CS195-5 : Introduction to Machine Learning
Lecture 36

Greg Shakhnarovich

December 11, 2006

Announcements

- Wednesday 12/13: final review.
- Office hours:
  - Payman
  - Dan
  - Greg: by appointment (e-mail).

What have we not covered?

- Inference in graphical models
  - Exact: local message passing algorithm
  - Approximate: sampling, loopy belief propagation.
- Undirected models: Markov random fields
- Example: image models

Structure learning

- We have assumed that the structure (edges) of the GM is given.
- We can learn it from data
  - A model selection task.
  - The best explanation: fully connected graph.
  - Need to penalize it by model complexity.
- (Non-linear) manifold learning
  - PCA recovers an "interesting" linear subspace of the data.
  - Many methods target non-linear subspaces.

Semi-supervised learning

- Small labeled data set $\sim p(\mathbf{x}, y) = p(\mathbf{x})p(y \mid \mathbf{x})$;
- Large set of unlabeled data $\sim p(\mathbf{x})$.
- We can use the unlabeled data to improve the model/estimates
  – Estimate density, and use the result to assign weights to labeled examples.
  – *Transduction*: predict the labels for the unlabeled data, and re-train the classifier pretending these are correct.

Active learning

- We are allowed to query the label of unlabeled examples.
- Labeling is expensive.
  - Recall: in linear regression,

$$\hat{\mathbf{w}} \sim \mathcal{N}\left(\mathbf{w}; \mathbf{w}^*, \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}\right)$$

  - Basic idea: query examples whose label will contribute most to your ability to predict future labels.

Online learning

- We observe examples in order, and start learning right away
- With each example (or small batch of examples) need to update the model
  - Often need to make predictions quickly!
- Applications:
  - Financial time series prediction
  - Adaptive systems
  - Robot exploration of environment

Reinforcement learning

- "Reinforcement learning is learning what to do–how to map situations to actions–so as to maximize a numerical reward signal." [Sutton & Barto]

- Main elements:
  - *Actions* that can be taken.
  - *Policy*: mapping from state of the environment to action.
  - *Reward* function: mapping state-action pairs to value.

- Objective: through trial and error, learn a policy that will maximize epected reward in the long run.

- Examples: many. E.g., inverted helicopter flight.

Theory

- What is learnable with a particular family of classifiers?

Theory

- What is learnable with a particular family of classifiers?

- Probably Approximately Correct (PAC) framework:
  - We select from a set $\mathcal{H}$ a hypothesis h^* that achieves zero training error.
  - How large should N be so that with probability at least $1 - \delta$, the expected risk of h^* is no more than ϵ?