Operator Induced Multi-Task Gaussian Processes for Solving Differential Equations

Arman Melkumyan
Australian Centre for Field Robotics
The University of Sydney
Sydney, NSW 2006, Australia
a.melkumyan@acfr.usyd.edu.au

1 Introduction

Ordinary and partial differential equations are extensively used in different branches of science and engineering to model wide variety of phenomena, such as diffusion, stability, wave propagation, population growth and chemical reactions, to mention just a few. For most practical problems these differential equations cannot be solved analytically and numerical techniques must be employed.

This paper develops methodology for applying mutli-task Gaussian processes to numerical solution of ordinary (ODEs) and partial (PDEs) differential equations. For different classes of ODEs and PDEs a detailed evaluation of the accuracy of the proposed methodology is presented by comparing the obtained numerical solutions with the corresponding exact analytical ones.

2 Differential Equations and Operator Induced Multi-Task GPs

This section briefly reviews differential equations and Gaussian processes (GPs) and then combines them to introduce operator induced multi-task Gaussian processes.

2.1 Differential Equations

Consider the linear differential equation

$$L_{1,\mathbf{x}}\left[u\left(\mathbf{x}\right)\right]=h_1\left(\mathbf{x}\right)\ \text{in the region}\ \Omega_1 \tag{1}$$

with the boundary and/or initial conditions

$$L_{l,\mathbf{x}}\left[u\left(\mathbf{x}\right)\right]=h_l\left(\mathbf{x}\right)\ \text{in the region}\ \Omega_l\ \text{where}\ l=2:M. \tag{2}$$

In Eq. (2) the regions $\Omega_2,...,\Omega_M$ usually represent the boundaries of Ω_1. The task is to find a function $u\left(\mathbf{x}\right)$ which satisfies the differential equation (1) and boundary/initial conditions (2).

2.2 Multi-Task Gaussian Processes

Consider the supervised learning problem of estimating M tasks $\mathbf{y}^*$ for a query point $\mathbf{x}^*$ given a set X of inputs $\mathbf{x}_{11},\dots,\mathbf{x}_{N_1 1},\mathbf{x}_{12},\dots,\mathbf{x}_{N_2 2},\dots,\mathbf{x}_{1M},\dots,\mathbf{x}_{N_M M}$ and corresponding noisy outputs $\mathbf{y}=\left(y_{11},\dots,y_{N_1 1},y_{12},\dots,y_{N_2 2},\dots,y_{1M},\dots,y_{N_M M}\right)^T$, where $\mathbf{x}_{il}$ and y_{il} correspond to the ith input and output for task l respectively, and N_l is the number of training examples for task l.

Introducing the multi-task covariance function

$$\text{cov}\left[f_l(\mathbf{x}),f_k(\mathbf{x}')\right]=K_{lk}(\mathbf{x},\mathbf{x}'), \tag{3}$$

inference can be computed using the conventional GP equations for mean and variance [1], [2]:

$$\bar{f}_l(\mathbf{x}^*)=\mathbf{k}_l^T K_{lk}^{-1}\mathbf{y} \quad \text{and} \quad \mathbb{V}[f_l(\mathbf{x}^*)]=\mathbf{k}_{l*}-\mathbf{k}_l^T K_{lk}^{-1}\mathbf{k}_l, \tag{4}$$

where

$$\mathbf{k}_l = [k_{1l}(\mathbf{x}^*, \mathbf{x}_{11}) \dots k_{1l}(\mathbf{x}^*, \mathbf{x}_{N_1 1}) \quad \dots \quad k_{Ml}(\mathbf{x}^*, \mathbf{x}_{1M}) \dots k_{Ml}(\mathbf{x}^*, \mathbf{x}_{N_M M})]^T. \tag{5}$$

Similarly, learning can be performed by maximizing the log marginal likelihood

$$\mathcal{L}(\theta) = -0.5 \log |K_{lk}| - \mathbf{y}^T K_{lk}^{-1} \mathbf{y} - 0.5(N_1 + N_2 + \dots + N_M) \log 2\pi \tag{6}$$

where θ is a set of hyper-parameters.

2.3 Operator Induced Multi-Task Gaussian Processes

Combining linear operators from differential equations with the multi-task GPs from machine learning introduce Operator Induced Multi-Task Gaussian Processes as follows:

Definition 1 *M task Gaussian process with the multi-task covariance functions $K_{ij}(\mathbf{x}, \mathbf{x}')$ induced by linear operators $L_{l,\mathbf{x}}[f(\mathbf{x})]$ from a single-task covariance function $K_0(\mathbf{x}, \mathbf{x}')$ according to the formula*

$$K_{lq}(\mathbf{x}, \mathbf{x}') = L_{q,\mathbf{x}'}[L_{l,\mathbf{x}}[K_0(\mathbf{x}, \mathbf{x}')]], \ l, q = 1 : M \tag{7}$$

is called an operator induced multi-task Gaussian process (OMGP). The operators $L_{l,\mathbf{x}}$ are called inducing operators and $K_0(\mathbf{x}, \mathbf{x}')$ is called basic covariance function.

The next section describes how to use OMGPs for numerical solution of differential equations.

3 Numerical Solution of Differential Equations

To find an approximate numerical solution of the continuous equations Eq. (1)-(2) one can replace them with the corresponding discrete equations:

$$f_l(\mathbf{x}_{il}) = h_l(\mathbf{x}_{il}), \ l = 1 : M, i = 1 : N_l \tag{8}$$

where $\mathbf{x}_{1l}$, $\mathbf{x}_{2l}$, ..., $\mathbf{x}_{N_l l} \in \Omega_l$, and $f_l(\mathbf{x})$ is defined as

$$f_l(\mathbf{x}) = L_{l,\mathbf{x}}[u(\mathbf{x})]. \tag{9}$$

Assuming that $u(\mathbf{x})$ is a GP with the covariance function $K_0(\mathbf{x}, x')$ and using the linearity of the operators $L_{l,\mathbf{x}}$, from Eq. (9) it follows that $f_l(\mathbf{x})$, $l = 1 : M$, comprise a multi-task GP with the following multi-task covariances:

$$\text{cov}[f_l(\mathbf{x}), f_q(\mathbf{x}')] = L_{q,\mathbf{x}'}[L_{l,\mathbf{x}}[K_0(\mathbf{x}, \mathbf{x}')]], \ l, q = 1 : M \tag{10}$$

Using the Definition 1 and Eqs. (10) and (7) one has that

$$\text{cov}[f_l(\mathbf{x}), f_q(\mathbf{x}')] = K_{lq}(\mathbf{x}, \mathbf{x}'), \ l, q = 1 : M \tag{11}$$

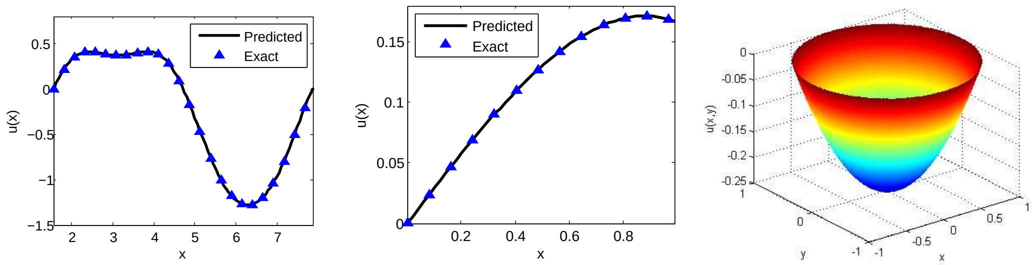
where the multi-task covariance functions $K_{lq}(\mathbf{x}, \mathbf{x}')$ are defined in Eq. (7).

Within the multi-task Gaussian processes framework Eq. (11) represents the covariance functions and Eq. (8) represents a set of noise-free observations. From Eqs. (11) and (8) it now follows that the task of finding an approximate solution of differential equations (1)-(2) can be solved by using multi-task GPs described in section 2.2 where the observations y_{il} are given by Eq. (8), i.e. $y_{il} = h_l(\mathbf{x}_{il})$. Substituting $y_{il} = h_l(\mathbf{x}_{il})$ into Eq. (6) and using the multi-task covariance function Eq. (7) one can maximize the log marginal likelihood to learn the hyper-parameters θ. Once the learning stage is completed, the values of the approximate solution of Eqs. (1)-(2) at arbitrary points can be inferred using Eqs. (3)-(5), (7).

Next section applies this technique to particular problems and evaluates the accuracy of the resulted numerical solutions.

4 Experimental Evaluation

In this section the proposed methodology is applied to numerical solution of ordinary and partial differential equations for which exact solutions are known. The results of numerical computations are compared to the corresponding exact solutions to quantify the mean absolute errors of the obtained numerical solutions. The squared exponential covariance function [1] is chosen as the basic covariance function $K_0(\mathbf{x}, \mathbf{x}')$ for all the experiments presented in this paper.



(a) First order differential equation (12)-(13) with variable coefficients

(b) Second order differential equation (15)-(17) with variable coefficients

(c) Second order elliptic partial differential equation (19)-(20)

Figure 1: Predicted and exact solutions for different kinds of differential equations

4.1 First Order ODE with Variable Coefficients

Consider the following differential equation with variable coefficients

$$u'(x) + u(x)\sin x = \sin^3 x \quad \text{in the region} \quad x \in (0.5\pi, 2.5\pi) \tag{12}$$

with the condition

$$u(\pi/2) = 0. \tag{13}$$

The exact solution of Eqs. (12)-(13) is

$$u(x) = \sin^2 x - 2\cos x + e^{\cos x} - 2. \tag{14}$$

The proposed OMGP method with 11 discretisation points (resulting in 11 data points for the GP) is applied to numerical solution of Eqs. (12)-(13). The results are presented in Fig. 1a. Comparison of the obtained numerical solution with the exact solution Eq. (14) in the interval of interest $x \in (0.5\pi, 2.5\pi)$ results in a mean absolute error of 3.8479E-006.

4.2 Second Order ODE with Variable Coefficients

The problem of a massive rotating rod can be described by the following differential equation [3]:

$$\frac{d^2u}{dr^2} + \frac{1}{r}\frac{du}{dr} - \frac{u}{r^2} = -\frac{\rho\omega^2}{\lambda + 2\mu} r \quad \text{in the region} \quad r \in (0,\ R) \tag{15}$$

and the following boundary conditions:

$$u(0) = 0, \ \ \sigma_{rr}(R) = 0. \tag{16}$$

$\sigma_{rr}(r)$ in Eq. (16) is the elastic radial stress which can be expressed by the following equation:

$$\sigma_{rr}(r) = \lambda\left(\frac{u}{r} + \frac{du}{dr}\right) + 2\mu\frac{du}{dr}. \tag{17}$$

The exact solution of Eqs. (15)-(17) is

$$u(r) = \frac{1}{8}\frac{\rho\omega^2}{\lambda + 2\mu} r\left(R^2\frac{2\lambda + 3\mu}{\lambda + \mu} - r^2\right). \tag{18}$$

For numerical computations it is assumed that $R = 1$, $\rho\omega^2/(\lambda + 2\mu) = 1$ and $\lambda = 2\mu$. Eq. (15) is discretized by seven equidistant points located in the region $r \in (0,\ R)$. The resulted numerical solution is presented in Fig. 1b and has a mean absolute error of 9.663E-11.

4.3 Second Order Elliptic PDE

Consider the Laplace partial differential equation representing the state of equilibrium in solids

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 1 \quad \text{in the region} \quad x^2 + y^2 < 1 \tag{19}$$

with the boundary condition

$$u\,(x,y) = 0 \quad \text{on the circle} \quad x^2 + y^2 = 1. \tag{20}$$

The exact solution of Eqs. (19)-(20) is

$$u\,(x,y) = 0.25\left(x^2 + y^2 - 1\right). \tag{21}$$

For discretization of Eq. (19) nine points are chosen on a regular grid inside the region of the unit circle $x^2 + y^2 < 1$. Additionally, four points are chosen on the circle $x^2 + y^2 = 1$ to discretize the boundary condition Eq. (20). This results into a dataset with just 13 points for multi-task GP learning and inference. Fig. 1c demonstrates the resulting numerical solution for Eqs. (19)-(20) which is visually indistinguishable from the plot of the exact solution Eq. (21). The mean absolute error of the numerical solution is equal to 1.5251E-011.

4.4 Second Order Hyperbolic PDE

Consider the following partial differential equation representing non-stationary wave propagation in elastic media

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{4}\frac{\partial^2 u}{\partial t^2} = 0 \quad \text{in the region} \quad x \in (0,\, \pi)\,, \tag{22}$$

together with the boundary conditions

$$u\,(0,\, t) = u\,(\pi,\, t) = 0 \quad \text{on the interval} \;\; t \in (0, \pi) \tag{23}$$

and the initial conditions

$$u\,(x,\, 0) = 0.1\text{sin}^3 x,\; \frac{\partial u}{\partial t}\,(x,\, 0) = 0 \quad \text{on the interval} \;\; x \in (0,\, \pi)\,. \tag{24}$$

The exact solution of Eqs. (22)- (24) is

$$u\,(x,\, t) = (3\cos{(2t)}\,\sin x - \cos{(6t)}\,\sin{(3x)})/40. \tag{25}$$

The region $x \in (0,\, \pi)$, $t \in (0, \pi)$ of Eq. (22) is discretized via a 15x15 regular grid of points. Each of the intervals of the boundary conditions Eq. (23) and the initial conditions Eq. (24) is discretized using 10 points. Comparison with the exact solution Eq. (29) shows that for this case the mean absolute error of the numerical solution is 1.587E-6.

5 Conclusions

Operator Induced Multi-Task Gaussian Processes (OMGPs) are introduced and a method is developed allowing application of OMGPs to numerical solution of differential equations. It is demonstrated that using the proposed approach both ordinary and partial differential equations with boundary and initial conditions can be solved with high accuracy.

Results of numerical evaluations for different classes of differential equations are presented. It is demonstrated that even with very small number of discretization points (i.e vary small datasets for the OMGP) the proposed methodology can lead to numerical solutions almost identical to the corresponding exact analytical ones.

Acknowledgments

This work has been supported by the Rio Tinto Centre for Mine Automation and the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

References

[1] Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press (2006)

[2] Bonilla, E.V., Chai, K.M., and Williams, C.K.I.: Multi-task gaussian process prediction. In J. C. Platt, D. Koller, Y. Singer, and S. Roweis, editors, NIPS, pages 153–160. MIT Press (2009)

[3] Teodor, M.A., Ardeshir, G.: Theory of Elasticity for Scientists and Engineers. Birkhauser Boston (2000)