Imperial College London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Gaussian Processes for Optimal Sensor Position

Background & Progress Report

Author:
Adrian LÖWENSTEIN

Supervisor:
Rossella ARCUCCI
Miguel MOLINA-SOLANA

Submitted in partial fulfillment of the requirements for the MSc degree in Computing (Machine Learning) of Imperial College London

June 2019

Contents

Chapter 1

Introduction

1.1 Summary

Gaussian processes (GP) have been widely used since the 1970's in the fields of geostatistics and meteorology. Current applications are in diverse fields including sensor placement In this project, we propose the employment of a GP model to calculate the optimal spatial positioning of sensors to study and collect air pollution data in big cities. We will then validate the results by means of a data assimilation software with the data at the proposed positions.

1.2 Data

London South Bank University (LSBU) air pollution data (velocity, tracer)

Chapter 2

Background

In this chapter we will cover the literature and the theory that will be used throughout the project. First we will review the context of the project and how it fits into the **MAGIC** project. Then our focus goes to the definition of **Gaussian Processes** (GP) and how they are used in the context of geospatial data. Furthermore, the use of GP relies heavily on **Covariance** matrixes which needs to be estimated. Those tools enable us to create **optimisation** algorithms for the position of sensors. Finally we will quickly explore the concepts of **Data Assimilation** (DA) that will be used to validate the results of the optimisation.

2.1 The MAGIC Project

This work is done in the context of the **Managing Air for Green Inner Cities** project. This is a multidisciplinary project and has for objective to find solutions to the pollution and heating phenomenons in cities. Traditionally, urban environmental control relies on polluting and energy consuming heating, ventilation and cooling (HVAC) systems. The usage of the systems increases the heat and the pollution levels, inducing an increased need for the HVAC. The MAGIC project aims at breaking this vicious circle and has for objective to provide tools to make possible the design of cities acting as a natural HVAC system.

This has been extensively discussed by Song et al. (2018). For this purpose, integrated management and decision-support system is under development. It includes a variety of simulations for pollutants and temperature at different scales; a set of mathematical tools to allow fast computation in the context of real-time analysis; and cost-benefit models to asses the viability of the planning options and decisions.

As explained by Song et al. (2018), the test site which has been selected to conduct the study is a real urban area located in London South Bank University (LSBU) in Elephant and Castle, London. In order to investigate the effect of ventilation on the cities problem, researchers in the MAGIC project have created simulations and experiments both in outdoor and indoor conditions, on the test site. They used wind tunnel experiments and computational fluid dynamics (CFD) to simulate the outdoor environment. Further works include the development of reduced-order modelling (ROM) in order to make faster the simulations while keeping a high level of accuracy (Arcucci et al., 2018).

Another key research direction in the use Data Assimilation (DA) and more specifically Variational DA (VarDA) for assimilating measured data in real time and allowing better prediction of the model in the near future (Arcucci et al., 2018). The further use of those method would be the optimisation of the position of the sensors which provide information for the VarDA.

2.2 Gaussian Process and Sensor Optimisation

In this chapter we will review Gaussian Processes (GP) which are probabilistic models for spatial predictions based on observations assumption.

As explained by Rasmussen and Williams (2006, p. 29), the history of Gaussian Processes goes back at least as far as the 1940s. A lot of usages were developed in various fields. Notably for predictions in spatial statistics (Cressie, 1991). Applied in particular in Geostatistics with methods known as *kringing*, and in Meteorology. Gradually GP started to be be used in more general cases for regression. Nowadays it is used in the context of Machine Learning.

A for sensor optimisation, we will follow the approach that was developed by Krause et al. (2008). This method relies on GP for finding a near-optimal solution to the problem of placing sensors.

2.2.1 Sensor Data Modelling

2.2.1.1 Multivariate Gaussian Distribution

GP will serve as a basis tool in our project. In the space we are monitoring we have a certain number of sensors measuring a certain quantity, such as temperature, pressure, speed of the wind or the concentration of a pollutant at a given position. We assume that the measured quantity has a *multivariate Gaussian joint distribution* between each point of the space. The associated random variable is $\mathcal{X}_{\mathcal{V}}$ for the set of locations $\mathcal{V}$ we would have the following distribution : $P(\mathcal{X}_{\mathcal{V}} = \boldsymbol{x}_{\mathcal{V}}) \sim \mathcal{N}(\mu_{\mathcal{V}}, \Sigma_{\mathcal{V}\mathcal{V}})$, or explicitly :

$$P(\mathcal{X}_{\mathcal{V}} = \boldsymbol{x}_{\mathcal{V}}) \frac{1}{(2\pi)^{n/2}|\Sigma_{\mathcal{V}\mathcal{V}}|} \exp^{-\frac{1}{2}(\boldsymbol{x}_{\mathcal{V}}-\mu_{\mathcal{V}})^T \Sigma_{\mathcal{V}\mathcal{V}}^{-1}(\boldsymbol{x}_{\mathcal{V}}-\mu_{\mathcal{V}})} \tag{2.1}$$

2.2.1.2 Prediction with Gaussian Processes

Let us still consider that we have the set of locations $\mathcal{V}$ and a set of sensors $\mathcal{A}$. In order to predict the quantity at positions were we have no sensors ($\mathcal{V} \backslash \mathcal{A}$) we can use a Gaussian Process. This GP is associated with a **mean function** $\mathcal{M}(\cdot)$ and a symmetric positive-definite **kernel function** $\mathcal{K}(\cdot,\cdot)$. We will denote the mean function values for a set of positions $\mathcal{A}$ by $\mu_{\mathcal{A}}$ and the kernel function values, or covariance matrix, between those points by $\Sigma_{\mathcal{A}}$. More detailed definitions are available in Rasmussen and Williams (2006, p. 13-16).

For a set of observations $\boldsymbol{x}_{\mathcal{A}}$ at positions $\mathcal{A}$ we can express for a finite set of other positions $\mathcal{V}\backslash\mathcal{A}$ the conditional distribution of those values. This means that we are able, for each point $y \in \mathcal{V}\backslash\mathcal{A}$, to predict the mean and the variance of $\boldsymbol{x}_y$. Using conditional distribution for the Multivariate Gaussian Distribution (Deisenroth et al., 2018, p. 193), we are able to express the following :

$$P(\mathcal{X}_y|\boldsymbol{x}_{\mathcal{A}}) = \mathcal{N}(\mu_{y|\mathcal{A}}, \Sigma_{y|\mathcal{A}}) \tag{2.2}$$
$$\mu_{y|\mathcal{A}} = \mu_y + \Sigma_{y\mathcal{A}}\Sigma_{\mathcal{A}\mathcal{A}}^{-1}(y - \mu_y) \tag{2.3}$$
$$\Sigma_{y|\mathcal{A}} = \Sigma_{yy} - \Sigma_{y\mathcal{A}}\Sigma_{\mathcal{A}\mathcal{A}}^{-1}\Sigma_{\mathcal{A}y} \tag{2.4}$$

An important point to notice is that the predicted covariance for the point y is not dependent of the values measured at $\mathcal{A}$, this is really useful because if allows us to define the uncertainty at y without using actual measurements.

For the rest of this section we assume that we have at our disposal a good estimate a of the covariance matrix between each point. In practise this is not that obvious as we will see in section 2.3. The following is valid for any covariance matrix that is symmetric and positive-definite.

2.2.2 Sensor Position Optimisation

Now that we have modelled the relationship between the positions with and without sensors we can establish an algorithm that was developed by Krause et al. (2008). The process of placing sensors in an optimal way is called in spatial statistics, *sampling* or *experimental design*

We want to place a number of k sensors positions indexed by $\mathcal{V}$

2.2.2.1 Placement Criterion

Entropy Criterion

Mutual Information Criterion

2.2.2.2 Algorithm

2.2.2.3 Improvements over the Algorithm

Address issue of matrix inversion with algorithm from (Rasmussen and Williams, 2006, p. 19)

2.2.2.4 Near-Optimality of the Algorithm

2.3 Covariance Estimation

We have seen how GP could be used for the optimisation of the position of sensors. In order to have good results we need to have a good estimate of the kernel function between the points of our space.

2.3.1 Stationnary Covariance

Isotropic kernels

2.3.2 Covariance Estimation

Sample variance = bad estimator in high dimensions

Chapter 3

Progress

This chapter su

3.1 Data Exploration

3.2 Covariance Estimation

3.3 Sensor Optimisation

Cressie (1991) Arcucci et al. (2018)

Chapter 4

Further Developements

Bibliography

Arcucci, R., Pain, C., and Guo, Y.-K. (2018). Effective variational data assimilation in air-pollution prediction. *Big Data Mining and Analytics*, 1(4):297–307. pages 3, 6

Cressie, N. A. C. (1991). *Statistics for spatial data*. Wiley series in probability and mathematical statistics. Wiley, New York. pages 3, 6

Deisenroth, M. P., Faisal, A., and Ong, C. S. (2018). Mathematics for Machine Learning. page 43. pages 4

Krause, A., Singh, A., and Guestrin, C. (2008). Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. page 50. pages 3, 4

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass. OCLC: ocm61285753. pages 3, 4

Song, J., Fan, S., Lin, W., Mottet, L., Woodward, H., Davies Wykes, M., Arcucci, R., Xiao, D., Debay, J.-E., ApSimon, H., Aristodemou, E., Birch, D., Carpentieri, M., Fang, F., Herzog, M., Hunt, G. R., Jones, R. L., Pain, C., Pavlidis, D., Robins, A. G., Short, C. A., and Linden, P. F. (2018). Natural ventilation in cities: the implications of fluid mechanics. *Building Research & Information*, 46(8):809–828. pages 2