Optimizing the ln likelihood function in George

author: Karen Y. Ng

Kernel used:

$$k(r) = \Sigma(r^2) + c\delta_{ij}$$

where $\Sigma(r^2)$ is `ExpSquaredKernel` or its derivatives.
The second term is the `WhiteKernel` in George

Parametrization used in `George`

See `george_examples/basic_properties_of_george.ipynb`

$$k(r^2) = \lambda^{-1} \exp\left(-\frac{r^2}{2l^2}\right)$$

Parametrization for better visualization...

It is hard to visualize the likelihood surface in the original scale, so we perform a log transformation

$$a = \log_{10}(\lambda^{-1})$$
$$b = \log_{10}(l^2)$$

To have:

$$k(r^2) = 10^a \exp\left(-\frac{r^2}{2 \times 10^b}\right)$$

How new parametrization relates to ρ

$$\begin{aligned} b &= \log_{10}(l^2) && (1)\\ &= \log_{10}(1/\beta) && (2)\\ &= -\log_{10}\left(-\frac{1}{4}\ln\rho\right) && (3)\end{aligned}$$

Note on transformation of variables

Jacobian needed to preserve the area of integrated PDF

$$f_y(\vec{y}) = f_x(\vec{x})|\det(J)|$$

Only when the transformed variable is the one that we integrate with respect to.

Since our likelihood can be written as

$$L(\lambda^{-1}, \beta|\vec{x}) = P(\vec{x}|\lambda^{-1}, \beta)$$

it shows that we are not actually integrating w.r.t. our parameters but integrating our variables $\vec{x}$. No Jacabian is needed.

Optimizing the GP lnlikelihood