OM-S20-13: Eigen Value Problems in Optimization

C. V. Jawahar

IIIT Hyderabad

http://preon.iiit.ac.in/om_quiz

28 Feb 2020

Eigen Values and Eigen Vectors

Eigen vectors are vectors x that do not get "rotated" by A, only "stretched" (by a factor of λ)

$$Ax = \lambda x$$

Eigen values of A can be obtained by solving the characteristic equation $|A - \lambda I| = 0$. Eigen vectors form the null space of matrix $(A - \lambda I)$

Interesting properties:

- $A^2 x = \lambda^2 x$
- $\prod_i \lambda_i = |A|$
- $\sum_i \lambda_i = tr(A)$
- Every vector is an eigen vector of I
- Symmetric matrices have real eigen values
- **Spectral Theorem:** Symmetric matrices $S = Q \Lambda Q^T$
- Generalized eiegen value problem of two symmetric matrices A and B.

$$Ax = \lambda Bx$$

Presence of EV in OM

1. Many optimization problems take the form:

$$\min_x \;\; z = x^T A x, \mathrm{s.t.} \|x\|^2 = 1$$

   Solution gives us $Ax = \lambda x, z^* = \lambda$

2. Maximum variance line fitting:
   Let $y_1, \dots y_N$ be the points and w be the direction. Variance on x should be maximized with a constraint on $||w|| = 1$. Also $x = w^T(y - b)$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \frac{1}{N}\sum_i x_i)^2 + \lambda(w^T w - 1)$$

3. Least norm solution to $Ax = 0$

$$\min_x \;\; \|Ax\|^2, \mathrm{s.t.} \|x\|^2 = 1$$

   Solution: $(A^T A)x = \lambda x, z^* = \lambda$

Applications: PCA

- Objective function of PCA is defined as:
  Given data matrix X: (N x M), find a direction u, such that the variance of the projection of X on u is maximized (Max information captured).
- Projection of X on u is Xu
- Mean of projections is $\bar{X}u$

$$\max_u \|(X - \bar{X}).u\|^2, \|u\|^2 = 1$$

The constraint comes because u is only a direction.

$$\max_u u^T \Sigma u, u^T u = 1$$

$$\max_u u^T \Sigma u - \lambda(u^T u - 1)$$

$$\Sigma u = \lambda u, Z^* = \lambda$$

Simple Graph Partitioning/Clustering

- **Matrices:** $\#edges = m$; $\#vertices = n$
  - Incidence matrix(J): $m \times n$
  - Degree Matrix: Diagonal(D; $n \times n$
  - Adjacency (A): $n \times n \in \{0, 1\}$
  - Laplacian(L): $n \times n$ $L = D - A$; $L = J^T J$
  - Affinity/Weight (A/W) $n \times n$
- A: Affinity matrix; A_{ij} affinity of ith and jth node (eg. $A_{ij} = e^{-d(x_i, x_j)}$)
- ω: vector with ω_i being the "membership" of i vertex/sample into the cluster.
- Optimization problem:

$$\text{Maximize} \quad \omega^T A \omega \quad \text{such that} \quad \omega^T \omega = 1$$

- ω is the eigen vector of A and characterize a good cluster.

Spectral Graph Partitioning/Clustering

- A be the adjacency matrix. Eg. (i) Fully connected with all nodes have degree d; (ii) two separate components of each have degree d_1.
- $Ax = y$; y_i is the sum of neighbours of node i.
- Case 1: $Ax = \lambda x$; $x = [1, 1, \ldots, 1]^T$ and $\lambda = d$
- Case 2: $x' = [1, , \ldots 1, , 0 \ldots 0]^T$ and $\lambda' = d_1$ and $x'' = [0, \ldots, 0, 1, 1, \ldots 1]^T$ and $\lambda'' = d_1$
- $Lx = 0$ is true for a constant vector say $x_i = 1$ vector, with $\lambda = 0$. (Why? D_{ii} is y_i)
- **Fiedler vector:** Eigen vector corresponding to the smallest positive (second smallest) eigen value.
- x_2 is orthogonal to x_1. Some elements of x_2 are positive and some negative. They characterize the two clusters.