Explicit Link Between Periodic Covariance Functions and State Space Models

Arno Solin and Simo Särkkä

(AISTATS 2014)

Discussion by: Piyush Rai

December 12, 2014
Overview

- An explicit connection between GP regression with periodic covariance functions and state-space models
- Based on expanding the periodic covariance function into a series of stochastic resonators
- Allows scaling up GP regression with periodic covariance functions to large data sets
- Proposed method also extended to GPs quasi-periodic covariance functions
Gaussian Process Regression

The kernel view

Given: \( n \) training examples \( \{(t_k, y_k)\}, k = 1, \ldots, n \)

\[
y_k = f(t_k) + \epsilon_k
\]
\[
f(.) \sim \mathcal{GP}(0, k(t, t'))
\]
\[
\epsilon_k \sim \mathcal{N}(0, \sigma_n^2)
\]

Prior assumptions about \( f \) (e.g., smoothness, periodicity, etc.) encoded in the covariance function \( k(t, t') \)

Can be solved in closed form but naïve solution is expensive: \( \mathcal{O}(n^3) \) complexity at test time

The state-space view

Consider an \( m \)-th order SDE

\[
\frac{df(t)}{dt} = Ff(t) + Lw(t)
\]
\[
y_k = Hf(t_k) + \epsilon_k
\]

where \( f(t) \) contains derivatives of \( f(t) \) up to order \( m - 1 \) and \( w(t) \) is the white noise process with spectral density \( Q_c \)

Model defined by \( F, L, Q_c, \) stationary covariance \( P_\infty \), and the observation model \( H \)

Solved using Kalman filtering and has \( \mathcal{O}(nm^3) \) time-complexity
Covariance and Spectral Density

- Shown here: a stationary covariance function (Matérn) and its spectral density

\[ k(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \exp(-i\omega\tau) \, d\omega. \]

- This equivalence enables transforming the GP into a state-space model and solve the problem more efficiently in \( O(n) \) time
GP Regression (the naïve way)
GP Regression (the naïve way)
GP Regression (via filtering and smoothing)
GP Regression (via filtering and smoothing)
How to establish the GP vs state-space model equivalence when the GP covariance function is periodic?
Periodic Covariance Functions

Start off with the squared exponential:

\[ k(x, x') = \sigma^2 \exp\left( -\frac{\|x - x'\|^2}{2\ell^2} \right) \]

Polar coordinates:

\[ x(t) = \begin{pmatrix} \cos(\omega_0 t) \\ \sin(\omega_0 t) \end{pmatrix} \]

The *canonical* periodic covariance:

\[ k_p(t, t') = \sigma^2 \exp\left( -\frac{2\sin^2(\omega_0 \frac{t-t'}{2})}{\ell^2} \right) \]
Periodic Covariance Functions

- Not amenable to the state-space transformation which requires the spectral density to be approximated by rational functions
**State-Space Formulation**

- Fourier series representation

\[ k_p(\tau) = \sum_{j=1}^{\infty} q_j^2 \cos(j\omega_0\tau) \]

- Each periodic term \( j \) can be constructed as solution of a second-order ODE

\[
\begin{pmatrix}
\dot{x}_j(t) \\
\dot{y}_j(t)
\end{pmatrix} =
\begin{pmatrix}
0 & -\omega_0j \\
\omega_0j & 0
\end{pmatrix}
\begin{pmatrix}
x_j(t) \\
y_j(t)
\end{pmatrix}
\]

with \((x_j(0), y_j(0))^\top \sim \mathcal{N}(0, q_j^2 I)\)

- One way to determine the coefficients \( q_j^2 \) is via projection to the cosine basis

\[
q_j^2 = \omega_0 \int_{-\pi/\omega_0}^{+\pi/\omega_0} k_p(\tau) \cos(j\omega_0\tau) d\tau
\]

but there are other way too
State-Space Formulation

- The state-space model  \( \frac{df(t)}{dt} = Ff(t) + Lw(t) \) will have block-diagonal matrices \( F, L, \) and \( P_{\infty} \) with blocks being (for \( j = 1, \ldots, J \))

\[
F_j^p = \begin{pmatrix} 0 & -\omega_0 j \\ \omega_0 j & 0 \end{pmatrix}, \quad L_j^p = I_2, \quad P_{\infty, j}^p = q_j^2 I_2
\]

and the measurement model matrix \( H \) in \( y_k = Hf(t_k) + \epsilon_k \) is a block-row vector of \( H_j^p = (1 \ 0) \). The diffusion part is zero (deterministic model).

- The spectral (variance) coefficients

\[
q_j^2 = \frac{2l_j(l^{-2})}{\exp(l^{-2})}, \quad \text{for} \quad j = 1, 2, \ldots
\]

and \( q_0^2 = I_0(l^{-2})/\exp(l^{-2}) \) where \( I_\alpha(z) \) is the modified Bessel function

- Taking the first \( J \) term of the series gives an approximation and this approximation converges uniformly to the actual covariance as \( J \to \infty \).
Approximated Covariance Functions (J=0)

Covariance function

Spectral density
Approximated Covariance Functions (J=1)

Covariance function

Spectral density
Approximated Covariance Functions (J=2)

Covariance function

Spectral density
Approximated Covariance Functions (J=3)

Covariance function

Spectral density
The shape of the periodic effect may change with time

Modeled using a product of a truly periodic covariance function $k_p(t, t')$ and another covariance function $k_q(t, t')$ with long characteristic length-scale

$$k(t, t') = k_p(t, t')k_q(t, t')$$
Quasi-Periodic Covariance Functions

Covariance function

Spectral density
Quasi-Periodic Covariances: State-Space Form

- Have state-space representations for both quasi and periodic parts
- Set up the state-space such as the feedback matrices of both parts commute (i.e., $F_p F_q = F_q F_p$)
- Properties of Kronecker product help accomplish this
- The joint model for the quasi-periodic product of two covariance functions can be written in a block-form

\[
\begin{align*}
F_j &= F^q \otimes I_2 + I_q \otimes F^p_j, \\
L_j &= L^q \otimes L^p_j, \\
Q_{c,j} &= Q^q_c \otimes q^2_j I_2, \\
P_{\infty,j} &= P^q_{\infty} \otimes P^p_{\infty,j}, \\
H_j &= H^q \otimes H^p_j
\end{align*}
\]
Experiments: Computational Complexity

![Graph showing computational time versus number of data points for Full GP solution and State space solution.](image)
Experiments: CO₂ Concentration

- Observations are CO₂ concentrations across time \( (n = 2227) \)
- GP covariance function

\[
k(t, t') = k_{SE}(t, t') + k_p(t, t')k_{\nu=3/2}(t, t') + k_{\nu=3/2}(t, t')
\]

- Converted to state-space, hyperparams optimized w.r.t. marginal likelihood
Experiments: Daily Births between 1969-1988

- Observations are number of daily births between 1969-1988 ($n = 7305$)
- The model: Matern ($\nu = 7/2$) for the slow trend, Matern ($\nu = 3/2$) for faster variation, Quasi-periodic (yearly) with Matern ($\nu = 3/2$) damping, Quasi-periodic (weekly) with Matern ($\nu = 3/2$) damping
- Converted to state-space, hyperparams optimized w.r.t. marginal likelihood
Established connections between periodic covariance functions and state-space models

The connection allows using efficient sequential inference methods (from state-space modeling) to solve periodic GP regression problem

Approximation error due to truncation available in closed-form