Lecture 1: Gaussian Random Element

Tianpei Xie

Jul. 8th., 2015

Contents

1 Gaussian Vector and its Distributions

1.1 Univariate Case

- **Definition** (***Gaussian Random Variable***)
  Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a measurable space, where $\mathcal{B}(\mathbb{R})$ is *the Borel* $\sigma-$*algebra* on $\mathbb{R}$. A *real-valued random variable* X is ***Normally distributed*** or ***Gaussian*** with expectation $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, if its ***distribution density*** with respect to Lebesgue measure is
  $$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

- **Remark** The followings are properties to the ***Gaussian distribution***

  1. The c.d.f. for ***the standard Normal distribution*** $\mathcal{N}(0,1)$ is
  $$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-u^2/2)du$$
  2. $p(x)$ is ***unimodal***, ***symmetric*** about the mean μ and it is ***uniformly bounded*** on $\mathbb{R}$. which has a ***unique maximum*** $\frac{1}{\sqrt{2\pi\sigma^2}}$ at the mean $x = \mu$.
  3. The Normal distribution has ***super-exponential decay tail***; that is, when x moves away from μ, $p(x)$ decreases *monotonically* and *very fast*.
  4. The ***barycenter*** *(or the center of gravity)* of $\mathcal{N}(\mu, \sigma^2)$ is $x = \mu$ due to $\int (x-\mu)p(x)dx = 0$; and the ***second central moment*** $\int (x-\mu)^2 p(x)dx = \sigma^2$.
  5. The ***characteristic function (Fourier transforms)*** and ***moment generating function (Laplace transforms)***
  $$\mathcal{F}\{p\} = \mathbb{E}_p\left[\exp(i\omega x)\right] = \exp\left(i\mu\omega - \frac{1}{2}\omega^2\sigma^2\right)$$
  $$\mathcal{L}\{p\} = \mathbb{E}_p\left[\exp(sx)\right] = \exp\left(s\mu + \frac{1}{2}s^2\sigma^2\right)$$
  6. $\mathcal{N}(\mu_1, \sigma_1^2) * \mathcal{N}(\mu_2, \sigma_2^2) = \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$, where $*$ is the ***convolution operation***. In other words, the family $\left\{\mathcal{N}(\mu, \sigma^2)\right\}$ is ***stable*** with respect to convolutions
  $$\mathcal{P}_1 * \mathcal{P}_2(A) \equiv \int_r \mathcal{P}_1(A-r)\mathcal{P}_2(dr),\ A \in \mathcal{B}^1.$$
  7. The ***Gaussian measure*** is ***convex***. (Note not the density function $p(x)$ but the measure $d\mathcal{P} = p(x)dx$). That is, for any sets $A, B \in \mathcal{B}(\mathbb{R})$, and each $\gamma \in [0,1]$,
  $$\gamma g(\mathcal{P}(A)) + (1-\gamma)g(\mathcal{P}(B)) \le g(\mathcal{P}(\gamma A + (1-\gamma)B))$$
  where $g : \mathbb{R}_+ \to \mathbb{R}_+$ is a normalizing function. For Gaussian measure, $g = \Phi^{-1}$ the inverse c.d.f.

1.2 Multivariate Case

- **Definition** (***Standard Gaussian Random Vector***)
  *A **random vector*** $X = (X_j)_{j=1}^n \in \mathbb{R}^n$ is called ***standard Gaussian***, if its components are ***independent*** and have *a standard normal distribution.* The ***distribution*** of X has a *density*

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^n} \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right), \quad \text{for } \boldsymbol{x} \in \mathbb{R}^n. \tag{1}$$

- **Definition** (***Gaussian Random Vector***)
  *A **random vector*** $Y \in \mathbb{R}^n$ is called ***Gaussian***, if it can be represented as $Y = a + LX$, where X is a *standard Gaussian vector*, $\boldsymbol{a} \in \mathbb{R}^n$, and $L : \mathbb{R}^n \to \mathbb{R}^n$ is a ***linear mapping***.

  Equivalently,

  Definition (***Gaussian Random Vector***)
  *A **random vector*** $Y \in \mathbb{R}^n$ is called ***Gaussian***, if $\langle v\,,\, Y\rangle$ is a *Normal random variable* for ***each*** $v \in \mathbb{R}^n$.

- **Definition** (***Covariance Operator** for **Gaussian Random Vector***)
  Given a Gaussian random vector $X = (X_j)_{j=1}^n$, define the ***covariance operator*** as a ***linear mapping*** $K_X : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\mathrm{cov}(\langle u\,,\, X\rangle\,, \langle v\,,\, X\rangle) = \langle u\,,\, K_X(v)\rangle\,.$$

  The *matrix representation* of K_X is called a ***covariance matrix***

$$\boldsymbol{K} = [K(i,j)]_{i,j=1}^n \in \mathbb{R}^{n\times n}, \qquad \text{where } K(i,j) = \langle e_i\,,\, K_X(e_j)\rangle\,.$$

- **Remark** (***The Covariance Operator is Self-Adjoint and Positive***)
  The covariance operator K is ***self-adjoint*** ($K_X^* = K_X$), ***positive semi-definite*** $K \succeq 0$. This is due to the *symmetry* and *positive definiteness* property of *inner product.*

  Equivalently, the covariance matrix $\boldsymbol{K}$ is ***symmetric***, ***positive semi-definite***.

- **Remark** (***Density for Multivariate Gaussian***)
  In the case, when the linear mapping L is ***invertible (non-degenerate)***, the ***multivariate Normal distribution*** $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K})$ can be defined via *its density function* w.r.t. the Lebesgue measure on $\mathbb{R}^n$

$$p(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \det(|\Sigma|)}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{K}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right) \tag{2}$$

- **Remark** The expression for density in (2) holds only if the linear operator L is invertible; that is, the general definition used is the linear projection definition. [Lifshits, 2013]

  If L is singular, K is singular, i.e., $\det K = 0$; there is *no proper density expression* as (2). On the other hand, for every K *nonnegative definite*, $L = K^{1/2}$ exists and is nonnegative definite as well.

- **Remark** (***The Characteristic Function of Multivariate Gaussian***)
  The characteristic functions of $\mathcal{N}(\boldsymbol{\mu}, K)$ is determined by its one-dimension projection

$$
\begin{aligned}
\varphi(\boldsymbol{v}) &= \int \exp\left(i \left\langle \boldsymbol{x}\,,\, \boldsymbol{v} \right\rangle\right) \mathcal{P}(d\boldsymbol{x}) \qquad (3)\\
&= \int \exp\left(ir\right) \mathcal{P}^{\boldsymbol{v}}(dr)\\
&= \exp\left(i\mu^{v}\omega - \frac{1}{2}\sigma^{2}(\boldsymbol{v})\omega^{2}\right)\bigg|_{\omega=1}\\
&= \exp\left(i \left\langle \boldsymbol{\mu}\,,\, \boldsymbol{v} \right\rangle - \frac{1}{2}\left\langle K\boldsymbol{v}\,,\, \boldsymbol{v} \right\rangle\right)
\end{aligned}
$$

  The equation (3) is known as the *characteristic functional* of measure $\mathcal{P}$.

  Use the affine mapping $\boldsymbol{\mu} + L\mathcal{P}_0$, the characteristic functional is given by

$$
\begin{aligned}
\varphi(\boldsymbol{v}) &= \int \exp\left(i \left\langle \boldsymbol{\mu} + L\boldsymbol{x}\,,\, \boldsymbol{v} \right\rangle\right) \mathcal{P}_0(d\boldsymbol{x})\\
&= \exp\left(i \left\langle \boldsymbol{\mu}\,,\, \boldsymbol{y} \right\rangle\right) \int \exp\left(i \left\langle L\boldsymbol{x}\,,\, \boldsymbol{v} \right\rangle\right) \mathcal{P}_0(d\boldsymbol{x})\\
&= \exp\left(i \left\langle \boldsymbol{\mu}\,,\, \boldsymbol{y} \right\rangle\right) \int \exp\left(i \left\langle \boldsymbol{x}\,,\, L^{*}\boldsymbol{v} \right\rangle\right) \mathcal{P}_0(d\boldsymbol{x})\\
&= \exp\left(i \left\langle \boldsymbol{\mu}\,,\, \boldsymbol{y} \right\rangle - \frac{1}{2}\left\langle L^{*}\boldsymbol{v}\,,\, L^{*}\boldsymbol{v} \right\rangle\right)\\
&= \exp\left(i \left\langle \boldsymbol{\mu}\,,\, \boldsymbol{y} \right\rangle - \frac{1}{2}\left\langle LL^{*}\boldsymbol{v}\,,\, \boldsymbol{v} \right\rangle\right)\\
&= \exp\left(i \left\langle \boldsymbol{\mu}\,,\, \boldsymbol{y} \right\rangle - \frac{1}{2}\left\langle K\boldsymbol{v}\,,\, \boldsymbol{v} \right\rangle\right)
\end{aligned}
$$

  And the density is computed, for L invertible, by change of variable for $\boldsymbol{y} = \boldsymbol{\mu} + L\boldsymbol{x}$

$$
\begin{aligned}
p_{\boldsymbol{\mu},K}(\boldsymbol{y}) &= \left|\det L\right|^{-1} p(\boldsymbol{x})\\
&= (2\pi)^{n/2} \left|\boldsymbol{K}\right|^{-1/2} \exp\left(-\left\langle K^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\,,\, \boldsymbol{y}-\boldsymbol{\mu} \right\rangle/2\right)
\end{aligned}
$$

- **Proposition 1.1** ***(Existence and Uniqueness of Gaussian Distribution)****[Lifshits, 2013]*
  Let $\mathcal{P}$ be a Gaussian distribution in $\mathbb{R}^n$. Then the mean value $\boldsymbol{\mu}$ and the covariance operator K of the measure $\mathcal{P}$ exist and are ***uniquely*** *defined. The operator K is* ***symmetric*** *and* ***positive definite****.*

- **Proposition 1.2** ***(Gaussian Random Vector from Kernel)*** *[Lifshits, 2013]*
  Assume $\boldsymbol{\mu} \in \mathbb{R}^n$ and $K : \mathbb{R}^n \to \mathbb{R}^n$ is nonnegative definite linear operator. Then there exists a unique Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, K)$ with mean $\boldsymbol{\mu}$ and covariance operator K. The characteristic functional of $\mathcal{N}(\boldsymbol{\mu}, K)$ has the form of (3). If the operator K is non-singular, the distribution $\mathcal{N}(\boldsymbol{\mu}, K)$ is absolutely continuous with respect to the Lebesgue measure, and its density is of form (2). There are no other Gaussian distribution in $\mathbb{R}^n$, except for the form $\mathcal{N}(\boldsymbol{\mu}, K)$.

2 Gaussian Random Element

2.1 Gaussian Random Element in Topological Vector Space

- **Definition** (***Random Element in Topological Vector Space***)
  Let $(\Omega, \mathscr{F}, \mathcal{P})$ be *a probability space*, $(\mathcal{X}, \mathscr{B})$ be a ***topological vector space*** with *σ-algebra $\mathscr{B}$*. A ***random element*** in $\mathcal{X}$ is *a $\mathscr{F}/\mathscr{B}$-measurable function* $X : \Omega \to \mathcal{X}$ so that

  $$X^{-1}(A) \in \mathscr{F}, \quad \forall A \in \mathscr{B}.$$

  We write $X \in \mathcal{X}$.

- **Definition** (***Duality***)
  Let $\mathcal{X}^*$ be *the* ***dual*** *space of* $\mathcal{X}$, i.e. *the space of* ***bounded linear functional*** *on* $\mathcal{X}$.

  We denote $\langle f\,,\,x \rangle$ the ***duality*** *between the spaces X and X^**, i.e.

  $$\langle f\,,\,x \rangle := f(x), \quad \forall f \in X^*, x \in X.$$

  Note that we ***do not confuse this notation with inner product***. In inner product $\langle x\,,\,y \rangle$ both arguments are from *the same space*.

- **Definition** (***Gaussian Random Element in Topological Vector Space***)
  A ***random element*** *$X \in \mathcal{X}$ is called* <u>***Gaussian***</u>, if

  $$\langle f\,,\,X \rangle := f(X)$$

  is a ***Normal random variable***, for all $f \in \mathcal{X}^*$.

- **Definition** (***Expectation***)
  A vector $a \in \mathcal{X}$ is called <u>***expectation***</u> *of a random element $X \in \mathcal{X}$* , if

  $$\mathbb{E}\left[\langle f\,,\,X \rangle\right] = \langle f\,,\,a \rangle$$

  for all $f \in \mathcal{X}^*$. We write $a = \mathbb{E}\left[X\right]$.

- **Definition** (***Covariance Operator***)
  A ***linear operator*** $K : \mathcal{X}^* \to \mathcal{X}$ is called <u>***covariance operator***</u> of *a random vector $X \in \mathcal{X}$*, if

  $$\operatorname{cov}(\langle f\,,\,X \rangle\,, \langle g\,,\,X \rangle) = \langle f\,,\,K\,g \rangle\,.$$

  for all $f, g \in X^*$. We write $K = \operatorname{cov}(X)$.

  Remark (***Covariance*** *as* ***Function-Valued Linear Transformation*** *on* ***Dual Space***)
  The covariance operator $K : \mathcal{X}^* \to \mathcal{X}$ acts on linear functional on $\mathcal{X}$ and returns an element (function) in $\mathcal{X}$

  $$f(Kg) := \operatorname{cov}(f(X),\, g(X))$$

- **Remark** (***Covariance Operator is Self-Adjoint and Positive***)
  Covariance operator *is* <u>***self-adjoint***</u>, due to symmetric property of covariance in $\mathbb{R}$.

  $$\langle f\,,\,Kg \rangle = \langle g\,,\,Kf \rangle\,, \quad \forall f, g \in X^*,$$

and it is ***positive (semi-definite)***, i.e.

$$\langle f\,,\, Kf\rangle = \mathrm{var}(f(X)) \geq 0, \quad \forall f \in X^{*}.$$

- **Remark** (***Topological Constraints*** *on $\mathcal{X}$ for* ***Gaussian Element***) [Lifshits, 2012]
  From the definition of Gaussian element, we see that it only makes sense when *the space of continuous linear functionals on $\mathcal{X}$ is **rich** enough.* For example, if $\mathcal{X}^{*} = \{0\}$, then any vector satisfies this definition rendering it *senseless.*

  Therefore, usually *one of* ***three situations*** of increasing generality is considered.

  1. $\mathcal{X}$ is a ***separable Banach space***, for example, $\mathcal{C}[0,1]$, $L^{p}[0,1]$ etc;
  2. $\mathcal{X}$ is a ***complete separable locally convex metrizable topological vector space***, for example, $\mathcal{C}[0,\infty)$, $\mathbb{R}^{\infty}$ etc.
  3. $\mathcal{X}$ is a ***locally convex topological vector space*** and a vector X is such that its distribution is a ***Radon measure***.

  In cases (1) and (2) *every* ***finite measure*** *is a* ***Radon measure***, thus case (3) is the most general one. These assumptions are called *usual assumptions* in [Lifshits, 2012, 2013]

- **Proposition 2.1** *(**Existence of Covariance Operator**) [Lifshits, 2013]*
  *Under usual assumptions on $\mathcal{X}$, every Gaussian random element in $\mathcal{X}$ possesses an **expectation** and a **covariance operator**. In other words, the distribution of Gaussian elements in $\mathcal{X}$ is of the form $\mathcal{N}(a,K)$.*

- **Remark** (***Distribution and Characteristic Function of Gaussan Random Element***)
  The pair (a,K) determines *the* ***distribution*** *of a Gaussian variable* $\langle f\,,\, x\rangle$ as

  $$\mathcal{N}(\langle f\,,\, a\rangle\,, \langle f\,,\, Kf\rangle),$$

  and we find ***the characteristic function*** of $\langle f\,,\, x\rangle$

  $$\begin{aligned}\varphi(\langle f\,,\, X\rangle) &= \mathbb{E}\left[\exp\left\{i\omega\,\langle f\,,\, x\rangle\right\}\right]\\ &= \exp\left(i\omega\,\langle f\,,\, a\rangle - \frac{1}{2}\omega^{2}\,\langle f\,,\, Kf\rangle\right)\\ &:= \exp\left(i\omega f(a) - \frac{1}{2}\omega^{2} f(Kf)\right)\end{aligned}$$

  Any *Radon distribution* in $\mathcal{X}$ is *determined* by *its characteristic function.* Therefore, distribution $\mathcal{N}(a,K)$ is ***unique***.

2.2 Examples of Gaussian Random Elements

- **Example** (***Standard Gaussian Measure in*** $\mathbb{R}^{\infty}$)
  Consider the space $\mathcal{X} = \mathbb{R}^{\infty}$ of *all countable infinite sequence* $(x_1, x_2, \ldots)$ equipped with the ***product topology***. The product topology induces a metric as

  $$\rho(\{x_n\}_{n=1}^{\infty}\,, \{y_n\}_{n=1}^{\infty}) = \sup_{n}\left\{\frac{\min|x_n - y_n|\,, 1}{n}\right\}.$$

$\mathbb{R}^{\infty}$ is a ***complete separable metric space*** under the product topology. The dual space $\mathcal{X}^* = c_0$ is *the space of sequences* $(f_1, f_2, \ldots)$ *with* $f_n = 0$ *for all but finite number of* n. The duality

$$\langle f\,,\,x\rangle = \sum_{n=1}^{\infty} f_n x_n < \infty.$$

Consider a sequence of *i.i.d.* $\mathcal{N}(0,1)$-*distributed random variables* as a vector $X \in \mathcal{X}$, i.e. $X := (X_n)_{n=1}^{\infty}$, $X_n \sim \mathcal{N}(0,1)$. Due to *stability* of normal distribution, for any $f \in \mathcal{X}^*$ the random variable

$$\langle f\,,\,X\rangle = \sum_{n=1}^{\infty} f_n X_n \sim \mathcal{N}(0,\sigma^2)$$

where $\sigma^2 = \sum_{n=1}^{\infty} f_n^2 < \infty$. Therefore, X is a ***Gaussian element***. It is clear that $\mathbb{E}\left[X\right] = 0$.

Embedding operator serves as ***covariance operator*** for X, i.e.

$$K = \iota : c_0 \hookrightarrow \mathbb{R}^{\infty}.$$

To show that

$$\begin{aligned}
\mathrm{cov}\left(\langle f\,,\,X\rangle\,,\langle g\,,\,X\rangle\right) &= \mathbb{E}\left[\langle f\,,\,X\rangle\,\langle g\,,\,X\rangle\right] \\
&= \mathbb{E}\left[\left(\sum_{n=1}^{\infty} f_n X_n\right)\left(\sum_{n=1}^{\infty} g_n X_n\right)\right] \\
&= \mathbb{E}\left[\sum_{n,m=1}^{\infty} f_n g_m X_n X_m\right] \\
&= \sum_{n,m=1}^{\infty} f_n g_m \mathbb{E}\left[X_n X_m\right] = \sum_{n,m=1}^{\infty} f_n g_m\, \delta_{n,m} \\
&= \sum_{n=1}^{\infty} f_n g_n := \langle f\,,\,Kg\rangle
\end{aligned}$$

We call *the distribution of* X *a* ***standard Gaussian measure in*** $\mathbb{R}^{\infty}$. ∎

- **Example** (***Gaussian Elements in a Hilbert space*** $\mathcal{H}$) [Lifshits, 2012]
  Let $\mathcal{X} = \mathcal{H}$ be a ***separable Hilbert space*** whose inner product will be denoted by $\langle\cdot\,,\,\cdot\rangle_{\mathcal{H}}$. By the Riesz representation theorem, we can identify its dual space $\mathcal{H}^*$ with $\mathcal{H}$, i.e. for each $f \in \mathcal{H}^*$, there exists a unique $x_f \in \mathcal{H}$ such that

  $$\langle f\,,\,x\rangle = f(x) = \langle x\,,\,x_f\rangle_{\mathcal{H}}\,, \quad \forall x \in \mathcal{H}.$$

  Define $h : \mathcal{H}^* \to \mathcal{H}$ as an *isometric isomorphism* that maps $f \mapsto x_f$.

  In order to construct a Gaussian element in $\mathcal{H}$, consider a ***complete orthonormal basis*** $\{\varphi_n\}_{n=1}^{\infty}$ on $\mathcal{H}$, a sequence of ***independent*** $\mathcal{N}(0,1)$***-distributed random variables*** $\{\xi_n\}_{n=1}^{\infty}$, and a sequence of ***non-negative numbers*** $\{\sigma_n\}_{n=1}^{\infty}$ satisfying assumption $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$ so that the series

  $$\sum_{n=1}^{\infty} \sigma_n \xi_n(\omega) \varphi_n$$

is ***convergent*** in $\|\cdot\|_{\mathcal{H}}$***-norm almost surely*** in $\mathcal{H}$. Define a *random element* $X : \Omega \to \mathcal{H}$ as the limit of the series

$$X = \sum_{n=1}^{\infty} \sigma_n \xi_n \varphi_n \tag{4}$$

This representation is called ***Karhunen-Loève expansion***.

For any linear functional $f \in \mathcal{H}^*$, we can its corresponding vector $x_f = h(f) \in \mathcal{H}$ and $x_f = \sum_{n=1}^{\infty} f_n \varphi_n$. Thus the random variable

$$\begin{aligned}
\langle f \,,\, X \rangle = \langle X \,,\, x_f \rangle_{\mathcal{H}} &= \left\langle \sum_{n=1}^{\infty} \sigma_n \xi_n \varphi_n \,,\, \sum_{n=1}^{\infty} f_n \varphi_n \right\rangle_{\mathcal{H}} \\
&= \sum_{n,m=1}^{\infty} \sigma_n \bar{f}_m \xi_n \left\langle \varphi_n \,,\, \varphi_m \right\rangle_{\mathcal{H}} \\
&\text{by orthonormal } \left\langle \varphi_n \,,\, \varphi_m \right\rangle_{\mathcal{H}} = \delta_{n,m} \\
&= \sum_{n=1}^{\infty} \sigma_n \bar{f}_n \xi_n \sim \mathcal{N}(0, \sigma^2)
\end{aligned}$$

where $\sigma^2 := \sum_{n=1}^{\infty} \sigma_n^2 f_n^2 \le (\sum_{n=1}^{\infty} \sigma_n^2) \sup_n |f_n|^2 < \infty$. Therefore, X is a ***Gaussian random element*** in $\mathcal{H}$ and $\mathbb{E}[X] = 0$. In order to find ***the covariance operator*** of X, let us compute

$$\begin{aligned}
\operatorname{cov}\left(\langle f \,,\, X\rangle \langle g \,,\, X\rangle\right) &= \mathbb{E}\left[\langle f \,,\, X\rangle \langle g \,,\, X\rangle\right] \\
&= \mathbb{E}\left[\left(\sum_{n=1}^{\infty} \sigma_n \bar{f}_n \xi_n\right)\left(\sum_{n=1}^{\infty} \sigma_n \bar{g}_n \xi_n\right)\right] \\
&= \sum_{n,m=1}^{\infty} \bar{f}_n \bar{g}_m \sigma_n \sigma_m \mathbb{E}\left[\xi_n \xi_m\right] \\
&\text{since } \mathbb{E}\left[\xi_n \xi_m\right] = \delta_{n,m} \\
&= \sum_{n=1}^{\infty} \sigma_n^2 \bar{f}_n \bar{g}_n = \langle f \,,\, Kg \rangle
\end{aligned}$$

By plugging in the basis, we have

$$K : g \to \sum_{n=1}^{\infty} \sigma_n^2 g_n \varphi_n = \sum_{n=1}^{\infty} \sigma_n^2 \left\langle g \,,\, \varphi_n \right\rangle \varphi_n \tag{5}$$

$$\Rightarrow \widetilde{K} = K \circ h^{-1} = \sum_{n=1}^{\infty} \sigma_n^2 \left\langle \cdot \,,\, \varphi_n \right\rangle_{\mathcal{H}} \varphi_n \tag{6}$$

Therefore σ_n^2 and φ_n are the ***eigenvalues*** and ***eigenfunctions*** of $\widetilde{K} = K \circ h^{-1}$ and $\widetilde{K}$ is a ***positive, compact operator*** on $\mathcal{H}$ since $\operatorname{tr}(\widetilde{K}) = \sum_{n=1}^{\infty} \sigma_n^2 < \infty$.

One can show that ***any Gaussian element in a Hilbert space admits a representation*** (4) [Lifshits, 2012]. This means that *a* ***Gaussian distribution*** *with* ***covariance operator*** K ***exists*** ***if and only if*** the induced linear operator $\widetilde{K} = K \circ h^{-1} \in \mathcal{L}(\mathcal{H})$ is a ***self-adjoint, positive***, ***trace-class operator*** (which is ***compact***). ∎

- **Remark** (***Equivalent Definition of Covariance Operator on Hilbert Space***)
In the previous example, we see that ***the covariance operator on Hilbert space*** can be equivalently *defined* via linear operator $\widetilde{K} : \mathcal{H} \to \mathcal{H}$ so that

$$\text{cov}\left(\langle f_h\,,\, X\rangle_{\mathcal{H}}\,,\, \langle g_h\,,\, X\rangle_{\mathcal{H}}\right) = \left\langle \widetilde{K} f_h\,,\, g_h \right\rangle_{\mathcal{H}}.$$

Note that $\widetilde{K} \succeq 0$ is ***self-adjoint*** and ***positive*** and it has ***finite trace*** $\text{tr}(\widetilde{K})$ so it is ***trace-class operator*** which is ***compact***. And, conversely, for any ***positive trace-class operator*** $K \in \mathcal{B}_1(\mathcal{H})$, there exists ***Gaussian element*** *in* $\mathcal{H}$ *with distribution* $\mathcal{N}(0, K)$.

- **Remark** (***Identity Operator is Not Covariance Operator on Hilbert Space***)
For identity operator $I : \mathcal{H} \to \mathcal{H}$, we see that its trace $\text{tr}\,(I) = \infty$, this means that it *does not admit* a Gaussian distribution as $\mathcal{N}(0, I)$ on infinite dimensional space $\mathcal{H}$. In fact, we can see that $\mathbb{E}\left[|X(t)|^2\right] = \infty$.

2.3 Gaussian Random Process

- **Definition** (***Random Process***)
Let $(\Omega, \mathscr{F}, \mathcal{P})$ be a probability space and T be a parametric set called ***index set***. *A* ***random process*** X on T is *a family of random variables* $X(t, \omega), t \in T$, defined on the ***common*** *probability space* $(\Omega, \mathscr{F}, \mathcal{P})$. For each $\omega \in \Omega$,

$$X(\omega) := \{X_t(\omega) : t \in T\}$$

is called a ***sample function*** of (X_t) and if T is one-dimensional, they are often called ***sample paths*** *of the process* (X_t).

- **Remark** Determined by index set T, we have:

  1. if $T \subset \mathbb{R}$, $\{X_t\}_{t\in T}$ is called a ***random process***.
  2. if $T \subset \mathbb{R}^n$, $\{X_t\}_{t\in T}$ is called a ***random field***.
  3. if $T = \mathbb{N}$, $\{X_t\}_{t\in T}$ is called a ***random sequence***.

- **Definition** (***Gaussian Random Process***)
A process $(X_t)_{t\in T}$ is called ***Gaussian*** if for any $t_1, \ldots, t_n \in T$ the *distribution of the random vector*

$$(X(t_1)\,, \ldots, X(t_n))$$

is a ***Gaussian distribution*** in $\mathbb{R}_n$.

The properties of a ***Gaussian process*** are ***completely determined*** by its ***expectation*** $\mathbb{E}\left[X(t)\right], t \in T$, and ***covariance*** $\text{cov}(X(s), X(t)), s, t \in T$.

- **Remark** (***Gaussian Random Process as Gaussian Element on Function Space***)
Consider *the topological vector space* $\mathcal{X} \subset \mathbb{R}^T$ as *a* ***function space*** *on* T, then the ***Gaussian random element*** in $\mathcal{X}$ is a *Gaussian process*:

$$\begin{aligned} &X : \Omega \to \mathcal{X} \subset \mathbb{R}^T \\ &\Rightarrow X(\omega)(t) = X(\omega, t), \forall t \in T \end{aligned}$$

- **Definition** (***Continuous Sample Path***)
If T is a ***topological space***, we say that $\{X_t\}_{t \in T}$ has ***continuous sample paths***, if the function $X(\cdot, \omega)$ is ***continuous*** *on* T for $\mathcal{P}$*-almost every* $\omega \in \Omega$.

2.4 Examples of Gaussian Random Processes

- **Example** (***Continuous Sample Path Gaussian Process***) [Lifshits, 2012]
Let T be a ***compact metric space***, let $\mathcal{X} = \mathcal{C}(T)$ denote *the* ***Banach space*** *of all* ***continuous functions*** *on* T *equipped with supremum norm*

$$\|x\|_{\infty} := \sup_{t \in T} |x(t)|$$

and with the corresponding ***topology of uniform convergence***. By *Riesz-Markov theorem*, the ***dual space*** $\mathcal{X}^* = \mathcal{M}(T)$ is a *space of* ***signed Radon measures of finite variations*** *on* T. The duality is given by

$$\langle \mu \,,\, f \rangle = \int_T f \, d\mu, \quad \forall f \in \mathcal{X}, \forall \mu \in \mathcal{M}(T) = \mathcal{X}^*.$$

Let $\{X(t), t \in T\}$, be a ***Gaussian random process*** with ***continuous sample paths*** on the parametric set T. It is *completely characterized* by the functions

$$a(t) := \mathbb{E}\left[X(t)\right], \quad K(s,t) := \mathrm{cov}\left(X(s)\,,\, X(t)\right).$$

Then we can view at $X := \{X(t), t \in T\}$ as a ***Gaussian random element*** of $\mathcal{X}$. The ***expectation*** of X is computed as

$$\mathbb{E}\left[X\right] = a := (a(t))_{t \in T},$$

and the ***covariance operator*** $K : \mathcal{M}(T) \to \mathcal{C}(T)$ can be calculated by

$$(K\nu)(s) = \int_T K(s,t)\nu(dt). \tag{7}$$

This is because

$$\begin{aligned}
\mathrm{cov}\left(\langle \mu \,,\, X\rangle \,,\, \langle \nu \,,\, X\rangle\right) &= \mathbb{E}\left[\langle \mu \,,\, (X-a)\rangle \langle \nu \,,\, (X-a)\rangle\right] \\
&= \mathbb{E}\left[\int_T (X-a) d\mu \int_T (X-a) d\nu\right] \\
&= \mathbb{E}\left[\int_{T \times T} (X(s)-a(s))(X(t)-a(t))\mu(ds)\nu(dt)\right] \\
&= \int_T \int_T \mathbb{E}\left[(X(s)-a(s))(X(t)-a(t))\right] \mu(ds)\nu(dt) \\
&= \int_T \left(\int_T K(s,t)\nu(dt)\right) \mu(ds) := \langle \mu \,,\, K\nu \rangle,
\end{aligned}$$

thus we have (7). ∎

- **Example** (***Wiener Process***) [Lifshits, 2012]
  We will now consider $T = [0, 1]$ and $\mathcal{X} = \mathcal{C}[0, 1]$ with *dual* $\mathcal{M}[0, 1]$. Define a *Gaussian element* composed of *the sample paths of a* ***Wiener process***

$$\mathcal{W} := \mathcal{W}(t), \quad 0 \leq t \leq 1,$$

i.e. of a process satisfying assumptions

$$\mathbb{E}\left[\mathcal{W}(t)\right] = 0, \quad \mathbb{E}\left[\mathcal{W}(s)\mathcal{W}(t)\right] = \min\left\{s, t\right\}.$$

It is just a special case of previous example, so we can find the expectation of $\mathcal{W}$ by

$$\mathbb{E}\left[\langle \mu \,,\, \mathcal{W} \rangle\right] = \mathbb{E}\left[\int_{[0,1]} \mathcal{W} d\mu\right] = \int_0^1 \mathbb{E}\left[\mathcal{W}(t)\right] \mu(dt) = 0$$

we have $\mathbb{E}\left[\mathcal{W}\right] = 0$. Moreover, the ***covariance operator*** $K : \mathcal{M}([0, 1]) \to \mathcal{C}([0, 1])$

$$\begin{aligned}(K\nu)(s) &= \int_0^1 K(s, t)\nu(dt) \\ &= \int_0^1 \min\left\{s, t\right\} \nu(dt). \qquad \blacksquare\end{aligned}$$

Remark Finally, we recall *the properties of Wiener process* $\mathcal{W}(t)$: [Lifshits, 2012]

1. It is 1/2-***self-similar***, i.e. for any $c > 0$ the process

$$Y(t) := \frac{\mathcal{W}(ct)}{\sqrt{c}}$$

   is also a ***Wiener process***;

2. It has ***stationary increments***;

3. It has ***independent increments***;

4. It is a ***Markov process***;

5. It admits ***time inversion***: the process

$$Z(t) := t\mathcal{W}\left(\frac{1}{t}\right)$$

   is also a ***Wiener process***.

References

Mikhail Lifshits. Lectures on gaussian processes. In *Lectures on Gaussian Processes*, pages 1–117. Springer, 2012.

Mikhail Antolevich Lifshits. *Gaussian random functions*, volume 322. Springer Science & Business Media, 2013.