Gaussian Process Convolution Model

Wessel Bruinsma

University of Cambridge, CBL

20 December 2019

$$f \sim \mathcal{GP}(0,\ \ k\ \ ).$$

$$f \sim \mathcal{GP}(0, ???).$$

$$f \sim \mathcal{GP}(0, ???).$$

AA^{T} is P.S.D.

$$f \sim \mathcal{GP}(0, ???).$$

AA^{T} is P.S.D.

"hh^{T}" $= h * Rh$ is P.S.D.

$$(Rh)(t) = h(-t)$$

$$f \sim \mathcal{GP}(0, ???).$$

AA^T is P.S.D.

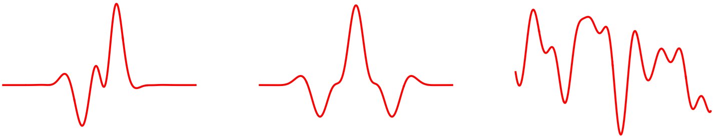
"hh^T"$= h * Rh$ is P.S.D.

$$(Rh)(t) = h(-t)$$

$$h \sim \mathcal{GP}(0, k_h),$$

$$f \mid h \sim \mathcal{GP}(0, h * Rh).$$

$h \sim \mathcal{GP}(0, k_h), \qquad k \,|\, h = h * Rh, \qquad f \,|\, k \sim \mathcal{GP}(0, k)\colon$



Gaussian Process Convolution Model

Model (GPCM (Tobar et al., 2015))

$$h \sim \mathcal{GP}(0, k_h), \qquad f \mid h \sim \mathcal{GP}(0, h * Rh).$$

Gaussian Process Convolution Model

Model (GPCM (Tobar et al., 2015))

$$h \sim \mathcal{GP}(0, k_h), \qquad f \mid h \sim \mathcal{GP}(0, h * Rh).$$

$$x \sim \mathcal{N}(0, I) \implies Ax \sim \mathcal{N}(0, AA^\mathsf{T})$$

Gaussian Process Convolution Model

Model (GPCM (Tobar et al., 2015))

$$h \sim \mathcal{GP}(0, k_h), \qquad f \mid h \sim \mathcal{GP}(0, h * Rh).$$

$$x \sim \mathcal{N}(0, I) \implies Ax \sim \mathcal{N}(0, AA^{\mathsf{T}})$$
$$x \sim \mathcal{GP}(0, \delta) \implies \text{“}hx\text{”} \sim \mathcal{GP}(0, \text{“}hh^{\mathsf{T}}\text{”})$$

Gaussian Process Convolution Model

Model (GPCM (Tobar et al., 2015))

$$h \sim \mathcal{GP}(0, k_h), \qquad f \mid h \sim \mathcal{GP}(0, h * Rh).$$

$$x \sim \mathcal{N}(0, I) \implies Ax \sim \mathcal{N}(0, AA^{\mathsf{T}})$$

$$x \sim \mathcal{GP}(0, \delta) \implies h * x \sim \mathcal{GP}(0, h * Rh)$$

Gaussian Process Convolution Model

Model (GPCM (Tobar et al., 2015))

$$h \sim \mathcal{GP}(0, k_h), \qquad f \mid h \sim \mathcal{GP}(0, h * Rh).$$

$$x \sim \mathcal{N}(0, I) \implies Ax \sim \mathcal{N}(0, AA^\mathsf{T})$$
$$x \sim \mathcal{GP}(0, \delta) \implies h * x \sim \mathcal{GP}(0, h * Rh)$$

Model (GPCM (Tobar et al., 2015), Equivalent Formulation)

$$h \sim \mathcal{GP}(0, k_h), \qquad x \sim \mathcal{GP}(0, \delta), \qquad f \mid h, x = h * x.$$

Inference

Model (GPCM (Tobar et al., 2015), Equivalent Formulation)

$$h \sim \mathcal{GP}(0, k_h), \qquad x \sim \mathcal{GP}(0, \delta), \qquad f \mid h, x = h * x.$$

Inference

Model (GPCM (Tobar et al., 2015), Equivalent Formulation)

$$h \sim \mathcal{GP}(0, k_h), \qquad x \sim \mathcal{GP}(0, \delta), \qquad f \mid h, x = h * x.$$

- Joint distribution:

$$p(f, h, \underset{\uparrow}{u}, x, \underset{\uparrow}{z}) = p(f \mid h, x) p(h \mid u) p(u) p(x \mid z) p(z).$$

inducing points for h and x resp.

Inference

Model (GPCM (Tobar et al., 2015), Equivalent Formulation)

$$h \sim \mathcal{GP}(0, k_h), \qquad x \sim \mathcal{GP}(0, \delta), \qquad f \mid h, x = h * x.$$

- Joint distribution:

$$p(f, h, \underset{\uparrow}{u}, x, \underset{\uparrow}{z}) = p(f \mid h, x) p(h \mid u) p(u) p(x \mid z) p(z).$$

inducing points for h and x resp.

- Approximate posterior:

$$q(f, h, u, x, z) = p(f \mid h, x) p(h \mid u) q(u) p(x \mid z) q(z).$$

Extension: Improved Inference

- Mean-field approximate posterior:

$$q(f, h, u, x, z) = p(f \mid h, x)p(h \mid u)q(u)p(x \mid z)q(z).$$

Extension: Improved Inference

- Mean-field approximate posterior:

$$q(f, h, u, x, z) = p(f \mid h, x)p(h \mid u)q(u)p(x \mid z)q(z).$$

- Structured mean-field approximate posterior:

$$q(f, h, u, x, z) = p(f \mid h, x)p(h \mid u)p(x \mid z)q(u, z).$$

Extension: Improved Inference

- Mean-field approximate posterior:

$$q(f, h, u, x, z) = p(f \mid h, x)p(h \mid u)q(u)p(x \mid z)q(z).$$

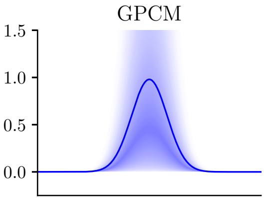
- Structured mean-field approximate posterior:

$$q(f, h, u, x, z) = p(f \mid h, x)p(h \mid u)p(x \mid z)q(u, z).$$

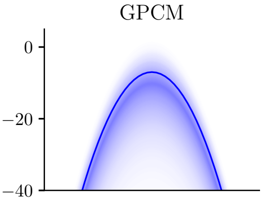
- MCMC to sample from q^*.

Extension: Causality

Prior over kernel:

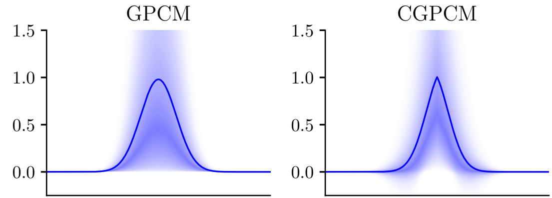


Prior over PSD:

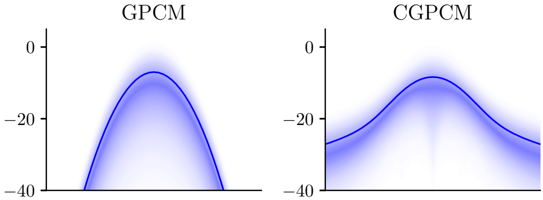


Extension: Causality

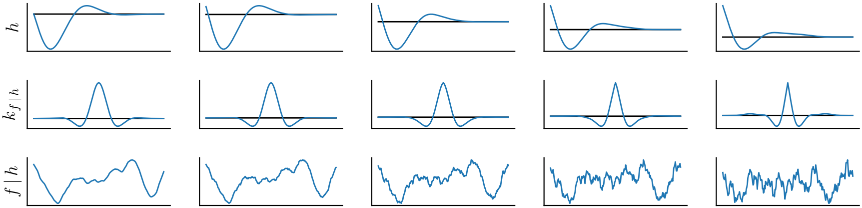
Prior over kernel:



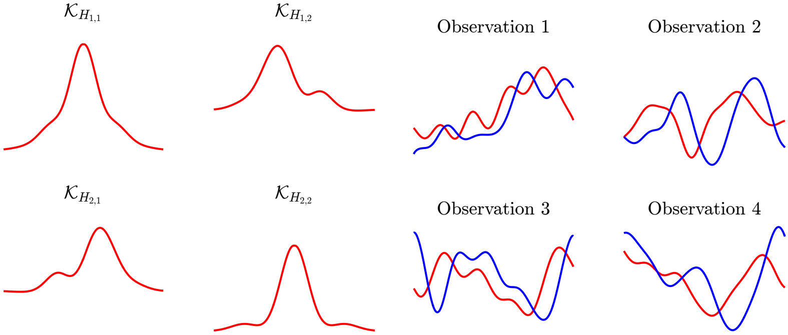
Prior over PSD:



Extension: Causality



Extension: Multiple Outputs



But what about the *kernel of the kernel*?

But what about the *kernel of the kernel*?

And the *kernel of the kernel of the kernel*?

But what about the *kernel of the kernel*?

And the *kernel of the kernel of the kernel*?

And the *kernel of the kernel of the kernel of the kernel*?

And the *kernel of the kernel of the kernel of the kernel of the kernel*?

$$\vdots$$

And the *kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel of the kernel*?

Extension: Deep Kernel Model

Model (N-Deep Kernel Model)

$$
\begin{aligned}
h_0 &\sim \mathcal{GP}(0, k_h), \\
h_1 \,|\, h_0 &\sim \mathcal{GP}(0, h_0 * R h_0), \\
&\vdots \\
h_N \,|\, h_{N-1} &\sim \mathcal{GP}(0, h_{N-1} * R h_{N-1}), \\
f \,|\, h_N &= h_N.
\end{aligned}
$$

Extension: Deep Kernel Model