

AE8803: Machine Learning with Gaussian Processes



Proposal



Proposal

Instructions

This may be done as a group project or individually. The size of the group must not exceed two. You need to formulate a problem for which a Gaussian process model is a reasonable modelling choice. As such your problem can take the form of regression, classification, design of experiment, or even unsupervised learning. You are permitted to use any open-source python platform for your chosen problem, however, you may also choose to code up your implementation from scratch.

The data that underscores your problem should be open-source; you are not permitted to work on problems of a proprietary nature. It is completely acceptable to have a problem that is aligned with your research interests, although your work will be judged by what you do during this project and not prior research.

A standard Gaussian process workflow assumes a prior based on a 1D RBF or Matern or periodic kernel. It also assumes that both prior and likelihood are Gaussian. You must articulate how your problem and the solution you are proposing requires either (i) a richer likelihood function, or (ii) a more complex covariance function.

Proposal

The first part of this assignment is a proposal. The proposal must address the following questions:

- What is input / output sought? [5]
- What type of kernels will be considered? [5]
- What are 1-2 relevant papers? [5]
- What is particularly novel about the implementation (prior, likelihood, kernel, mean function, etc.)? [3]
- What is the data source being used? [2]

Kindly ensure that your proposal has a list of references. The main body of the proposal (without references) should not exceed one page. Once submitted, you will receive feedback on your proposal

which will help with your project.

Due: Monday 8th April 15:30 on Canvas

Possible ideas

Below you will find a non-exhaustive list of ideas that would satisfy the points above:

1. Design a Gaussian process model for a 2D/3D velocity field based on sparse velocity measurements. Describe and code how you would make the covariance functions (a) divergence free and (b) curl free. For training and testing data you may consider analytical equations such as the Taylor Green vortex system.
2. Design a Gaussian process model for sensor data that exhibits some symmetry that is captured in the construction of the covariance function. If you plan to work on the JetBot project, then your proposal should focus on velocity magnitude measurements taken at several locations within a rectangular area. All measurements are at the same height and the objective is to use as few measurements as possible to estimate the velocity magnitude field.
3. Design a Gaussian process model that is constructed across a sphere, or more generally, over a space that is not a hypercube. If using polar or spherical coordinates ensure relevant $[0, 2\pi]$ periodicity. One example would be modelling temperature or pressure across the globe using sparse measurements from ECMWF or NOAA.
4. Linear operators acting on a Gaussian process yield another Gaussian process. Build a Gaussian process that leverages this for both derivative and integral operations, i.e., you observe both integral (e.g., location) and derivative (e.g., acceleration) data for a Gaussian process model (e.g., velocity).
5. Any Gaussian process model with multiple output quantities, where one needs to account for correlations between the outputs.
6. A Gaussian process model that forecasts the latitude, longitude and altitude of flights departing a given airport. A project like this may leverage historical flight trajectory data found in websites such as Flightradar24.
7. Gaussian process models where one also wants to reduce the dimensionality associated with the inputs; this can take the form of a dimension reducing subspace or even a deep neural network kernel.