COT 6936: Topics in Algorithms

Giri Narasimhan

ECS 254A / EC 2443; Phone: x3748

giri@cs.fiu.edu

http://www.cs.fiu.edu/~giri/teach/COT6936_S10.html

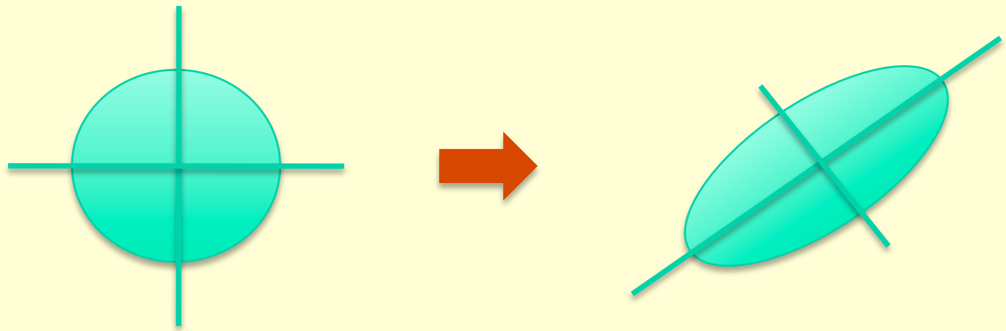
https://online.cis.fiu.edu/portal/course/view.php?id=427

Spectral Methods

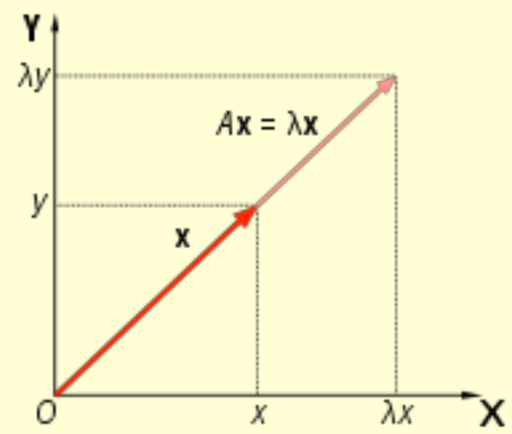
- Graph Connectivity problems
  - Google Page Rank
- Graph Partitioning problems
  - Clustering (even linearly non-separable case)
- Markov Chain Mixing problems
  - Random walks in graphs

Matrices and Eigenvalues

- Array of values
- Linear Transformation



- Eigenvalues and Eigenvectors
  - $Ax = \lambda x$
  - Under transformation A, eigenvectors only experience change in magnitude, not direction
  - $A = Q \Lambda Q^{-1}$



Graph Bisection

- Construct adjacency matrix *A*
- Construct Laplacian L = D – *A*
  - D = diagonal matrix with degrees along diagonal
- L is positive semi-definite (PSD); has non-neg eigenvalues; has smallest eigenvalue = 0
- Second eigenvector provides information about bisection.
  - Signs of 2nd eigenvector give a good bisection
  - Extreme case: Connected components have constant values in 2nd eigenvector

Graph Bisection (Continued)

- Eigenvalues indicate strength of bisection
- How to get bisections with n/2 vertices?
  - Use median value in second eigenvector
- How to get k partitions?
  - Perform bisections recursively
  - Use more eigenvectors

Spectral Clustering: Strategy

- Given data points and a distance function, construct a weighted graph
- Let A be its adjacency matrix; let D be diagonal matrix with degrees along diagonal
- Construct Laplacian L (PSD, non-neg eigenv.)
  - Unnormalized: $L = D - A$
  - Normalized symmetric: $L = D^{-1/2} L D^{1/2}$
  - Random Walk: $L = D^{-1} L$
- Matrix L_k has cols = first k eigenvectors of L
- Cluster rows of L_k

Spectral Clustering

- Need distance measure (need not be a metric), i.e., triangle inequality not needed
- Not Model-based
- Global method
- Turns discrete problem into continuous

Randomized Algorithm for MAX 3-SAT

- Assume each clause has 3 distinct literals
- Randomly assign 0/1 to all variables
  - Each clause is satisfied with prob 7/8
  - Expected number of clauses = 7m/8
  - There exists a truth assignment that satisfies 7m/8 clauses
- **Problem**:
  - How can we find a satisfying truth assignment with at least 7m/8 clauses satisfied?

Derandomization

- Consider randomized algorithm from slide 8
- $E[S\varphi] = \frac{1}{2} E[S\varphi \mid x_1 = T] + \frac{1}{2} E[S\varphi \mid x_1 = F]$
- $E[S\varphi \mid x_1 = T]$, $E[S\varphi \mid x_1 = F]$ can be computed in polynomial time. WHY?
- If $(E[S\varphi \mid x_1 = T] \geq E[S\varphi \mid x_1 = F])$, then

  $E[S\varphi \mid x_1 = T] \geq E[S\varphi] \geq 7m/8$
- Set $x_1 = T$, and reduce φ to φ'.
- Find value for x_2 and so on.

How to compute the expected values

- $E[X] = \Sigma_i P(C_i = T)$
- For example, let
  - $C_i = (x_1 \vee -x_2 \vee x_3)$
- $P(C_i = T | x_1 = T) = 1$
- $P(C_i = T | x1 = F) = 1 - P(C_i = F | x_1 = F) = \frac{3}{4}$