gg-GP

Rob McCulloch

November 13, 2019

Griddy Gibbs Inference for Gaussian Process Parameters

Read 15.1 to 15.2.5 (inclusive) of "Machine Learning, A Probabalistic Perspective" by Kevin Murphy. You can find pdf on the internet.

The goal is to use simple Gibbs sampling to infer the parameters of a simple Gaussian Process model.

Our model is

$$y_i = f(x_i) + \epsilon_i$$

To keep things simple let's assume $x_i \in R$.

where $\epsilon_i \sim N(0, \sigma_y^2)$ and

$$(f(x_1), f(x_2), \ldots, f(x_N))' \equiv f(x) \sim N(0, K), \quad K_{ij} = \kappa(x_i, x_j).$$

Then,

$$Y \mid x \sim N(0, K_y), \quad K_y = K + \sigma_y^2 I.$$

Now let κ depend on some parameters so that we have $\kappa(x_i, x_j; \gamma)$ and hence

$$K_y = K(\gamma) + \sigma_y^2 I.$$

The example in Murphy is

$$\kappa(x_i, x_j; \sigma_f, l) = \sigma_f^2 \exp(-\frac{1}{2l^2} (x_i - x_j)^2)).$$

so $\gamma = (\sigma_f, l)$.

Letting $\theta = (\gamma, \sigma_y)$ We have

$$K_y(\theta) = K(\gamma) + \sigma_y^2 I.$$

Then $Y \mid x \sim N(0, K_y(\theta))$ and

$$\log(p(y \mid x, \theta)) = -\frac{1}{2} y' K_y(\theta)^{-1} y - \frac{1}{2} \log(|K_y(\theta)|) - \frac{N}{2} \log(2\pi).$$

which is Equation 15.2.2 of Murphy.

Now the idea of the project is :

- use lines 1,2, and 6 of Algorithm 15.1 in Murphy to compute $\log(p(y \mid x, \theta))$.
- use a simple grid of values for each element of θ.
- use something like the Griddy Gibbs sampler to obtain posterior inference for θ.

That is, fix all of the elements of θ but one: $\theta = (\theta_i, \theta_{-i})$. Fix θ_{-i} and then evaluate

$$\log(p(\theta_i \,|\, \theta_{-i}, y)) \propto \log(p(y \,|\, x, \theta)) + \log(p(\theta))$$

on a grid of values for θ_i. Here $p(\theta)$ is the prior.

Alternatively, you could use a simple MH step instead of the griddy Gibbs for the each θ_i.

Note, Algorithm 15.1 just says that to compute $\log(p(y \,|\, x, \theta))$ let L be the Cholesky root of $K_y(\theta)$. Then use the Cholesky to compute $|K_y(\theta)|$ and $K_y(\theta)^{-1}y \equiv \alpha$.

That is, If $K_y = LL'$, then

$$\alpha = K_y^{-1}y = (LL')^{-1}y = (L^{-1})'L^{-1}y.$$

and

$$\log(|K_y|^{1/2}) = \sum \log(L_{ii}).$$

Use real and simulated data to see how your procedure works.