Assignment 6

5, 5.2 Since our equation is quadratic, all that needs to be shown is the point found is a minimum and not a maximum. We can do this by taking looking at the second derivatives.

Looking at the j'th dimention

$f'(z_j) = \sum -2(a_{ji} - z_i)$

$f''(z_j) = 2m$

Since the second derivatives are all positive, this is a local minimum

5, 5.3 1 $m * \log(x) = \log(a_1 * a_2 * \ldots * a_m)$

$\log(x) = \frac{1}{m}\log(a_1 * a_2 * \ldots * a_m)$

$x = 10^{\frac{1}{m}}(a_1 * a_2 * \ldots * a_m)$

5, 5.3 2 I was mostly just typing to myself while I was thinking about the equation, looking at the individual pieces of the summation, and I forgot to delete that line before I submitted it.

5, 5.4 I understand your point. I wasn't specifically ignoring the other partials, just didn't think about them at the time…which I'm not sure is any better…

6.2 First show that if v' is orthogonal to x, then $|c - v| \leq |c - v'|$

We know $c = v + xc^T x,\ \ v = c - xc^T x,\ \ {v'}^T x = 0,\ \ x^T v' = 0,\ \ x^T x = 1$

Assume to the contrary $|c - v| > |c - v'|$, then $|c - v|^2 > |c - v'|^2$

=> $|xc^T x|^2 > |(v - v') + xc^T x|^2$ (by substitution)

$> \left((v - v') + xc^T x\right) \cdot ((v - v') + xc^T x)$

$> (v - v') \cdot (v - v') + 2(v - v') \cdot (xc^T x) + (xc^T x) \cdot (xc^T x)$

$> |v - v'|^2 + |xc^T x|^2 + 2v^T xc^T x - 2{v'}^T xc^T x$

$> |v - v'|^2 + |xc^T x|^2 + 2x^T vc^T x - 0$

$2x^T vc^T x = 2x^T (c - xc^T x)c^T x = 2x^T cc^T x - 2x^T xc^T xc^T x$

$$2x^T cc^T x - 2x^T cc^T x = 0$$
$$> |v - v'|^2 + |xc^T x|^2$$

Since $|v - v'|^2 \geq 0$, this is a contradiction and so $|c - v| \leq |c - v'|$

6.3 Next, show that for any vector w, $\sum (a_i - c)^T (c - w)$ from $i = 1$ to m is equal to 0.

$$\sum (a_i - c)^T (c - w) = \sum (a_i - c) \cdot (c - w) = \sum (a_i \cdot c - a_i \cdot w - c \cdot c + c \cdot w) =$$

$$(\sum a_i \cdot c - mc \cdot c) - (\sum w \cdot a_i - mc \cdot w) =$$

$$\frac{1}{m}\left(\left(c \cdot \frac{a_1 + a_2 + \cdots + a_m}{m}\right) - c \cdot c\right) - \frac{1}{m}\left(\left(w \cdot \frac{a_1 + a_2 + \cdots + a_m}{m}\right) - w \cdot c\right) =$$

$$\frac{1}{m}(c \cdot c - c \cdot c) - \frac{1}{m}(w \cdot c - w \cdot c) = 0$$

Finally we can use these two properties to show $\sum |a_i - v|^2 \leq \sum |a_i - v'|^2$ from $i = 1$ to m, for each v' such that $x^T v' = 0$

i. $\sum |a_i - v'|^2 = \sum |a_i - c|^2 + 2 \sum (a_i - c)^T (c - v') + \sum |c - v'|^2$
ii. $\sum |a_i - v|^2 = \sum |a_i - c|^2 + 2 \sum (a_i - c)^T (c - v) + \sum |c - v|^2$

Since $\sum |a_i - c|^2$ is constant, $2 \sum (a_i - c)^T (c - v)$ and $2 \sum (a_i - c)^T (c - v') = 0$, we are left comparing $+\sum |c - v'|^2$ and $\sum |c - v|^2$.

As shown in part 6.2, $|c - v| \leq |c - v'|$, so $|c - v|^2 \leq |c - v'|^2$ and $\sum |c - v|^2 \leq \sum |c - v'|^2$

Thus $\sum |a_i - v|^2 \leq \sum |a_i - v'|^2$ from $i = 1$ to m, for each v' such that $x^T v' = 0$

So $v = c - xc^T x$ minimizes $\sum |a_i - z|^2$ subject to $z^T x = 0$.

6.4 Simplify $\sum b_i b_i^T$ from $i = 1$ to m

Each $b_i b_i^T$ is an $n\ x\ n$ symmetric matrix. Therefore their sum is also a symmetric matrix.

We can write $\sum b_i b_i^T = BB^T = A$, where A is an $n\ x\ n$ symmetric matrix, I believe this was shown in assignment 2

6.5 Solve max $x^T \left(\sum b_i b_i^T\right) x$ from $i = 1$ to m subject to $x^T x = 1$

This problem becomes maximizing $x^T A x$ subject to $x^T x = 1$ where A is a symmetric matrix defined above

Also shown in assignment 2, since A is symmetric it has all real eigenvalues.

Furthermore, a symmetric matrix A can be written as $O^T DO$ where O is an orthogonal matrix where the columns are the eigenvectors of A, and the diagonals elements of D are the are the eigenvalues of A (source: wolfram's description of symmetric matrices)

Now the equation becomes max $x^T O^T DOx = (Ox)^T D(Ox)$

Since O is orthogonal, $O^T O = I$, so $(Ox)^T(Ox) = x^T O^T Ox = x^T x = 1$, meaning $|Ox|^2 = |x|^2$ and we can just try and maximize $y^T Ay$, where $y = Ox$ and then backtrack to find x

Rewrite $y^T Ay = \sum y_i \lambda y_i = \sum \lambda_i y_i^2$ from $i = 1$ to n

The constraint $y^T y = 1$ means that $\sum y_i^2 = 1$ from $i = 1$ to n

This equation is maximized when there is a 1 multiplied with the largest λ_i value, or the standard basis vector with a 1 in the dimension with the largest eigenvalue.

O is an orthogonal matrix of the eigenvector of A columns, so to create this maximum scenario, x must be the largest (normalized) eigenvector of A. This maximizes the dot product of x and the row in O that will be multiplied by the largest eigenvalue value of A.