Gaussian Process Dynamical Models

JM Wang, DJ Fleet, A Hertzmann

Dan Grollman, RLAB 3/21/2007

Foci

- Want low-dimensional representations of high-dimensional data
  - Find the latent space
  - Dimensionality reduction
- Want to model *dynamics* of the system
  - How it evolves over time
  - Not just distance in input space

Domain

- Human motion (gait)
  - 62 dimensions, video
  - Regular DR treats each frame (pose) separately
  - Want to include how poses relate over time

Non-linear Dynamical Model

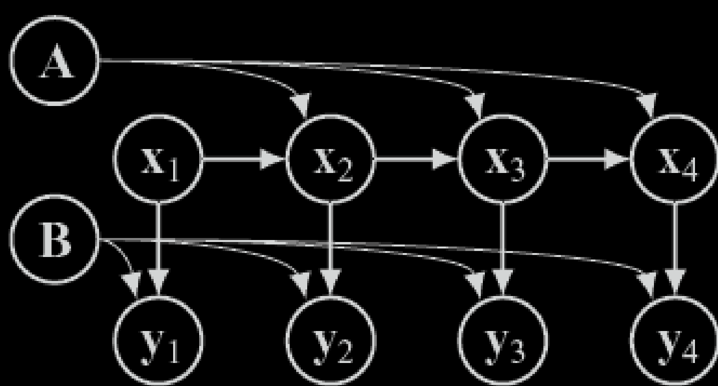
- State evolution in latent space

$$\mathbf{x}_t = f\left(\mathbf{x}_{t-1};\mathbf{A}\right)+\mathbf{n}_{x,t} \qquad f\left(\mathbf{x};\mathbf{A}\right)=\sum_i \mathbf{a}_i\phi_i\left(\mathbf{x}\right)$$

- Observation (latent->observed)

$$\mathbf{y}_t = g\left(\mathbf{x}_t;\mathbf{B}\right)+\mathbf{n}_{y,t} \qquad g\left(\mathbf{x}_t;\mathbf{B}\right)=\sum_j \mathbf{b}_j\psi_j\left(\mathbf{x}\right)$$

- Model:



Issues

- Need parameters A and B and basis functions!
  - and enough data to constrain them
- Why not marginalize them out?
  - Can do it in closed form

Latent->observed mapping

- isotropic Gaussian prior on B, marinalize over g:

$$p\left(\mathbf{Y} \mid \mathbf{X}, \overline{\beta}\right) = \frac{|\mathbf{W}|^N}{\sqrt{(2\pi)^{ND} |\mathbf{K}_{\mathbf{Y}}|^D}} \exp\left(-\frac{1}{2} \operatorname{tr}\left(\mathbf{K}_{\mathbf{Y}}^{-1} \mathbf{Y} \mathbf{W}^2 \mathbf{Y}^T \right)\right)$$

- Use RBF for kernel

Latent dynamics

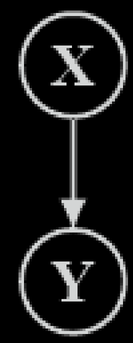
- Use 1st order Markov, marginalize over A with isotropic Gaussian prior:

$$p(\mathbf{X} \mid \overline{\alpha}) = p(\mathbf{x}_1) \frac{1}{\sqrt{(2\pi)^{(N-1)d} \mid \mathbf{K}_{\mathbf{X}} \mid^d}} \exp\left(-\frac{1}{2} \mathrm{tr}\left(\mathbf{K}_{\mathbf{X}}^{-1} \mathbf{X}_{out} \mathbf{X}_{out}{}^T \right) \right)$$

- use linear + RBF kernel

Final Model

- All coordinates are jointly correlated, as are poses.



- Simple, no?

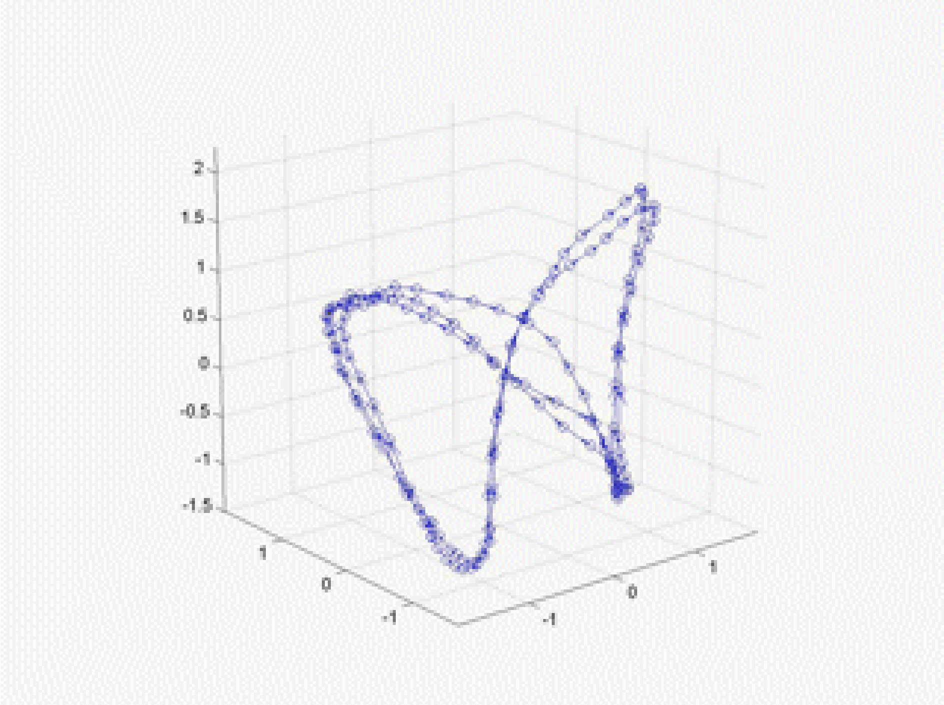
Learning

- All you have is Y, the observed variables (62 D human pose information)
  - minimize negative log-posterior numerically:

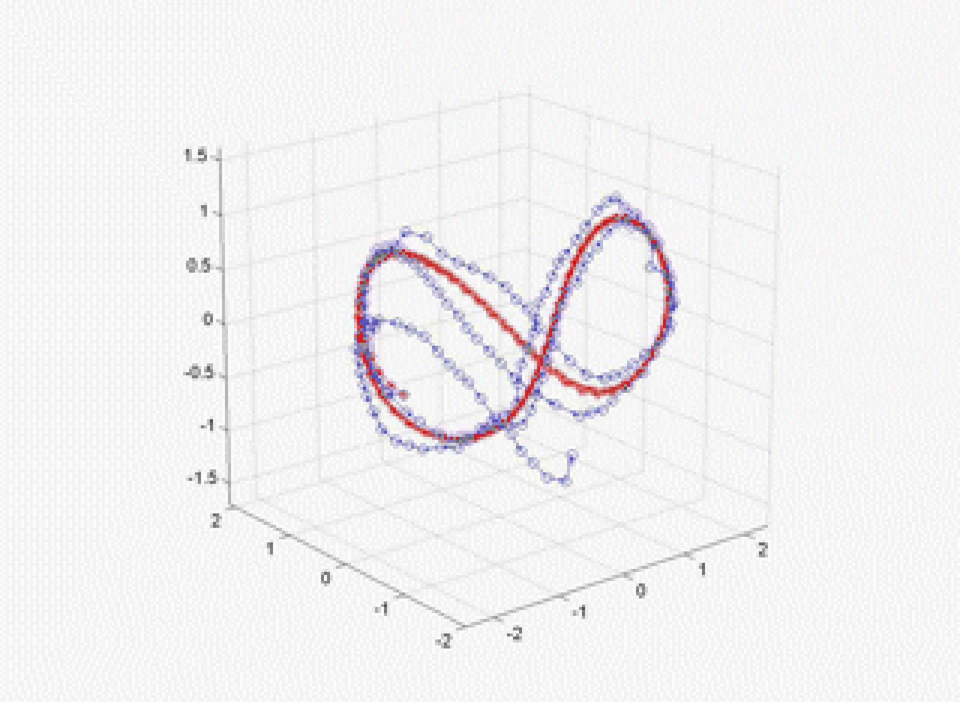
    $L = -\ln p\left(\mathbf{X}, \overline{\alpha}, \overline{\beta} \mid \mathbf{Y}\right)$
  - Initialize latent coordinates with PCA
    - 3D, because 2D is unstable

Results

PCA



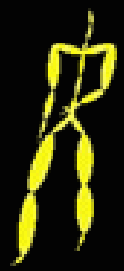
GPDM



New motion generation

Original

New



Other uses

- Forecasting
  - Use mean-prediction (motion generation) to look ahead
- Missing Data
  - Use mean-prediction to fill in gaps
    - Must increase uncertainty in training data via downsampling

Points / Issues

- Dimensionality of latent space is set via PCA space.
  - What if unknown?
- Scalability
  - Because of correlation between points, difficult to make sparse.

Want more?

- http://www.dgp.toronto.edu/~jmwang/gpdm