PROBABILISTIC ESTIMATION OF CHIRP INSTANTANEOUS FREQUENCY USING GAUSSIAN PROCESSES

arXiv:2205.06306

Zheng Zhao

Uppsala Universitet

2 June 2022

Introduction

Chirp signals are defined by

$$\alpha(t)\sin\left(\phi_0 + 2\pi\int_0^t f(s)\,\mathrm{d}s\right), \tag{1}$$

where

- α: instantaneous amplitude.
- ϕ_0: initial phase.
- f: instantaneous frequency (IF)

Introduction

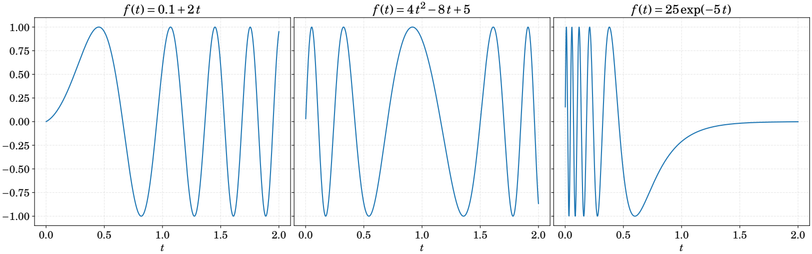


Figure: Some chirp examples with linear, quadratic, and geometric IFs.

Problem formulation

Suppose that we have chirp measurements $\{y_k\}_{k=1}^T$ from

$$\begin{aligned} Y_k &:= Y(t_k) = \alpha(t_k)\sin\Big(\phi_0 + 2\pi\int_0^{t_k} f(s)\,\mathrm{d}s\Big) + \xi_k, \\ \xi_k &\sim \mathrm{N}(0, \Xi). \end{aligned} \tag{2}$$

The goal is to estimate f from the data $y_{1:T} := \{y_k\}_{k=1}^T$.

Moreover, the amplitude α is assumed unknown.

Problem formulation

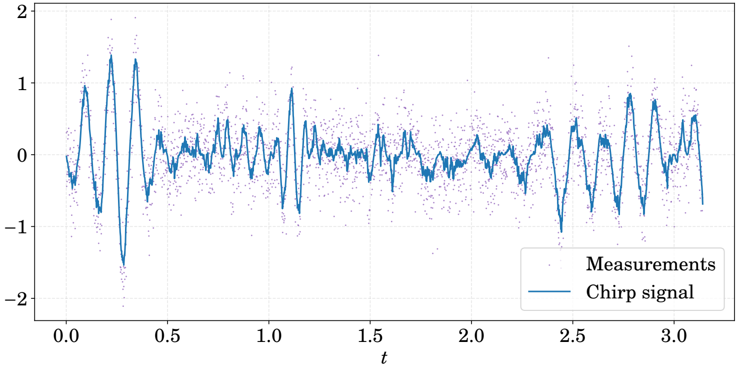


Figure: An example of chirp signals that we aim to dealing with. This chirp is contaminated by unknown random noises, making it hard to infer its frequency.

Solution in GP

Classical problem for > 60 years. A plethora of classical methods:

- Hilbert transform
- First-moment power spectrum
- Polynomial regression
- Adaptive notch filter
- ...

Since we are members of the Bayesian cult, we would like to put a prior on f then solve its posterior distribution.

Solution in GP

We could model the unknown f by a Gaussian process (GP), in the way that

$$\begin{aligned} f(t) &:= g(V(t)), \\ V(t) &\sim \mathrm{GP}\big(0, C_V(t,t')\big), \end{aligned} \tag{3}$$

where g is any positive bijection.

We select/design/hand-craft C_V based on the information (e.g., continuity) we know about the true IF.

For example, if the true IF is continuously differentiable, then we can choose C_V from that of Matérn $3/2$.

Solution in GP

GPs are function-valued random variables distributed according to the Gaussian measure.

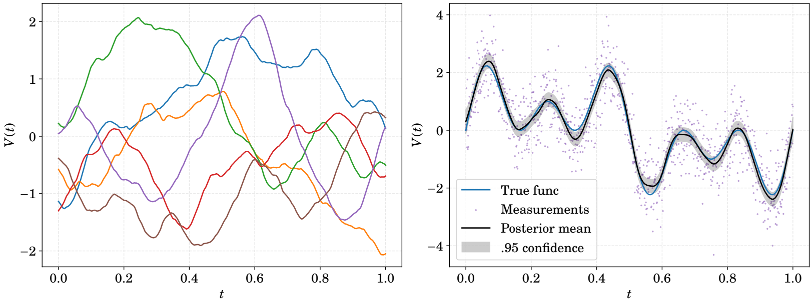


Figure: Left: six GP samples of V according to a Matérn 3/2 C_V. Right: a GP regression example under this C_V.

Solution in GP

Putting it altogether, we have the following GP regression model

$$
\begin{aligned}
V(t) &\sim \mathrm{GP}\big(0, C_V(t,t')\big), \\
Y_k &= \alpha(t_k) \sin\left(\phi_0 + 2\pi \int_0^{t_k} g(V(s))\,\mathrm{d}s\right) + \xi_k
\end{aligned}
\tag{4}
$$

from which we want to obtain

$$
p_{V(t) \mid Y_{1:T}}(v \mid y_{1:T}), \quad \forall t \in [0, \infty). \tag{5}
$$

But it's kinda hard ... not to mention that α is unknown! Seems dead end.

Is there any way to "eliminate" α, $\sin(\cdots)$, and $\int \cdots$ in the model?

Solution in GP

Yep, we could put a prior on the chirp signal too! For instance, find another GP

$$X(t) \sim \mathrm{GP}\big(m_X(t;f), C_X(t,t';f)\big), \tag{6}$$

such that the samples of it are valid chirps as per Equation (1).

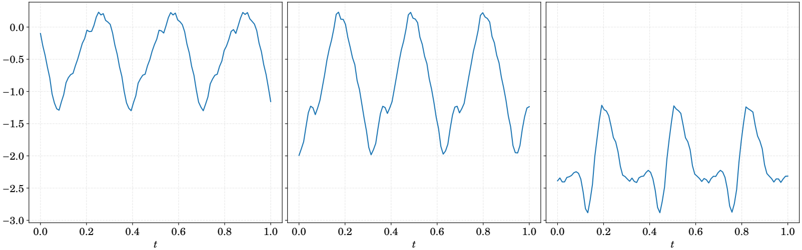
Question is how to find such a pair of m_X and C_X. Since chirps are periodic, m_X and C_X should have periodic structures as well.

Solution in GP

GP guys will immediately recall the periodic covariance function

$$C_X(t,t';f) = \sigma^2 \, \mathrm{e}^{-2\sin\left(\pi f\,(t-t')\right)^2/\ell^2}. \tag{7}$$

Samples from X using this cov function more or less look like this:



Looks okay-ish to model chirps, but f must be constant, or C_X is positive definite no more. Kinda no go.

Solution in GP

Suppose that we have found a valid and meaningful pair (m_X, C_X).

Then the "supposedly-working" GP regression model reads

$$\begin{aligned} V(t) &\sim \mathrm{GP}\big(0, C_V(t,t')\big), \\ X(t) \mid V(t) &\sim \mathrm{GP}\big(m_X(t;V), C_X(t,t';V)\big), \\ Y_k &= H X(t) + \xi_k. \end{aligned} \tag{8}$$

- $HX(t)$ stands for the chirp.
- X is parametrised by the IF $f(t) = g(V(t))$.

Solution in GP

But it's still not so easy to obtain $p_{V(t) \mid Y_{1:T}}(v \mid y_{1:T})$:

1. V and X are jointly not a GP. Have to use, e.g., MCMC, VI.
2. Computation is a problem! Need to solve many matrix inversions of dimension T, but signals are usually lengthy ...

Why not use sparse (pseudo-points) methods to approximate the full-rank covariance matrices? This is standard for solving large-scale GP regression problems.

Unfortunately, the sparse approximations have side effects in this chirp application, because they introduce down-samplings!

Solution in GP

Bear in mind: specifying the GP mean and cov functions is not the only way to construct a GP.

A number of alternatives, for instance,

- Stochastic (partial) differential equations.
- Precision matrices (i.e., GMRFs).
- Couplings
- ...

Solution in SDE

For this chirp application, we will use the SDE construction. Reasons being

- Signals are temporal.
- Markov property for cheap computation (linear in time) hence, no T-huge cov matrix inversion.
- No need to explicitly select/design/hand-craft the mean and cov functions.

All we need to do is to replace the GPs in Equation (8) with SDE-GPs.

Solution in SDE

The SDE-GP prior X for chirp signals we use is

$$\begin{aligned} \mathrm{d}X(t) &= \begin{bmatrix} -\lambda & -2\pi f \\ 2\pi f & -\lambda \end{bmatrix} X(t)\,\mathrm{d}t + b\,\mathrm{d}W_X(t), \\ X(0) &\sim \mathrm{N}\big(m_0^X, P_0^X\big), \end{aligned} \tag{9}$$

Recall that solutions to linear SDEs are GPs with implicitly defined mean and cov functions.

To understand that this X is a reasonable prior for chirp signals, let's take a look at its mean and cov functions.

Solution in SDE

X's mean m_X indeed carries (damped) chirps, in the sense that

$$\mathbb{E}[X(t)] = \begin{bmatrix} \alpha\, \mathrm{e}^{-\lambda t} \cos(\phi_0 + 2\pi f t) \\ \alpha\, \mathrm{e}^{-\lambda t} \sin(\phi_0 + 2\pi f t) \end{bmatrix} \tag{10}$$

which is the solution to a harmonic differential equation.

$$\mathrm{d}X(t) = \begin{bmatrix} -\lambda & -2\pi f \\ 2\pi f & -\lambda \end{bmatrix} X(t)\, \mathrm{d}t. \tag{11}$$

SOLUTION IN SDE

As for X's covariance function C_X, please refer to the paper to see its formulation. It plots like this:

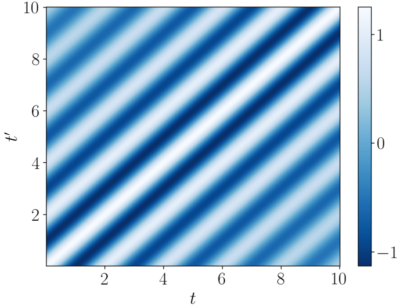


Figure: $C_X(t,t')$ evaluated at Cartesian $[0,10]\times[0,10]$. Parameters are $f = 0.5$ Hz, $\lambda = 0.1$, $b = 0.5$, and $P_0^X = 1.25 I_2$. It is clear to see the periodic structure, and the fading effect on the anti-diagonal due to the damping.

Solution in SDE

Recall our prior

$$\begin{aligned} \mathrm{d}X(t) &= \begin{bmatrix} -\lambda & -2\pi f \\ 2\pi f & -\lambda \end{bmatrix} X(t)\,\mathrm{d}t + b\,\mathrm{d}W_X(t), \\ X(0) &\sim \mathrm{N}\big(m_0^X, P_0^X\big), \end{aligned} \tag{12}$$

We can immediately replace the constant f with a time-varying one, for example, a GP! No need to worry about the positive definiteness of C_X.

Solution in SDE

What does the cov function C_X look like if we use a time-varying f?

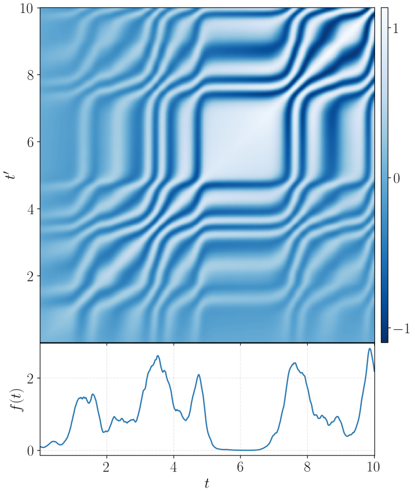


Figure: See that the periodicity changes over time driven by the value of f.

Solution in SDE

We then need to construct V as SDE-GP as well. For the sake of pedagogy, we use the Matérn 3/2 GP.

$$
\begin{aligned}
f(t) &= g(V(t)), \\
\mathrm{d}\begin{bmatrix} V(t) \\ \frac{\mathrm{d}V(t)}{\mathrm{d}t} \end{bmatrix} &= M \begin{bmatrix} V(t) \\ \frac{\mathrm{d}V(t)}{\mathrm{d}t} \end{bmatrix} \mathrm{d}t + L\,\mathrm{d}W_V(t), \\
V(0) &\sim \mathrm{N}\big(0, P_0^V\big)
\end{aligned}
\tag{13}
$$

where

$$
M = \begin{bmatrix} 0 & 1 \\ -3/\ell^2 & -2\sqrt{3}/\ell \end{bmatrix}, \quad L = \begin{bmatrix} 0 \\ 2\sigma(\sqrt{3}/\ell)^{3/2} \end{bmatrix}, \quad P_0^V = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 3\sigma^2/\ell^2 \end{bmatrix}.
$$

This is same as with saying

$$
V(t) \sim \mathrm{GP}\big(0, C_{\mathrm{Mat.3/2}}(t,t')\big). \tag{14}
$$

Solution in SDE

Define $U(t) \coloneqq \begin{bmatrix} X(t) & V(t) & \frac{\mathrm{d}V(t)}{\mathrm{d}t} \end{bmatrix}$. Putting all together, our chirp IF estimation model is

$$
\begin{aligned}
\mathrm{d}U(t) &= A\big(U(t)\big)\,\mathrm{d}t + B\,\mathrm{d}W(t), \\
U(0) &\sim p_{U(0)}(u), \\
Y_k &= H\,U(t_k) + \xi_k,
\end{aligned}
\tag{15}
$$

$$
A(U(t)) \coloneqq \begin{bmatrix} -\lambda & -2\pi g(V(t)) & 0 & 0 \\ 2\pi g(V(t)) & -\lambda & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -3/\ell^2 & -2\sqrt{3}/\ell \end{bmatrix} U(t)
$$

$$
B \coloneqq \begin{bmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2\sigma(\sqrt{3}/\ell)^{3/2} \end{bmatrix}, \quad H \coloneqq \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}.
$$

Solution in SDE

To see that this prior U is a suitable model for chirp signals, let's draw some samples from $H\,U$.

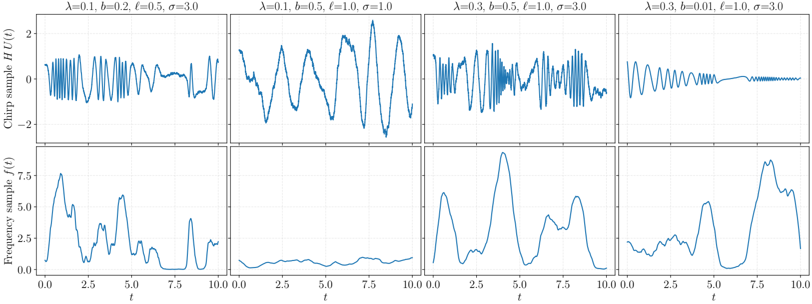


Figure: It can generate a rich variety of randomised chirp signals by tuning the model parameters.

Solution in SDE

Solving the posterior density

$$p_{U(t) \mid Y_{1:T}}(u \mid y_{1:T}), \quad \forall t \in [0, \infty) \tag{16}$$

is essentially a (continuous-discrete) stochastic filtering and smoothing problem.

A plenty of solvers ...

- Gaussian filters and smoothers (e.g., EKFSs, UKFSs).
- Particle filters and smoothers.
- ...

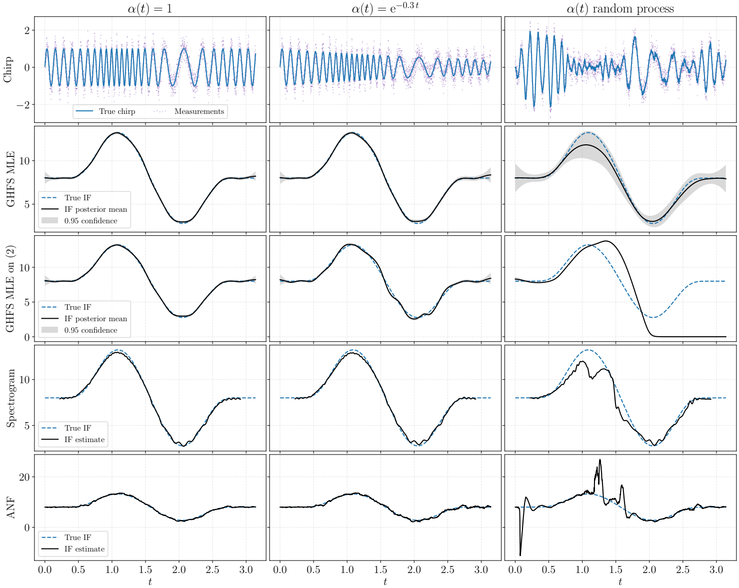
Experiments

Synthetic test model

$$f(t) = a\,b\cot(t)\csc(t)\,\mathrm{e}^{-b\csc(t)} + c, \quad t \in (0,\pi),$$

$$Y_k = \alpha(t_k)\sin\left(2\pi\int_0^{t_k} f(s)\,\mathrm{d}s\right) + \xi_k, \quad \xi_k \sim \mathrm{N}(0, 0.1),$$

Experiments



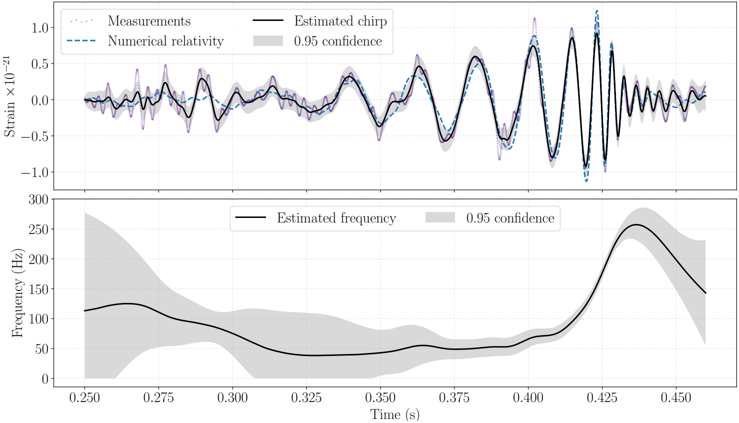
Experiments

RMSE ($\times 10^{-1}$)	$\alpha(t) = 1$	$\alpha(t) = \mathrm{e}^{-0.3t}$	$\alpha(t)$ is a random process
Hilbert transform	7.13 ± 2.35	11.74 ± 11.06	54.63 ± 25.58
Spectrogram	1.53 ± 0.08	1.82 ± 0.18	8.17 ± 4.31
Polynomial MLE	8.87 ± 0.09	8.90 ± 0.13	10.01 ± 4.33
ANF	2.13 ± 0.16	3.05 ± 0.31	37.77 ± 23.57
EKFS MLE old model	1.09 ± 0.20	19.53 ± 18.14	41.46 ± 19.48
GHFS MLE old model	0.67 ± 0.17	3.84 ± 7.95	39.47 ± 19.36†
EKFS MLE	0.70 ± 0.17	0.98 ± 0.24	6.37 ± 7.04
GHFS MLE	$\mathbf{0.65} \pm 0.16$	$\mathbf{0.93} \pm 0.24$	5.36 ± 5.59
CD-EKFS MLE	1.53 ± 0.67	2.86 ± 2.26	6.05 ± 6.55
CD-GHFS MLE	0.72 ± 0.18	1.16 ± 0.34	$\mathbf{4.91} \pm 3.74$‡

† and ‡ encounter 1 and 15 NaN numerical errors, respectively.

Real application

Frequency estimation of gravitational wave GW150914.



Tack!

Scan to access the code and preprint.
`https://github.com/spdes/chirpgp`