Gaussian process regression models under linear inequality conditions

A. F. López-Lopera
Mines Saint-Étienne (EMSE), France.

Supervisor(s): O. Roustant (EMSE, France), F. Bachoc (Institut de Mathématiques de Toulouse, France) and N. Durrande (EMSE, France; PROWLER.io, Cambridge, UK)

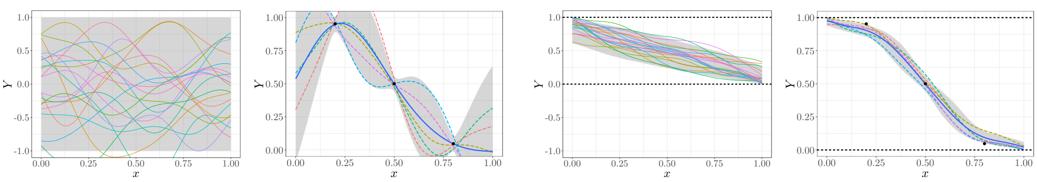
Ph.D. expected duration: Oct. 2016 - Sep. 2019

Adress: Institut Henri FAYOL – 158 cours Fauriel F-42023 Saint-Étienne, France.

Email: andres-felipe.lopez@emse.fr

Abstract:

Taking into account inequality constraints (e.g. boundedness, monotonicity, convexity) into Gaussian process (GP) models can lead to more realistic predictions guided by the physics of data [6, 4]. Figure 1 compares two models that either ignore or take into account both boundedness (i.e. $0 \leq y(x) \leq 1$, for $x \in [0, 1]$) and monotonicity constraints (i.e. $y(x) \geq y(x')$, if $x \geq x'$).



(a) unconstrained GP (b) constrained GP

Figure 1: GP regression models under (a) no constraints and (b) boundedness and monotonicity constraints. Each panel shows: (left) samples from the different types of Gaussian priors, and (right) the resulting GP regression model conditioned on three observations (dots).

We aim at investigating a GP framework that can account for inequality constraints. Our main contributions are threefold.

First, building on the approach proposed in [6], we introduced in [4] a full Gaussian-based framework to satisfy a set of linear inequality constraints. The benefit of using the finite-dimensional representation of [6] leads to satisfy the inequalities everywhere in the input space. Furthermore, it was proved in [2] that the resulting posterior mode is the optimal constrained interpolation function in the reproducing kernel Hilbert space. Due to the truncated Gaussianity of the posterior, its distribution can be approximated via Monte Carlo or Markov chain Monte Carlo. We investigated several samplers in examples on both synthetic and real-world data, under different types of constraints. We found that the Hamiltonian Monte Carlo (HMC)-based sampler from [7] achieves the best trade-off between running time and effective sample rates.

Despite the promising results in [4], our experiments were limited up to 2D problems due to the tensor structure of our framework. This brings us to our second contribution, where various alternatives have been explored for going to higher dimensions and for a high number of observations. In the first direction, we introduced noise for the relaxation of the interpolation constraints. This also relaxed the constraints of the HMC sampler improving its efficiency. As a result, we were now able to use our framework in 5D spaces [3]. Moreover, since the computational complexity here depends on the number of basis functions rather than the observations, we can

handled thousands of observations. In a second direction, we considered specific assumptions on the target function that are suitable in high dimensions. In particular, we adapted our framework to additive functions, where sampling from the posterior distribution in high dimensions can be achieved through sampling in lower dimensional spaces (e.g. 1D spaces by assuming first-order additivity) [5]. Figure 2 plots a 5D example for a first-order additive target function satisfying different types of inequality constraints per dimension.

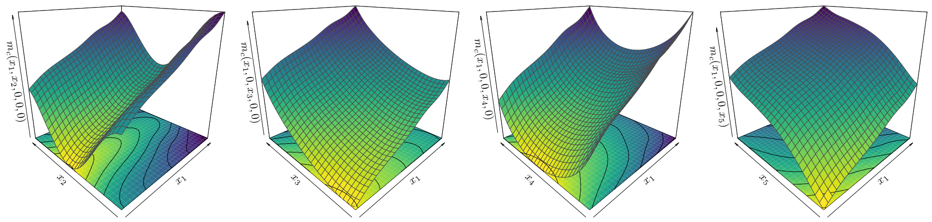


Figure 2: Additive GP regression model for $\mathbf{x} \mapsto 2x_1 + \cos(6x_2) + 2x_3^2 + 4(x_4 - 0.5)^2 + 2\arctan(2x_5)$ for $\mathbf{x} \in [0,1]^5$. The constrained predictive mean is shown satisfying: monotonicity constraints across the first and fifth dimensions, and convexity constraints across the third and fourth dimensions. No constraints were imposed across the second dimension.

Third, we considered the problem of estimating the covariance parameter under inequality constraints. We studied the properties of both unconstrained and constrained maximum likelihood (ML) estimators. Under fixed-domain asymptotics, we showed that, loosely speaking, any consistency result for the (unconstrained) ML is preserved for the constrained ML when adding boundedness, monotonicity and convexity conditions [4]. We also showed that the constrained ML estimator (cMLE), conditionally to the fact that the GP satisfies those constraints, has the same asymptotic distribution as the unconditional MLE [1].

References

[1] F. Bachoc, A. Lagnoux, and A. F. López-Lopera. Maximum likelihood estimation for Gaussian processes under inequality constraints. *ArXiv e-prints*, April 2018.

[2] X. Bay, L. Grammont, and H. Maatouk. Generalization of the Kimeldorf-Wahba correspondence for constrained interpolation. *Electron. J. Stat.*, 10(1):1580–1595, 2016.

[3] A. F. López-Lopera, F. Bachoc, N. Durrande, and O. Roustant. Approximating Gaussian process emulators with inequality constraints and noisy observations (submitted). 2018.

[4] A. F. López-Lopera, F. Bachoc, N. Durrande, and O. Roustant. Finite-dimensional Gaussian approximation with linear inequality constraints. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1224–1255, 2018.

[5] A. F. López-Lopera, N. Durrande, F. Bachoc, and O. Roustant. Additive Gaussian processes under inequality constraints (in preparation). 2019.

[6] H. Maatouk and X. Bay. Gaussian process emulators for computer experiments with inequality constraints. *Mathematical Geosciences*, 49(5):557–582, 2017.

[7] Ari Pakman and Liam Paninski. Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians. *Journal of Computational and Graphical Statistics*, 23(2):518–542, 2014.

Short biography – A. F. López-Lopera received the BEng and MEng degrees in electrical engineering from the *Universidad Tecnológica de Pereira*, Colombia. Currently, he is a PhD student in applied mathematics at EMSE, France. In the thesis, *Metamodelling under inequality constraints*, GP-based models are explored in two main directions: the estimation under inequality constraints, and the extension to higher dimensions. His PhD is funded by the chair in applied mathematics OQUAIDO.