Weak Probability Distributions on Reproducing Kernel Hilbert Spaces

Kuelbs, Larkin and Williamson

Jon Cockayne

25 July 2017 (“Larkin Day”)

Motivation

Given $\{(x_i, y_i)\}$, $i = 1, \ldots, M$, recover $f(x)$.

$$y_i = f(x_i) + \xi$$

Motivation

Given $\{(x_i, y_i)\}$, $i = 1, \ldots, M$, recover $f(x)$.

$$y_i = f(x_i) + \xi$$

Kuelbs et. al. argue that the right way to do this is to use probability.

Motivation

Given $\{(x_i, y_i)\}$, $i = 1, \ldots, M$, recover $f(x)$.

$$y_i = f(x_i) + \xi$$

Kuelbs et. al. argue that the right way to do this is to use probability.

This paper provides a rigorous review of the theory of Gaussian measures on function spaces.

Desiderata

1. **Favour smoother curves over rougher ones.**
2. **Assign equal probability to curves of equal norm.**

Kuelbs et. al. argue that this implies a Gaussian distribution.

The Familiar Choice

Define

$$\nu\left(\{h \in H : (\langle y_1, h\rangle, \ldots, \langle y_n, h\rangle) \in E\}\right) \propto \int_E \exp\left(-\frac{1}{2} x^T \Sigma^{-1} x\right) \mathrm{d}x$$

for y_i **linearly independent functionals in** H^* and $E \subset \mathbb{R}^n$, with

$$\Sigma_{ij} = \langle y_i, y_j \rangle$$

The Familiar Choice

Define

$$\nu\left(\{h \in H : (\langle y_1, h\rangle, \ldots, \langle y_n, h\rangle) \in E\}\right) \propto \int_E \exp\left(-\frac{1}{2}x^T\Sigma^{-1}x\right) \mathrm{d}x$$

for y_i **linearly independent functionals in** H^* and $E \subset \mathbb{R}^n$, with

$$\Sigma_{ij} = \langle y_i, y_j\rangle$$

It is shown that the extension of the algebra generated by these cylinder sets to Borel sets of H is not countably additive.

Enlarging the Space

Idea: enlarge H by completing with respect to a measurable (semi)-norm $\|\cdot\|_1$.

Definition

A (semi-)norm $\|\cdot\|_1$ is called a *measurable (semi-)norm* (?wrt ν?) if, for every $\epsilon > 0$ there exists a finite-dimensional projection P_0 such that, for all finite-dimensional P orthogonal to P_0

$$\nu(\{h \in H : \|Ph\|_1 > \epsilon\}) < \epsilon$$

Idea: enlarge H by completing with respect to a measurable (semi)-norm $\|\cdot\|_1$.

Definition

A (semi-)norm $\|\cdot\|_1$ is called a *measurable (semi-)norm* (?wrt ν?) if, for every $\epsilon > 0$ there exists a finite-dimensional projection P_0 such that, for all finite-dimensional P orthogonal to P_0

$$\nu(\{h \in H : \|Ph\|_1 > \epsilon\}) < \epsilon$$

...in other words I can “absorb” an arbitrary amount of probability mass into P_0 with respect to these sets.

Enlarging the Space

Define μ by:

$$\mu(\{b \in B : (y_1(b), \ldots, y_n(b)) \subset E\}) = \nu(\{h \in H : ((y_1|H)(b), \ldots, (y_n|H)(b)) \subset E\})$$

Enlarging the Space

Define μ by:

$$\mu(\{b \in B : (y_1(b), \ldots, y_n(b)) \subset E\}) = \nu(\{h \in H : ((y_1|H)(b), \ldots, (y_n|H)(b)) \subset E\})$$

Theorem (Gross's Theorem)

If H is a real separable Hilbert space and B is the completion of H under the measurable norm $\|\cdot\|_1$ then μ is countably additive on the cylinder sets of B.

Definition (Tame Function)

A function f on H is called a *tame function* if there exists a finite-dimensional projection P so that $f(h) = f(Ph)$.

Integration wrt ν

- Extend tame functions on H to B by the map $\Gamma : H \to B$.

Integration wrt ν

- Extend tame functions on H to B by the map $\Gamma : H \to B$.
- It is shown that $\Gamma(f)$ can be integrated w.r.t. μ when f is tame.

Integration wrt ν

- Extend tame functions on H to B by the map $\Gamma : H \to B$.
- It is shown that $\Gamma(f)$ can be integrated w.r.t. μ when f is tame.
- Extend functions on H to B via

$$\Gamma(f) = \tilde{f} = \lim_{n\to\infty} \Gamma\left(f \circ P_n\right)$$

where P_n is a sequence of finite-dimensional projections increasing strongly to the identity.

Integration wrt ν

- Extend tame functions on H to B by the map $\Gamma : H \to B$.
- It is shown that $\Gamma(f)$ can be integrated w.r.t. μ when f is tame.
- Extend functions on H to B via

$$\Gamma(f) = \tilde{f} = \lim_{n\to\infty} \Gamma\left(f \circ P_n\right)$$

  where P_n is a sequence of finite-dimensional projections increasing strongly to the identity.
- The authors then prove that *under certain regularity conditions*

$$\int_B \Gamma(f)(x)\mu(\mathrm{d}x) = \lim_{n\to\infty} \int_H f \circ P_n(h)\nu(\mathrm{d}h)$$

Role of the Reproducing Kernel

Suppose H is an RKHS with reproducing kernel k, and let $\|h\|_1 = \sup_{x \in D} h(x)$. Let δ_t denote the evaluation functional at t. Then $\{\delta_t(b) : t \in D\}$ is a GP and

$$\mathbb{E}(\delta_t \delta_s) = k(t, s)$$

Summary

- Rigorous construction of Gaussian measures on Hilbert spaces.
- Rigorous construction of integrals w.r.t. these measures.