Optimization Theory - Bonus Problem

Dustin Leatherman

April 8, 2020

Contents

1 Problem Statement **1**
1.1 Hints 2

2 Solution **2**

1 Problem Statement

Let $A \in R^{m \times N}$, where $m \leq N$.

The matrix A has m rows and N columns.

$$A = [\vec{a}_1 | \vec{a_2} | \vec{a_3} | ... | \vec{a_N}]$$

Suppose each $\|\vec{a_j}\|_2 = 1$ for $j = 1, 2, 3, ..., N$

Let μ be the coherence of A where

$$\mu = \underset{j \neq k}{max} |\langle a_j, a_k \rangle|$$

Show that

$$N^2 \leq m(M + (N^2 - N)\mu^2)$$

By rearranging terms, this can also be expressed as

$$\frac{N - m}{m(N - 1)} \leq \mu^2$$

1.1 Hints

Let $H = AA^T$ and $G = A^T A$.

Let $tr(H)$ be the trace of H.

The following facts are useful:

1. $tr(H) \leq \sqrt{m}\sqrt{tr(HH^T)}$
2. $tr(BB^T) = tr(B^T B)$

2 Solution

Let

$$A^T A = \begin{bmatrix} \langle \vec{a}_1, \vec{a}_1 \rangle & \langle \vec{a}_1, \vec{a}_2 \rangle & ... & \langle \vec{a}_1, \vec{a}_N \rangle \\ \langle \vec{a}_2, \vec{a}_1 \rangle & \langle \vec{a}_2, \vec{a}_2 \rangle & ... & \langle \vec{a}_2, \vec{a}_N \rangle \\ ... & ... & ... & ... \\ \langle \vec{a}_N, \vec{a}_1 \rangle & ... & ... & \langle \vec{a}_N, \vec{a}_N \rangle \end{bmatrix}$$

$$tr(A^T A) = tr(G) = \sum_{i=1}^{N} \langle \vec{a}_i, \vec{a}_i \rangle = \sum_{i=1}^{N} \|a_{ii}\|_2^2 \tag{1}$$

Let the dot product be represented as

$$\langle \vec{a}, \vec{a} \rangle = \vec{a}^T a = \|a\|_2^2$$

Since $\|a_j\| = 1$, then

$$\sum_{i=1}^{N} \|a_{ii}\|_2^2 = N \tag{2}$$

Based on Hint (2),

$$tr(AA^T) = tr(A^T A) = tr(H) = tr(G) = N$$

Let

$$H^T = (AA^T)^T = A^T A = G$$

and

$$G^T = (A^T A)^T = AA^T = H$$

Then $tr(HH^T) = tr(G^T G)$

Given that the trace is defined as the sum of the diagonal elements, then

$$tr(G^T G) = \sum_{i=1}^{N} \langle \vec{a}_i, \vec{a}_i \rangle \tag{3}$$

Using (2) and the definitions for G^T, H^T, Hint (1) can be rewritten as:

$$\begin{aligned} tr(H) \leq& \sqrt{m}\sqrt{tr(HH^T)} \\ N \leq& \sqrt{m}\sqrt{tr(HH^T)} \\ N^2 \leq& m\ tr(HH^T) \\ N^2 \leq& m\ tr(G^T G) \end{aligned} \tag{4}$$

Since this is an inequality, substituting $tr(G^T G)$ for a larger value means that the inequality still holds. Consider the sum of dot products squared. It can be said that $tr(B) \leq \sum_{i,j}^{N} |\langle \vec{b}_i, \vec{b}_j \rangle|^2$ for some square matrix B. Since the trace is the sum of diagonal elements, this can be rewritten as

$$tr(B) \leq tr(B) + \sum_{i \neq j}^{N} |\langle \vec{b}_i, \vec{b}_j \rangle|^2 \tag{5}$$

(4) can be further reduced using (5)

$$\begin{aligned} N^2 \leq& m\ tr(G^T G) \\ N^2 \leq& m(tr(G) + \sum_{i \neq j}^{N} \langle \vec{a}_i, \vec{a}_j \rangle) \\ \frac{N^2}{m} \leq& N + \sum_{i \neq j}^{N} |\langle \vec{a}_i, \vec{a}_j \rangle|^2 \\ \frac{N^2}{m} - N \leq& \sum_{i \neq j}^{N} |\langle \vec{a}_i, \vec{a}_j \rangle|^2 \\ \rightarrow \frac{N(N-m)}{m} \leq& \sum_{i \neq j}^{N} |\langle \vec{a}_i, \vec{a}_j \rangle|^2 \end{aligned} \tag{6}$$

By intuition, the average of a set of numbers is less than the maximum value in that set.

$$\frac{1}{N}\sum_{i=1}^{N} a_i \leq max\ a_i \tag{7}$$

Since $\sum_{i\neq j}^{N} |\langle \vec{a}_i, \vec{a}_j \rangle|^2$ does not include the diagonal elements, there are $N(N-1)$ elements. Thus we can rewrite (6) as

$$\frac{N(N-m)}{m}\frac{1}{N(N-1)}\sum_{i\neq j}^{N} |\langle \vec{a}_i, \vec{a}_j \rangle|^2 \leq \max_{i\neq j} |\langle \vec{a}_i, \vec{a}_j \rangle|^2 \tag{8}$$

Since $\sum_{i\neq j}^{N} |\langle \vec{a}_i, \vec{a}_j \rangle|^2$ can never be negative, its presence in the inequality is not required and can be removed.

Thus

$$\frac{N-m}{m(N-1)} \leq \max_{i\neq j} |\langle a_i, a_j \rangle|^2 = \mu^2 \tag{9}$$

[End]