DRAFT VERSION APRIL 6, 2017
Preprint typeset using LATEX style AASTeX6 with modifications

FASTER GAUSSIAN PROCESS MODELING

TBD

ABSTRACT

Much faster solves using Cholesky solver.

Keywords: methods: data analysis — methods: statistical — asteroseismology — stars: rotation — planetary systems

1. INTRODUCTION

Gaussian Processes (GPs; Rasmussen & Williams 2006). We made this faster.

2. CHOLESKY METHOD

2.1. *Factorization*

Semi-separable

$$\begin{aligned}
u_{n,j}^{(1)} &= a_j\,e^{-c_j\,t_n}\,\cos(d_j\,t_n) + b_j\,e^{-c_j\,t_n}\,\sin(d_j\,t_n) && (1)\\
u_{n,j}^{(2)} &= a_j\,e^{-c_j\,t_n}\,\sin(d_j\,t_n) - b_j\,e^{-c_j\,t_n}\,\cos(d_j\,t_n) && (2)\\
v_{m,j}^{(1)} &= e^{c_j\,t_m}\,\cos(d_j\,t_n) && (3)\\
v_{m,j}^{(2)} &= e^{c_j\,t_m}\,\sin(d_j\,t_n) && (4)
\end{aligned}$$

$$\begin{aligned}
K &= \mathrm{diag}(\Sigma) + \mathrm{tril}(U\,V^{\mathrm{T}}) + \mathrm{triu}(V\,U^{\mathrm{T}}) && (5)\\
&= L\,\mathrm{diag}(D)\,L^{\mathrm{T}} && (6)
\end{aligned}$$

where

$$L = I + \mathrm{tril}(U\,X^{\mathrm{T}}) \qquad (7)$$

Pre-condition:

$$\phi_{n,j} = e^{-c_j\,(t_n - t_{n-1})} \qquad (8)$$

and $\phi_{1,j} = 1$.

$$\begin{aligned}
\tilde{u}_{1,n,j} &= a_j\,\cos(d_j\,t_n) + b_j\,\sin(d_j\,t_n) && (9)\\
\tilde{u}_{2,n,j} &= a_j\,\sin(d_j\,t_n) - b_j\,\cos(d_j\,t_n) && (10)\\
\tilde{v}_{1,m,j} &= \cos(d_j\,t_n) && (11)\\
\tilde{v}_{2,m,j} &= \sin(d_j\,t_n) && (12)\\
\tilde{X}_{k,n,j} &= e^{-c_j\,t_n}\,X_{k,n,j} && (13)
\end{aligned}$$

Recursion:

$$
\begin{aligned}
S_n^{(j,k)} &= \mathrm{diag}(\phi_n)\,\left[S_{n-1}^{(j,k)} + D_{n-1}\,\tilde{X}_{j,n-1}^{\mathrm{T}}\,\tilde{X}_{k,n-1}\right]\,\mathrm{diag}(\phi_n) && (14)\\
D_n &= \Sigma_n - \tilde{u}_{1,n}\,S_n^{(1,1)}\,\tilde{u}_{1,n}^{\mathrm{T}} - 2\,\tilde{u}_{1,n}\,S_n^{(1,2)}\,\tilde{u}_{2,n}^{\mathrm{T}} - \tilde{u}_{2,n}\,S_n^{(2,2)}\,\tilde{u}_{2,n}^{\mathrm{T}} && (15)\\
D_n\,\tilde{X}_{1,n} &= \tilde{v}_{1,n} - \tilde{u}_{1,n}\,S_n^{(1,1)} - \left[S_n^{(1,2)}\,\tilde{u}_{2,n}^{\mathrm{T}}\right]^{\mathrm{T}} && (16)\\
D_n\,\tilde{X}_{2,n} &= \tilde{v}_{2,n} - \tilde{u}_{2,n}\,S_n^{(2,2)} - \tilde{u}_{1,n}\,S_n^{(1,2)} && (17)
\end{aligned}
$$

where the matrices $S_1^{(j,k)}$ are $J \times J$ matrices of zeros. The computational cost of this section is $\mathcal{O}(N\,J^2)$.

2.2. *Solving*

To solve $L\,\boldsymbol{z} = \boldsymbol{y}$,

$$
\begin{aligned}
f_{k,n} &= \mathrm{diag}(\phi_n)\,\left[f_{k,n-1} + \tilde{X}_{k,n-1}^{\mathrm{T}}\,z_{n-1}\right] && (18)\\
z_n &= y_n - [\tilde{u}_{1,n}\,f_{1,n} + \tilde{u}_{2,n}\,f_{2,n}] && (19)
\end{aligned}
$$

where $f_{k,1} = 0$. Start from $n = 1$ and repeat. The computational cost of this section is $\mathcal{O}(N\,J)$.

To solve $L^{\mathrm{T}}\,\boldsymbol{z} = \boldsymbol{y}$,

$$
\begin{aligned}
f_{k,n} &= \mathrm{diag}(\phi_{n+1})\,\left[f_{k,n+1} + \tilde{u}_{k,n+1}^{\mathrm{T}}\,z_{n+1}\right] && (20)\\
z_n &= y_n - \left[\tilde{X}_{1,n}\,f_{1,n} + \tilde{X}_{2,n}\,f_{2,n}\right] && (21)
\end{aligned}
$$

where $f_{k,N} = 0$. Start from $n = N$ and repeat. The computational cost of this section is $\mathcal{O}(N\,J)$.

2.3. *Dot products*

To compute the dot product $L^{\mathrm{T}}\,\boldsymbol{z} = \boldsymbol{y}$,

$$
\begin{aligned}
f_{k,n} &= \mathrm{diag}(\phi_{n+1})\,\left[f_{k,n+1} + \tilde{u}_{k,n+1}^{\mathrm{T}}\,z_{n+1}\right] && (22)\\
y_n &= z_n + \left[\tilde{X}_{1,n}\,f_{1,n} + \tilde{X}_{2,n}\,f_{2,n}\right] && (23)
\end{aligned}
$$

where $f_{k,N} = 0$.

To compute the dot product $L\,\boldsymbol{z} = \boldsymbol{y}$

$$
\begin{aligned}
f_{k,n} &= \mathrm{diag}(\phi_n)\,\left[f_{k,n-1} + \tilde{X}_{k,n-1}^{\mathrm{T}}\,z_{n-1}\right] && (24)\\
y_n &= z_n + [\tilde{u}_{1,n}\,f_{1,n} + \tilde{u}_{2,n}\,f_{2,n}] && (25)
\end{aligned}
$$

To compute the dot product $K\,\boldsymbol{z} = \boldsymbol{y}$, take two passes

$$
\begin{aligned}
f_{k,n} &= \mathrm{diag}(\phi_n)\,\left[f_{k,n-1} + \tilde{v}_{k,n-1}^{\mathrm{T}}\,z_{n-1}\right] && (26)\\
y_n &= \Sigma_n\,z_n + \tilde{u}_{1,n}\,f_{1,n} + \tilde{u}_{2,n}\,f_{2,n} && (27)
\end{aligned}
$$

then

$$f_{k,n} = \mathrm{diag}(\phi_{n+1}) \left[f_{k,n+1} + \tilde{u}_{k,n+1}^{\mathrm{T}} \, z_{n+1} \right] \tag{28}$$

$$y_n = y_n + [\tilde{v}_{1,n} \, f_{1,n} + \tilde{v}_{2,n} \, f_{2,n}] \tag{29}$$

REFERENCES

Rasmussen, C. E., & Williams, K. I. 2006,
Gaussian Processes for Machine Learning
(MIT Press)