Stochastic Partial Differential Equations

Based on Lectures by Dr. A. Chandra and Dr. G. Cannizzaro

November 27, 2019

Contents

1. Gaussian Measure Theory

1.1. Introduction

The driving motivation throughout this chapter is defining a Banach space valued Wiener process and defining how to integrate against it. Recall a $\mathbb{R}^n$-valued stochastic process $(W_t)_{t\geq 0}$ is a Wiener process if and only if it is a continuous Gaussian process with mean 0 and covariance function $C(s,t) = s \wedge t$. This is a nice short definition, and it's easy to take for granted the amount of measure theory it took to get to this definition. Let's spell some of those steps out.

1. We developed tools like mean, covariance and Fourier transforms.

2. We defined what a Gaussian random variable was and proved properties of Gaussian random variables.

3. We showed existence of a Gaussian process with the correct covariance structure.

4. We proved a continuity criterion in order to modify our Gaussian process to a continuous Gaussan process.

For a Banach space $\mathcal{B}$-valued stochastic process, we currently have none of these tools. The goal of this chapter is to remedy that.

1.2. Analysis of $\mathcal{B}$-valued random variables

1.2.1. Two approaches

Throughout these notes $\mathcal{B}$ be a real Banach space. Usually $\mathcal{B}$ will also be seperable, but we'll try and say explicitly when this is the case. Let μ be a Borel probability measure on $\mathcal{B}$ and X be a $\mathcal{B}$-valued random variable such that $X \sim \mu$. We'd like recover tools like means and Fourier transforms. Let's start with the mean, say m. From the finite dimensional case it's tempting to define

$$m := \int_{\mathcal{B}} x\,\mu(\mathrm{d}x). \tag{1.1}$$

However this integral cannot be understand as a Lebesgue integral. After all, x doesn't take values in $\mathbb{R}^n$. At this point we have two options.

Option One: The Bochner Integral

We could try and rigorously define this integral. The method of integration used for $\mathcal{B}$-valued functions are Bochner integrals. The construction of the Bochner integral follows the same method as the Lebesgue integral; we define the integral for a class of simple functions and extend the integral by approximation. Rather than going into details on the construction, we'll list off some of the properties of the Bochner integral. Firstly let $(E, \mathcal{E}, \nu)$ be a probability space and $f : E \to \mathcal{B}$ be measurable and $\mathcal{B}$ be seperable[1]. Then f is integrable if and only if

$$\int_E \|f(x)\|_E \, \nu(\mathrm{d}x) < \infty. \tag{1.2}$$

The Bochner integral behaves nicely with bounded linear operators. Let $\hat{\mathcal{B}}$ be another Banach space and $L : \mathcal{B} \to \hat{\mathcal{B}}$ be a bounded linear operator. Then the Bochner integral can be exchanged with L, i.e. for all integrable $f : E \to \mathcal{B}$,

$$\int_E L(f(x)) \, \nu(\mathrm{d}x) = L\left(\int_E f(x) \, \nu(\mathrm{d}x) \right). \tag{1.3}$$

Hence if we assume $\int_E \|x\| \, \mu(\mathrm{d}x) < \infty$ then the mean

$$\bar{m}_\mu := \int_E x \, \mu(\mathrm{d}x) \tag{1.4}$$

can be defined using the Bochner integral. This approach is valuable in how it carries the intuition from the finite dimensional case. In addition, it makes clear that the mean is an element of $\mathcal{B}$ without the use of reflexivity (as we will see in the next approach).

Option Two: One-Dimensional Marginals

Alternatively, we've already developed a rich theory of Lebesgue integration for $\mathbb{R}$-valued functions. What if we instead used this theory? This would require a 'natural' way to map $\mathcal{B}$ to $\mathbb{R}$. Since $\mathcal{B}$ is a normed vector space, we have exactly that: the space of bounded linear functionals $\mathcal{B}^*$. For any $l \in \mathcal{B}^*$, $\mathbb{E}[l(X)]$ is just

---

[1]Seperability of $\mathcal{B}$ is need to ensure f is the pointwise limit of simple functions, i.e. strong measurablity. If we take strong measurability as an additional assumption on f then seperability of $\mathcal{B}$ can be relaxed

given by a Lebesgue integral. So to define the mean, just gather all these values for diffent $l \in \mathcal{B}^*$. Following this idea, we define the mean as $m : \mathcal{B}^* \to \mathbb{R}$ where

$$m(l) := \mathbb{E}[l(X)] = \int_{\mathcal{B}} l(x)\, \mu(\mathrm{d}x). \tag{1.5}$$

How can we recover a mean $\bar{m}_\mu \in \mathcal{B}$? Suppose that $\int_E \|x\|\, \mu(\mathrm{d}x) < \infty$. We can then prove that $m_\mu : \mathcal{B}^* \to \mathbb{R}$ is a bounded linear functional, in other words $m_\mu \in \mathcal{B}^{**}$. At this point we could assume that $\mathcal{B}$ were reflexive. Then we are guaranteed there exists $\bar{m}_\mu \in \mathcal{B}$ such that

$$m_\mu(l) = l(\bar{m}_\mu) \quad \forall l \in \mathcal{B}^*. \tag{1.6}$$

However, this is unnecessary supposing $\mathcal{B}$ is seperable. From our first approach of using Bochner integrals, we know in fact that $\int_\mu \|x\|\, \mu(\mathrm{d}x) < \infty$ is a sufficient condition to define $\bar{m}_\mu$ as in. Then since the Bochner integral can be exchanged with any $l \in \mathcal{B}^*$, the property in holds. This shows two things. Firstly, it shows our two approaches recover the same mean. Secondly, it shows reflexivity is an extraneous assumption.

The value of this second approach really shines when the concept we're describing becomes more intangible than a mean, for example when defining Gaussianity on a Banach space.

1.2.2. Random Variables on Banach Spaces

Seen as we're working with $l(X)$ so much, it's useful to have a notation for its law.

Definition 1.1 (*Pushforward measures*)

Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. Let μ be a measure on $(E, \mathcal{E})$ and $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ be measurable. Then the pushforward of μ be f is the measure $f_*\mu$ on $(F, \mathcal{F})$ where

$$f_*\mu(A) := \mu(f^{-1}(A)). \tag{1.7}$$

This is defined so that $l(X) \sim l_*\mu$. It is also sensible to have a name for the laws $\{l_*\mu : l \in \mathcal{B}^*\}$.

Definition 1.2 (*One-dimensional marginals*)

Let μ be a probability measure on $\mathcal{B}$. Then the one-dimenional margins of μ are the laws $l_*\mu$ for all $l \in \mathcal{B}^*$.

The next proposition shows that the one-dimensional marginals carry enough information to uniquely characterise a measure μ.

Proposition 1.3 (*Sufficiency of one-dimensional marginals*)

> *Let μ and ν be Borel probability measures on a seperable Banach space $\mathcal{B}$. Suppose $l_*\mu = l_*\nu$ for all $l \in \mathcal{B}^*$. Then $\mu = \nu$.*

Proof. We start in finite dimensions, suppose $\mathcal{B} = \mathbb{R}^n$. Then considering Fourier transforms. All $\xi \in \mathbb{R}^n$ defines a bounded linear functional $l_\xi : \mathbb{R}^n \to \mathbb{R}^n$ by $l_\xi(x) = \xi \cdot x$. Therefore

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{i\xi \cdot x} \mu(\mathrm{d}x) = \int_{\mathbb{R}} e^{it} \, (l_\xi)_* \mu(\mathrm{d}t) \tag{1.8}$$

$$= \int_{\mathbb{R}} e^{it} \, (l_\xi)_* \nu(\mathrm{d}t) = \hat{\nu}(\xi). \tag{1.9}$$

Since Fourier transforms uniquely determine probability measures, $\mu = \nu$.

For infinite dimensions, the cylindrical sets of $\mathcal{B}$ are given by

$$\mathrm{Cyl}(\mathcal{B}) := \{T^{-1}(A) : T : \mathcal{B} \to \mathbb{R} \text{ linear and bounded, } A \subset \mathbb{R}^n \text{ Borel}\}. \tag{1.10}$$

This is a subset of the Borel σ-algebra on $\mathcal{B}$ by continuity of $\mathcal{B}$. This is a π-system since

$$T^{-1}(A) \cap S^{-1}(B) = (T \oplus S)^{-1}(A \times B). \tag{1.11}$$

By the finite dimensional case we just proved, $T_*\mu$ is uniquely determined by

$$\{L_*(T_*\mu) : L \in (\mathbb{R}^n)^*\} = \{(L \circ T)_*\mu : L \in (\mathbb{R}^n)^*\} \subset \{l_*\mu : l \in \mathcal{B}^*\}. \tag{1.12}$$

Thus μ and ν agree on $\mathrm{Cyl}(\mathcal{B})$ and therefore $\sigma(\mathrm{Cyl}(\mathcal{B}))$. The topology on any seperable metric space is generated by countable unions of balls, thus it suffices to show $B_r(x_0) \in \mathrm{Cyl}(\mathcal{B}))$ for any $x_0 \in \mathcal{B}, r > 0$. There exists linear functionals $l_n \in \mathcal{B}^*$ such that

$$\|x\| = \sup_n l_n(x). \tag{1.13}$$

Thus

$$B_r(x_0) = \bigcap_n \{x \in \mathcal{B} : l_n(x - x_0) \le r\} \in \mathrm{Cyl}(\mathcal{B}). \tag{1.14}$$

$\square$

Next we formalise what we said in the introduction.

Definition 1.4 (*Mean Operator*)

> *Let μ be a probability measure on $\mathcal{B}$ and suppose $l_*\mu$ has finite mean for all $l \in \mathcal{B}^*$.*

Then the mean operator of μ is a linear operator $\mu : \mathcal{B}^* \to \mathbb{R}$ where

$$m_\mu(l) := \int_\mathcal{B} l(x)\,\mu(dx). \tag{1.15}$$

μ is centered if $m_\mu \equiv 0$.

Similarly we can define the covariance.

Definition 1.5 (*Covariance Operator*)

Let μ be a probability measure on $\mathcal{B}$ and suppose $l_*\mu$ has finite second moment for all $l \in \mathcal{B}^*$. Then the covariance operator of μ is the bilinear operator $C_\mu : \mathcal{B}^* \times \mathcal{B}^* \to \mathbb{R}$ where

$$C_\mu(l,l') := \int_\mathcal{B} l(x)l'(x)\mu(dx). \tag{1.16}$$

Remark 1. Note that C_μ is a symmetric, positive semi-definite, bilinear operator.

We'll leave the question of whether or not these operators are bounded till later.

Definition 1.6 (*Fourier Transform*)

The Fourier transform of a probabilty measure μ on $\mathcal{B}$ is given by $\hat{\mu} \in \mathcal{B}^* \to \mathbb{C}$ where

$$\hat{\mu}(l) = \int_\mathcal{B} e^{il(x)}\mu(dx). \tag{1.17}$$

Remark 2. Unlike the mean, the Fourier transform of a measure of $\mathbb{R}^n$ is normally a function to begin with, just a function $\mathbb{R}^n \to \mathbb{R}$ rather than $(\mathbb{R}^n)^* \to \mathbb{R}$. This is a use of the Hilbert space structure on $\mathbb{R}^n$. For a general Hilbert space $\mathcal{H}$, let μ be a probabilty measure on $\mathcal{H}$. We can then instead define $\hat{\mu} : \mathcal{H} \to \mathbb{R}$ by

$$\hat{\mu}(\xi) := \int_\mathcal{H} e^{i\langle \xi, x\rangle}\mu(\mathrm{d}x). \tag{1.18}$$

This is then equivalent to our definition since by the Riesz representation theorem the map $\xi \to l_\xi$ where $l_\xi(x) = \langle \xi, x\rangle$ is an isomorphism between $\mathcal{H}$ and $(\mathcal{H})^*$.

In the finite dimensional case, Fourier transforms uniquely determine distributions. This holds true in seperable Banach spaces.

Proposition 1.7 (*Uniqueness of Fourier transforms*)

Let μ and ν be probability measures on a seperable Banach space $\mathcal{B}$. If $\hat{\mu} = \hat{\nu}$ then

$$\mu = \nu.$$

Proof. $\hat{\mu}$ uniquely determines the (classical) Fourier transform of $l_*\mu$ since

$$\widehat{l_*\mu}(\xi) = \hat{\mu}(\xi l). \tag{1.19}$$

From the finite dimensional case we know $\widehat{l_*\mu}$ uniquely determines $l_*\mu$. Therefore $\hat{\mu}$ uniquely determines the one-dimensional marginals of μ, and thus μ itself by proposition 1.3 on page 6. $\qquad\square$

1.3. Gaussian measures

We've now got our tools ready. Now we'll move onto the next step, defining what a Gaussian random variable is.

1.3.1. Basic definition

As aluded to in the section on analysis of $\mathcal{B}$-valued random variables, there is not an immediately obvious definition for what a Gaussian measure is on $\mathcal{B}$ using Bochner integrals. We can't really hope for a density function after all, there's no generalisation of the Lebesgue measure to general Banach spaces. However we do know what Gaussian measures are on $\mathbb{R}$, so we can use our approach of considering one-dimensional marginals.

Definition 1.8 (*Gaussian measures*)

A Borel probability measure μ on $\mathcal{B}$ is a Gaussian measure if $l_*\mu$ is a Gaussian measure. Note we consider Dirac measures to be variance 0 Gaussian measures in this definition.

Gaussian measures on $\mathbb{R}^n$ are uniquely characterised by their covariance and mean. This is true since they uniquely determine the Fourier transform. This remains true in infinite dimensions.

Proposition 1.9 (*Fourier transform of Gaussian measures*)

Let μ be a Gaussian measure on $\mathcal{B}$. Then

$$\hat{\mu}(l) = \exp\left\{ im_\mu(l) - \frac{1}{2}C_\mu(l,l) \right\}. \tag{1.20}$$

Corollary 1.10 (*Gaussian measures are determined by mean and covariance*)

A Gaussian measure on a seperable Banach space is uniquely determined by it's mean

and covariance operators.

We've already stated that covariance operators need to be positive definite, symmetric and bilinear. But is the converse true, can we get an existence result? Speficially given such an operator, can we necessarily find a centered Gaussian measure with that covariance? Unfortunately no, and we'll see one reason why in the next subsection.

1.3.2. Integrability of Gaussian Measures

It's hopefully somewhat intuitive that boundedness of the covariance operator will be related to boundedness of the second moment of the corresponding measure. In finite dimensions, Gaussian random variables are nice because of the rapid tail decay of the probability density function. We now show a similar statement is true for general Gaussian measures. Suprisingly, this will come from rotational invariance of Gaussian measure.

Definition 1.11 (*Rotational invariance*)

Let μ be a measure on $\mathcal{B}$. For each $\varphi \in \mathbb{R}$, define $R_\varphi : \mathcal{B} \times \mathcal{B} \to \mathcal{B} \times \mathcal{B}$ by

$$R_\varphi(x, y) = (x \sin \varphi + y \cos \varphi, x \cos \varphi - y \sin \varphi). \tag{1.21}$$

Then μ is invariant under rotations of φ if $(R_\varphi)_*(\mu \otimes \mu) = \mu \otimes \mu$.

Proposition 1.12 (*Rotational invariance of centered Gaussians*)

Centered Gaussian measures on a seperable Banach space $\mathcal{B}$ are invariant under rotations by any angle.

Proof. Check the Fourier transform of the rotated measure is equal to the fourier transform of the original measure. $\square$

The following theorem shows as long as a measure is invariance under rotations of $\frac{\pi}{4}$ then it has double exponential tail decay.

Theorem 1.13 (*Fernique*)

Let μ a probability measure on $\mathcal{B}$. Suppose μ is invariant under rotations by $\frac{\pi}{4}$. Then there exists $\alpha > 0$ such that $\|x\|$ has doubly exponential moments, i.e.

$$\int e^{\alpha \|x\|^2} \mu(dx) < \infty. \tag{1.22}$$

Proof. Take $t, \tau > 0$. Then

$$\mu(\|x\| \leq \tau)\mu(\|x\| > t) = \mu \otimes \mu \left(\frac{\|x + y\|}{\sqrt{2}} > t, \frac{\|x + y\|}{\sqrt{2}} > \tau \right) \tag{1.23}$$

$$\leq \mu \otimes \mu \left(\|x\| > \frac{t - \tau}{2}, \|x\| > \frac{t - \tau}{2}) \right) \tag{1.24}$$

$$= \mu \left(\|x\| > \frac{t - \tau}{2} \right)^2 \tag{1.25}$$

INSERT REST OF PROOF $\qquad\qquad\square$

Fernique's theorem allows us to not only show all moments of a Gaussian measure are finite, it allows us to bound the higher moments using the first moment.

Proposition 1.14 (*Moment bounds on Gaussians*)

Let μ be a centered Gaussian on $\mathcal{B}$. Then there exists $\alpha, K > 0$ such that for all $n \in \mathbb{N}$

$$\int_{\mathcal{B}} |x|^{2n} \, \mu(dx) \leq n! K \alpha^{-n} M^{2n}. \tag{1.26}$$

where $M = \int_{\mathcal{B}} \|x\| \, \mu(dx)$.

Proof. INSERT PROOF $\qquad\qquad\square$

Let's return to our question of showing existence of a Gaussian measure given a covariance operator. We already know they must be symmetric and PSD. The following theorem shows that they must further be bounded operators.

Proposition 1.15 (*Boundedness and representation of covariance operators*)

Let μ be a centered Gaussian measure on a seperable Banach space. Then there exists $\|C_\mu\| < \infty$ such that $C_\mu(l, l') \leq \|C_\mu\| \|l\| \|l'\|$. Moreover there exists continuous operator $\hat{C}_\mu : \mathcal{B}^* \to \mathcal{B}$ such that

$$C_\mu(l, l') = l' \left(C_\mu(l) \right). \tag{1.27}$$

Proof. The first part of the proof is clear by setting $\|C_\mu\|$ to $\int_{\mathcal{B}} \|x\|^2 \, \mu(dx)$ and employing boundedness of l and l'.

For the second part, since

$$\int_{\mathcal{B}} \|xl(x)\| \, \mu(dx) \leq \|l\| \int_{\mathcal{B}} \|x\|^2 dx < \infty \tag{1.28}$$

we can formalise

$$\hat{C}_\mu(l) = \int_{\mathcal{B}} xl(x)\,\mu(\mathrm{d}x) \tag{1.29}$$

as a Bochner integral. Then since the Bochner integral can be exchanged with continuous linear functionals

$$l'(\hat{C}_\mu(l)) = l'\left(\int_{\mathcal{B}} xl(x)\,\mu(\mathrm{d}x)\right) = \int_{\mathcal{B}} l'(xl(x))\,\mu(\mathrm{d}x)$$

$$= \int_{\mathcal{B}} l(x)l'(x)\mu(\mathrm{d}x) = C_\mu(l, l') \tag{1.30}$$

as required. $\qquad\square$

So now let's add boundedness to our list of assumptions. Is that enough to characterize covariance operators of Gaussian measures? Unfortunately the answer is still no. In fact for a general Banach space there is no nice characterisation. However for Hilbert spaces we can give an actual answer.

1.3.3. Characterisation of Gaussian measures on Hilbert spaces

As promised, we now characterise the covariance operators of Gaussian measures in Hilbert spaces. In proposition 1.15 on the previous page we've already shown there exists $\hat{C}_\mu : \mathcal{H}^* \to \mathcal{H}$ such that $C_\mu(l, l') = l'(\hat{C}_\mu(l))$ for all $l, l' \in \mathcal{B}^*$. Then by the Riesz representation, there exists $\tilde{C}_\mu : \mathcal{H} \to \mathcal{H}$ such that

$$C_\mu(\langle \cdot, x \rangle, \langle \cdot, y \rangle) = \langle \tilde{C}_\mu x, y \rangle \tag{1.31}$$

We will state our characterisation in terms of $\tilde{C}_\mu$. Firstly let's recall some definitions from functional analysis.

Definition 1.16 (*Positive definiteness and trace-class*)

Let $\mathcal{H}$ be a seperable Hilbert space and let $L : \mathcal{H} \to \mathcal{H}$ be linear.

1. *L is positive definite if L is self-adjoint (symmetric) and $\langle Lh, h \rangle > 0$ for all non-zero $h \in \mathcal{H}$*

2. *L is trace-class if for any orthonormal basis (e_n) of $\mathcal{H}$ there exists a sequence of numbers $\{\lambda_n\}$ such that $\sum_n |\lambda_n| < \infty$ and*

$$\langle x, Lx \rangle = \sum_n \lambda_n \langle x, e_n \rangle^2. \tag{1.32}$$

The trace of L is then $\operatorname{Tr} L := \sum_n \lambda_n$.

You can prove it is sufficient to check the trace-class property for any single orthonormal basis in order for it to hold for all orthonormal bases.

Trace-class is a stronger property than boundedness, in particular there exists operators that are bounded and not trace class. Hence we see symmetry, positive semi-definitess, and boundedess was truly insufficient to characterize Gaussian covariances.

Proposition 1.17 (*Characterization of Gaussian covariances in Hilbert spaces*)

Let $\mathcal{H}$ be a seperable Hilbert space and $\tilde{C} : \mathcal{H} \to \mathcal{H}$. Then the following are equivalent

1. *There exists a Gaussian measure μ on $\mathcal{H}$ such that*

$$C_\mu(\langle \cdot, x \rangle, \langle \cdot, y \rangle) = \langle \tilde{C}x, y \rangle \quad \forall x, y \in \mathcal{H}. \tag{1.33}$$

2. *$\tilde{C}$ is a linear operator that is positive semi-definite and trace-class.*

Proof. (1) $\implies$ (2): Bilinerity, symmetry and positive semi-definiteness of $\hat{C}_\mu$ translates to linearity, self-adjointness and positive semi-definiteness of $\tilde{C}$ respectively. To show trace-class fix any orthonormal basis $\{e_n\}$. Then

$$\mathrm{Tr}(\tilde{C}) = \sum_n \langle \tilde{C}e_n, e_n \rangle = \sum_n C_\mu(\langle \cdot, e_n \rangle, \langle \cdot, e_n \rangle) = \sum_n \int_\mathcal{H} \langle x, e_n \rangle^2 \, \mu(\mathrm{d}x)$$

$$= \int_\mathcal{H} \sum_n \langle x, e_n \rangle^2 \, \mu(\mathrm{d}x) = \int_\mathcal{H} \|x\|^2 \, \mu(\mathrm{d}x) < \infty \tag{1.34}$$

using proposition 1.14 on page 10.

(2) $\implies$ (1): Trace-class operators are compact. The spectral theorem for compact, self-adjoint operators says there is an orthonormal basis $\{e_n\}$ of $\mathcal{H}$ consisting of eigenvectors of $\tilde{C}$, say $\tilde{C}(e_n) = \lambda_n e_n$, and that the λ_n are real. Furthermore $\lambda_n \geq 0$ for all n as $\tilde{C}$ is positive semi-definite. In addition

$$\sum_n \lambda_n = \sum_n \langle \tilde{C}e_n, e_n \rangle = \mathrm{Tr}(\tilde{C}) < \infty. \tag{1.35}$$

There exists a probability space on which we can define $X_1, X_2, \ldots$ a sequence of i.i.d. $N(0,1)$ random variables. Then let μ_n be the law of $\sum_{i=1}^n \sqrt{\lambda_i} X_i e_i$. We can show μ_n converges weakly to a limit μ and that μ is a Gaussian measure with the required covariance. $\square$

1.3.4. Abstract Wiener Spaces and Cameron-Martin Theory - what is a standard normal?

1.4. Existence of Gaussian Processes

Continuity Criterions
   This section focuses on Gaussian measures on $C([0,1]^d, \mathbb{R})$ under the supremum norm. Hence it makes sense to talk about the regularity of the paths of a Gaussian measure.
   The following is both an existence and uniqueness result about Gaussian measures with sufficiently smooth paths.

Theorem 1.18 (*Kolmogorov Continuity Criterion*)

> *For $d \geq 1$, let $C : [0,1]^d \times [0,1]^d$ be a function[2] such that:*
>
> 1. *For any finite collection of points $\{x_1, \ldots, x_n\} \subset [0,1]^d$ the matrix*
>
> $$C_{ij} := C(x_i, x_j) \tag{1.36}$$
>
>   *is symmetric, positive semi-definite.*
>
> 2. *There exists $\alpha > 0$ and $K > 0$ such that*
>
> $$C(x,x) - 2C(x,y) + C(y,y) \leq K|x-y|^{2\alpha}. \tag{1.37}$$
>
> *Then there exists a unique Gaussian measure μ on $C([0,1]^d, \mathbb{R})$ such that*
>
> $$C_\mu(\delta_x, \delta_y) := \int_{C([0,1]^d, \mathbb{R})} f(x) f(y) \, \mu(df) = C(x,y) \tag{1.38}$$
>
> *for all $x, y \in [0,1]^d$ and $\mu(C^\beta([0,1]^d, \mathbb{R})) = 1$ for all $\beta \in (0, \alpha)$.*

   Before we prove it, let's understand this theorem and give some intuition to the proof. Firstly, note C is a function $[0,1]^d \times [0,1]^d \to \mathbb{R}$ rather than $C([0,1]^d, \mathbb{R})^* \times C([0,1]^d, \mathbb{R})^* \to \mathbb{R}$. This can be understand as specifying the covariance structure of only $\{(\delta_x)_* \mu : x \in [0,1]^d\}$ rather than for all one-dimensional marginals. We'll first show this is enough to characterise and ensure Gaussianity of μ.

Proposition 1.19 (*Gaussianity via weak-$*$ dense marginals*)

> *Let μ be a Borel probability measure on $\mathcal{B}$ and let $\mathcal{A}$ be weak-* dense in $\mathcal{B}^*$.*
>
> 1. *If $l_* \mu$ is Gaussian for all $l \in \mathcal{A}$ then μ is Gaussian.*
>
> 2. *If $\mathcal{B}$ is seperable then $\{l_* \mu : l \in \mathcal{B}^*\}$ uniquely determines μ.*

Proof. Given any $l \in \mathcal{B}^*$, take a sequence $l_n \in \mathcal{A}$ such that $l_n \xrightarrow{w*} l$ as $n \to \infty$. Let $X \sim \mu$. Then weak-$*$ convergence is exactly pointwise convergence therefore $l_n(X) \to l(X)$ almost surely. Thus $(l_n)_* \mu \Rightarrow l_* \mu$ by Skorokhod's representation theorem. Then since the weak limit of Gaussians is Gaussian, $l_* \mu$ is Gaussian. Hence μ is Gaussian.

The uniqueness statement follows from proposition 1.3 on page 6 since the one-dimensional marginals are uniquely determined as weak limits of sequences in $\{l_* \mu : l \in \mathcal{A}\}$. $\qquad\square$

Thus we need linear combinations of Dirac measures to be weak-$*$ dense. This is true, and we'll shove the proof into the appendix.

Proposition 1.20 (*Weak-$*$ density of Dirac measures*)

> *Suppose that $\mathcal{D} \subset [0,1]^d$ is dense in $[0,1]^d$. Then*
>
> $$\mathcal{A} := \mathrm{span}\{\delta_x : x \in \mathcal{D}\} \tag{1.39}$$
>
> *is weak-$*$ dense in $C([0,1]^d, \mathbb{R})^*$.*

Secondly, let's understand the continuity statement better. Suppose we've already proven there exists a Gaussian measure μ on $C([0,1]^d, \mathbb{R})$ such that $C(x,y) = C_\mu(\delta_x, \delta_y)$ for all $x, y \in [0,1]^d$. Rather than working directly with μ, it's more intuitive to work with a continuous random process $F : \Omega \to C([0,1]^d, \mathbb{R})$ that has law μ. We'll use the notation $F_x(\omega) := F(\omega)(x)$. Then $C(x,y) = C_\mu(\delta_x, \delta_y) = \mathrm{Cov}(F_x, F_y)$. Thus the condition (2) on C is equivalent to requiring

$$\|F_x - F_y\|_2 := \sqrt{\mathbb{E}[(F_y - F_y)^2]} \leq K|x-y|^\alpha \quad \forall x, y \in [0,1]^d. \tag{1.40}$$

So condition (2) gives α-Hölder continuity of the map $[0,1]^d \to L^2(\Omega)$ where $x \mapsto F_x$. We'd like this to instead be a pathwise statement, for fixed ω is the map $x \mapsto F_x(\omega)$ Hölder continuous? While we can't get α-Hölder continuity, the the Kolmogorov Continuity Criterion says we can get arbitary close. In particular for almost every ω, $x \mapsto F_x(\omega)$ is β-Hölder continuous for all $\beta \in (0, \alpha)$.

It should be noted that condition (2) isn't only used to show pathwise Hölder continuity, it is used to ensure there even exists a Gaussian measure μ on the space $C([0,1]^d, \mathbb{R})$ such that $C_\mu(\delta_x, \delta_y) = C(x,y)$ in the first place. We know Gaussian covariances are at least continuous so some continuity criterion is certainly needed. Maybe if we didn't need Hölder continuity we can give a weaker condition, but something needs to replace condition (2) to ensure existence.

Proof of KCC. By Kolmogorov's extension theorem, there exists a stochastic pro-

cess $F : \Omega \to \mathbb{R}^{[0,1]^d}$ such that for all $x_1, \ldots, x_n \in [0,1]^d$,

$$(F_{x_1}, \ldots, F_{x_n})^T \sim N(0, \Sigma) \tag{1.41}$$

where $\Sigma_{i,j} = C(x_i, x_j)$. Condition (1) ensures that Σ is a valid covariance matrix. However Kolmogorov's extension theorem only gives that $F : \Omega \to \mathbb{R}^{[0,1]^d}$ is measurable with respect to the product σ-algebra on $\mathbb{R}^{[0,1]^d}$. This is a weak σ-algebra, meaning we have to be careful about what events we consider.

We want to determine Hölder continuity of $x \mapsto F_x(\omega)$. Fix $\beta \in (0, \alpha)$. Then $f : [0,1]^d \to \mathbb{R}$ would be β-Hölder continuous if

$$\sup_{x,y \in [0,1]^d, x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|^\beta} < \infty. \tag{1.42}$$

However the quantity on the left is not a measurable function of F, so we approximate it. Let $\mathcal{D}$ be the grid of dyadic points in $[0,1]^d$. Note $\mathcal{D}$ is countable and dense in $[0,1]^d$. We consider the approximation $M_\beta(f) : \mathbb{R}^{[0,1]^d} \to \mathbb{R}$ by

$$M_\beta(f) = \sup_{x,y \in \mathcal{D}, x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|^\beta}. \tag{1.43}$$

The function inside the supremum is measurable with respect to the product σ-algebra since it is a pointwise evaluation. Then M_β is measurable with respect to the product σ-algebra as a countable supremum of measurable functions. If $M_\beta(f) < \infty$ then f is β-Hölder continuous along the dyadic rationals, we want to extend this to the entirety of f. To assist with this we can define the regulariser $\text{Reg}_\beta : \mathbb{R}^{[0,1]^d} \to C([0,1], \mathbb{R})$ by

$$\text{Reg}_\beta(f)(x) := \begin{cases} \lim_{s \in \mathcal{D}, s \to x} f(s) & \text{if } M_\beta(f) < \infty \\ 0 & \text{otherwise.} \end{cases} \tag{1.44}$$

Claim 1. $\text{Reg}_\beta : \mathbb{R}^{[0,1]^d} \to C([0,1]^d, \mathbb{R})$ *is well defined* (*in particular the limit in the definition exists and is independent of the sequence taken*) *and*

1. *For all* $f \in \mathbb{R}^{[0,1]^d}$, $\text{Reg}_\beta(f) \in C^\beta([0,1]^d, \mathbb{R})$

2. $\text{Reg}_\beta(f)(x) = f(x) \mathbb{1}_{\{M_\beta(f) < \infty\}}$ *for all* $x \in \mathcal{D}$

3. Reg_β *is measurable with respect to the product σ-algebra on* $\mathbb{R}^{[0,1]^d}$ *and the Borel σ-algebra on* $C([0,1]^d, \mathbb{R})$ (*under supremum norm*).

15

Proof. We'll only show (3) here. Since the Borel algebra on $C([0,1]^d, \mathbb{R})$ is generated by balls, it suffices to check that $\{f \in \mathbb{R}^{[0,1]^d} : \|\mathrm{Reg}_\beta(f) - f_0\|_\infty < r\}$ is measurable for all $f_0 \in C([0,1]^d, \mathbb{R})$ and $r > 0$. By continuity of $\mathrm{Reg}_\beta(f)$ and density of $\mathcal{D}$,

$$\|\mathrm{Reg}_\beta(f) - f_0\|_\infty = \sup_{x \in \mathcal{D}} |\mathrm{Reg}_\beta f(x) - f_0(x)| \tag{1.45}$$

$$= \sup_{x \in \mathcal{D}} |f(x)\mathbb{1}_{\{M_\beta(f) < \infty\}} - f_0(x)|. \tag{1.46}$$

Thus

$$\{\|\mathrm{Reg}_\beta(f) - f_0\|_\infty < r\} = \bigcap_{x \in \mathcal{D}} \left\{ |f(x)\mathbb{1}_{\{M_\beta(f) < \infty\}} - f_0(x)| < r \right\} \tag{1.47}$$

$$= \left[\{M_\beta(f) = \infty\} \cap \{\|f_0\|_\infty < r\} \right]$$
$$\cup \left[\{M_\beta(f) < \infty\} \cap (\cap_{x \in \mathcal{D}} \{|f(x) - f_0(x)| < r\}) \right]. \tag{1.48}$$

The event $\{\|f_0\|_\infty < r\}$ is an artifact of our construction and is always equal to $\varnothing$ or $\mathbb{R}^{[0,1]^d}$, thus is measurable. Since M_β is measurable, $\{M_\beta(f) < \infty\}$ and $\{M_\beta(f) = \infty\}$ is measurable. $\{|f(x) - f_0(x)| < r\}$ depends only on pointwise evaluations of f thus is measurable. Hence $\{\|\mathrm{Reg}_\beta(f) - f_0\| < r\}$ is measurable as a countable intesection and union of measurable sets. ∎

Hence $\hat{F} = \mathrm{Reg}_\beta(F)$ is a continuous stochastic process $\Omega \to C([0,1]^d, \mathbb{R})$. There's no point defining $\mathrm{Reg}_\beta(F)$ if it doesn't carry over properties of F, so the first thing we'll do is show the two processes are modifications of each other. Rigorously we show that for all $x \in [0,1]^d$, $\mathrm{Reg}_\beta(F)_x = F_x$ almost surely[3]. We do this is two steps, firstly for $x \in \mathcal{D}$ and then for all $x \in [0,1]^d$. When $x \in \mathcal{D}$, $\mathrm{Reg}_\beta(F)_x = F_x \mathbb{1}_{\{M_\beta(F) < \infty\}}$. Thus it suffices to show $M_\beta(F) < \infty$ almost surely.

Claim 2. $M_\beta(F) < \infty$ *almost surely.*

Proof. We can actually show the stronger statement $\mathbb{E}[M_\beta(F)] < \infty$. INSERT REST OF PROOF. ∎

Now we extend this to all $x \in [0,1]^d$.

Claim 3. $\mathbb{P}(\mathrm{Reg}_\beta(F)_x = F_x) = 1$ *for all $x \in [0,1]^d$.*

---

[3]Note this means $\mathbb{P}(\mathrm{Reg}_\beta(F)_x = F_x) = 1$ for all x, not $\mathbb{P}(\mathrm{Reg}_\beta(F)_x = F_x \ \forall x) = 1$

Proof. For any $x \in [0,1]^d$, let $x_n \to x$ where x_n is a sequence in $\mathcal{D}$. Then by condition (2) on C

$$\|F_{x_n} - F_x\|_2^2 = \text{Var}(F_{x_n} - F_x) \tag{1.49}$$

$$= C(x_n, x_n) - 2C(x_n, x) + C(x, x) \tag{1.50}$$

$$\leq K\|x - x_n\|^\alpha \xrightarrow{n \to \infty} 0. \tag{1.51}$$

Thus $F_{x_n} \to F_x$ in L^2. So there is some subsequence along which $F_x = \lim_n F_{x_{k_n}}$ almost surely. By continuity $\text{Reg}_\beta(F)_x = \lim_n \text{Reg}_\beta(F)_{x_{k_n}}$ always. Thus almost surely

$$\text{Reg}_\beta(F)_x = \lim_n \text{Reg}_\beta(F)_{x_{k_n}} = \lim_n F_{x_{k_n}} = F_x. \tag{1.52}$$

$\blacksquare$

Let μ be the law of $\text{Reg}_\beta(F)$, this is a measure on $C([0,1]^d, \mathbb{R})$.

Claim 4. *μ satisfies the following properties.*

1. *μ is Gaussian.*

2. *For all $x, y \in [0,1]^d$, $C_\mu(\delta_x, \delta_y) = C(x, y)$.*

3. *$\mu(C^\beta([0,1]^d, \mathbb{R})) = 1$.*

4. *μ is uniquely determined by C and independent of the $\beta \in (0, \alpha)$ chosen in the construction.*

Proof. For any finite set $x_1, \ldots, x_n \in [0,1]^d$, almost surely

$$(\text{Reg}_\beta(F)_{x_1}, \ldots, \text{Reg}_\beta(F)_{x_n}) = (F_{x_1}, \ldots, F_{x_n}) \tag{1.53}$$

thus is jointly Gaussian. Thus for all $\lambda \in \mathbb{R}^n$, $(\sum_{i=1}^n \lambda_i \delta_{x_i})_* \mu$ is Gaussian. In other words if we define

$$\mathcal{A} := \text{span}\{\delta_x : x \in [0,1]^d\} \tag{1.54}$$

then $l_* \mu$ is Gaussian for all $l \in \mathcal{A}$. Since $\mathcal{A}$ is weak-$*$ dense in $C([0,1]^d, \mathbb{R})^*$ by proposition 1.20 on page 14, μ is Gaussian by proposition 1.19 on page 13.

For all $x, y \in [0,1]^d$ we have

$$C_\mu(\delta_x, \delta_y) = \mathbb{E}[\text{Reg}_\beta(F)_x \, \text{Reg}_\beta(F)_y] = \mathbb{E}[F_x F_y] = C(x, y). \tag{1.55}$$

By construction, $\text{Reg}_\beta(F)$ is β-Hölder continuous thus $\mu(C^\beta([0,1]^d, \mathbb{R})) = 1$.

Finally since C determines $l_* \mu$ for all $l \in \mathcal{A}$, μ is uniquely determined by proposition 1.19 on page 13. In particular C is independent of β thus μ is independent of β. $\blacksquare$

This finishes the proof. $\square$

1.4.1. Abstract Wiener Spaces

To motivate this section, let us try and recover a definition of a Wiener process in higher dimensions. First we define a Gaussian process. Fortunately this definition is a direct translation of the finite dimensional version.

Definition 1.21

Let $(B_t)_{t \geq 0}$ be a time indexed sequence of random variables such that B_t takes values in $\mathcal{B}$ a seperable Banach space. Then (B_t) is a Gaussian process if the law of

$$\left(B_{t_1}, \ldots, B_{t_n} \right) \tag{1.56}$$

is a Gaussian measure on $\mathcal{B}^n$ w.r.t. the product topology for all finite sets of times $t_1, \ldots, t_n$.

Now recall the definition of a Wiener process.

A. Cameron-Martin Theory: A stupidly abstract approach

The way I like to think about Cameron-Martin theory is perhaps a bit too abstract. Recall $\mathcal{B}^*$ can be 'embedded' into $\mathcal{L}^2(\mathcal{B}, \mu)$. Rigorously consider the map $j : \mathcal{B}^* \to \mathcal{L}^2$ (we're not using i since that's used later).

$$j(f) = [f] \tag{A.1}$$

which maps f to the equivalence class of f under almost sure equality. We're using $\mathcal{L}^2$ instead of L^2 to distinguish the space of equivalence classes from the space of functions. Then j is a continuous map and the RKHS $\mathcal{R}_\mu$ is $\overline{j(\mathcal{B})}$.

We then want embed $\mathcal{R}_\mu$ into $\mathcal{B}$. Embedding $\mathcal{R}_\mu$ into $\mathcal{B}^{**}$ is not too difficult. j has an adjoint map $j^* : (\mathcal{L}^2)^* \to \mathcal{B}^{**}$ where $j^*(\varphi) = \varphi \circ j$ for all $\varphi \in (\mathcal{L}^2)^*$. We don't know if j^* is injective. For that the following proposition from functional analysis helps.

Proposition A.1 (*Injective of the adjoint map*)

Let X and Y be Banach spaces and $j : X \to Y$ be a bounded linear map. Then the following are equivalent:

1. *$j^* : Y^* \to X^*$ is injective.*

2. *$j(X)$ is dense in Y.*

This tells us that $\mathcal{L}^2$ is too big a space to look at, instead we should restrict the range of j from $\mathcal{L}^2$ to R_μ. Then the above proposition tells us that $j^* : \mathcal{R}_\mu^* \to \mathcal{B}^{**}$ is an injection thus is an embedding of $\mathcal{R}_\mu^* \to \mathcal{B}^{**}$. Recall by Riesz representation there is a natural bijective linear isometry $\mathrm{eq} : \mathcal{R}_\mu \to \mathcal{R}_\mu^*$ where $\mathrm{eq}(f) := \langle \cdot, f \rangle_{L^2}$. Then $j^* \circ \mathrm{eq}$ embeds $\mathcal{R}_\mu$ into $\mathcal{B}^{**}$. Now let's unravel the action of this embedding, for $l \in \mathcal{B}^*$.

$$j^* \circ \mathrm{eq}(f)(l) = \langle j(l), f \rangle_{L^2} = \int_{\mathcal{B}} l(x) f(x)\, \mu(\mathrm{d}x) = l\left(\int_{\mathcal{B}} x f(x)\, \mu(\mathrm{d}x) \right). \tag{A.2}$$

The integral inside the last quantity is a well defined Bochner integral since $\|f\|_{L^2} < \infty$ guarantees that $x f(x)$ is integrable. This tells us that $j \circ \mathrm{eq}(\mathcal{B}^*) \subset$

$$\mathcal{B}^* \xrightarrow{\ j\ } \mathcal{R}_\mu \xrightarrow{\ \text{eq}\ } \mathcal{R}_\mu^* \xrightarrow{\ j^*\ } \text{ev}(\mathcal{H}_\mu)$$

Figure A.1.: Diagram of the maps involved in Cameron-Martin theory

$\text{ev}(\mathcal{B})$ thus we can embed $\mathcal{R}_\mu$ into $\mathcal{B}$ by $k := \text{ev}^{-1} \circ j^* \circ \text{eq}$. Removing the abstraction

$$k(f) = \int_\mathcal{B} x f(x)\, \mu(\mathrm{d}x). \tag{A.3}$$

Note that $k \circ j$ is exactly $\hat{C}_\mu$. Then k is a bijection onto it's image. This gives us the Cameron-Martin space.

Definition A.2

Let μ be a Gaussian measure on $\mathcal{B}$ with RKHS $\mathcal{R}_\mu$ and k as defined in this chapter. Then the Cameron-Martin space of μ is

$$\mathcal{H}_\mu := k(\mathcal{R}_\mu) \subset \mathcal{B}. \tag{A.4}$$

This is a Hilbert space equipped with inner product $\langle \cdot, \cdot \rangle_\mu$ inherited from the $\langle \cdot, \cdot \rangle_{L^2}$ inner product $\mathcal{R}_\mu$ by k, explicitly

$$\langle h, k \rangle_\mu := \langle k^{-1}(h), k^{-1}(k) \rangle_{L^2} = \int_\mathcal{B} h^*(x) k^*(x)\, \mu(dx) \tag{A.5}$$

where $h = \int_\mathcal{B} x h^*(x)\, \mu(dx)$ and $k = \int_\mathcal{B} x k^*(x)\, \mu(dx)$.

We've drawn the sequence of maps use in appendix A. The downside of this approach of constructing the Cameron-Martin space is in how long it takes to set up all the structure involved. However it does have benefits. For me, it justifies why we consider the image of $\hat{C}_\mu$. It's because, up to isometric isomorphisms, it is the adjoint map of the embedding of $\mathcal{R}_\mu$ into $\mathcal{L}^2$. It also allows us to reuse theorems from functional analysis. The following corollaries are immediate or can be proved very quickly from the construction of $\mathcal{H}_\mu$.

Corollary A.3

There exists $C \geq 0$ such that

$$\|h\|^2 \leq C \langle h, h \rangle_\mu, \quad \forall h \in \mathcal{H}_\mu. \tag{A.6}$$

Proof. By definition of the Cameron-Martin space and since k is bounded

$$\|h\|^2 \leq \|k\|_{\text{op}}^2 \|k^{-1}(h)\|_{L^2} := \|k\|^2 \langle h, h \rangle_\mu. \tag{A.7}$$

$\square$

Corollary A.4

There is a natural isometric isomorphism between $\mathcal{H}_\mu$ and $\mathcal{R}_\mu$.

Proof. k is the natural isometric isomorphism. $i = k^{-1} : \mathcal{H}_\mu \to \mathcal{R}_\mu$ is the notation used in Prof. Hairer's notes. $\square$

Corollary A.5

We have the alternate characterisation

$$\langle h, h \rangle_\mu = \sup \{ l(h) : C_\mu(h, h) \leq 1 \}, \quad \forall h \in \mathcal{H}_\mu. \tag{A.8}$$

Moreover $\mathcal{H}_\mu = \{ x \in \mathcal{B} : \langle h, h \rangle_\mu < \infty \}$ exactly.

Proof. Unravelling definitions gives

$$\langle h, h \rangle_\mu = \|k^{-1}(h)\|_{L^2} = \|(\text{eq}^{-1} \circ (j^*)^{-1} \circ \text{ev})(h)\|_{L^2} \tag{A.9}$$

$$= \|((j^*)^{-1} \circ \text{ev})(h)\|_{\text{op},L^2} = \|\text{ev}(h) \circ j^{-1}\|_{\text{op},L^2} \tag{A.10}$$

$$= \sup \left\{ (\text{ev}(h)(j^{-1}(f))) : f \in \mathcal{R}_\mu, \|f\|_{L^2} \leq 1 \right\} \tag{A.11}$$

$$= \sup \left\{ \text{ev}(h)(l) : l \in \mathcal{B}^*, C_\mu(l, l) \leq 1 \right\} \tag{A.12}$$

$$= \sup \left\{ l(h) : C_\mu(l, l) \leq 1 \right\} \tag{A.13}$$

where the second to last line uses density of $j(\mathcal{B}^*)$ in $\mathcal{R}_\mu$. $\square$

B. Miscellaneous Theorems

Proposition B.1 (*Seperable metric spaces have a countable base of balls*)

> Let (X, d) be a seperable metric space and $A \subset X$ be open. Let D be a countable dense set of X. Then
>
> $$A = \bigcup_{x \in D \cap A} \bigcup_{\substack{q \in \mathbb{Q} \\ B_q(x) \subset A}} B_q(x). \tag{B.1}$$
>
> In other words $\{ B_q(x) : q \in \mathbb{Q}, x \in D \}$ is a countable base of the topology on (X, d).

Proof. Fix any $y \in A$. Then by openess there exists $r > 0$ such that $B_r(y) \subset A$. By density there exists $x \in B_{r/4}(y) \cap D$. Pick rational $q \in \left(\frac{1}{4}r, \frac{3}{4}r \right)$. Then $y \in B_q(x)$ and by triangle inequality $B_q(x) \subset B_{\frac{r}{4}+q}(y) \subset B_r(y) \subset A$. $\square$

Proposition B.2 (*Approximation of norm in seperable spaces*)

> Suppose $\mathcal{B}$ is a seperable Banach space. Then there exists countable many linear functions $l_n \in \mathcal{B}^*$ such that
>
> $$\|x\| = \sup_n l_n(x) \quad \forall x \in \mathcal{B}. \tag{B.2}$$

Proof. Let $\mathcal{D} = \{x_n\}$ be a countable dense set of $\mathcal{B}$. Then by Hahn-Banach, for each x_n there exists $l_n \in \mathcal{B}^*$ such that $\|l_n\| \leq 1$ and $l_n(x_n) = \|x_n\|$. Since $\|l_n\| \leq 1$

$$\|x\| \geq \sup_n l_n(x) \quad \forall x \in \mathcal{B}. \tag{B.3}$$

To show equality, for any $x \in \mathcal{B}$ take a sequence $\{x_n\} \subset \mathcal{D}$ such that $x_n \to x$ in $\mathcal{B}$. Then

$$\begin{align}
|l_n(x) - \|x\|| &= |l_n(x) - \|x_n\| + \|x_n\| - \|x\|| \tag{B.4} \\
&= |l_n(x) - l_n(x_n) + \|x_n\| - \|x\|| \tag{B.5} \\
&\leq |l_n(x - x_n)| + |\|x_n\| - \|x\|| \tag{B.6} \\
&\leq \|x - x_n\| + |\|x_n\| - \|x\|| \xrightarrow{n \to \infty} 0. \tag{B.7}
\end{align}$$

Therefore $l_n(x) \to \|x\|$. Thus we have equality. $\square$

Proposition B.3 (*Almost sure convergence implies weak convergence of laws*)

> Let μ_n, μ be probability measures on a metric space S. Suppose there exists random variables $X_n, X : \Omega \to S$ defined on a common probability space such that $X_n \sim \mu_n$, $X \sim \mu$, and $X_n \to X$ almost surely. Then $\mu_n \Rightarrow \mu$.

Proof. Let $f : S \to \mathbb{R}$ be bounded and continuous. Then

$$\int_S f(s)\, \mu_n(\mathrm{d}x) = \mathbb{E}[f(X_n)] \to E[f(X)] = \int_S f(x)\, \mu(\mathrm{d}x) \qquad \text{(B.8)}$$

as $n \to \infty$ by the dominated convergence theorem. Thus $\mu_n \Rightarrow \mu$. $\qquad\square$

The converse is also true assuming μ has seperable support, that if we have weak convergence we can find representations which converge almost surely. This is the Skorokhod representation theorem.

Proposition 1.20 (*Weak-∗ density of Dirac measures*)

> Suppose that $\mathcal{D} \subset [0,1]^d$ is dense in $[0,1]^d$. Then
>
> $$\mathcal{A} := \mathrm{span}\{\delta_x : x \in \mathcal{D}\} \qquad \text{(1.39)}$$
>
> is weak-∗ dense in $C([0,1]^d, \mathbb{R})^*$.

Proof. Suppose $\mathcal{A}$ were not dense in $C([0,1]^d, \mathbb{R})^*$, say l_0 is not in the weak-* closure of $\mathcal{A}$. Then by the Hahn-Banach seperation, there exists a weak-* continuous linear functional $L : C([0,1]^d, \mathbb{R})^* \to \mathbb{R}$ that seperates the closed convex set $\bar{\mathcal{A}}$ and the compact convex set $\{l_0\}$, i.e. there exists $\alpha \in \mathbb{R}$ such that

$$L(l) \leq \alpha \leq L(l_0) \quad \forall l \in \bar{\mathcal{A}}. \qquad \text{(B.9)}$$

Then as $\mathcal{A}$ is a vector space, for all $l \in \mathcal{A}$ and $\lambda \in \mathbb{R}$ we have $\lambda L(l) = L(\lambda l) \leq \alpha$. Therefore $L(l) = 0$. Since L is continuous w.r.t. the weak-∗ topology there exists $f \in C([0,1]^d, \mathbb{R})$ such that $L(l) = l(f)$ for all $l \in \mathcal{B}^*$. Then $f(x) = L(\delta_x) = 0$ for all $x \in \mathcal{D}$. Thus $f \equiv 0$ by density of $\mathcal{D}$ and continuity of f. Hence $L \equiv 0$, contradiction. $\qquad\square$

When we use this in the proof of KCC we actually have $\mathcal{D} = [0,1]^d$ rendering density trivial.

Proposition B.4 (*Gaussians converge weakly to Gaussians*)

> Let μ_n be a sequence of Gaussian laws on $\mathcal{B}$. Suppose $\mu_n \Rightarrow \mu$. Then μ is Gaussian.

Proof. Start with $\mathcal{B} = \mathbb{R}$, suppose each $\mu_n \sim N(\theta_n, \sigma_n)$. (μ_n) is a tight sequence since it is weakly convergent. By tightness there exists an bounded interval $[K, -K]$ such that for all n, $\mu_n([-K, K]) > \frac{1}{2}$. Then $\theta_n \in [-K, K]$ for all n else over half the mass of some μ_n would lie outside $[-K, K]$. Similarly

$$\mu_n([-K, K]) = \int_{-K}^{K} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\theta_n)^2} dx \leq K\sqrt{\frac{2}{\pi\sigma_n^2}}. \tag{B.10}$$

Thus (σ_n) must be bounded else the above quantity could be reduced below $\frac{1}{2}$. Hence there exist a subsequence along which $\theta = \lim_n \theta_{k_n}$ and $\sigma = \lim_n \sigma_{k_n}$ both exist. Then since $x \mapsto e^{i\xi x}$ is a continuous bounded function, by definition of weak convergence

$$\hat{\mu}(\xi) = \lim_n \hat{\mu}_n(\xi) = \lim_n \hat{\mu}_{k_n}(\xi) = \lim_n e^{i\theta_n \xi - \frac{1}{2}\sigma_n^2 \xi^2} = e^{i\theta\xi - \frac{1}{2}\sigma^2}. \tag{B.11}$$

Since Fourier transforms uniquely determine probability measures, $\mu \sim N(\theta, \sigma^2)$.

For general $\mathcal{B}$, if $\mu_n \Rightarrow \mu$ then $l_* \mu_n \to l_* \mu$ for all $l \in \mathcal{B}^*$. Therefore $l_* \mu$ is Gaussian by the $\mathcal{B} = \mathbb{R}$ case. Hence μ is Gaussian. $\qquad\square$

Proposition B.5 (*Mutual Singularity and Total Variation*)

Let μ and ν be probability measures on the measurable space $(E, \mathcal{E})$. Recall the total variation

$$\|\mu - \nu\|_{TV} = \sup_{A \in \mathcal{E}} |\mu(A) - \nu(A)|. \tag{B.12}$$

Then the following are equivalent:

1. *μ and ν are mutually singular.*

2. *$\|\mu - \nu\|_{TV} = 1$.*

Proof. $(1) \implies (2)$ is clear. For the converse if $\|\mu - \nu\|_{TV} = 1$ then there exists $\{A_n\} \subset \mathcal{E}$ such that

$$\lim_n |\mu(A_n) - \nu(A_n)| = 1 \tag{B.13}$$

The sequence $(\mu(A_n))$ and $(\nu(A_n))$ are bounded so by Bolzano-Weierstrass we can pass to a subsequence along which $\alpha := \lim_n \mu(A_n)$ and $\beta := \lim_n \nu(A_n)$ both exist. Then $\alpha, \beta \in [0, 1]$ but also $|\alpha - \beta| = 1$. The only way this holds is if $\alpha = 0$ and $\beta = 1$ or $\alpha = 1$ and $\beta = 0$. WLOG we have the case $\alpha = 0$ and $\beta = 1$. In other words $\lim_n \mu(A_n) = 0$ so by passing to a further subsequence we can WLOG that

$$\mu(A_n) \leq 2^{-n}. \tag{B.14}$$

Let $A = \{A_n \text{ occurs infinitely often}\} = \cap_{n \geq 1} \cup_{k \geq n} A_k$. Then $\sum_n \mu(A_n) < \infty$ thus $\mu(A) = 0$ by the first Borel Cantelli lemma. Finally

$$\nu(A) = \lim_n \nu(\cup_{k \geq n} A_k) \geq \lim_n \nu(A_n) = 1. \qquad \text{(B.15)}$$

Thus $\mu(A) = 0$ and $\nu(A) = 1$, showing that μ and ν are mutually singular. $\quad \square$