4

Gaussian random variables

Having studied E-valued Gaussian sums of the form $\sum_{n=1}^{N} \gamma_n x_n$ in the previous lecture, we now turn to general theory of Gaussian random variables with values in a Banach space E. The results of this lecture will be important for the construction of an E-valued stochastic integral with respect to Brownian motion.

We start with a proof of the Fernique theorem on integrability of Gaussian random variables. This theorem makes it possible to investigate L^p-convergence of sequences of Gaussian random variables. As it turns out, every E-valued Gaussian random variable can be represented in a canonical way as an L^p-convergent (finite or infinite) sum $\sum_{n\geqslant 1} \gamma_n x_n$. This representation theorem permits us to extend the covariance domination principle and the Kahane-Khintchine inequality to arbitrary E-valued Gaussians.

4.1 Fernique's theorem

A real-valued random variable γ is called *Gaussian* if there exists a number $q \geqslant 0$ such that its Fourier transform is given by

$$\mathbb{E}\big(\exp(-i\xi\gamma)\big) = \exp(-\tfrac{1}{2}q\xi^2), \qquad \xi \in \mathbb{R}.$$

By uniqueness of Fourier transforms one deduces that $\gamma = 0$ almost surely if $q = 0$, and that γ has a distribution with density

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

if $q > 0$. It follows that $\mathbb{E}\gamma = 0$ and $\mathbb{E}\gamma^2 = q$, which means that γ is centred and has variance q. We call γ *standard Gaussian* if $q = 1$; this definition is consistent with the one given in Lecture 3.

Let E be a real Banach space.

Definition 4.1. *An E-valued random variable X is* Gaussian *if the real-valued random variable $\langle X, x^* \rangle$ is Gaussian for all $x^* \in E^*$.*

Much of the theory of Banach space-valued Gaussian random variables depends on a fundamental integrability result due to FERNIQUE. For its proof we need a lemma.

Lemma 4.2. *Let X and Y be independent and identically distributed E-valued Gaussian random variables. Then $U := (X+Y)/\sqrt{2}$ and $V := (X-Y)/\sqrt{2}$ are independent and have the same distribution as X and Y.*

Proof. Let μ be the common distribution of X and Y. Then $\widehat{\mu}(x^*) = \exp(-\frac{1}{2}q(x^*))$, where $q(x^*) = \mathbb{E}\langle X, x^* \rangle^2 = \mathbb{E}\langle Y, x^* \rangle^2$. Using the independence of X and Y we have

$$\begin{aligned}\mathbb{E}\exp(-i\langle U, x^*\rangle) &= \mathbb{E}\exp(-i\tfrac{1}{2}\sqrt{2}\langle X, x^*\rangle)\mathbb{E}\exp(-i\tfrac{1}{2}\sqrt{2}\langle Y, x^*\rangle)\\ &= \exp(-\tfrac{1}{4}q(x^*))\exp(-\tfrac{1}{4}q(x^*)) = \exp(-\tfrac{1}{2}q(x^*)).\end{aligned}$$

By the uniqueness theorem for the Fourier transform, this shows that U has the same distribution as X and Y. A similar computation shows that V has the same distribution as X and Y.

We will prove that U and V are independent by checking that $\mu_{(U,V)} = \mu \times \mu$, where $\mu_{(U,V)}$ is the distribution of the $E \times E$-valued random variable (U, V). Identifying $(E \times E)^*$ with $E^* \times E^*$ with pairing $\langle (x, y), (x^*, y^*) \rangle = \langle x, x^* \rangle + \langle y, y^* \rangle$, by the uniqueness theorem for the Fourier transform it is enough to prove that $\widehat{\mu}_{(U,V)}(x^*, y^*) = \widehat{\mu}(x^*)\widehat{\mu}(y^*)$ for all $x^*, y^* \in E^*$. But this follows from

$$\begin{aligned}\widehat{\mu}_{(U,V)}(x^*, y^*) &= \mathbb{E}\,\exp\Big(-i\big(\langle U, x^*\rangle + \langle V, y^*\rangle\big)\Big)\\ &= \mathbb{E}\,\exp\Big(-\tfrac{1}{2}i\sqrt{2}\big(\langle X, x^*+y^*\rangle + \langle Y, x^*-y^*\rangle\big)\Big)\\ &= \mathbb{E}\,\exp\big(-\tfrac{1}{2}i\sqrt{2}\langle X, x^*+y^*\rangle\big)\mathbb{E}\,\exp\big(-\tfrac{1}{2}i\sqrt{2}\langle Y, x^*-y^*\rangle\big)\\ &= \exp\big(-\tfrac{1}{4}q(x^*+y^*)\big)\exp\big(-\tfrac{1}{4}q(x^*-y^*)\big)\\ &= \exp\big(-\tfrac{1}{2}(q(x^*)+q(y^*))\big)\\ &= \widehat{\mu}(x^*)\widehat{\mu}(y^*).\end{aligned}$$

□

Theorem 4.3 (Fernique). *Let X be an E-valued Gaussian variable. There exists a constant $\beta > 0$ such that*

$$\mathbb{E}\exp(\beta\|X\|^2) < \infty. \tag{4.1}$$

Proof. On a possibly larger probability space, let X' be independent copy of X. For instance, identify X with the random variable $X(\omega_1, \omega_2) := X(\omega_1)$ on $\Omega \times \Omega$ and define X' on $\Omega \times \Omega$ by $X'(\omega_1, \omega_2) := X(\omega_2)$.

Fix $t \geqslant s > 0$. By the lemma,

$$\begin{aligned}
&\mathbb{P}\{\|X\| \leqslant s\} \cdot \mathbb{P}\{\|X'\| > t\} \\
&\qquad = \mathbb{P}\Big\{\Big\|\frac{X - X'}{\sqrt{2}}\Big\| \leqslant s\Big\} \cdot \mathbb{P}\Big\{\Big\|\frac{X + X'}{\sqrt{2}}\Big\| > t\Big\} \\
&\qquad \leqslant \mathbb{P}\Big\{\Big|\frac{\|X\| - \|X'\|}{\sqrt{2}}\Big| \leqslant s,\ \frac{\|X\| + \|X'\|}{\sqrt{2}} > t\Big\} \\
&\qquad \overset{(*)}{\leqslant} \mathbb{P}\Big\{\|X\| > \frac{t-s}{\sqrt{2}},\ \|X'\| > \frac{t-s}{\sqrt{2}}\Big\} \\
&\qquad = \mathbb{P}\Big\{\|X\| > \frac{t-s}{\sqrt{2}}\Big\} \cdot \mathbb{P}\Big\{\|X'\| > \frac{t-s}{\sqrt{2}}\Big\},
\end{aligned}$$

where in $(*)$ we used that the set

$$\big\{(\xi, \eta) \in \mathbb{R}_+^2 :\ |\xi - \eta| \leqslant s\sqrt{2} \ \text{ and } \ \xi + \eta > t\sqrt{2}\big\}$$

is contained in the set

$$\Big\{(\xi, \eta) \in \mathbb{R}_+^2 :\ \xi > \frac{t-s}{\sqrt{2}} \ \text{ and } \ \eta > \frac{t-s}{\sqrt{2}}\Big\}.$$

Hence, since X and X' have the same distribution,

$$\mathbb{P}\{\|X\| \leqslant s\}\mathbb{P}\{\|X\| > t\} \leqslant \Big(\mathbb{P}\Big\{\|X\| > \frac{t-s}{\sqrt{2}}\Big\}\Big)^2. \tag{4.2}$$

Choose $r > 0$ such that $\mathbb{P}\{\|X\| \leqslant r\} \geqslant \frac{2}{3}$. Define $t_0 := r$ and $t_n := r + \sqrt{2}t_{n-1}$ for $n \geqslant 1$. By induction it is checked that $t_n = r((\sqrt{2})^{n+1} - 1)/(\sqrt{2} - 1)$, so $t_n \leqslant r(\sqrt{2})^{n+4}$. Put

$$\alpha_n := \frac{\mathbb{P}\{\|X\| > t_n\}}{\mathbb{P}\{\|X\| \leqslant r\}}.$$

Note that $\alpha_0 \leqslant (1 - \frac{2}{3})/\frac{2}{3} = \frac{1}{2}$. From (4.2) with $s = r$, $t = t_{n+1}$ we obtain

$$\alpha_{n+1} \leqslant \Big(\frac{\mathbb{P}\{\|X\| > t_n\}}{\mathbb{P}\{\|X\| \leqslant r\}}\Big)^2 = \alpha_n^2.$$

Therefore $\alpha_n \leqslant \alpha_0^{2^n} \leqslant 2^{-2^n}$ and it follows that

$$\mathbb{P}\{\|X\| > t_n\} = \alpha_n \mathbb{P}\{\|X\| \leqslant r\} \leqslant 2^{-2^n}.$$

Using these estimates we obtain

$$\begin{aligned}\mathbb{E}(\exp(\beta\|X\|^2)) &\leqslant \mathbb{P}\{\|X\| \leqslant t_0\} \cdot \exp(\beta t_0^2) \\ &\qquad + \sum_{n=0}^{\infty} \mathbb{P}\{t_n < \|X\| \leqslant t_{n+1}\} \cdot \exp(\beta t_{n+1}^2) \\ &\leqslant \exp(\beta r^2) + \sum_{n=0}^{\infty} 2^{-2^n} \exp(\beta r^2 2^{n+5}) \\ &= \exp(\beta r^2) + \sum_{n=0}^{\infty} \exp(2^n[-\log 2 + 32\beta r^2]),\end{aligned}$$

and this sum converges if $\beta > 0$ is taken small enough. □

In what follows we need much less: it will suffice to know that $\mathbb{E}\|X\|^p < \infty$ for all $1 \leqslant p < \infty$.

As a simple corollary to Fernique's theorem we note that the expectation of a Gaussian random variable is well-defined. In fact we have the following result:

Corollary 4.4. *If X is E-valued Gaussian, then $\mathbb{E}X = 0$.*

Proof. For all $x^* \in E^*$ we have $\langle \mathbb{E}X, x^* \rangle = \mathbb{E}\langle X, x^* \rangle = 0$ and we may appeal to the Hahn-Banach theorem. □

4.2 The covariance operator

In order to characterise Gaussian variables in terms of their Fourier transforms we introduce the following terminology.

Definition 4.5. *A bounded operator $Q \in \mathscr{L}(E^*, E)$ is called*

- positive*, if $\langle Qx^*, x^* \rangle \geqslant 0$ for all $x^* \in E^*$;*
- symmetric*, if $\langle Qx^*, y^* \rangle = \langle Qy^*, x^* \rangle$ for all $x^*, y^* \in E^*$.*

Proposition 4.6. *For an E-valued random variable X the following assertions are equivalent:*

(1) *X is Gaussian;*
(2) *there exists a positive symmetric operator $Q \in \mathscr{L}(E^*, E)$ such that the Fourier transform of X is given by*

$$\mathbb{E}\exp(-i\langle X, x^* \rangle) = \exp\big(-\tfrac{1}{2}\langle Qx^*, x^* \rangle\big), \qquad x^* \in E^*.$$

The operator Q is uniquely determined by (2)*. Moreover,*

$$\mathbb{E}\langle X, x^* \rangle^2 = \langle Qx^*, x^* \rangle, \qquad x^* \in E^*.$$

Proof. (1)⇒(2): Since X is square integrable by Theorem 4.3, the random variable $\langle X, x^*\rangle X$ is integrable and we may define

$$Qx^* := \mathbb{E}\langle X, x^*\rangle X, \quad x^* \in E^*.$$

From $\langle Qx^*, y^*\rangle = \mathbb{E}\langle X, x^*\rangle\langle X, y^*\rangle$ we see that Q is positive and symmetric. Since $\langle X, x^*\rangle$ is Gaussian with variance $\mathbb{E}\langle X, x^*\rangle^2 = \langle Qx^*, x^*\rangle$, we have

$$\mathbb{E}\exp(-i\langle X, x^*\rangle) = \exp(-\tfrac{1}{2}\langle Qx^*, x^*\rangle).$$

(2)⇒(1): Replacing x^* by ξx^* in the assumption, we see that the Fourier transform of $\langle X, x^*\rangle$ equals

$$\mathbb{E}\exp(-i\xi\langle X, x^*\rangle) = \exp\bigl(-\tfrac{1}{2}\xi^2\langle Qx^*, x^*\rangle\bigr).$$

Thus $\langle X, x^*\rangle$ is Gaussian with variance $\langle Qx^*, x^*\rangle$.

If R is another positive symmetric operator satisfying condition (2), then $\langle Qx^*, x^*\rangle = \langle Rx^*, x^*\rangle$ for all $x^* \in E^*$. By polarisation this implies $\langle Qx^*, y^*\rangle = \langle Rx^*, y^*\rangle$ for all $x^*, y^* \in E^*$, and therefore $Q = R$. □

The operator Q is called the *covariance operator* of X. The reader is warned that not every positive symmetric operator $Q \in \mathscr{L}(E^*, E)$ is the covariance of an E-valued random variable X. This may happen even if E is a separable infinite-dimensional Hilbert space (see Exercise 2).

Corollary 4.7. *Every E-valued Gaussian random variable is symmetric.*

Proof. Just note that X and $-X$ have the same Fourier transforms. □

We proceed with two simple constructions to produce new Gaussian variables from old ones. The first asserts that sums of independent Gaussian variables are Gaussian.

Proposition 4.8. *Let $X_1, \dots, X_N$ be independent E-valued Gaussian random variables with covariance operators $Q_1, \dots, Q_N$. Then the sum $X := \sum_{n=1}^N X_n$ is Gaussian with covariance operator $Q = \sum_{n=1}^N Q_n$.*

Proof. For all $x^* \in E^*$ we have, by independence,

$$\begin{aligned}\mathbb{E}\exp(-i\langle X, x^*\rangle) &= \mathbb{E}\prod_{n=1}^N \exp(-i\langle X_n, x^*\rangle) = \prod_{n=1}^N \mathbb{E}(\exp(-i\langle X_n, x^*\rangle)) \\ &= \prod_{n=1}^N \exp(-\tfrac{1}{2}\langle Q_n x^*, x^*\rangle) = \exp(-\tfrac{1}{2}\langle Qx^*, x^*\rangle).\end{aligned}$$ □

Compositions of Gaussians with bounded operators are Gaussian again:

Proposition 4.9. *If X is E-valued Gaussian with covariance operator Q, and if $T \in \mathscr{L}(E,F)$ is a bounded operator, then TX is F-valued Gaussian with covariance operator TQT^*.*

Proof. This follows by computing the Fourier transform of TX:

$$\begin{aligned}\mathbb{E}(\exp(-i\langle TX, x^*\rangle)) &= \mathbb{E}(\exp(-i\langle X, T^*x^*\rangle)) \\ &= \exp(-\tfrac{1}{2}\langle QT^*x^*, T^*x^*\rangle)) = \exp(-\tfrac{1}{2}\langle TQT^*x^*, x^*\rangle)). \qquad \square\end{aligned}$$

As an application we prove next that if the E-valued random variables $X_1, \dots, X_N$ are *jointly Gaussian*, that is, if the E^N-valued random variable $(X_1, \dots, X_N)$ is Gaussian, then $X_1, \dots, X_N$ are independent if and only if they are *uncorrelated* in the sense that

$$\mathbb{E}\langle X_m, x^*\rangle\langle X_n, y^*\rangle = 0, \quad \forall m \neq n,\ x^*, y^* \in E^*.$$

Proposition 4.10. *Let $X_1, \dots, X_N$ be E-valued random variables such that the E^N-valued random variable $X = (X_1, \dots, X_N)$ is Gaussian. The following assertions are equivalent:*

(1) *$X_1, \dots, X_N$ are independent;*
(2) *$X_1, \dots, X_N$ are uncorrelated.*

Proof. We proceed in two steps.

Step 1 – First we consider the scalar case. Let $\gamma_1, \dots, \gamma_N$ be real-valued random variables such that the $\mathbb{R}^N$-valued random variable $\gamma = (\gamma_1, \dots, \gamma_N)$ is Gaussian. Note that each γ_n is Gaussian; this follows from Proposition 4.9 by applying coordinate projections. We shall prove that $\gamma_1, \dots, \gamma_N$ are independent if and only if $\gamma_1, \dots, \gamma_N$ are uncorrelated.

The 'only if' part follows from $\mathbb{E}\gamma_m\gamma_n = \mathbb{E}\gamma_m\mathbb{E}\gamma_n = 0$ for all $m \neq n$. For the 'if' part we note that if $\gamma_1, \dots, \gamma_N$ are uncorrelated, the covariance matrix of γ is diagonal: $Q = \operatorname{diag}(q_1, \dots, q_N)$ with $q_n = \mathbb{E}\,\gamma_n^2$. Then the Fourier transform of γ is given by

$$\begin{aligned}\mathbb{E}\big(\exp(-i\langle\gamma, \xi\rangle)\big) &= \exp(-\tfrac{1}{2}\langle Q\xi, \xi\rangle) = \exp\big(-\tfrac{1}{2}\sum_{n=1}^{N} q_n\xi_n^2\big) \\ &= \prod_{n=1}^{N} \exp(-\tfrac{1}{2}q_n\xi_n^2) = \prod_{n=1}^{N} \mathbb{E}\exp(-i\xi_n\gamma_n), \quad \xi \in \mathbb{R}^N.\end{aligned}$$

Let μ and μ_n denote the distributions of γ and γ_n, respectively. The above identity implies that μ and the product measure $\mu_1 \times \dots \times \mu_N$ have the same Fourier transform. Hence from Theorem 2.8 we deduce that $\mu = \mu_1 \times \dots \times \mu_N$. This implies that $\gamma_1, \dots, \gamma_N$ are independent.

Step 2 – Next we turn to the proof of the proposition. For all choices of $x_1^*, \dots, x_N^* \in E^*$ the $\mathbb{R}^N$-valued random variable $(\langle X_1, x_1^*\rangle, \dots, \langle X_N, x_N^*\rangle)$

is Gaussian by Proposition 4.9, since it is the image of $(X_1, \dots, X_N)$ under the linear transformation from E^N to $\mathbb{R}^N$, $(x_1, \dots, x_n) \mapsto (\langle x_1, x_1^* \rangle, \dots, \langle x_N, x_N^* \rangle)$.

(1)⇒(2): This implication follows from the corresponding implication in Step 1 since the independence of $X_1, \dots, X_n$ implies the independence of $\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle$.

(2)⇒(1): By Step 1, for all $x_1^*, \dots, x_N^* \in E^*$ the random variables $\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle$ are independent and therefore

$$\widehat{\mu_{(X_1,\dots,X_N)}}(x_1^*, \dots, x_N^*) = \mathbb{E} \exp(-i \sum_{n=1}^{N} \langle X_n, x_n^* \rangle) = \prod_{n=1}^{N} \mathbb{E} \exp(-i \langle X_n, x_n^* \rangle)$$
$$= \prod_{n=1}^{N} \widehat{\mu_{X_n}}(x_n^*) = \mu_{X_1} \widehat{\times \cdots \times} \mu_{X_N}(x_1^*, \dots, x_N^*).$$

Hence $\mu_{(X_1,\dots,X_N)} = \mu_{X_1} \times \cdots \times \mu_{X_N}$ by Theorem 2.8. □

The joint Gaussianity condition cannot be relaxed to Gaussianity of each of the X_n; see Exercise 1.

4.3 Series representation

The main result of this section states that every E-valued Gaussian random variable can be represented as a Gaussian sum of the form $\sum_{n \geqslant 1} \gamma_n x_n$, where $(\gamma_n)_{n \geqslant 1}$ is a Gaussian sequence and $(x_n)_{n \geqslant 1}$ is a (finite or infinite) sequence in E. This fact enables us to extend various results for Gaussian sums, such as the Kahane-Khintchine inequality, to arbitrary Gaussian random variables.

We start with a simple proposition stating that limits of Gaussian variables are Gaussian.

Proposition 4.11. *If $(X_n)_{n=1}^{\infty}$ is a sequence of E-valued Gaussian variables and X is a random variable such that*

$$\lim_{n \to \infty} \langle X_n, x^* \rangle = \langle X, x^* \rangle \quad \textit{in probability for all } x^* \in E^*,$$

then X is Gaussian. Its covariance operator $Q \in \mathscr{L}(E^*, E)$ is given by $\langle Q x^*, y^* \rangle = \lim_{n \to \infty} \langle Q_n x^*, y^* \rangle$ for $x^*, y^* \in E^*$.

Proof. Fixing $x^* \in E^*$, after passing to a subsequence we may assume that $\lim_{n \to \infty} \langle X_n, x^* \rangle = \langle X, x^* \rangle$ almost surely. Then, by the dominated convergence theorem,

$$\mathbb{E} \exp(-i\xi \langle X, x^* \rangle) = \lim_{n \to \infty} \mathbb{E} \exp(-i\xi \langle X_n, x^* \rangle) = \lim_{n \to \infty} \exp(-\tfrac{1}{2}\xi^2 \langle Q_n x^*, x^* \rangle).$$

Since each of the terms $\langle Q_n x^*, x^* \rangle$ is non-negative, this implies that the limit $q(x^*) := \lim_{n \to \infty} \langle Q_n x^*, x^* \rangle$ exists. From the resulting identity

$$\mathbb{E}\exp(-i\xi\langle X, x^*\rangle) = \exp(-\tfrac{1}{2}\xi^2 q(x^*))$$

we conclude that $\langle X, x^*\rangle$ is Gaussian for all $x^* \in E^*$. By definition this means that X is Gaussian.

Denote by Q the covariance operator of X. For all $\xi \in \mathbb{R}$,

$$\exp(-\tfrac{1}{2}\xi^2\langle Qx^*, x^*\rangle) = \mathbb{E}\exp(-i\xi\langle X, x^*\rangle) = \exp(-\tfrac{1}{2}\xi^2 q(x^*)).$$

From this we deduce that $\langle Qx^*, x^*\rangle = q(x^*) = \lim_{n\to\infty}\langle Q_n x^*, x^*\rangle$. Applying this to $x^* + y^*$ we find $\langle Qx^*, y^*\rangle = \lim_{n\to\infty}\langle Q_n x^*, y^*\rangle$ for all $x^*, y^* \in E^*$. □

For a Gaussian random variable X with covariance operator Q, we denote by H_X the closed linear subspace in $L^2(\Omega)$ spanned by the random variables $\{\langle X, x^*\rangle : \ x^* \in E^*\}$. The operator $i_X : H_X \to E$,

$$i_X\langle X, x^*\rangle := \mathbb{E}\langle X, x^*\rangle X = Qx^*, \tag{4.3}$$

is well-defined and bounded by Hölder's inequality and Fernique's theorem. Its adjoint is given by $i_X^* x^* = \langle X, x^*\rangle$. This leads to the factorisation

$$Q = i_X i_X^*. \tag{4.4}$$

Here, and in similar situations later on, we identify H_X and its dual H_X^* by means of the Riesz representation theorem. Since we are working over the real scalar field this identification is linear and should never lead to any confusion. For a generalisation of the factorisation (4.4) to arbitrary positive symmetric operators Q see Exercise 3.

Theorem 4.12 (Karhunen-Loève expansion). *Let X be an E-valued Gaussian random variable.*

(1) *The space H_X is separable.*
(2) *If $(\gamma_n)_{n\geqslant 1}$ is an orthonormal basis of H_X, then $(\gamma_n)_{n\geqslant 1}$ is a Gaussian sequence and*

$$\sum_{n\geqslant 1}\gamma_n i_X \gamma_n = X,$$

where convergence holds almost surely and in $L^p(\Omega; E)$ for all $1 \leqslant p < \infty$.

Proof. Define $\widetilde{H_X}$ as the closed linear subspace of $L^2(E, \mu_X)$ spanned by E^*; here we think of the functionals $x^* \in E^*$ as functions on E. In view of $\mathbb{E}\langle X, x^*\rangle^2 = \int_E \langle x, x^*\rangle^2\, d\mu_X(x)$, the mapping $x^* \mapsto \langle X, x^*\rangle$ extends uniquely to an isometry of Hilbert spaces $\widetilde{H_X} \simeq H_X$.

Let E_0 be a separable closed subspace of E containing the essential range of X. Then $\mu_X(E_0) = 1$ and therefore the identity mapping gives an isometry $L^2(E_0, \mu_X) \simeq L^2(E, \mu_X)$. Since the Borel σ-algebra $\mathscr{B}(E_0)$ is generated by a countable family of open sets (take a dense sequence $(x_n)_{n=1}^\infty$ in E_0 and consider the open balls $B(x_n, q)$ with rational $q > 0$), the space $L^2(E_0, \mu_X)$ is separable. It follows that $L^2(E,\mu_X)$ is separable and hence so is $\widetilde{H_X}$, it being a closed subspace of $L^2(E,\mu_X)$. It follows that H_X is separable.

Let $(\gamma_n)_{n\geqslant 1}$ be a (finite or countably infinite) orthonormal basis of H_X. Every random variable in H_X is Gaussian by Proposition 4.11. In particular, linear combinations of the γ_n are Gaussian, which means that all vectors $(\gamma_{n_1},\dots,\gamma_{n_N})$ are Gaussian as $\mathbb{R}^N$-valued random variables. Therefore Proposition 4.10 implies that the γ_n are independent.

For all $x^*\in E^*$ we have the identities

$$\sum_{n\geqslant 1}\gamma_n\langle i_X\gamma_n,x^*\rangle=\sum_{n\geqslant 1}\gamma_n\,\mathbb{E}\gamma_n\langle X,x^*\rangle=\langle X,x^*\rangle$$

in H_X, noting that the middle expression is the expansion of $\langle X,x^*\rangle$ with respect to the orthonormal basis $(\gamma_n)_{n\geqslant 1}$ of H_X. The result now follows from the Itô-Nisio theorem. □

For the readers familiar with weak*-topologies we sketch an alternative, more functional analytic proof of the separability of H_X. The dual E_0^* is weak*-separable, by the separability of E_0. Regarding i_X as a bounded injective operator from H_X to E_0, the adjoint i_X^* is weak*-continuous and maps E_0^* onto a weak*-separable and weak*-dense subspace of H_X. But the weak*-topology of the Hilbert space H_X is the same as the weak topology. By the Hahn-Banach theorem, the weak closure of $i_X^*E_0^*$ equals its strong closure, and the separability of H_X follows.

As an application of Theorem 4.12 we extend Theorem 3.12 to arbitrary E-valued Gaussian random variables.

Corollary 4.13 (Kahane-Khintchine inequality). *Let X be an E-valued Gaussian variable. Then for all $1\leqslant p,q<\infty$ we have*

$$(\mathbb{E}\|X\|^p)^{\frac{1}{p}}\leqslant K_{p,q}(\mathbb{E}\|X\|^q)^{\frac{1}{q}}.$$

Proof. For the special case where $X=\sum_{n=1}^N\gamma_n x_n$ this was proved in Theorem 3.12. The general case follows by combining this with the Karhunen-Loève expansion. □

4.4 Convergence

As an application of the Kahane-Khintchine inequality, we show next that if a sequence of Gaussian random variables converges in probability, then it converges in L^p for all $1\leqslant p<\infty$.

We start with a classical inequality for non-negative random variables.

Lemma 4.14 (Paley-Zygmund inequality). *Let ξ be a non-negative random variable. If $0<\mathbb{E}\xi^2\leqslant c(\mathbb{E}\xi)^2<\infty$ for some $c>0$, then for all $0<r<1$ we have*

$$\mathbb{P}\{\xi>r\mathbb{E}\xi\}\geqslant\frac{(1-r)^2}{c}.$$

Proof. Using the non-negativity of ξ we have

$$(1-r)\mathbb{E}\xi = \mathbb{E}(\xi - r\mathbb{E}\xi) \leqslant \mathbb{E}(1_{\{\xi > r\mathbb{E}\xi\}}(\xi - r\mathbb{E}\xi)) \leqslant \mathbb{E}(1_{\{\xi > r\mathbb{E}\xi\}}\xi)$$

and therefore, by the Cauchy-Schwarz inequality,

$$(1-r)^2(\mathbb{E}\xi)^2 \leqslant \big(\mathbb{E}(1_{\{\xi > r\mathbb{E}\xi\}}\xi)\big)^2 \leqslant \mathbb{E}1_{\{\xi > r\mathbb{E}\xi\}}\mathbb{E}\xi^2.$$

The result follows upon dividing both sides by $\mathbb{E}\xi^2$. □

Theorem 4.15. *For a sequence $(X_n)_{n=1}^\infty$ of E-valued Gaussian random variables the following assertions are equivalent:*

(1) *the sequence $(X_n)_{n=1}^\infty$ converges in probability to a random variable X;*
(2) *for some $1 \leqslant p < \infty$ the sequence $(X_n)_{n=1}^\infty$ converges in $L^p(\Omega; E)$ to a random variable X;*
(3) *for all $1 \leqslant p < \infty$ the sequence $(X_n)_{n=1}^\infty$ converges in $L^p(\Omega; E)$ to a random variable X.*

In this situation the limit random variable X is Gaussian.

Proof. Fix $1 \leqslant p < \infty$. It suffices to prove that $\lim_{n\to\infty} X_n = X$ in probability implies $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; E)$. Note that X is Gaussian by Proposition 4.11.

Step 1 - Fix $1 \leqslant q < \infty$. By Fernique's theorem we have $\mathbb{E}\|X_n\|^q < \infty$ for all $n \geqslant 1$. In this step we prove the uniform bound

$$\sup_{n\geqslant 1} \mathbb{E}\|X_n\|^q < \infty. \tag{4.5}$$

From the Paley-Zygmund inequality, for all $n \geqslant 1$ we obtain

$$\mathbb{P}\Big\{\|X_n\|^2 > \frac{1}{2}\mathbb{E}\|X_n\|^2\Big\} \geqslant \frac{1}{4K_{4,2}^4}, \tag{4.6}$$

where $K_{4,2}$ is the Kahane-Khintchine constant corresponding to $p = 4$ and $q = 2$. On the other hand, given $\varepsilon > 0$, for any $r > 0$ we find an index $N \geqslant 1$ such that for all $n \geqslant N$,

$$\begin{aligned}&\mathbb{P}\big\{\|X_n\|^2 > r\big\} \\ &\quad\leqslant \mathbb{P}\Big\{\|X\| > \frac{1}{2}\sqrt{r}\Big\} + \mathbb{P}\Big\{\|X_n - X\| > \frac{1}{2}\sqrt{r}\Big\} \leqslant \mathbb{P}\Big\{\|X\| > \frac{1}{2}\sqrt{r}\Big\} + \varepsilon.\end{aligned}$$

Thus for large enough $r_0 > 0$ we find an index $N_0 \geqslant 1$ such that for $n \geqslant N_0$,

$$\mathbb{P}\big\{\|X_n\|^2 > r_0\big\} < 2\varepsilon.$$

If for some subsequence we had $\lim_{k\to\infty} \mathbb{E}\|X_{n_k}\|^2 = \infty$, then for all sufficiently large k we would obtain

$$\mathbb{P}\big\{\|X_{n_k}\|^2 > \frac{1}{2}\mathbb{E}\|X_{n_k}\|^2\big\} \leqslant \mathbb{P}\big\{\|X_{n_k}\|^2 > r_0\big\} < 2\varepsilon,$$

contradicting (4.6). We conclude that $\sup_{n\geqslant 1}\mathbb{E}\|X_n\|^2 < \infty$. Now (4.5) follows from the Kahane-Khintchine inequality.

Step 2 - Fix $1 \leqslant p < q < \infty$. By Step 1, the triangle inequality in $L^q(\Omega; E)$, and a scaling argument we may assume that $\sup_{k\geqslant 1}\mathbb{E}\|X_k - X\|^q \leqslant 1$. Using this together with Hölder's inequality (with $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$), for fixed $\varepsilon > 0$ we obtain

$$\begin{aligned}\mathbb{E}\|X_k - X\|^p &= \mathbb{E}(1_{\{\|X_k - X\|\leqslant\varepsilon\}}\|X_k - X\|^p) + \mathbb{E}(1_{\{\|X_k - X\|>\varepsilon\}}\|X_k - X\|^p)\\ &\leqslant \varepsilon^p + \mathbb{E}(1_{\{\|X_k - X\|>\varepsilon\}}\|X_k - X\|^p)\\ &\leqslant \varepsilon^p + (\mathbb{P}\{\|X_k - X\| > \varepsilon\})^{\frac{p}{r}}.\end{aligned}$$

Since $\lim_{k\to\infty} X_k = X$ in probability, it follows that

$$\limsup_{k\to\infty}\mathbb{E}\|X_k - X\|^p \leqslant \varepsilon^p.$$

This being true for all $\varepsilon > 0$ we arrive at $\limsup_{k\to\infty}\mathbb{E}\|X_k - X\|^p = 0$. □

4.5 Exercises

1. This exercise presents an example of two uncorrelated Gaussian random variables which are not independent. This shows that the joint Gaussianity condition in Proposition 4.10 cannot be omitted.

   Let γ be a standard Gaussian random variable on a probability space $(\Omega_1, \mathscr{F}_1, \mathbb{P}_1)$ and let r be a Rademacher variable on a probability space $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$. Define the random variables φ_1 and φ_2 on the product space $(\Omega, \mathscr{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathscr{F}_1 \times \mathscr{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$ by

   $$\varphi_1(\omega_1, \omega_2) = \gamma(\omega_1), \quad \varphi_2(\omega_1, \omega_2) = \gamma(\omega_1)r(\omega_2).$$

   a) Show that φ_1 and φ_2 are Gaussian.
   b) Show that φ_1 and φ_2 are uncorrelated.
   c) Show that φ_1 and φ_2 fail to be independent.

      Hint: Consider, for instance, the events $\{|\varphi_1| \leqslant 1\}$ and $\{|\varphi_2| \leqslant 1\}$.

2. In this exercise we prove SAZANOV*'s theorem*: a bounded linear operator Q on a separable Hilbert space H with inner product $[\cdot,\cdot]$ is a Gaussian covariance operator if and only if Q is positive, self-adjoint and the sum $\sum_{n=1}^{\infty}[Qh_n, h_n]$ converges for some (equivalently, for every) orthonormal basis $(h_n)_{n=1}^{\infty}$ of H.

   a) Suppose Q satisfies the conditions of the Sazanov theorem, let $(h_n)_{n=1}^{\infty}$ be an orthonormal basis of H, and put $x_n := Q^{\frac{1}{2}}h_n$. Show that the Gaussian sum $\sum_{n=1}^{\infty}\gamma_n x_n$ converges in $L^2(\Omega; H)$ and defines a Gaussian H-valued random variable with covariance Q.

b) Suppose conversely that X is an H-valued Gaussian random variable with covariance operator Q. Then Q is positive and symmetric. Show that if $(h_n)_{n=1}^{\infty}$ is any orthonormal basis for H, then

$$\sum_{n=1}^{\infty}[Qh_n, h_n] = \mathbb{E}\|X\|^2.$$

c) Deduce that the identity operator on a separable infinite-dimensional Hilbert space fails to be a Gaussian covariance operator.

3. (!) The identity (4.4) shows that every Gaussian covariance operator can be written as $Q = TT^*$ for a suitable operator T from a Hilbert space into E. In this exercise we generalise this observation to arbitrary positive symmetric operators.
Let $Q \in \mathscr{L}(E^*, E)$ be positive and symmetric.
a) Show that the formula

$$[Qx^*, Qy^*] := \langle Qx^*, y^* \rangle$$

defines an inner product on the range of Q.

The Hilbert space completion of the range of Q with respect to this inner product is denoted by H_Q.
b) Show that the identity mapping $Qx^* \mapsto Qx^*$ extends uniquely to a bounded operator i_Q from H_Q into E.
c) Prove the identity

$$i_Q i_Q^* = Q.$$

d) Prove the statement concerning Q in the proof of Theorem 3.9.
Hint: Consider an orthonormal basis of the (finite-dimensional) Hilbert space H_Q.

4. Suppose that X is an E-valued Gaussian random variable with covariance operator Q. We compare the mappings $i_X : H_X \to E$ defined by (4.3) and $i_Q : H_Q \to E$ of the previous exercise.
a) Show that the mapping $\langle X, x^* \rangle \mapsto i_Q^* x^*$ extends uniquely to an isometry from H_X onto H_Q.
b) Prove that $i_X(H_X) = i_Q(H_Q)$ and show that X takes its values in $\overline{i_X(H_X)} = \overline{i_Q(H_Q)}$ almost surely.

5. Let Q be a positive self-adjoint operator on a Hilbert space H and let $\sqrt{Q}$ be its unique positive square root.
a) Show that the range of $\sqrt{Q}$ is a Hilbert space with respect to the norm

$$\|\sqrt{Q}h\| := \inf\{\|h'\| : \ h' \in H, \ \sqrt{Q}h' = \sqrt{Q}h\}.$$

b) Show that the identity mapping $Qh \mapsto Qh$ extends uniquely to an isometry

$$H_Q \simeq \mathrm{range}(\sqrt{Q}),$$

where H_Q is defined as in the previous two exercises.

Notes. A comprehensive treatment of the theory of Gaussian variables is given in Bogachev [8]. See also the monographs of Janson [53], Vakhania, Tarieladze, Chobanyan [105], and the older lecture notes of Kuo [62].

Theorem 4.3 is a celebrated result due to Fernique [39]. By a (non-trivial) modification of the proof one obtains the following stronger result: if $\mathscr{X}$ is a uniformly tight family of E-valued Gaussian random variables, then there exist constants $\beta > 0$ and $C > 0$ such that

$$\mathbb{E}(\exp(\beta\|X\|^2) \leqslant C \qquad \forall X \in \mathscr{X}.$$

Using powerful concentration of measure inequalities it can be shown that the supremum of all admissible constants β for which the conclusion of Fernique's theorem holds is equal to $1/2\sigma^2(X)$, where

$$\sigma^2(X) = \sup_{\|x^*\|\leqslant 1} \mathbb{E}\langle X, x^*\rangle^2.$$

We refer to Kwapień and Woyczyński [65], Ledoux [68], and Ledoux and Talagrand [69] for expositions of this result and further reading.

The proof of Theorem 4.15 is taken from Rosiński and Suchanecki [96].

For more on the Karhunen-Loève expansion of Gaussian variables we recommend [65]. The convergence of the series can be alternatively deduced from the martingale convergence theorem for Banach space-valued martingales, but we have chosen not to do so here in order to keep the presentation self-contained.

A Borel measure μ on a Banach space E is called *Gaussian* if it is the distribution of an E-valued Gaussian random variable X, or equivalently, if the image measure $\langle \mu, x^*\rangle$ are Gaussian on $\mathbb{R}$ for all $x^* \in E^*$ (to see that the latter implies the former consider the random variable $X(x) := x$ on the probability space (E, μ)). The *covariance operator* of μ is then defined as the covariance operator Q of X. In view of the identities $\langle Qx^*, x^*\rangle = \mathbb{E}\langle X, x^*\rangle^2 = \int_E \langle x, x^*\rangle^2 \, d\mu(x)$ this is well-defined. For the sake of unity of presentation we have stated all results in terms of random variables. Some results, such as Theorem 4.3 and Propositions 4.6 and 4.9, can equally well be formulated in terms of Gaussian measures.

Exercise 2 c) tells us that on an infinite-dimensional Hilbert space H there is no *standard* Gaussian measure, that is, a Gaussian measure whose covariance operator is the identity operator. More can be said, however. Let us call a subset C of H *cylindrical* if it is of the form

$$C = \{h \in H : \ ([h, h_1], \dots, [h, h_n]) \in B\}$$

for certain $h_1, \dots, h_n \in H$ and a Borel set B in $\mathbb{R}^n$. More generally, cylindrical sets in Banach spaces can be defined by replacing the role of the h_j by functionals x_j^*. We have already used cylindrical sets in the proof of the uniqueness theorem for the Fourier transform (Theorem 2.8). The cylindrical sets form an algebra of sets in H. It can be shown that there exists a unique finitely additive measure γ_H on this algebra with the property that the restrictions of γ_H to finite-dimensional subspaces of H are standard Gaussian measures.

The pair (i_Q, H_Q) constructed in Exercise 3 is called the *reproducing kernel* associated with Q. The operator $i_Q : H_Q \to E$ is in fact injective, and the factorisation $Q = i_Q i_Q^*$ is minimal in the following sense: if H is a Hilbert space and $T : H \to E$ is a bounded operator such that $Q = TT^*$, then there exists a bounded surjection $P : H \to H_Q$ such that $T = i_Q P$. For more information on reproducing kernel Hilbert spaces as well as an explanation of the terminology we refer the interested reader to SCHWARTZ [98] and the book by VAKHANIA, TARIELADZE, CHOBANYAN [105].