

# Gaussian processes and gravitational wave detection

Robin Hankin<sup>1</sup>

<sup>1</sup>Auckland University of Technology

SKA Symposium 2018

## Outline

- 1 The Gaussian process

## setup

- Consider a random function  $\eta(x)$ , usually  $\eta: \mathbb{R}^n \rightarrow \mathbb{R}$
- difficult or expensive to evaluate; examples include climate models and nuclear test results
- Typically  $n \simeq 20$
- Bayesian view is to consider this a *random function*\*

## setup

- Consider a random function  $\eta(x)$ , usually  $\eta: \mathbb{R}^n \rightarrow \mathbb{R}$
- difficult or expensive to evaluate; examples include climate models and nuclear test results
- Typically  $n \simeq 20$
- Bayesian view is to consider this a *random function*\*
- But this isn't to say we know nothing about it
- We might have evaluated the function before at nearby values
- or have expert judgement available.

# Gaussian

## Inputs

$$\mathbb{E}\eta(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \boldsymbol{\beta}$$

$$\text{Cov}(\eta(\mathbf{x}), \eta(\mathbf{x}')) = \mathbf{c}(\mathbf{x} - \mathbf{x}')$$

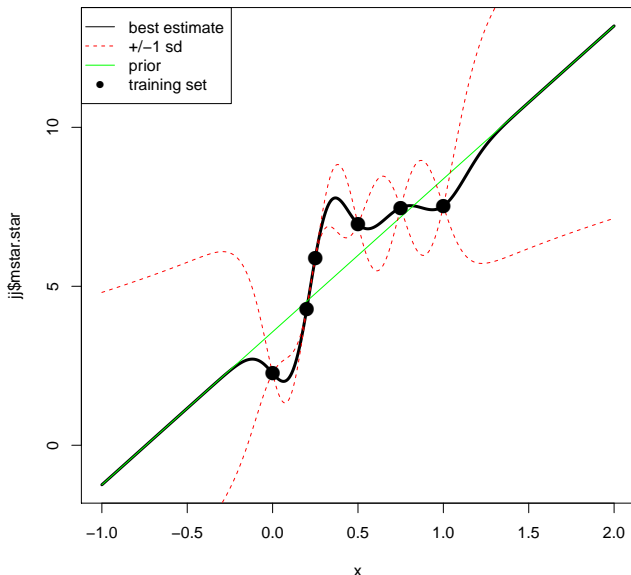
got to be careful with the form of  $\mathbf{c}(\cdot, \cdot)$ ...

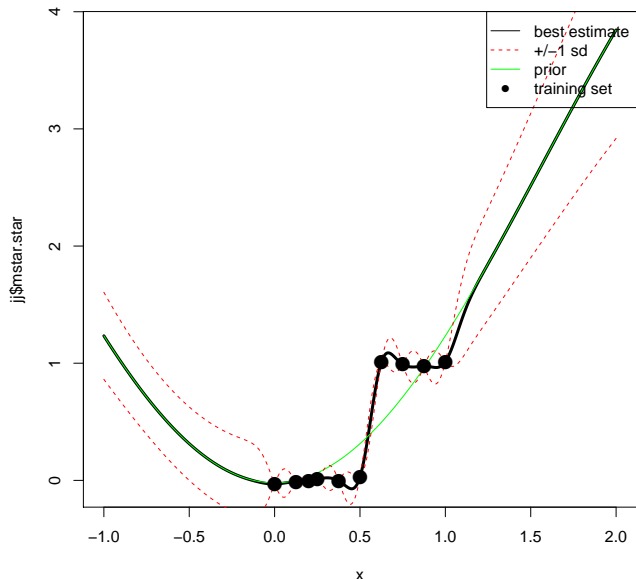
## problem

$$\text{Cov}(\eta(\mathbf{x}_1), \dots, \eta(\mathbf{x}_n)) = \sigma^2 \mathbf{A}$$

must be positive definite

Function  $\mathbf{c}(\cdot, \cdot)$  must be the characteristic function of a symmetric PDF





Then  $z_1 | z_2 = y$ , that is the distribution of  $\eta(x)$  conditional on a “training set” of code evaluations  $y = (\eta(x_1), \dots, \eta(x_n))^T$ , is multivariate Gaussian with mean

$$\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(a - \mu_2)$$

and variance-covariance matrix

$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$



Usually condition on code observations evaluated on a design matrix:

$$H = \begin{pmatrix} h(x_1) \\ \vdots \\ h(x_n) \end{pmatrix}$$

recalling that  $\mathbb{E}\eta(x) = h(x)^T \beta$  we can estimate  $\beta$  with

$$\hat{\beta} = \left( H^T A^{-1} H \right)^{-1} H^T A^{-1} y$$

where  $A$  is the variance matrix of code observations. If  $t(x)$  is the covariance between  $x$  and the training set  $x_1, \dots, x_n$ , it turns out that

$$\mathbb{E}\eta(x) = h(x)^T \hat{\beta} + t(x)^T A^{-1} (y - H\beta)$$

This is the *emulator*. It provides a cheap approximation to the expensive-to-evaluate function  $\eta(x)$  and also furnishes an estimate of its own error.

Bochner's theorem states that the covariance function must be the characteristic function of a symmetric probability Borel measure  $X$ :

For any random variable  $X$ :

$$\phi_X(t) = \mathbb{E}e^{it^T X}$$

Usually take a Gaussian distribution.

This gives

$$c(x, x') = e^{-(x-x')^T B(x-x')}$$

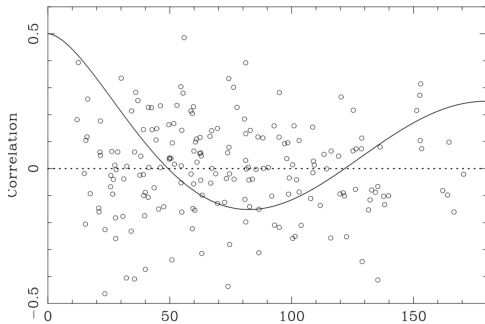
where  $B$  is a positive-definite matrix that measures the roughness of the function  $\eta(\cdot)$ .

Situation is more complicated in the case of  $\mathbb{S}^2$ . Can use the natural embedding  $\mathbb{S}^2 \subset \mathbb{R}^3$ , but this sucks.

Schoenberg 1938 shows that all positive definite functions on  $\mathbb{S}^2$  are of the form

$$\psi(\theta) = \sum_{n=0}^{\infty} \frac{b_{n,2}}{n+1} P_n(\cos \theta)$$

where  $P()$  are Legendre polynomials and the  $b_{n,2}$  are numerical coefficients.



5

$$\text{Cor}(i, j) = \frac{1}{3} + \frac{1 - \cos \gamma_{ij}}{2} \log \left( \frac{1 - \cos \gamma_{ij}}{2} \right) - \frac{1}{6} \cdot \frac{1 - \cos \gamma_{ij}}{2}$$