A Hybrid Gaussian Process Convolution Neural Network Approach to Single Image Super Resolution

Steven I. Reeves

August 1, 2019

Abstract

In this manuscript we present a new interpolatation scheme for Adaptive Mesh Refinement (AMR) for

1 Introduction

2 Gaussian Process Modeling

The method we are presenting is based on Gaussian Process Modeling, and in this section we give a brief overview on constructing a Gaussian Process Model. Gaussian Processes are a family of stochastic processes in which any finite collection of random variables sampled from this process are joint normally distributed. In a more general sense, Gaussian Processes sample functions from an infinite dimensional function space. The interpolating routine described in detail in section **??**, will be drawn from a data-informed distribution space trained on the coarse grid data.

To construct a Gaussian Process model, one needs to specify a *prior probability distribution* for the function space. Samples (function values evaluated at known locations) are then used to update this prior probability distribution, and using Bayes' Theorem a *posterior probability distribution* is generated given the prior and the samples.

2.1 A statistical introduction to Gaussian Processes

The construction of the posterior probability distribution over the function space is the heart of Gaussian Process Modeling. We can draw functions from this data adjusted space to generate an interpolating model. Specifically, the posterior may be used to probabilistically predict the value of a function at points where the function has not been previously sampled. We use the *posterior mean* for our AMR prolongator.

A Gaussian Process is a collection of random variables, in which any finite collection has a joint Gaussian distribution [**?**][**?**]. GPs can be fully defined by two functions: a mean function $\bar{f}(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})]$ and a covariance function that generates a symmetric, positive-definite kernel $K(\mathbf{x}, \mathbf{y}) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$.

We denote functions f drawn from a GP with mean function $\bar{f}(\mathbf{x})$ and covariance $K(\mathbf{x}, \mathbf{y})$ as $f \sim \mathcal{GP}(\bar{f}, K)$. Analogous to finite-dimensional distributions we write the covariance as

$$K(\mathbf{x}, \mathbf{y}) = \mathbb{E}\left[\left(f(\mathbf{x}) - \bar{f}(\mathbf{x})\right)\left(f(\mathbf{y}) - \bar{f}(\mathbf{y})\right)\right] \tag{1}$$

where $\mathbb{E}$ is with respect to the GP distribution.

One controls the GP by specifying both $\bar{f}(\mathbf{x})$ and $K(\mathbf{x}, \mathbf{y})$, typically as some hyper-parameterized functions. These hyper-parameters allow us to give the "character" of functions generated by the posterior (i.e. length scales, differentiability). Suppose we have a given GP, and N locations $\mathbf{x}_i, i = 1, \ldots, N$ at which we collect samples $f(\mathbf{x}_i$, then we can calculate the likelihood $\mathcal{L}$ – the probability of the data given the GP model. Let $\mathbf{f} = [f(\mathbf{x}_1, \ldots, f(\mathbf{x}_N)]^T$ then the likelihood is

$$\mathcal{L} \equiv P(\mathbf{f}|\mathcal{GP}(\bar{f}, K)) = (2\pi)^{-N/2} \det|\mathbf{K}|^{-1/2} \exp\left[-\frac{1}{2}\left(\mathbf{f} - \bar{\mathbf{f}}\right)\mathbf{K}\left(\mathbf{f} - \bar{\mathbf{f}}\right)\right] \tag{2}$$

where $\mathbf{K}$ is a matrix generated by $K_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j)$, $i, j = 1, \ldots, N$, and $\bar{\mathbf{f}} = [\bar{f}(\mathbf{x}_1), \cdots \bar{f}(\mathbf{x}_N)]$. Using these samples we can make a probabilistic statement about the value of the function $f \sim \mathcal{GP}(\bar{f}, K)$ at a new point $\mathbf{x}_*$. That is, we can model the value of $f(\mathbf{x}_*)$ using this GP model. This is especially import for Adaptive Mesh Refinement, as we need to construct data at a finer resolution.

An application of Bayes' Theorem along with the conditioning property, directly onto the joint Gaussian prior given the samples $\mathbf{f}$ give the posterior distribution of the predicted value, f_* given $\mathbf{f}$.

$$P(f_*|\mathbf{f}) = (2\pi U^2)^{-1/2} \exp\left[-\frac{(f_* - \bar{f}_*)^2}{2U^2}\right]. \tag{3}$$

Perhaps more important for this application, the posterior PDF gives a new *posterior mean function*

$$\tilde{f}_* \equiv \bar{f}(\mathbf{x}_*) + \mathbf{k}_*^T \mathbf{K}^{-1} \cdot (\mathbf{f} - \bar{\mathbf{f}}) \tag{4}$$

and *posterior covariance*

$$U^2 \equiv k_{**} - \mathbf{k}_*^T \mathbf{K}^{-1} \cdot \mathbf{k}_*. \tag{5}$$

In the application of AMR prolongation, we assume that the coarse grid data is perfectly sampled – that is, there is no *uncertainty*. Because of this, we do not need the posterior covariance, and the posterior mean will become the basis for our interpolation.

In the next subsections we describe two modeling schemes for AMR prolongation. The first is a description for prolongating pointwise states, and the second will illustrate the method for volume-averaged quantities.

References