

On Gaussian Process multiple-fold cross-validation

In this talk I will give an overview of some recent results pertaining to the fast calculation of Gaussian Process multiple-fold cross-validation residuals and their covariances, as well as to GP Parameter Estimation via related approaches. At first, the focus will be put on results from (arXiv:2101.03108 [stat.ME], joint work with Cedric Schärer), where fast Gaussian process leave-one-out formulae are generalized to multiple-fold cross-validation. We will illustrate how resulting covariances affect model diagnostics and also how accounting for dependency between such residuals affect cross-validation-based estimation of the scale parameter. Numerical experiments will highlight the accuracy of our approach as well as the speed-ups that it enables. Next (joint work with Athénaïs Gautier and Cédric Travelletti), we will build on the latter generalizations to further study the impact of data partitioning on covariance hyperparameter fitting. In turn, we will present an application of the considered formulae and criteria to a case where observations involve linear forms in the underlying GP, illustrating on an inverse problems from geosciences how the way to group observations in cross-validation can affect range estimation.