Graphical Models
NME 130 Discussion session

Andreas Krause
California Institute of Technology

rsrg@caltech
..where theory and practice collide

New GM class CS/CNS/EE 155 Fall '09

- **Still in design, I'd love to get feedback..**
- Key Theme: Global insight from local observations
  - Segment image based on noisy pixel values
  - Parse sentence based on noisy syntax
  - Detect faults based on noisy sensors
  - …
- Bayesian perspective:
  Reality = collection of random vars $P(X_1, \ldots, X_n)$
- Need to
  - Specify / represent distribution P → **Modeling**
  - Compute conditional probabilities $P(X_1 \mid X_3)$ → **Inference**
  - Estimate distribution P from data → **Learning**

Syllabus

- Modeling and representation (5 lectures)
- Inference (6 lectures)
- Learning GMs from data (5 lectures)
- Applications and case studies (2 lectures)
- “Research frontier” (3 lectures)

Syllabus

- Modeling and representation
  - Conditional independence = factorization (Markov property,…)

    $P(X_1,\ldots,X_n) = 1/Z\ \Psi_1(X_1,X_3) \ldots \Psi_m(X_7,X_3,X_n)$
  - Graph representations of cond.indep. (=factorization)
    - Directed (acyclic) graphs (Bayes nets, DBNs)
    - Undirected graphs (MRFs, CRFs), Hammersley-Clifford thm etc.
  - Factor types: Multinomial, Gaussians, exponential family, …
  - Examples: Regression, Naïve Bayes, Markov chains, HMMs, Kalman filters, mixture models (clustering…), PCA
- Inference
- Learning
- Applications
- "Research frontier"

Syllabus

- Modeling
- Inference
  - Exact (3 lectures)
    - Low treewidth: Variable elimination, junction tree inference
    - High treewidth: Alg. curcuits, MAP in MRFs (Graph cuts, etc.)
  - Approximate (3 lectures)
    - Sampling based (MCMC, Gibbs, particle filtering) [exact in limit]
    - Message passing (loopy BP, EP; convergence)
    - Optimization based (Variational Inference, mean field; ADF)
- Learning
- Applications
- "Research frontier"

Syllabus

- Modeling
- Inference
- Learning
  - Parameter learning (2 lectures)
    - Fully observable (MLE, MAP)
    - Partially observable (EM)
  - Structure learning (3 lectures)
    - Chao-Liu for polytrees, TANs
    - Combinatorial (search, latent var discovery) + Relaxations (L1)
- Applications
- "Research frontier"

Syllabus

- Modeling
- Inference
- Learning
- Applications
  - Computer vision (make3d, image segmentation, …)
  - Networks (sensor nets, network tomography)
  - Recommender systems (matrix factorization)
  - CompBio (gene regulatory networks, phylogeny, …)
  - NLP (probabilistic parsing, statistical translation, …)
  - Diagnostics / decision theory (value of information, Bayesian experimental design, …)
- "Research frontier"

Syllabus

- Modeling
- Inference
- Learning
- Applications
- "Research frontier"
  - Relational models (probabilistic databases, …)
  - Probabilistic logic (Markov logic networks, …)
  - Nonparametric Bayesian methods (Gaussian processes, Dirichlet processes, …)
  - Connections to game theory? Graphical games, …

Connections to NME 130

- Info-theory, Coding (turbo codes, BP)
  - Model specified in advance, not learnt from data
- Stochastic optimal control
  - MDPs/POMDPs as controlled Markov chains / HMMs
- Dynamical systems
  - Inference in hybrid systems
  - Filtering
- Robustness
  - Bayesian model averaging
  - Robust Bayesian analysis
- Synthesis / Hard limits
  - Computational and statistical complexity
- …

All this in 6 Lectures??

- GMs = Probability/Statistics meets algorithms/optimization
- Modeling (2 lectures)
  - Conditional independence = factorization; "structured" (high-dim) distributions;
  - Factor graphs (capture both Bayes nets and MRFs)
- Inference (2 lectures)
  - Inference as optimization (variational methods, BP as special case)
  - Sampling (Gibbs, MCMC)
- Learning (2 lecture)
  - Parameter: (EM, IPF)
  - Structure: Combinatorial optimization (structure search, Chao-Liu, …)

Discussion

- Other connections?
- Overlap with other courses?