Tuesday, March 31

- James Enouen (USC)
  - ↳ Estimating Shapley values (my research interest)
  - ↳ Thursday 12pm in Roberts North 15

Today

Suyog Soti (Databricks) !! :)

↳ Fairview → CU Boulder → Google → Databricks

↳ real load balancing

Thursday, April 2

Power Method

Motivation: Compute X_k in $\tilde{O}(ndk)$ time instead of $O(nd^2)$ time

$$X^T X = \sum_{i=1}^{d} v_i v_i^T \sigma_i^2 = A = \sum_{i=1}^{d} v_i v_i^T \lambda_i$$

Goal: Compute z s.t. $\|z - v_i\|_2 \le \epsilon$ in $O\left(nd \frac{\log d/\epsilon}{\gamma}\right)$ time

$$z^{(0)} \sim \mathcal{N}(0, I_d)$$

$$z^{(0)} = c_1^{(0)} v_1 + c_2^{(0)} v_2 + \dots + c_d^{(0)} v_d$$ since $v_1, \dots, v_d$ is a basis

for $t = 1, \dots, m$

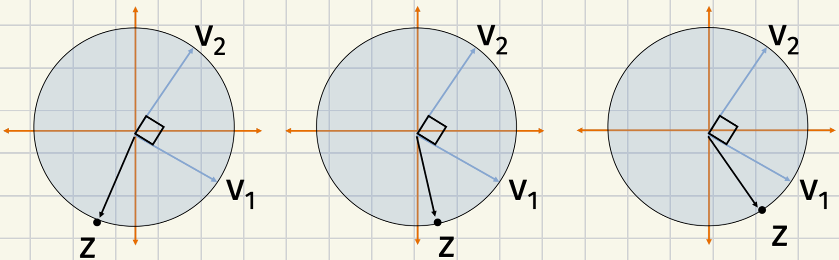
$$n^{(t)} = \|z^{(t)}\|_2$$

$$z^{(t+1)} = A z^{(t)} // n^{(t)}$$

Return $z^{(m)}$ where $m = \frac{\log(d^{3.5}/\epsilon)}{\gamma}$

Geometric Intuition

0 iterations 1 iterations 2 iterations



$$z^{(q)} = \frac{1}{\prod_t \eta^{(t)}} A^m z^{(0)} =$$

$$z^{(q)} = \frac{1}{\prod_t \eta^{(t)}} A^m z^{(0)}$$

Fact: $\left|\frac{c_j^{(0)}}{c_1^{(0)}}\right| \le d^3$ whp

Fact: $(1-\gamma)^{1/\gamma} \le 1/e$

$$\gamma = \frac{\lambda_1 - \lambda_2}{\lambda_1} \le \frac{\lambda_1 - \lambda_j}{\lambda_1}$$

$$\left|\frac{c_j^{(q)}}{c_1^{(q)}}\right| = \left|\frac{c_j^{(0)}}{c_1^{(0)}}\right| \left(\frac{\lambda_j}{\lambda_1}\right)^q \le d^3 \left(1 - \frac{\lambda_1 - \lambda_j}{\lambda_1}\right)^m \le d^3 \left(\left(1-\gamma\right)^{1/\gamma}\right)^{\log(d^3 \sqrt{d/\epsilon})} \le \sqrt{\epsilon/d}$$

Since $\sum_{i=1}^{d} (c_i^{(q)})^2 = 1 = (c_1^{(q)})^2 \left(1 + \sum_{i=2}^{d} \left|\frac{c_i^{(q)}}{c_1^{(q)}}\right|^2\right) \le (c_1^{(q)})^2 (1 + \epsilon/2)$

$\Rightarrow |c_1^{(q)}| \ge 1 - \epsilon$ then $\| z^{(q)} - v_1 \|_2^2 = 2 - 2\langle z^{(q)}, v_1 \rangle \overset{\text{WLOG}}{\le} 2 - 2(1-\epsilon) \le 2\epsilon$

hiding some constants

Krylov Supspace

Power method just uses last vector. Can we do better?

$$K = \{ z^{(0)}, A z^{(0)}, A^2 z^{(0)}, \ldots, A^{m-1} z^{(0)} \}$$

Lanczos Method

$z^{(0)} \sim \mathcal{N}(0, I)$ initial start

Compute $Q \in \mathbb{R}^{d \times m}$, an orthonormal basis for K

Form $T = Q^T A Q$ (T is tridiagonal, prove for hw)

Let z be top eigenvector of T

Return $y^* = Q z$

Analysis:

1. $y^* = Qz$ is the best vector in $\mathcal{K}$ i.e.,

$$y^* = \underset{y}{\operatorname{argmin}} \|X - Xyy^T\|_F^2$$

$$= \underset{y}{\operatorname{argmin}} \|X\|_F^2 - \|Xyy^T\|_F^2 =$$

2. There is a very "good" vector w in $\mathcal{K}$. All w in $\mathcal{K}$

$$w = \alpha_1 z^{(0)} + \alpha_2 A z^{(0)} + \ldots + \alpha_m A^{m-1} z^{(0)}$$

$$= p(A)\, z^{(0)}$$

$$= \sum_{i=1}^{d} p(\lambda_i)\, v_i v_i^T z^{(0)}$$

$$= \sum_{i=1}^{d} p(\lambda_i)\, v_i\, c_i^{(0)}$$

Goal: Find $p(\lambda)$ s.t. $p(\lambda_j)$ small for $j \neq i$ but $p(\lambda_1)$ large

Answer: Chebyshev polynomial!

Claim: There is degree $m = O(\sqrt{1/\gamma} \log \frac{d}{\epsilon})$ Chebyshev p where $p(1) = 1$ but $|p(\lambda)| \leq \epsilon$ for $0 \leq \lambda \leq 1 - \gamma$

Efficient Lanczos

$$T = \begin{bmatrix} T_{11} & T_{12} & 0 & 0 & 0 \\ T_{21} & T_{22} & T_{23} & 0 & 0 \\ 0 & T_{32} & T_{33} & T_{34} & 0 \\ 0 & 0 & T_{43} & T_{44} & T_{45} \\ 0 & 0 & 0 & T_{54} & T_{55} \end{bmatrix}$$

Goal: Build Q and T efficiently

$$K = \{ z^{(0)}, Az^{(0)}, A^2 z^{(0)}, \dots, A^{m-1} z^{(0)} \}$$

$$\left.\begin{array}{l} q_1, \dots, q_t, q_{t+1}, \dots, q_m \text{ are orthonormal cols. of } Q \\ q_t \text{ in the span of } z^{(0)}, Az^{(0)}, \dots, A^{t-1} z^{(0)} \end{array}\right\} \begin{array}{l} \text{Since } A \text{ symmetric,} \\ T = Q^T A Q \text{ is tridiagonal} \end{array}$$

Naive (Gram-Schmidt)

$q_1 = z^{(0)}$

for $t = 1, \dots, m-1$:

$$q_t = A^{t-1} z^{(0)} - \sum_{i=1}^{t-1} \langle A^{t-1} z^{(0)}, q_i \rangle q_i$$

$$q_t = q_t / \|q_t\|_2$$

Clever

$q_1 = z^{(0)}$

for $t = 1, \dots, m-1$:

$$q_t = A^{t-1} z^{(0)} - \langle q_{t-1}, A^{t-1} z^{(0)} \rangle q_{t-1} - \langle q_{t-2}, A^{t-1} z^{(0)} \rangle q_{t-1}$$

$$q_t = q_t / \|q_t\|_2$$

Why is this enough?

Hint: Inspect $AQ = QT$