CMPSCI 711: More Advanced Algorithms

Vectors 7: Subspace Embeddings and Regression

Andrew McGregor

Last Compiled: April 26, 2018

Subspace Embedding

Recall: Exists distribution D over $\Pi \in \mathbb{R}^{k \times n}$ where $k = O(\epsilon^{-2} \log m)$ such that for any $v_1, \ldots v_m \in \mathbb{R}^n$, with probability $\geq 1 - \delta$,

$$\forall i, j, \quad \|\Pi v_i\|_2^2 = (1 \pm \epsilon)\|v_i\|_2^2 \qquad \text{and} \qquad \|\Pi(v_i - v_j)\|_2^2 = (1 \pm \epsilon)\|v_i - v_j\|_2^2$$

Definition

Let $E \subseteq \mathbb{R}^n$ be a linear subspace of dimension d. We say Π is a subspace embedding for E, if for any unit $x \in E$, $\|\Pi x\|_2^2 = 1 \pm \epsilon$

We'll prove existence of low-dimensional subspace embedding via γ-nets.

Theorem

We say $M = \{y_1, y_2, \ldots\}$ is a γ-net for E if for every unit $x \in E$ there exists $y \in M$ such that

$$\|y - x\|_2 \leq \gamma \ .$$

There exists a γ-net for E of size at most $(1 + 2/\gamma)^d$.

Proof of Theorem: Bounding Size of γ-Net

- Construct a γ-net N for $\mathbb{R}^d$ of size at most $(1+2/\gamma)^d$:
  - While there exists a unit $x \in \mathbb{R}^d$ that is distance greater than γ from all points in N, add x to N.
  - Balls of radius $\gamma/2$ centered at each point in N are disjoint A ball centered at origin of radius $1+\gamma/2$ covers all these $|N|$ balls. Hence

$$|N| \leq \frac{\text{volume of ball of radius } (1+\gamma/2)}{\text{volume of ball of radius } \gamma/2} = \frac{(1+\gamma/2)^d}{(\gamma/2)^d} \leq (1+2/\gamma)^d .$$

- Let A be a matrix whose columns are orthornormal basis for E. Then $M = \{Ax : x \in N\}$ is a γ-net for E of size at most $(1+2/\gamma)^d$:
  - Pick arbitrary unit $z \in E$. Let $x \in \mathbb{R}^d$ be unit vector with $z = Ax$.
  - Let $x' \in N$ such that $\|x - x'\|_2 \leq \gamma$ and $y = Ax' \in M$ then

$$\|z - y\|_2 = \|Ax - Ax'\|_2 = \|x - x'\|_2 \leq \gamma$$

  where second inequality follows since columns of A are orthonormal.

Preserving Distances for Net Implies

Theorem

Let M be a $1/2$-net for E. If for all $y, y' \in M$,

$$\|\Pi y\|_2^2 = 1 \pm \epsilon \qquad \textit{and} \qquad \|\Pi(y - y')\|_2^2 = \|y - y'\|_2^2(1 \pm \epsilon)$$

then Π is a subspace embedding for E.

- Pick arbitrary unit $z \in E$.
- Lemma 1: $z = z_1 + z_2 + \dots$ where $z_i/\|z_i\|_2 \in M$ and $\|z_i\|_2 \leq 2\gamma^{i-1}$.
- Lemma 2: For all i, j

  $$\|\Pi z_i\|_2^2 = \|z_i\|_2^2 \pm \epsilon \|z_i\|_2^2 \quad \text{and} \quad \langle \Pi z_i, \Pi z_j \rangle = \langle z_i, z_j \rangle \pm O(\epsilon)\|z_i\|_2 \|z_j\|_2$$

- Note $\sum_{j \geq 1} \|z_j\|_2^2 = O(1)$ and $\sum_{j \geq 1} \|z_j\|_2 = O(1)$ and so,

$$\begin{aligned}
\|\Pi z\|_2^2 &= \|\Pi \sum_i z_i\|_2^2 \\
&= \sum_i \|\Pi z_i\|_2^2 + 2\sum_{i \neq j} \langle \Pi z_i, \Pi z_j \rangle \\
&= \sum_i \|z_i\|_2^2 + \epsilon \sum_i \|z_i\|_2^2 + 2\sum_{i \neq j} \langle z_i, z_j \rangle \pm O(\epsilon) \sum_{i \neq j} \|z_i\|_2 \|z_j\|_2 \\
&= \|z\|_2^2 + O(\epsilon)
\end{aligned}$$

Proof of Lemma 1

Can write $z = z_1 + z_2 + \ldots$ where $\frac{z_i}{\|z_i\|_2} \in M$ and $\|z_i\|_2 \leq 2\gamma^{i-1}$.

- Let $z_1 \in M$ such that $\|z - z_1\|_2 \leq \gamma$ and note $\|z\|_2 = 1 < 2$.
- Suppose we have chosen $z_1, \ldots, z_{i-1}$ such that

$$\alpha_i := \|z - z_1 - \ldots - z_{i-1}\|_2 \leq \gamma^{i-1}$$

- Pick $y \in M$ with

$$\|(z - z_1 - \ldots - z_{i-1})/\alpha_i - y\|_2 \leq \gamma$$

and let $z_i = \alpha_i y$. Then

$$\alpha_{i+1} := \|z - z_1 - \ldots - z_{i-1} - z_i\|_2 \leq \gamma\alpha_i \leq \gamma^i$$

and so

$$\|z_i\| \leq \|z - z_1 - \ldots - z_{i-1} - z_i\|_2 + \|z - z_1 - \ldots - z_{i-1}\|_2 \leq \gamma^i + \gamma^{i-1} \leq 2\gamma^{i-1} \,.$$

Proof of Lemma 2

- Let $y = z_i/\|z_i\|_2$. Then $\|\Pi z_i\|_2^2 = \|z_i\|_2^2(1 \pm \epsilon)$ because,

$$\frac{\|\Pi z_i\|_2^2}{\|z_i\|_2^2} = \|\Pi y\|_2^2 = 1 \pm \epsilon$$

- Let $y' = z_j/\|z_j\|_2$. Note that

$$\begin{aligned}
\langle \Pi y, \Pi y' \rangle &= \frac{1}{2}\left(\|\Pi y\|_2^2 + \|\Pi y'\|_2^2 - \|\Pi(y - y')\|_2^2\right) \\
\langle y, y' \rangle &= \frac{1}{2}\left(\|y\|_2^2 + \|y'\|_2^2 - \|y - y'\|_2^2\right)
\end{aligned}$$

and corresponding terms on right hand side differ by $O(\epsilon)$,

$$\langle \Pi y, \Pi y' \rangle \langle y, y' \rangle \pm O(\epsilon)$$

Finishing Up

- There exists $\Pi \in \mathbb{R}^{k \times n}$ where

$$k = O(\epsilon^{-2} \log |M|) = O(\epsilon^{-2}(d)$$

  such that for any $y, y' \in M$,

$$\|\Pi y\|_2^2 = (1 \pm \epsilon)\|y\|_2^2 \quad \text{and} \quad \|\Pi(y - y')\|_2^2 = (1 \pm \epsilon)\|y - y'\|_2^2$$

- Previous theorem establishes this is subspace embedding for E.

Application: Regression

- Given $A \in \mathbb{R}^{n \times d}$ and $b \in \mathbb{R}^n$, we want to find $x \in \mathbb{R}^d$ such that $Ax \approx b$, in particular,

$$x_{opt} = \operatorname{argmin}_x \|Ax - b\|_2$$

- Let E be the $d+1$ dimensional subspace spanned by columns of A and b. And let Π be a subspace embedding for E. Let

$$\tilde{x} = \operatorname{argmin}_x \|\Pi Ax - \Pi b\|_2$$

- Then

$$\|\Pi A\tilde{x} - \Pi b\|_2^2 \leq \|\Pi Ax_{opt} - \Pi b\|_2^2 \leq (1+\epsilon)\|Ax_{opt} - b\|_2^2(1+\epsilon)$$

  and

$$\|\Pi A\tilde{x} - \Pi b\|_2^2 \geq (1-\epsilon)\|A\tilde{x} - b\|_2^2$$