Problems with `GPvecchia` variance estimates

Marcin Jurek

October 8, 2020

Consider a zero-mean Gaussian random field $x(\cdot)$ defined on the $[0,1]^2$ domain and evaluated on a regular grid of dimensions $n_x \times n_y$. Let $\mathbf{x}$ be a vector where each entry corresponds to the value of x at a certain point of the gird. Assume that x has an exponential covariance function with range parameter λ and scale σ^2, i.e.

$$cov\left(x(s_1), x(s_2)\right) = \sigma^2 e^{-d(s_1,s_2)/\lambda}$$

where $d(s_1, s_2)$ is the Euclidian distance between locations s_1 and s_2. This lets us write the distribution of $\mathbf{x}$ as

$$\mathbf{x} \sim \mathcal{N}(0, \sigma^2\Sigma),$$

where we can write the i, j-th element of Σ as $\Sigma[i,j] = e^{-d(s_i,s_j)/\lambda}$.

Using this notation we can write that $\Sigma^{-1/2}\mathbf{x} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$. This means, we can estimate the variance parameter, σ^2 as

$$\hat{\sigma}^2 = \frac{1}{n-1}\mathbf{x}^\top\Sigma^{-1}\mathbf{x} \tag{1}$$

where $n = n_x n_y$.

We also assume that we observe the values of the process at certain grid locartions and we label the vector of these observations $\mathbf{y}$. If we assume that the observations contain some iid Gaussian measurement error, we can write $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$, with $\mathbf{v} \sim \mathcal{N}(0, \rho^2\mathbf{I})$ and $\mathbf{H}$ being generated from the identity matrix by removing the rows corresponding to grid points with no data.

`GPvecchia` exports the function `calculate_posterior_VL()` which returns, among other things, $\hat{\mathbf{x}} = E(\mathbf{x}|\mathbf{y})$, the approximate posterior mean of $\mathbf{x}$ given the data. If $\hat{\mathbf{x}}$ is a good approximation of $\mathbf{x}$, then one might expect that using it in 1 would produce a good approximation of σ^2. In other words, if we define

$$\hat{\hat{\sigma}}^2 = \frac{1}{n-1}\hat{\mathbf{x}}^\top\Sigma^{-1}\hat{\mathbf{x}},$$

then we would expect $\hat{\hat{\sigma}}^2$ to be close to σ^2.

It turns out, however, that this is not the case, **even if the measurement error variance ρ^2 is very small**. For example in the attached file we used $\rho^2 = 10^{-8}$ and $\lambda = 0.2$ and we assumed that half of the process is observed at half of all grid points. Our simulations also show that when we do not use the Vecchia approximation and calculate $\hat{\sigma}^2$ instead, it seems to be a an unbiased estimate of σ^2