Midterm 1 Review

Test format

- 1 hour test
- Multiple choice & short answer questions
- No calculator
- Closed book, closed notes

Supervised Learning

Regression (Output $\in \mathbb{R}$)

Classification (Output in a finite set)

Hypothesis : $\theta^T x$.

Learning : Identifying hypothesis

Cost function J

Computation Objective : Find the hypothesis in our set that minimizes the cost on the training set.

Key : Identify cost function that is "continuous / differentiable"

(Stochastic) Gradient Descent

(Stochastic) Gradient Descent

Start with an initial θ

Repeat for some number of iterations

$$\theta \leftarrow \theta - \alpha \nabla_\theta J(\theta)$$

Linear Regression : (S)GD on least squares

Perceptron : SGD on a particular loss fn

Logistic : (S)GD on a particular cost fn.

Softmax : (S) GD on another cost fn.

→ Alternate Solution : Solve Normal eqns.

Generalized Linear Models

- $p(y | x; \theta) \sim$ Exponential Family

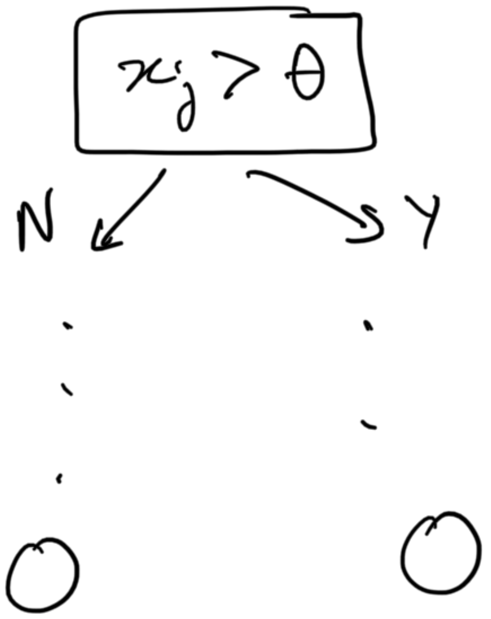
$$p(y|\eta) = b(y)\, e^{\eta^T T(y) - a(\eta)}$$

- Compute $E[\underbrace{T(y)}_{=y}]$

- Finding that maximizes the likelihood of the training set.

$$\theta_j \leftarrow \theta_j - \alpha \sum_{i=1}^{n} \left(h_\theta(x^{(i)}) - y^{(i)} \right) x_j^{(i)}$$

Decision Trees



Cons : Overfitting data.

---

k - Nearest Neighbors

→ k neighbors of test data in training set.

$$h_\theta(x) = \begin{cases} -1 & \theta^T x \leq T \\ +1 & \theta^T x > T \end{cases}$$

$$J(\theta) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left[h_\theta(x^{(i)}) \neq y^{(i)} \right]$$

$$p(y; \eta) = b(y)\, e^{\eta^T T(y) - a(\eta)}$$

$$p(y;\lambda) = \frac{\lambda^y}{y!} e^{-\lambda}$$

Observation : η is same fn of λ.

Term $\frac{\lambda^y}{y!}$ – depends on λ & y

$\forall\ z.\quad z = e^{\ln z}$

$$p(y;\lambda) = \underbrace{\frac{1}{y!}}\ e^{y\ln\lambda - \lambda}$$

$b(y) = \frac{1}{y!}$

$$\eta^T T(y) = y \ln \lambda$$

$$\eta = \ln\lambda$$

$$T(y) = y$$

$$a(\eta) = \lambda = e^{\eta}$$

$$p(y;\phi) = \begin{cases} \phi & y=1 \\ 1-\phi & y=0 \end{cases}$$

$$p(y;\phi) = \phi^y (1-\phi)^{1-y}$$