

03

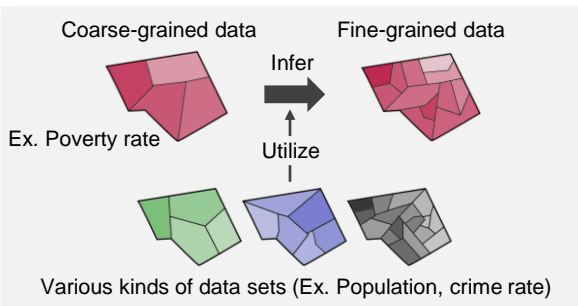
Refinement of spatially aggregated data

Multivariate Gaussian processes for spatially aggregated data

Abstract

Spatial data collected from cities are often aggregated into geographical partitions (e.g., districts). We propose a probabilistic model for **refining coarse-grained aggregated data** by utilizing multiple aggregated data sets with various granularities. Our model is based on multivariate Gaussian processes (GPs), in which dependences between data sets are established by linearly mixing some independent latent GPs. We newly **introduce an observation model with spatial aggregation processes**, which allows us to use multiple aggregated data sets for the refinement task even if they have various granularities. Our model can be used for predicting data values with arbitrary fine granularity; it is **useful for finding key pin-point regions (e.g., poverty area) in a city**, efficiently. In the future, we will extend the model to handle data gathered from multiple cities simultaneously.

Problem : refinement of spatially aggregated data



◆ Task

Predicting fine-grained data by utilizing multiple aggregated data sets with various granularities.

◆ Idea

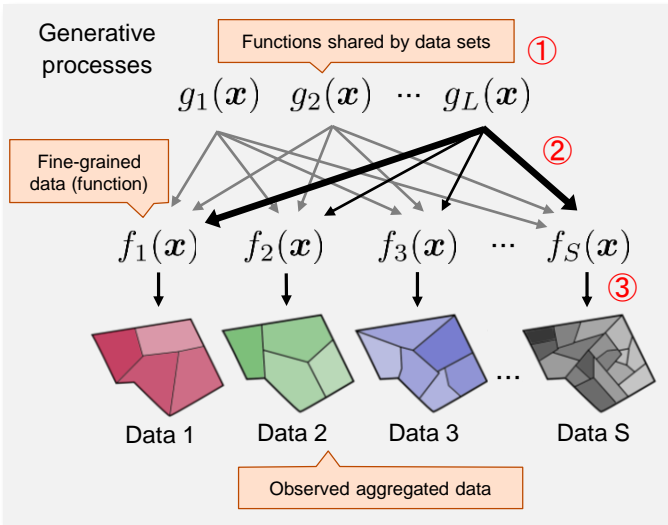
Interpolating coarse-grained data by using fine-grained data that have spatial correlation similar to target data.

◆ Difficulty

It is not straightforward to evaluate the similarity between aggregated data sets whose granularity is different.

Proposal : spatially aggregated Gaussian processes

We design generative processes of multiple aggregated data sets and train the model from observation data.



◆ Point ① : Spatial interpolation

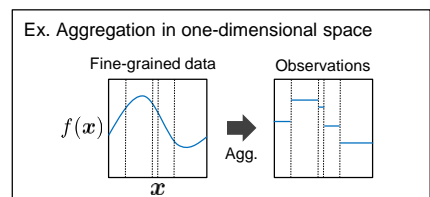
Assume the underlying smooth functions (i.e., GPs).

◆ Point ② : Dependences between data sets

Share the functions $g_l(x)$ among multiple data sets on the basis of the similarity of spatial patterns.

◆ Point ③ : Spatial aggregation processes

Average fine-grained data values over regions



References

- [1] Y. Tanaka, T. Tanaka, T. Iwata, T. Kurashima, M. Okawa, Y. Akagi, H. Toda, "Spatially aggregated Gaussian processes with multivariate areal outputs," *Proc. 33rd Conference on Neural Information Processing Systems (NeurIPS)*, pp. 3000-3010, 2019.

Contact

Yusuke Tanaka Email: cs-openhouse-ml@hco.ntt.co.jp
Service Evolution Laboratory

