \cap \left\{ \left\| x_n + \sum_{j=n+1}^N r_j x_j \right\| > r \right\} \right)$$
$$= \mathbb{P}(A_n \cap \left\{ \left\| x_n + (S_N - S_n) \right\| > r \right\}),$$

and similarly $\mathbb{P}\{\|S_N - S_{n-1}\| > r\} = \mathbb{P}\{\|x_n + (S_N - S_n)\| > r\}$. Hence, by the independence of A_n and $S_N - S_n$,

$$\begin{split} \mathbb{P}(A_n \cap \{ \|S_N\| > 2r\}) &\leqslant \mathbb{P}(A_n \cap \{ \|S_N - S_{n-1}\| > r\}) \\ &= \mathbb{P}(A_n) \mathbb{P}\{ \|x_n + (S_N - S_n)\| > r\} \\ &= \mathbb{P}(A_n) \mathbb{P}\{ \|S_N - S_{n-1}\| > r\} \leqslant 2\mathbb{P}(A_n) \mathbb{P}\{ \|S_N\| > r\}, \end{split}$$

where the last step follows from Lévy's inequality after changing the order of summation. Summing over n = 1, ..., N and using Lévy's inequality once more we obtain

$$\mathbb{P}\{\|S_N\| > 2r\} = \sum_{n=1}^{N} \mathbb{P}(A_n \cap \{\|S_N\| > 2r\}) \leq 2\sum_{n=1}^{N} \mathbb{P}(A_n) \mathbb{P}\{\|S_N\| > r\}$$
$$= 2\mathbb{P}\Big\{\max_{1 \leq n \leq N} \|S_n\| > r\Big\} \mathbb{P}\{\|S_N\| > r\} \leq 4[\mathbb{P}\{\|S_N\| > r\}]^2. \ \Box$$

We are now ready to prove the following result, which is the Banach space generalisation due to KAHANE of a classical result for scalar random variables of KHINTCHINE.

Theorem 3.11 (Kahane-Khintchine inequality - Rademacher sums). For all $1 \leq p, q < \infty$ there exists a constant $K_{p,q}$, depending only on p and q, such that for all finite sequences $x_1, \ldots, x_N \in E$ we have

$$\left(\mathbb{E}\left\|\sum_{n=1}^{N}r_{n}x_{n}\right\|^{p}\right)^{\frac{1}{p}} \leq K_{p,q}\left(\mathbb{E}\left\|\sum_{n=1}^{N}r_{n}x_{n}\right\|^{q}\right)^{\frac{1}{q}}.$$

Proof. By Hölder's inequality it suffices to consider the case p > 1 and q = 1. Fix vectors $x_1, \ldots, x_N \in E$. Writing $X_n = r_n x_n$ and $S_N = \sum_{n=1}^N X_n$, we

may assume that $\mathbb{E}||S_N|| = 1$.

Let $j \ge 1$ be the unique integer such that $2^{j-1} . By successive$ applications of Lemma 3.10 for r > 0 we have

$$\mathbb{P}\{\|S_N\| > 2^j r\} \leqslant 4^{2^j - 1} (\mathbb{P}\{\|S_N\| > r\})^{2^j}.$$

Chebyshev's inequality gives $r\mathbb{P}\{||S_N|| > r\} \leq \mathbb{E}||S_N|| = 1$. Hence,

$$\mathbb{E}||S_N||^p = \int_0^\infty pt^{p-1} \mathbb{P}\{||S_N|| > t\} dt$$

$$= 2^{jp} \int_0^\infty pr^{p-1} \mathbb{P}\{||S_N|| > 2^j r\} dr$$

$$\leq 2^{jp} 4^{2^j - 1} \int_0^\infty pr^{p-1} (\mathbb{P}\{||S_N|| > r\})^{2^j} dr$$

$$\leq (2p)^p 4^{2p-1} \int_0^\infty pr^{p-1} (\mathbb{P}\{||S_N|| > r\})^p dr$$

$$\leq (2p)^p 4^{2p-1} \int_0^\infty p \mathbb{P}\{||S_N|| > r\} dr$$

$$\leq (2p)^p 4^{2p-1} p.$$

The best possible constants $K_{p,q}$ in this inequality are called the *Kahane-Khintchine constants*. Note that $K_{p,q} = 1$ if $p \leq q$ by Hölder's inequality. The bound on $K_{p,1}$ produced in the above proof is not the best possible: for instance it is known that $K_{p,1} = 2^{1-\frac{1}{p}}$; see the Notes at the end of the lecture.

By an application of the central limit theorem, the Kahane-Khintchine inequality extends to Gaussian sums:

Theorem 3.12 (Kahane-Khintchine inequality - Gaussian sums). For all $1 \leq p, q < \infty$ and all finite sequences $x_1, \ldots, x_N \in E$ we have

$$\left(\mathbb{E}\left\|\sum_{n=1}^{N}\gamma_{n}x_{n}\right\|^{p}\right)^{\frac{1}{p}} \leqslant K_{p,q}\left(\mathbb{E}\left\|\sum_{n=1}^{N}\gamma_{n}x_{n}\right\|^{q}\right)^{\frac{1}{q}},$$

where $K_{p,q}$ is the Kahane-Khintchine constant.

Proof. Fix k = 1, 2, ... and define $\varphi_n^{(k)} := \frac{1}{\sqrt{k}} \sum_{j=1}^k r_{nk+j}$. For each k we have

$$\left(\mathbb{E}\left\|\sum_{n=1}^{N}\varphi_{n}^{(k)}x_{n}\right\|^{p}\right)^{\frac{1}{p}} = \left(\mathbb{E}\left\|\sum_{n=1}^{N}\sum_{j=1}^{k}r_{nk+j}\frac{x_{n}}{\sqrt{k}}\right\|^{p}\right)^{\frac{1}{p}}$$
$$\leq K_{p,q}\left(\mathbb{E}\left\|\sum_{n=1}^{N}\sum_{j=1}^{k}r_{nk+j}\frac{x_{n}}{\sqrt{k}}\right\|^{q}\right)^{\frac{1}{q}} = K_{p,q}\left(\mathbb{E}\left\|\sum_{n=1}^{N}\varphi_{n}^{(k)}x_{n}\right\|^{q}\right)^{\frac{1}{q}}.$$

The proof is completed by passing to the limit $k \to \infty$ and using the central limit theorem.

The attentive reader has noticed that we are cheating a bit in the above proof, as the usual formulation of the central limit theorem only asserts that $\lim_{k\to\infty}(\varphi_1^{(k)},\ldots,\varphi_N^{(k)}) = (\gamma_1,\ldots,\gamma_N)$ in distribution, that is,

$$\lim_{k \to \infty} \mathbb{E}f(\varphi_1^{(k)}, \dots, \varphi_N^{(k)}) = \mathbb{E}f(\gamma_1, \dots, \gamma_N)$$

for all bounded continuous functions $f : \mathbb{R}^N \to \mathbb{R}$. We will show next how, in the present situation, the convergence of the L^r -norms (with r = p, q) of the sums can be deduced from this. The main idea is contained in the next lemma.

Lemma 3.13. Suppose $\varphi_0, \varphi_1, \ldots$ and φ are \mathbb{R}^N -valued random variables such that for all bounded continuous functions $f : \mathbb{R}^N \to \mathbb{R}$ we have

$$\lim_{k \to \infty} \mathbb{E}f(\varphi_k) = \mathbb{E}f(\varphi).$$

Let $\Phi : \mathbb{R}^N \to \mathbb{R}$ be a Borel function such that $\sup_{k \ge 1} \mathbb{E} |\Phi(\varphi_k)| < \infty$ and $\mathbb{E} |\Phi(\varphi)| < \infty$. If $g : \mathbb{R}^N \to \mathbb{R}$ is a continuous function satisfying

$$|g(t)| \leq |c(t)| |\Phi(t)|, \quad t \in \mathbb{R}^N,$$

where $c: \mathbb{R}^N \to \mathbb{R}$ is a bounded function satisfying $\lim_{|t|\to\infty} |c(t)| = 0$, then

$$\lim_{k \to \infty} \mathbb{E}g(\varphi_k) = \mathbb{E}g(\varphi).$$

Proof. Let $g_R := g \cdot 1_{\{|g| < R\}} + R \cdot 1_{\{g \ge R\}} - R \cdot 1_{\{g \le -R\}}$ denote the truncation of g at the levels $\pm R$. By assumption we have

$$\lim_{k \to \infty} \mathbb{E}g_R(\varphi_k) = \mathbb{E}g_R(\varphi).$$
(3.5)

Furthermore, by dominated convergence,

$$\lim_{R \to \infty} \mathbb{E}g_R(\varphi) = \mathbb{E}g(\varphi).$$
(3.6)

Fix $\varepsilon > 0$ and choose $R_0 > 0$ so large that $\sup_{|t|>R_0} |c(t)| < \varepsilon$. Choose $R_1 > 0$ so large that $|g(t)| > R_1$ implies $|t| > R_0$. Then, for all $R \ge R_1$,

$$\sup_{k \ge 0} \mathbb{E}|g(\varphi_k) - g_R(\varphi_k)| \le \sup_{k \ge 0} \mathbb{E}(1_{\{|g| > R\}}(\varphi_k)|g(\varphi_k)|) \le \sup_{k \ge 0} \mathbb{E}(1_{\{|g| > R\}}(\varphi_k)|c(t)||\varPhi(\varphi_k)|) \le \varepsilon \sup_{k \ge 0} \mathbb{E}|\varPhi(\varphi_k)|,$$
(3.7)

Combined with (3.6) and (3.5), this gives the desired result.

Now we can finish the proof of Theorem 3.12:

Lemma 3.14. With the notations of Theorem 3.12, for all $1 \leq r < \infty$ and $x_1, \ldots, x_N \in E$ we have

$$\lim_{k \to \infty} \mathbb{E} \Big\| \sum_{n=1}^{N} \varphi_n^{(k)} x_n \Big\|^r = \mathbb{E} \Big\| \sum_{n=1}^{N} \gamma_n x_n \Big\|^r,$$

where $\gamma_1, \ldots, \gamma_N$ are independent standard Gaussian variables.

Proof. Without loss of generality we may assume that $\max_{1 \leq n \leq N} ||x_n|| \leq 1$. We fix $1 \leq r < \infty$ and check the condition of Lemma 3.13 for the functions $\Phi : \mathbb{R}^N \to \mathbb{R}$ and $g : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\Phi(t) := \exp\Big(\sum_{n=1}^{N} |t_n| ||x_n||\Big), \qquad g(t) := \Big\|\sum_{n=1}^{N} t_n x_n\Big\|^r,$$

where $\varphi_k := (\varphi_1^{(k)}, \dots, \varphi_N^{(k)})$ and $\varphi := (\gamma_1, \dots, \gamma_N)$.

If φ is a symmetric real-valued random variable, then

$$\mathbb{E} \exp(|\varphi|) = \mathbb{E} \exp(-1_{\{\varphi < 0\}}\varphi) + \mathbb{E} \exp(1_{\{\varphi \ge 0\}}\varphi)$$
$$= \mathbb{E} \exp(1_{\{-\varphi < 0\}}\varphi) + \mathbb{E} \exp(1_{\{\varphi \ge 0\}}\varphi) \leqslant 2\mathbb{E} \exp(\varphi).$$

Hence, since $\max_{1 \leq n \leq N} ||x_n|| \leq 1$,

$$\begin{split} \mathbb{E}\,\varPhi(\varphi_k) &\leqslant \prod_{n=1}^N \mathbb{E}\exp(|\varphi_n^{(k)}|) \leqslant 2^N \prod_{n=1}^N \mathbb{E}\exp(\varphi_n^{(k)}) \\ &= 2^N \prod_{n=1}^N \prod_{j=1}^k \mathbb{E}\exp\left(\frac{r_{nk+j}}{\sqrt{k}}\right) = 2^N \left(\frac{1}{2}\exp\left(\frac{1}{\sqrt{k}}\right) + \frac{1}{2}\exp\left(\frac{-1}{\sqrt{k}}\right)\right)^{kN} \\ &= 2^N \mathscr{O}\left(1 + \frac{1}{2k}\right)^{kN} = 2^N \exp(N/2) \cdot \mathscr{O}(1) \quad \text{as } k \to \infty. \end{split}$$

3.3 Exercises

1. Let $(X_n)_{n=1}^N$ be a sequence of independent symmetric *E*-valued random variables, and let $(r_n)_{n=1}^N$ be a Rademacher sequence which is independent of $(X_n)_{n=1}^N$. Prove that the sequences $(X_n)_{n=1}^N$ and $(r_n X_n)_{n=1}^N$ are identically distributed.

Hint: As in the proof of Theorem 3.2 it may be assumed that $(X_n)_{n=1}^N$ and $(r_n)_{n=1}^N$ are defined on distinct probability spaces. Use Fubini's theorem together with the result of Exercise 2.2.

Remark: This technique for introducing Rademacher variables is known as *randomisation*. It enables one to apply inequalities for Rademacher sums in E to sums of independent symmetric random variables in E.

- 2. (!) Let $(r'_n)_{n=1}^{\infty}$ and $(r''_n)_{n=1}^{\infty}$ be independent Rademacher sequences on probability spaces $(\Omega', \mathscr{F}', \mathbb{P}')$ and $(\Omega'', \mathscr{F}'', \mathbb{P}')$. Prove that on the product $(\Omega, \mathscr{F}, \mathbb{P}) = (\Omega' \times \Omega'', \mathscr{F}' \otimes \mathscr{F}'', \mathbb{P}' \otimes \mathbb{P}'')$, the sequence $(r'_m r''_n)_{m,n=1}^{\infty}$ consists of Rademacher variables, but as a (doubly indexed) sequence it fails to be a Rademacher sequence (that is, the random variables $r'_m r''_n$ fail to be independent).
- 3. (!) We continue with the notations of the previous exercise. Prove that for $1 \leq p < \infty$ the following version of the contraction principle holds for double Rademacher sums in the spaces $L^p(A)$, where (A, \mathscr{A}, μ) is a σ -finite measure space: there exists a constant $C_p \geq 0$ such that for all finite sequences $(f_{mn})_{m,n=1}^N$ in $L^p(A)$ and all scalars $(a_{mn})_{m,n=1}^N$ we have

$$\mathbb{E} \Big\| \sum_{m,n=1}^{N} a_{mn} r'_m r''_n f_{mn} \Big\|^p \leqslant C_p^p \big(\max_{1 \leqslant m,n \leqslant N} |a_{mn}|^p \big) \mathbb{E} \Big\| \sum_{m,n=1}^{N} r'_m r''_n f_{mn} \Big\|^p.$$

40 3 Sums of independent random variables

Hint: Proceed in three steps: (i) the result holds for $E = \mathbb{R}$ with exponent 2; (ii) the result holds for $E = \mathbb{R}$ with exponent p; (iii) the result holds for $E = L^p(A)$ with exponent p.

4. Let $1 \leq p \leq 2$. A Banach space *E* is said to have *type p* if there exists a constant $C_p \geq 0$ such that for all finite sequences x_1, \ldots, x_N in *E* we have

$$\left(\mathbb{E}\left\|\sum_{n=1}^{N}r_{n}x_{n}\right\|^{2}\right)^{\frac{1}{2}} \leqslant C_{p}\left(\sum_{n=1}^{N}\|x_{n}\|^{p}\right)^{\frac{1}{p}}.$$

Let $2 \leq q \leq \infty$. The space *E* is said to have *cotype q* if there exists a constant $C_q \geq 0$ such that for all finite sequences x_1, \ldots, x_N in *E* we have

$$\left(\sum_{n=1}^{N} \|x_n\|^q\right)^{\frac{1}{q}} \leqslant C_q \left(\mathbb{E} \left\|\sum_{n=1}^{N} r_n x_n\right\|^2\right)^{\frac{1}{2}}.$$

For $q = \infty$ we make the obvious adjustment in the second definition. Prove the following assertions:

- a) Every Banach space has type 1 and cotype ∞ (accordingly, a Banach space is said to have *non-trivial type* if it has type $p \in (1, 2]$ and *finite cotype* if it has cotype $q \in [2, \infty)$).
- b) Every Hilbert space has type 2 and cotype 2.
- c) If a Banach space has type p for some $p \in [1, 2]$, then it has type p' for all $p' \in [1, p]$; if a Banach space has cotype q for some $q \in [2, \infty]$, then it has cotype q' for all $q' \in [q, \infty]$.
- d) Let $p \in [1, 2]$. Prove that if E has type p, then the dual space E^* has cotype $p', \frac{1}{p} + \frac{1}{p'} = 1$.

Hint: For each $x_n^* \in E^*$ choose $x_n \in E$ of norm one such that $||x_n^*|| \ge \frac{1}{2} |\langle x_n, x_n^* \rangle|$. Then use Hölder's inequality to the effect that for all scalar sequences $(b_n)_{n=1}^N$ one has

$$\Big(\sum_{n=1}^{N} |b_n|^{p'}\Big)^{\frac{1}{p'}} = \sup\Big\{\sum_{n=1}^{N} a_n b_n : \Big(\sum_{n=1}^{N} |a_n|^p\Big)^{\frac{1}{p}} \leqslant 1\Big\}.$$

Remark: The analogous result for spaces with cotype fails. Indeed, the reader is invited to check that l^1 has cotype 2 while its dual l^{∞} fails to have non-trivial type.

5. Let $p \in [1, 2]$. Prove that a Banach space E has type p if and only if it has *Gaussian type* p, that is, if and only if there exists a constant $C \ge 0$ such that for all finite sequences x_1, \ldots, x_N in E we have

$$\left(\mathbb{E}\left\|\sum_{n=1}^{N}\gamma_{n}x_{n}\right\|^{2}\right)^{\frac{1}{2}} \leqslant C\left(\sum_{n=1}^{N}\|x_{n}\|^{p}\right)^{\frac{1}{p}}.$$

Hint: One direction follows from Corollary 3.6. For the other direction use a randomisation argument.

Remark: The corresponding assertion for cotype is also true but much harder to prove; see the Notes.

Notes. The results of this lecture are classical and can be found in many textbooks. Our presentation borrows from ALBIAC and KALTON [1] and DI-ESTEL, JARCHOW, TONGE [35]. Both are excellent starting points for further reading.

The Kahane contraction principle is due to KAHANE [54], who also extended the classical scalar Khintchine inequality to arbitrary Banach spaces. It is an open problem to determine the best constants $K_{p,q}$ in the Kahane-Khintchine inequality; a recent result of LATALA and OLESZKIEWICZ [67] asserts that the constant $K_{p,1} = 2^{1-\frac{1}{p}}$ is optimal for $1 \leq p \leq 2$.

For a proof of Theorem 3.7 see, e.g., [35]. The proofs of Theorems 3.9 and 3.11 are taken from ALBIAC and KALTON [1]. The central limit argument in Lemma 3.14 is adapted from TOMCZAK-JAEGERMANN [102].

The contraction principle for double Rademacher sums of Exercise 3 has been introduced by PISIER [92]. This property, nowadays known under the rather unsuggestive name 'property (α)' plays an important role in many advanced results in Banach space-valued harmonic analysis. It can be shown that the Rademachers can be replaced by Gaussians without changing the class of spaces under consideration. Not every Banach space has property (α); a counterexample is the space c_0 .

The notions of type and cotype were developed in the 1970s by MAUREY and PISIER. As we have seen in Exercise 4, Hilbert spaces have type 2 and cotype 2. A celebrated theorem of KWAPIEŃ [64] asserts that Hilbert spaces are the only spaces with this property: a Banach space E is isomorphic to a Hilbert space if and only if E has type 2 and cotype 2. Another class of spaces of which the type and cotype can be computed are the L^p -spaces. For the interested reader we include a proof that the spaces $L^p(A)$, with $1 \leq p < \infty$ and $(A, \mathscr{A}, \mu) \sigma$ -finite, have type min $\{p, 2\}$. A similar argument can be used to prove that they have cotype max $\{p, 2\}$.

Let $f_1, \ldots, f_N \in L^p(A)$ and put $r := \min\{p, 2\}$. Using the Fubini theorem, the scalar Kahane-Khintchine inequality, the type p inequality, Hölder's inequality, and the triangle inequality in $L^{\frac{p}{r}}(A)$, we obtain

423 Sums of independent random variables

$$\begin{split} \left(\mathbb{E} \left\| \sum_{n=1}^{N} r_{n} f_{n} \right\|_{L^{p}(A)}^{p} \right)^{\frac{1}{p}} &= \left(\int_{A} \mathbb{E} \left| \sum_{n=1}^{N} r_{n} f_{n}(\xi) \right|^{p} d\mu(\xi) \right)^{\frac{1}{p}} \\ &\leq K_{p2} \Big(\int_{A} \left(\mathbb{E} \left| \sum_{n=1}^{N} r_{n} f_{n}(\xi) \right|^{2} \right)^{\frac{p}{2}} d\mu(\xi) \Big)^{\frac{1}{p}} \\ &= K_{p2} \Big(\int_{A} \left(\sum_{n=1}^{N} |f_{n}(\xi)|^{2} \right)^{\frac{p}{2}} d\mu(\xi) \Big)^{\frac{1}{p}} \\ &\leq K_{p2} \Big(\int_{A} \left(\sum_{n=1}^{N} |f_{n}(\xi)|^{r} \right)^{\frac{p}{r}} d\mu(\xi) \Big)^{\frac{1}{p}} \\ &= K_{p2} \Big\| \sum_{n=1}^{N} |f_{n}|^{r} \Big\|_{L^{\frac{p}{r}}(A)} \\ &\leq K_{p2} \Big(\sum_{n=1}^{N} \|f_{n}\|^{r} \Big\|_{L^{\frac{p}{r}}(A)} \Big)^{\frac{1}{r}} \\ &= K_{p2} \Big(\sum_{n=1}^{N} \|f_{n}\|_{L^{p}(A)}^{r} \Big)^{\frac{1}{r}}. \end{split}$$

An application of the Kahane-Khintchine inequality for $L^{p}(A)$ to replace the L^p -moment in the left hand side by the L^2 -moment finishes the proof.

It was noted in Exercise 4 that if E has type p, then E^* has cotype p' (where $\frac{1}{p} + \frac{1}{p'} = 1$) and that the analogous duality result for cotype fails. It is a deep result of PISIER [93] that if E has cotype $q \in [2, \infty)$ and non-trivial type, then E^* has type q', $\frac{1}{q} + \frac{1}{q'} = 1$. The fact that a Banach space has cotype q if and only if it has Gaussian

cotype q can be deduced from a deep result of MAUREY and PISIER (see [1, Chapter 11]) which gives a purely geometric characterisation of type and cotype. For the details we refer to [35].