Gaussian processes and gravitational wave detection

Robin Hankin[1]

[1] Auckland University of Technology

SKA Symposium 2018

Outline

setup

- Consider a random function $\eta(x)$, usually $\eta\colon \mathbb{R}^n \longrightarrow \mathbb{R}$
- 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 $\eta\colon \mathbb{R}^n \longrightarrow \mathbb{R}$
- 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

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

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

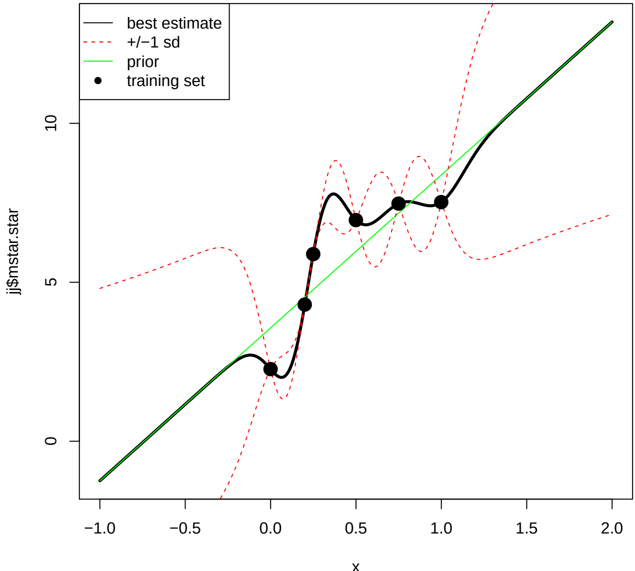
got to be careful with the form of $c(\cdot,\cdot)$...

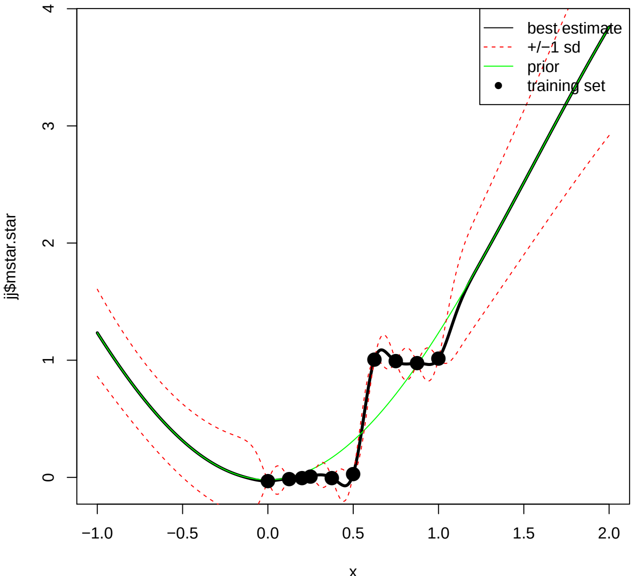
problem

$$\mathrm{Cov}(\eta(x_1), \ldots, \eta(x_n)) = \sigma^2 A$$

must be positive definite

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


best estimate
+/−1 sd
prior
training set
10
5
0
jj$mstar.star
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
x


best estimate
+/−1 sd
prior
training set
jj$mstar.star
x
0
1
2
3
4
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

Then $z_1|z_2 = y$, that is the distribution of $\eta(x)$ conditional on a "training set" of code evaluations $y = (\eta(x_1), \ldots, \eta(x_n))^T$, 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 β 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 $x_1, \ldots, x_n$, 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

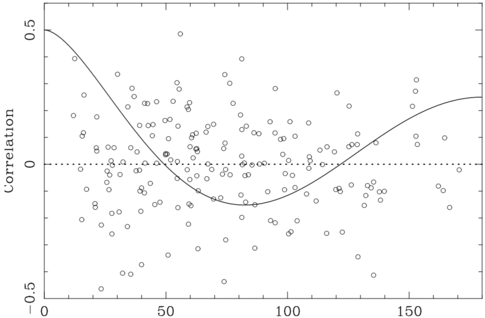
$$c(x, x') = e^{-(x-x')^T B(x-x')}$$

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.


Correlation
0.5
0
−0.5
0
50
100
150

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