Lecture 04/03/16

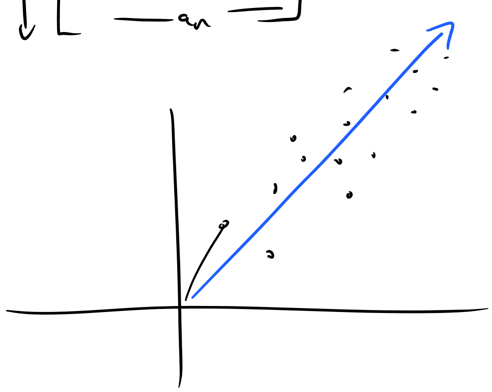
Friday, March 4, 2016 2:56 PM

Recap of Course

Problems	Algorithms	Techniques
Graph Partitioning	Hashing	Concentration (Markov, Chernoff)
{ Min Cut, Multiway Cut }	Johnson-Lindenstrauss	Matrix Approximation
Max Flow	SVD → data compression; → noise reduction	Matrix Perturbation
Facility location	Power Iteration	Balls & Bins
Machine Scheduling	LP rounding ↳ Primal only ↳ Primal-dual	Gaussian Random Variable (χ^2 distribution)
Compressed Sensing	Probability amplification by "anding" (LSH)	Duality
		Sparsity of BFS of LP
		RIP of matrices
		Union bound over ∞ vectors
		Extremal arguments
		Epsilon-Net

SVD

$$
\begin{array}{c c}
 & \xleftarrow{\quad d \quad} \\
n \updownarrow & \begin{bmatrix} — a_1 — \\ — a_2 — \\ \vdots \\ — a_n — \end{bmatrix}
\end{array}
$$



v_1 = 1st singular vector

= direction of highest projection^2 of the n points

= min squared projection distance.

v_2 = 2nd singular vector

= direction of highest projection in subspace orthogonal to v_1

} right singular vectors

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T \qquad \sigma_i u_i = A v_i$$

① the k-right singular vectors span the best "k-dimensional" subspace to minimize projection^2 distance. $\forall k$.

e.g.: best-fit plane = span $\{v_1, v_2\}$

↳ greedy is optimal for SVD!

$v_1 \; v_2 \ldots \; v_r$ ← One ordering, SVD finds it!

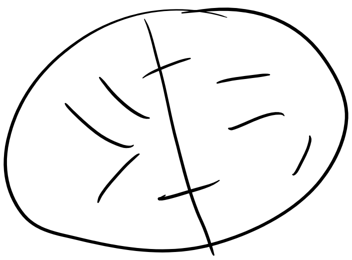
If I knew that inherently, data is 3-d + noise,

$$\boxed{A_3 = \sum_{i=1}^{3} \sigma_i u_i v_i^T}$$

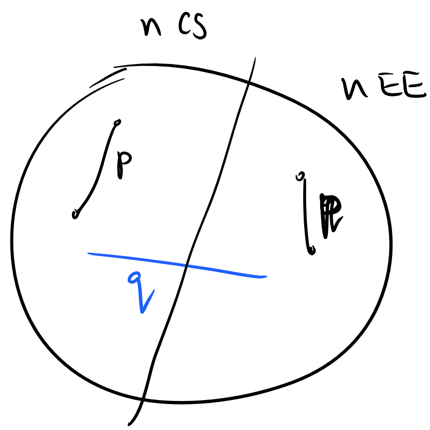
$A \qquad A_1 \quad A_2 \quad A_3 \quad A_4 \ldots$

which $\|A - A_k\|$ is small? — k is a good dimension for dim sequence.

Stochastic Block Models



← generated by an unknown partitioning



$p > q$

From graph, can we recover the CS vs EE partitioning?

Ans. depends on p & q. (eg) If $p = q$, not possible

how separated p & q are.

Counting # common neighbors → Combinatorial Algo?

More robust Algo, SVD.

$$A(G) = \begin{bmatrix} 1 \,\&\, 0 \\ \text{w.p. } p \text{ or } q \end{bmatrix} \xleftarrow{\text{sampled from}} M = \left[\begin{array}{c|c} p\;p\;p & q\;q\;q \\ p\;p & q \\ \hline q\;q\;q & p\;p\;p \\ q & p\;p \end{array}\right]$$

A is sampled from M.

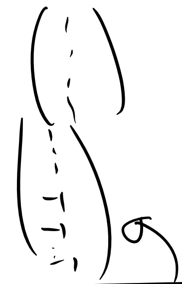
(can view it as "noisy form" of M)

M is rank 2

↳ A_2 = best rank 2 approximation of A (using SVD)

every point = (2-d vector)
cluster these vectors

In M, the first eigenvector



2^{nd} eigenvector

Algo $\boxed{\text{Cluster } 2^{nd} \text{ eigenvector of } A}$

Pf Matrix perturbation theory
$R = A - M$, we showed that $\|R\|_2$ is small

largest absolute value of eigenvectors ← Chernoff bound for matrices.

$\boxed{\text{eigenvectors of } A \text{ \& } M \text{ are close by}}$!

θ_i = angle b/w i^{th} e.v of A & i^{th} e.v of M

$$\sin 2\theta_i \le \frac{2\|R\|}{gap(i)}$$

gap(i)

λ_{i-1} λ_i = i^{th} e.v of M λ_{i+1}

$\boxed{LP}$ ← Motivated
← Defined
← Equational form
General Form
BFS = (vertex soln. = Extreme point)
Simplex Algo (sufficient to explore BFS)
Duality

Machine Scheduling → 2-approx for scheduling on unrelated M/c.

In optimal BFS,
non-zero variables ≤ # linearly independent tight constraints

Scheduling problem,
~~Strath~~ ≤ Non-zero variables in any extreme point

Scheduling problem,

~~Start~~ ← Non-zero variables in <u>any</u> extreme point

looked like a tree + one cycle

Using tree structure, , max load increases by ≤ 1 job for each machine

---

Compressed Sensing → RIP (gaussian concentration)
→ LP optimization is easy.

$y = A x$ underdetermined system

$m \ll n$

+ assumption that $\|x\|_0 \leq k$

$\Rightarrow$ I can recover <u>uniquely</u> x from y.

A matrix = Sensing matrix = "how to mix the input coordinates" in our measurements.

If A is $\delta_{2k} \leq 1/3$ RIP,

then LP can recover x from y

$$\min \|x\|_1 \leftarrow \text{linear objective}$$

$$Ax = y \leftarrow \text{linear constraints}$$

Random gaussian A has $\delta_{2k} \leq 1/3$ RIP

as long as $m \geq \Omega(k \log n)$

JL like proof

Union bound over all unit x, $\|x\|_0 \leq k$.

Union bound over an <u>ϵ-net</u> N

From N, all vectors are also well-behaved
