Latent Gaussian Process

Friday, March 15, 2024 12:19 PM

https://onlinelibrary.wiley.com/doi/10.1002/nme.5751
Leveraging the nugget parameter for efficient Gaussian process modeling

Latent map Gaussian processes for mixed variable metamodeling

https://www.sciencedirect.com/science/article/pii/S004578252100459X

Only has multiple input but single output

Recently, Zhang et al. [32] developed a novel method that projects each qualitative input to a distinct continuous latent space which allows to directly use the corresponding latent variables in the GP's covariance function. This method builds latent variable Gaussian processes (LVGPs) and has been shown to consistently outperform prior methods. From the standpoint of converting qualitative inputs to quantitative ones, LVGPs are similar to the second approach discussed above. However, instead of relying on domain knowledge and manual conversion, LVGPs employ the training data and maximum likelihood estimation (MLE) and hence are much more versatile, efficient, and generally more accurate.

https://bpb-us-e2.wpmucdn.com/faculty.sites.uci.edu/dist/1/863/files/2023/01/LMGP-Paper.pdf

Mapping categorical variables into a latent space has a strong justification because in all physical systems with such inputs, there exist some underlying numerical features that characterize the levels of each categorical variable.

How come no iteration through l or l'

3.1. Unrestrictive covariance (UC)

One popular strategy for GP modeling with categorical variables, introduced by Qian et al. [46], assumes a correlation function with the following form:

$$r(\boldsymbol{w}, \boldsymbol{w}') = \prod_{i=1}^{d} \tau_{l_i l_i'}^{i} \times exp\left\{-(\boldsymbol{x} - \boldsymbol{x}')^T 10^{\Omega} (\boldsymbol{x} - \boldsymbol{x}')\right\}, \quad (9)$$

[1806.07504] A Latent Variable Approach to Gaussian Process Modeling with Qualitative and Quantitative Factors (arxiv.org)
LVGP: Latent Variable Gaussian Process Modeling with Qualitative and Quantitative Input Variables version 2.1.5 from CRAN (rdrr.io)

https://www.nature.com/articles/s41598-020-60652-9 (the first author is the same)
We first present a solar cell design problem, in which both the light scattering structure pattern and material selection are optimized simultaneously, in contrast to the original design problem that only considered tuning the scattering pattern represented by quantitative variables2,33.

The source code of this work will be made available upon request to the corresponding author.

https://asmedigitalcollection.asme.org/mechanicaldesign/article/145/3/031701/1147205/Descriptor-Aided-Bayesian-Optimization-for-Many
The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

!!!Gaussian Process Latent Variable Model — Pyro Tutorials 1.9.0 documentation

Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data (neurips.cc) (2003)

[2202.12979] Generalised Gaussian Process Latent Variable Models (GPLVM) with Stochastic Variational Inference (arxiv.org) (2022)

Gaussian Process Latent Variable Models (GPLVM) with SVI — GPyTorch 0.1.dev97+gf73fa7d documentation (Gpytorch)

Gaussian Process Latent Variable Model — Pyro Tutorials 1.9.0 documentation

Andy Jones (andrewcharlesjones.github.io)

Gaussian process latent variable models

Bayesian Gaussian process latent variable model (Bayesian GPLVM) — GPflow 2.4.0 documentation(Gpflow)

2011.03715.pdf (arxiv.org)
Bayesian Nonparametric Dimensionality Reduction of Categorical Data for Predicting Severity of COVID-19 in Pregnant Women (no code implementation)
Bayesian Nonparametric Dimensionality Reduction of Categorical Data for Predicting Severity of COVID-19 in Pregnant Women - PMC (nih.gov)

From Probabilistic PCA to the GPLVM (gregorygundersen.com)

Gaussian process latent variable model factorization for context-aware recommender systems - ScienceDirect
Our method can flexibly deal with both real-valued and categorical contexts.

A review on Gaussian Process Latent Variable Models - ScienceDirect

5.1. GPLVM for discrete variables

The conventional GPLVM and its extensions are suitable for analysis of continuous data. However, in many machine learning task such as natural language processing and medical diagnosis, users often get discrete variables. Although [78], has proposed a GPLVM for the estimation of multivariate categorical data, its inference is based on the variational approximation and sampling approaches which have a high computational complexity. To overcome the difficulty, we mainly should consider two factors: the construction of likelihood and the inference method of latent variables, which deserve more attentions.

!!!Latent Gaussian Processes for Distribution Estimation of Multivariate Categorical Data (arxiv.org)
GitHub - yaringal/CLGP: Categorical Latent Gaussian Process

https://www.scopus.com/record/display.uri?origin=citedby&eid=2-s2.0-85185007529&noHighlight=false&relpos=1

https://proceedings.neurips.cc/paper_files/paper/2018/file/165a59f7cf3b5c4396ba65953d679f17-Paper.pdf
In this paper, we are interested in the heterogeneous case for which the outputs are a mix of continuous, categorical, binary or discrete variables with different likelihood functions.

1806.07504.pdf (arxiv.org)
A Latent Variable Approach to Gaussian Process Modeling with Qualitative and Quantitative Factors
Yichi Zhang1 , Siyu Tao1 , Wei Chen1 , and Daniel W. Apley

!!!This is a Python implementation of the Latent Variable Gaussian Process (LVGP) model Zhang et al. (https://arxiv.org/abs/1806.07504). Their R package is available at https://cran.r-project.org/web/packages/LVGP/index.html.
GitHub - balaranjan/LVGP

Latent map Gaussian processes for mixed variable metamodeling

Nicholas Oune, Ramin Bostanabad*
https://bpb-us-e2.wpmucdn.com/faculty.sites.uci.edu/dist/1/863/files/2023/01/LMGP-Paper.pdf

https://github.com/Bostanabad-Research-Group/GP-Plus/blob/main/Examples/05_Example_1Dsin.ipynb