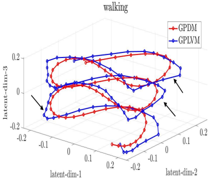
Gaussian Process Methods for Dimension Reduction

Applied to Human Motion Capture Data



Sai Satya Charan Malladi

Dec 16, 2021

AEROSP 567 - Project 4

Department of Aerospace Engineering
University of Michigan, Ann Arbor

Introduction

Paper Details

- Implementing a part of the paper "Gaussian Process Dynamical Models" by Wang et.al[1]. The authors propose a dimension reduction method and apply it to human-motion data.
- Tasks like video-based people tracking and data-driven animation require good statistical models for human motion. Difficult as human pose parameterized by $>$ 60 parameters and human motion is complex.
- Activity-specific human motion has smaller intrinsic dimension. So the motion is modeled by a dynamical process on a low-dimensional latent space (3-dimensional) and the pose (62-dimensional) is generated by a observation process from the latent space.

Primary Goal

- Implement Gaussian Process Dynamical Model (GPDM) and Gaussian Process Latent Variable Method (GPLVM) to test the claim that "GPDM generates a smoother latent map than GPLVM". Smoother maps help in better human motion reconstruction and hence are desirable.

[1] Jack M Wang, David J Fleet, and Aaron Hertzmann. "Gaussian process dynamical models". In: *NIPS*. vol. 18. Citeseer. 2005, p. 3.

Gaussian Process Dynamical Models (GPDM)

Let $\boldsymbol{Y} = [\boldsymbol{y_1} \ldots \boldsymbol{y_T}]$ be T observations, each J-dimensional, $\boldsymbol{X} = [\boldsymbol{x_1} \ldots \boldsymbol{x_T}]$ be D-dimensional latent variables. Consider the following Markov dynamics,

$$\begin{aligned} \boldsymbol{x}_t &= \boldsymbol{f}\left(\boldsymbol{x}_{t-1}; \boldsymbol{A}\right) + \boldsymbol{n}_{x,t} \qquad & \boldsymbol{f}(\boldsymbol{x}; \boldsymbol{A}) &= \sum_{k=1}^{K} \boldsymbol{a}_k \phi_{\boldsymbol{k}}(\boldsymbol{x}) \\ \boldsymbol{y}_t &= \boldsymbol{g}\left(\boldsymbol{x}_t; \boldsymbol{B}\right) + \boldsymbol{n}_{y,t} \qquad & \boldsymbol{g}(\boldsymbol{x}; \boldsymbol{B}) &= \sum_{m=1}^{M} \boldsymbol{b}_m \psi_{\boldsymbol{m}}(\boldsymbol{x}) \end{aligned} \tag{1}$$

Wang's modeling assumption is that each row of $\boldsymbol{B}$ and $\boldsymbol{A}$ i.e $\boldsymbol{b_j}$ ans $\boldsymbol{a_d}$ are i.i.d Gaussian. This idea is used to integrate out $\boldsymbol{B}$ and $\boldsymbol{A}$ in 1 to give us expressions for $\boldsymbol{P(Y \mid X)}$ and $\boldsymbol{P(X)}$. This helps us write the inference equation,

$$\boldsymbol{P(X \mid Y)} \propto \boldsymbol{P(Y \mid X)P(X)} \tag{2}$$

The latent variables $\boldsymbol{X}$ and the hyperparameters (of kernels corresponding to $\phi_{\boldsymbol{k}}(\boldsymbol{x})$, $\psi_{\boldsymbol{m}}(\boldsymbol{x})$) are obtained by maximizing the log-posterior. The only change in the formulation for Gaussian Process Latent Variable Method (GPLVM) is that, $\boldsymbol{x}_t$ does not vary according to its own dynamics so the latent variables $\boldsymbol{X}$ and the hyperparameters are found by maximizing log-likelihood $\boldsymbol{P(Y \mid X)}$.

The data for the implementation is obtained from CMU motion capture database. The data is a 62 dimensional time series for each motion.

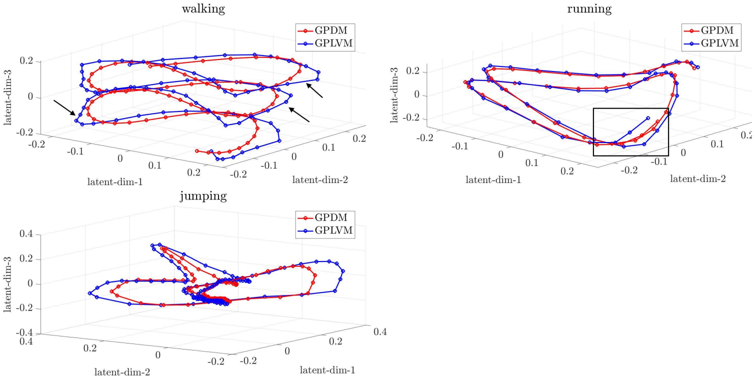


Figure 1: Comparison of GPDM and GPLVM

For 3D people tracking and computer animation, it is desirable to generate new motions efficiently. Drawing samples $\tilde{\mathbf{X}}_{1:T}^{(j)} \sim p\left(\tilde{\mathbf{X}}_{1:T} \mid \mathbf{x}_0, \mathbf{X}, \mathbf{Y}\right)$ using MCMC methods is computationally expensive. The authors propose mean-prediction to generate sample motions efficiently. In mean prediction, the next time step $\tilde{\mathbf{x}}_t$ is conditioned on $\tilde{\mathbf{x}}_{t-1}$

$$\tilde{\mathbf{x}}_t \sim \mathcal{N}\left(\mu_X\left(\tilde{\mathbf{x}}_{t-1}\right); \sigma_X^2\left(\tilde{\mathbf{x}}_{t-1}\right)\mathbf{I}\right)$$

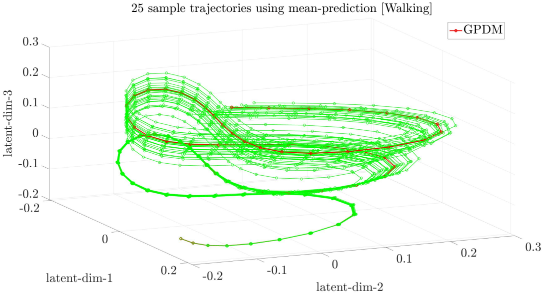


Figure 2: 25 sample trajectories using mean-prediction [Walking]

Conclusion

- Through this project I have explored a new application of Gaussian process for dimension reduction.
- Concepts like marginalisation, Bayesian inference, dynamical systems, maximisation of log-like-hood taught in class were helpful for understanding math behind the method proposed in the paper.
- The results obtained from my implementation for "walking" motion are found to be inline with the trends reported in the paper.
- The optimization process was found to be very sensitive to initial guess of hyperparameters.