Review Questions

Lyle Ungar
What loss functions have we used?

◆ And when is each good?
What loss functions have we used?

- L0, L1, L2
- Log likelihood (MLE, MAP)
- Hinge
- Logistic
- Exponential
- Entropy (MDL), KL-divergence

**Boosting**: \( \exp(-y_if_{\alpha}(x_i)) \)

**Logistic**: \( \log(1 + \exp(-y_if_{w}(x_i))) \)
What optimization methods have we used?

◆ And when is each good?
What optimization/search methods have we used?

- Stochastic gradient descent
  - Mini-batch
- Stepwise/streamwise feature selection
- Streaming methods
  - LMS
  - Perceptron
- Quadratic optimization
- Power method
- Lagrange multipliers/KKT
- EM
What types of duality have we seen?

- Between features and observations
  - $X'X$ or $XX'$
- Lagrange duality