University of Edinburgh
INFR11156: Algorithmic Foundations of Data Science (2019)
Solutions of the Coursework

Student ID Number: s1555638

Solutions for Part A

Theorem 1. For a suitable choice of k, with a random initial vector $x \in \mathbb{R}^n$ where every x_i is generated according to $N(0,1)$, with constant probability the output vector x_k satisfies the following

$$\frac{x_k^t B x_k}{x_k^t x_k} \geq (1-\epsilon)\frac{\lambda_1}{1+4n(1-\epsilon)^{2k}}$$

where λ_1 is the largest eigenvalue of matrix B.

Lemma 2. Let $v \in \mathbb{R}^n$ in a way that $||v|| = 1$. Create a vector $x \in \mathbb{R}^n$ such that $x_i \sim N(0,1)$.

$$\mathbb{P}[|\langle x, v\rangle| \geq \tfrac{1}{2}] \geq \Omega(1)$$

alongside with

$$\mathbb{P}[||x||^2 \leq 2n] \geq 1 - e^{\frac{-n}{8}}$$

where the latter states that the max square of the norm for the Gaussian random vector is $2n$ with high probability.

Lemma 3. Let $s_1, \cdots, s_n \sim N(0,1)$ be a set of independent normal random variables and $s_1 = \langle x_1, v_1\rangle$. Then following holds

$$\mathbb{P}[|\tfrac{1}{n}(s_1^2 + \cdots + s_n^2) - 1| \geq \epsilon] \leq e^{\frac{-n\epsilon^2}{8}}$$

Lemma 4. For every $k > 0$ and $\epsilon > 0$ if y is defined by $y = B^k x$ and $x \in \mathbb{R}^n$ where $|\langle x, v_1\rangle| \geq \frac{1}{2}$, the following applies with constant probability

$$\frac{y^T By}{yTy} \geq \frac{(1-\epsilon)\lambda_1}{1+4n(1-\epsilon)^{2k}}$$

Proof of Theorem 1.
By solving Lemma 2, it has been proven that a randomly sampled $x \in \mathbb{R}^n$, where $x_i \sim N(0,1)$ satisfies $|\langle x, v\rangle| \geq \frac{1}{2}$ for any $||v|| = 1$ with constant probability. Using the findings of Lemma 2, Lemma 4 states that

$$\frac{y^T By}{yTy} \geq \frac{(1-\epsilon)\lambda_1}{1+4n(1-\epsilon)^{2k}}$$

As such, Theorem 1 holds due to the fact that $||x||^2 = n$

Proof of Lemma 2.
Initially, we find that $\mathbb{E}[|\langle x, v\rangle|] = 0$ and
$\mathbb{E}[|\langle x, v\rangle^2|] = \sum_{i,j} x_i v_i x_j v_j = ||v||^2 = 1$ where Ω is the sample space of a Cumulative Distribution Function. $\langle x, v\rangle$ is a normal random variable due to the fact that it is a linear combination of independent normal random variables. This means that $\langle x, v\rangle \sim N(0,1)$. Draw a graph of the Cumulative

Distribution Function for the standard normal random variables. A statement can be made from observations that if $s \sim N(0,1)$, then

$$\left[|s| \geq \tfrac{1}{2}\right] \geq \Omega(1)$$

It is important to note for Lemma 4 that this inequality holds for $v = v_1$, where v_1 is eigenvector of eigenvalue λ_1, which means

$$\mathbb{P}\left[|\langle x, v_1\rangle| \geq \tfrac{1}{2}\right] \geq \Omega(1)$$

Additionally, the inequality with $\mathbb{P}\left[||x||^2 \leq 2n\right] \geq 1 - e^{\frac{-n}{8}}$ can be proved using the sum of independent random normal variables and concentrion bounds that come with them.

Proof of Lemma 3.
This lemma is proven using Markov inequality with Chernoff bound while using the moment generating function of χ^2 distributions. This has already been shown in *Lecture 3: High-Dimensional Spaces (2) Proof of Lemma 2, around the middle of (2).* As a result of the operations a new inequality is obtained

$$\mathbb{P}\left[\left|(x_1^2 + \cdots + x_n^2 - n)\right| \geq \epsilon\right] \leq e^{\frac{\epsilon^2}{8n}}.$$

Where $\epsilon = n$ inserting it in the inequality above proves the Lemma.

Proof of Lemma 4.
This identical lemma has been proven in *Lecture 6: Best-fit Subspaces and Singular Value Decomposition (2) Proof of Lemma 3. All four formulas are identical (1), (2), (3), $\square$.*

THIS IS THE END OF MY SOLUTIONS.