Debate Advice
Affirmative Team

• **Thesis Statement**: Every (good) paper has a thesis. What is the most provocative statement the paper is trying to make?

• **Spill the beans early**: Explain the key insight of the paper as soon as possible. Like on Slide 1.

• **Teaser**: Show an example result first so that it’s clear to everyone what the goal is. Also mention the input, clearly.

• **Equations**: Explain them carefully or don’t include them.
• **Examine the Thesis**: Is it sound? Does it overreach? Is it practically useful?

• **State a Core Objection**: State a Core Objection early in the process.

• **Explain limitations**: Usually in the paper. Assumptions, model restrictions, computation, etc.
Human Motion

How can we model human motion?
Why Model Motion?

Keyframing
Why Model Motion?

Labeling

$X_t^i$: Point Position at time $t$

$Z_t^i$: Point Label at time $t$
Why Model Motion?

3D Reconstruction

$$\mathbf{X}_t = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & \cdots & X_P \\ Y_1 & Y_2 & Y_3 & Y_4 & \cdots & Y_P \end{bmatrix}$$
Why Model Motion?

Action Recognition
Knowns and Unknowns

There is structure in human pose and motion

Keyframing \( t \)

Labeling

Reconstruction

Recognition
Lecture in One Slide

Notation

\[ X_{t+1} = D X_t \]

Dynamics

\[ X_t = B Y_t \]

Latent Variable Model

\[ Y_{t+1} = G Y_t \]

Latent Dynamics
Representing Pose

Configuration Space

\( D \text{ Features} \)

\( D \text{ Points} \)

\( D \text{ Joint Angles} \)

\( X_t \in \mathbb{R}^D \)

Observation
The Space of Actions

\[ p(X_1, \cdots, X_F) \]
Frame-wise Independent Human Motion

Data Space

\[ p(X_t, \cdots, X_{t+F}) = p(X_t)p(X_{t+1})p(X_{t+2}) \cdots p(X_{t+F}) \]
\[ p(X_1)p(X_2) \ldots p(X_F) \]
\[ p(X_t) = \mathcal{N}(X_t | \mu, \Sigma) \]
Autoregressive Models

First-order Markov Model

\[ p(X_1, \cdots, X_F) = p(X_1) \prod_{t=2}^{F} p(X_t | X_{t-1}) \]

\[ p(X_t | X_1, \cdots, X_{t-1}) = p(X_t | X_{t-1}) \]
First-Order Markov Models

Linear Dynamics

\[ p(X_t | X_{t-1}) \]

\[ X_t = f(X_{t-1}) \quad \text{Dynamical Function} \]

\[ X_{t+1} = DX_t \quad \text{Linear Dynamics} \]
Markov Models

First Order

\[ X_{t+1} = DX_t \quad \text{D} \in \mathbb{R}^{D \times D} \]

\[ X_{t+n} = D^n X_t \]
Markov Model

Reconstruction

\[
\begin{bmatrix}
\bar{X}_{t+1} \\
\bar{X}_{t+2} \\
\bar{X}_{t+3} \\
\vdots \\
\bar{X}_{t+F+1}
\end{bmatrix}
= D
\begin{bmatrix}
X_t \\
X_{t+1} \\
X_{t+2} \\
\vdots \\
X_{t+F}
\end{bmatrix}
\]
First-order AR

Reconstruction

- Ground truth
- Reconstructed
Autoregressive Models

First-order Markov Model

\[
p(X_1, \cdots, X_F) = p(X_1) \prod_{t=2}^{F} p(X_t | X_{t-1})
\]

\[
p(X_t | X_{t-1}) = \mathcal{N}(X_t | DX_{t-1}, \Sigma)
\]
First-Order AR

Prediction

\[
\begin{bmatrix}
\bar{X}_{t+1} \\
\bar{X}_{t+2} \\
\bar{X}_{t+3} \\
\vdots \\
\bar{X}_{t+F+1}
\end{bmatrix}
= 
\begin{bmatrix}
D \\
D^2 \\
D^3 \\
\vdots \\
D^F
\end{bmatrix}
X_t
\]

Observability Matrix
First-Order AR
Linear Prediction

\[
\begin{bmatrix}
X_{t+1} \\
\bar{X}_{t+2} \\
\bar{X}_{t+3} \\
\vdots \\
\bar{X}_{t+F+1}
\end{bmatrix}
= 
\begin{bmatrix}
D \\
D^2 \\
D^3 \\
\vdots \\
D^F
\end{bmatrix}
X_t
\]
First-Order AR

Prediction Generalization

\[
\begin{bmatrix}
\bar{X}_{t+k} \\
\bar{X}_{t+1+k} \\
\bar{X}_{t+2+k} \\
\vdots \\
\bar{X}_{t+F+k}
\end{bmatrix}
= 
\begin{bmatrix}
D \\
D^2 \\
D^3 \\
\vdots \\
D^F
\end{bmatrix}
X_{t+k}
\]
First Order AR
Prediction Generalization

\[
\begin{aligned}
\bar{X}_{t+k} &= \begin{bmatrix} X_{t+k} \\ \bar{X}_{t+k} \\ \bar{X}_{t+1+k} \\ \bar{X}_{t+2+k} \\ \vdots \\ \bar{X}_{t+F+k} \end{bmatrix} = \begin{bmatrix} D & D^2 & D^3 & \cdots & D^F \\ \end{bmatrix} \begin{bmatrix} X_{t+k} \\ \end{bmatrix}
\end{aligned}
\]
Autoregressive Models
First-order Markov Model

\[
p(X_t | X_{t-1}) = \mathcal{N}(X_t | DX_{t-1}, \Sigma)
\]
Ideas?
Autoregressive Models
Second-Order Markov Model

\[ p(X_1, \cdots, X_F) = p(X_1)p(X_2|X_1) \prod_{t=1}^{F} p(X_t|X_{t-1}, X_{t-2}) \]
$X_{t+1} = D \begin{bmatrix} X_{t-1} \\ X_t \end{bmatrix}$

$D \in \mathbb{R}^{D \times 2D}$
Considerations

AR systems

- **Complexity**: Curse of dimensionality, computation, compaction?

- **Predictive Precision**: How accurately does the model predict observations?

- **Generalization Ability**: How well does the model generalize to new data?
Considerations
Trade-offs

• **Memory**: How far back should you look? How much is it worth in extra dimensionality?

• **Linearity**: How much of a limitation is the linearity of the dynamical system?
Singular Values of D

Compaction across Actions (walk, sit, fall, etc.)
Ideas?
Pose Correlations

Latent Variable Models

• How many degrees of freedom are there *really*?

\[ X_t \sim \mathcal{N}(\mu, \Sigma) \]
Latent Variable Models

Linear Models

\[ w_1 + w_2 + w_3 + \ldots \]
Principal Component Analysis

Linear Projection

\[
\{X_n\} : \text{Training Data} \quad X_t \in \mathbb{R}^D
\]

\[
Y_t \in \mathbb{R}^M \quad M < D
\]
Probabilistic PCA

Distribution

\[ p(X_t|Y_t) \]

\[ X_t = BY_t + \mu + \epsilon \]

\[ p(Y_t) = \mathcal{N}(z|0, I) \]

\[ p(X_t|Y_t) = \mathcal{N}(X_t|BY_t + \mu, \sigma^2 I) \]
Probabilistic PCA

Generative View of PCA

\[ p(z) \]

\[ p(x|\hat{z}) \]

\[ \mu \]

\[ \mu \]

\[ p(x) \]
Probabilistic PCA
Sampling Standing Up

\[ X_t \sim \mathcal{N}(\mu, \Sigma) \]
Probabilistic PCA

Sampling Standing Up

\[ p(X_t|Y_t) = \mathcal{N}(X_t|BY_t + \mu, \sigma^2 I) \]
Graphical Model
Component Analysis

\[ X_t = B Y_t + \mu + \epsilon \]
Graphical Model
Component Analysis

\[ X_t = B Y_t + \mu + \epsilon \]
PCA Works!
Less than 5cm max-error with <30 components

Max. PCA Recon. Error per Action

- armmovement
- bend
- boxing
- climb
- jump
- run
- sit
- walk
- wave

Graph showing the maximum PCA reconstruction error per action against the number of PCA components (K).
Well...Somewhat...

PCA works when Action is Known

Performance of PCA with Data Diversity
Overcomplete Dictionaries
Ramakrishna et al. 2012

\[
\arg \min_{Y_t} \| X_t - B Y_t \|_2
\]

\[X_t = B Y_t\]

L1-norm “encourages” sparsity in \( Y \)
3D Reconstruction

Reprojection Error Decreases at Each Iteration

Iteration No.: 1
3D Reconstruction

3D Pose and Camera

Iteration No.: 1
Graphical Models

Dynamical Models

Latent Space

\[ Y_t \rightarrow Y_{t+1} \rightarrow Y_{t+2} \rightarrow \cdots \rightarrow Y_{t+F} \]

Data Space

\[ X_t \rightarrow X_{t+1} \rightarrow X_{t+2} \rightarrow \cdots \rightarrow X_{t+F} \]
Linear Dynamical System

Graphical Summary

\[ Y_t \rightarrow Y_{t+1} \rightarrow Y_{t+2} \rightarrow \cdots \rightarrow Y_{t+F} \]

\[ X_t \rightarrow X_{t+1} \rightarrow X_{t+2} \rightarrow \cdots \rightarrow X_{t+F} \]

\[ X_{t+1} = DX_t \quad X_t = BY_t \]

Linear Dynamics

Latent Variable Model
Model Reduction

Dynamics in the Latent Space

\[ X_{t+1} = DX_t \quad X_t = BY_t \]

\[ BY_t = DBY_t \]

\[ Y_t = B^T DBY_t \]

\[ Y_t = GY_t \]

\[ G = B^T DB \]
Model Reduction

Projected Dynamics

\[ G = B^T D B \]
Dynamical Models

Graphical Models

Latent Space

Data Space
\[ p(X_1, \cdots, X_F) = \int p(X_1, \cdots, X_F, Y_1, \cdots, Y_F) dY \]

\[ p(Y_t | Y_{t-1}) = \mathcal{N}(Y_t | G Y_{t-1}, \Gamma) \]

\[ p(X_t | Y_t) = \mathcal{N}(X_t | B Y_t, \Sigma) \]

\[ p(Y_1) = \mathcal{N}(Y_1 | \mu, V) \]
Nonlinear Dynamical Models?

• Linear-Gaussian Models work for individual activities

• Nonlinear Latent Variable Models:
  • Density Networks
  • Generative Topographic Mapping
  • Kernel PCA
  • Gaussian Process Latent Variable Models

• Nonlinear Dynamics:
  • Switching Linear Dynamical Models
  • Gaussian Process Dynamical Models
  • Sampling-based methods (i.e., Particle Filters)
Reading List

• Pavlovic et al. Learning Switching Linear Models of Human Motion
• Lawrewnce et al. Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data
• Fleet, Motion Models for People Tracking.
• Ramakrishna et al., Reconstructing 3D Human Pose from 2D Image Landmarks, 2012.