State Space Methods for Efficient Inference in Student-t Process Regression

Arno Solin and Simo Särkkä
Aalto University, Finland

Algorithm 1: Student-t filter.

\textbf{for} k = 1, 2, \ldots, n \textbf{do}
\begin{align*}
\text{Filter prediction:} \\
 m_{k|k-1} &= A_{k-1} m_{k-1} + \text{Cov}(k) \\
 v_{k|k-1} &= \text{vec}((k) - H(k)) \\
 &\text{where} f(t) = (f(t_1), f(t_2), \ldots, f(t_n)) \text{ holds the} \\
 &\text{stochastic processes, and} \text{Cov}(k) \text{ is a white noise} \\
 &\text{process with spectral density} \text{Q}_0 \text{ and initial state} \text{f}(0) = \mathbb{N}(0, \text{P}_0). \\
\end{align*}

\text{A TP can be constructed as a scale mixture of state space} \\
\text{form SDEs by setting the spectral density to} \text{Q}_0,\text{and} \text{using} \\
\text{the initial state} \text{f}(0) = \mathbb{N}(0, \text{P}_0). \text{ where} \gamma \text{ is an inverse} \\
\text{gamma random variable.} \text{The solution can be written out in closed-form} \\
\text{at the specified time points} t_k = 1, 2, \ldots \text{, as} f(t_k) = \text{L}_{k|k-1} \\
\text{such that} \text{f}_k = \mathbb{N}(0, \text{P}_k) \text{ and} \text{L}_{k|k-1} \text{,} \\
\text{where} \text{Q}_0 = \mathbb{N}(0, \text{Q}_0). \text{The} \\
\text{entangled noise model is included by} \\
\text{augmenting} \text{it into the state.}

Algorithm 2: Student-t smoother.

\begin{align*}
\text{Smooother prediction:} \\
 m_{k|n} &= A_{k|n} m_{k|k-1} + \text{Cov}(k) \\
 v_{k|n} &= \text{vec}((k) - H(k)) \\
 &\text{where} \text{Q}_0 = \mathbb{N}(0, \text{Q}_0). \text{The} \\
\end{align*}

\text{The filter gives the marginal likelihood for} \text{hyperparameter optimization.} \text{The} \\
\text{smoothing outcome corresponds to the naive TP} \text{regression result.}

\text{CONCLUSIONS}

\text{We have generalized the connection between} \text{Gaussian process regression and Kalman filtering to} \\
\text{more general statistical processes and non-Gaussian} \text{Bayesian filtering.} \text{This link enables the} \\
\text{use of efficient sequential inference methods to solve TP regression problems in} \text{C(n) time} \text{complexity.} \text{An} \\
\text{example implementation is available on the} \text{author web page:} \text{http://arno.solin.fi}

\text{REFERENCES}


Tracking of a Moving Vehicle

Interpolation of missing GPS observations by two dimensional GP regression (Gaussian smoothing) and TP regression (Student-t smoothing). The unknown ground truth is shown by dots and the colored patches illustrate the credible intervals up to 95%.

Stock Price Data

The log share price of Apple Inc. \(n = 8537\) modeled by GP/TP with a covariance function sum of a constant, linear, Matérn (smoothness 3/2), and exponential covariance function. The main difference comes from the different hyperparameters.

Computational Efficiency

Demonstration of the computational benefits of the state space model in solving a TP regression problem for a number of data points up to 10,000.