Exploring the Heat Equation with Gaussian Processes

Misha Padidar*

*padidar@umail.ucsb.edu,3280559

December 7, 2016

1 Introduction

Solving the heat equation is a well known problem with a defined analytic solution. This, however, makes it a perfect problem to employ novel methods of finding a solution. Rather than solving the heat equation analytically, we aim to solve the ill posed one dimensional inverse heat equation using a Posterior Gaussian Process(GP). More specifically, we wish to find the relationship between temperature at a position on a rod and the location of a heat source under the rod. A Gaussian Process is a distribution over functions, which uses data in conjunction with specified mean and kernel functions to create conditional mean and kernel functions; the distributions of which interpolate the data. In terms of our problem, we are using a posterior guassian process with measured temperature data, taken from a specified location on a rod, to create adjusted mean and kernel functions, which in turn yield heat source locations. We can then compare our estimates to the analytic solution to check the effectiveness of the method. To explore the solutions to the inverse heat equation, we shall take a look at the analytic solution to the heat equation, then "solve" the inverse heat equation with a GP, and conclude by analyzing the error in our methods.

2 The Heat Equation

The one dimensional heat equation is a partial differential equation given by

$$\frac{\partial^2 u}{\partial x^2} = D\frac{\partial^2 u}{\partial t^2} \tag{1}$$

where D is the thermal diffusivity. We approach the heat equation with the dirac delta initial condition $u(x, 0) = \delta(x)$, the solution of the which can be found via Fourier transforms to be

$$u(x,t) = \frac{1}{\sqrt{4Dt\pi}} \exp\left(\frac{-(y-x)^2)}{4Dt}\right) \tag{2}$$

where y is the location of the heat source and x is a location on the rod. Note that for our problem, we take our measurements at a fixed location $x = x_0$ and time $t = \tau$ such that $2D\tau = 1$, varying only the heat source y. So denote the temperature at (x_0, τ) as $T(y)$. This yields the equation

$$T(y) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(y-x_0)^2)}{2}\right) \tag{3}$$

Our problem however requires that we solve for the location of the heat source y rather than the temperature. Rearranging the solution to the heat equation gives us what would be the solution to the inverse heat equation given our measurement location and time

$$y(T) = x_0 + \sqrt{-2\ln(T\sqrt{2\pi})} \tag{4}$$

Now that we have found an analytic solution to our problem, let us explore the solution to our problem with a Gaussian Processes.

3 Solution with Gaussian Process

We aim to use temperature data measured at a fixed location on a rod to find the location of a heat source, y, at $t = 0$. To do this we shall use a posterior gaussian process on our temperature data to create adjusted mean and covariance functions, where functions from the distributions are possible solutions to the equation. However, rather than generating solutions with the adjusted mean and covariance sturucture we will simply use the adjusted mean as our solution, as it represents the average of all the possible functions that could be our solution. To use a GP to solve the problem at hand we must define initial kernel and mean functions. We choose mean zero as we have no good estimate what our mean of our data may be. Also, as shown in equation (7), this forces our adjusted mean to be based solely on the function values f and the covariance matrices. Our kernel will be defined as the standard mean squared kernel as we expect our data to be distributed normally.

$$m(x) = 0 \tag{5}$$

$$k(x, x') = \sigma_y \exp\left(\frac{-(x - x')^2}{2l^2}\right) + \sigma_n^2 \delta_{ii} \tag{6}$$

We define the hyperparameters ℓ, σ_n, σ_y as the length scale, the noise, and scaling factor, respectively. They are used to adjust the kernel function to decrease error from the true solution. Using the mean function we can construct a GP for our test points(T^*, y^*), given our training data (T, y). The GP is defined as

$$f^*|f \sim GP(m_D, k_D)$$

Where we define the adjusted mean m_D and adjusted kernel as k_D as

$$m_D(x) = m(x) + \Sigma(X, x)^T \Sigma^{-1}(f - m) \tag{7}$$

$$k_D(x, x') = k(x, x') - \Sigma(X, x)^T \Sigma^{-1} \Sigma(X, x') \tag{8}$$

Where $\Sigma, \Sigma_*, \Sigma_{**}$ are the training-training, training-test, and test-test covariance matrices respectively [1]. As the GP is a distribution over functions, we can use the adjusted mean and covariance functions to come up with an estimate for the test points, f^*, aka heat source location.

$$f^*|f \sim N(\mu_* + \Sigma_*^T \Sigma^{-1}(f - \mu), \Sigma_{**} - \Sigma_*^T \Sigma^{-1} \Sigma_*) \tag{9}$$

Now that we have an expression for the estimated heat source location using a GP, we can solve and test our results.

4 Results

Using mean and kernel functions from equations (5) and (6), respectively, with hyperparameters $\ell = 1, \sigma_y = 1, \sigma_n = 0.01$ we were able to run a posterior gaussian process to effectively solve for an inverse heat function. The training data, $(\bar{T}, \bar{y})$ was generated in reverse from equation (3) with measurement location, $x_0 = 3$, set of sources distributed evenly between $[3, 6.5]$ separated by $\Delta y = 0.05$ and noise σ_n^2 to account for natural error in measurements. Note that we only considered sources to the right of our measurement location for increased accuracy in the gaussian process. We also introduced a constant $\beta = 100$, to amplify the size of the data we were working with, as the GP was not fitting the data accurately with the given hyperparameters on the unit scale. All of our calculations were adjusted to account for β. The training temperatures are thus given by

$$\bar{T}_i = \frac{\beta}{\sqrt{2\pi}} \exp\left(\frac{-(\bar{y}_i - x_0)^2}{2}\right) + \sigma_n^2 \delta_{ii} \tag{10}$$

Our test points were distributed discretely from $T^* = [0, \max(\bar{T})], \Delta T^* = 0.1$ We then calculated the adjusted mean and kernel functions, m_D, k_D, using equations (7) and (8). The result of the gaussian process is shown in *figure 1.*

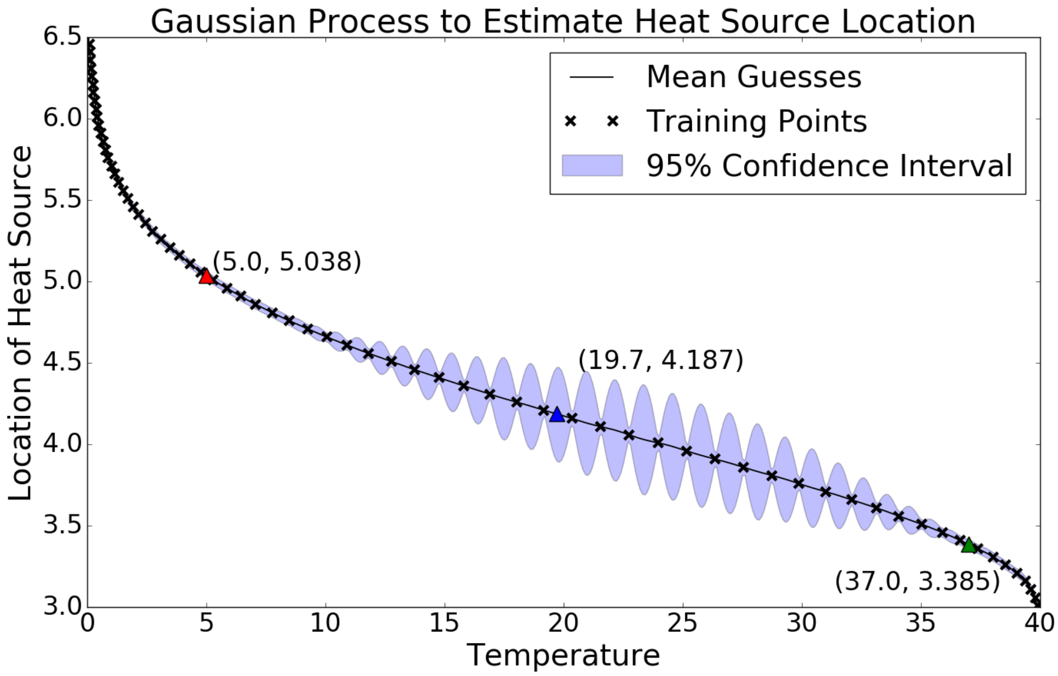

Figure 1: Estimating heat source locations on the rod based on temperature measurement locations taken at $x_0 = 3$ and $t = \tau$, where $2D\tau = 1$. Training data was generated using equation (10) with sources at $\bar{y} = [3, 6.5]$ separated by distance $\Delta y = 0.05$ and σ_n^2, and scalar $\beta = 100$. Hyperparameters $l = 1, \sigma_y = 1, \sigma_n = 0.01$

The figure above shows very little error in the gaussian process. This is due to the choice of mean and kernel functions, as well as the large sampling of data points. The relative errors can be seen below for selected data points.

Temperature	Analytic Heat Source	Estimated Heat Source	Rel. Error
5.0	5.03804	5.03789	2.91916e-05
19.7	4.18795	4.18875	1.90424e-04
37.0	3.38811	3.39095	8.38293e-04

Figure 2: Comparison of the estimated heat source location from a Gaussian Process to the analytic solution.

Although this Gaussian Process proved to be remarkably accurate, accuracy diminishes quickly as the sample size of training points decreases. Also note that performing a GP on the full bell curves causes significant distortion in the gaussian process, unless a significantly higher concentration of points is used. Nevertheless, error can be reduced through optimization of hyperparameters or more aptly choosing mean or kernel functions. However, the GP yielded such promising results that optimization was not necessary.

5 Closing Thoughts

Although the analytic problem of solving the inverse heat equation may be ill posed, the existence of a solution makes it a perfect problem for implementing the mathematically novel method of the Gaussian Process. Gaussian Processes allowed us to effectively estimate the location of a heat source given temperature measurements at a fixed time. A more accurate solution may have been found with optimization of hyperparameters, or by choosing more appropriate mean and kernel functions. However our minimal relative error, on the scale of 10^{-5}, showed that this was not necessary, and furthermore that the gaussian process has the capability for incredibly accurate interpolation. Nevertheless this problem could be explored further, as we only found the relationship between position, temperature and heat source location in one dimensional heat propogation. A complete analysis of the inverse heat equation should include the relationship between heat source location to temperature of the rod at *varying times*, and perhaps even consider multidimensional analogs of the problem. In any case, the Gaussian process has proved to be an imense method for solving some problems in the physical realm, and should be considered in furture study.

References

[1] Carl Edward Rasmussen. “Gaussian Processes in Machine Learning”. In: *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures*. Ed. by Olivier Bousquet, Ulrike von Luxburg, and Gunnar Rätsch. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 63–71. ISBN: 978-3-540-28650-9. DOI: 10.1007/978-3-540-28650-9_4. URL: http://dx.doi.org/10.1007/978-3-540-28650-9_4.