DS-GA 1014 - Homework 9

Eric Niblock

November 6th, 2020

1. (2 points). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. We assume that the minimum $m \stackrel{def}{=} min_{\overrightarrow{\mathbf{x}} \in \mathbb{R}^n} f(\overrightarrow{\mathbf{x}})$ of f on $\mathbb{R}^n$ is finite, and that the set of minimizers of f

$$\mathcal{M} \stackrel{def}{=} \{\overrightarrow{\mathbf{v}} \in \mathbb{R}^n | f(\overrightarrow{\mathbf{v}}) = m\}$$

is non-empty.

(a) Show that $\mathcal{M}$ is a convex set.

Since f is convex, we have that for any $\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}} \in \mathbb{R}^n$, and $\alpha \in [0, 1]$,

$$f(\alpha\overrightarrow{\mathbf{x}} + (1-\alpha)\overrightarrow{\mathbf{y}}) \leq \alpha f(\overrightarrow{\mathbf{x}}) + (1-\alpha)f(\overrightarrow{\mathbf{y}}) \tag{1}$$

Now suppose we choose two vectors from $\mathcal{M}$ and call them $\overrightarrow{\mathbf{v}}_1, \overrightarrow{\mathbf{v}}_2$. We know that $m = f(\overrightarrow{\mathbf{v}}_1) = f(\overrightarrow{\mathbf{v}}_2)$ since this is a condition of being in set $\mathcal{M}$. Furthermore, since f is convex, we have,

$$f(\alpha\overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) \leq \alpha f(\overrightarrow{\mathbf{v}}_1) + (1-\alpha)f(\overrightarrow{\mathbf{v}}_2) \tag{2}$$

$$f(\alpha\overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) \leq \alpha m + (1-\alpha)m \tag{3}$$

$$f(\alpha\overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) \leq m \tag{4}$$

However, $f(\alpha\overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) < m$ is not possible, since we defined m to be the minimum value of f for all $\overrightarrow{\mathbf{v}}$ (and the set of all $\overrightarrow{\mathbf{v}}$ includes $\alpha\overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2$). Hence, it must be the case that $f(\alpha\overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) = m$.

Note the original goal is to prove that $\mathcal{M}$ is a convex set, which is true if for any $\overrightarrow{\mathbf{v}}_1, \overrightarrow{\mathbf{v}}_2 \in S$, and $\alpha \in [0, 1]$, we have,

$$\alpha \overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2 \in S \tag{5}$$

The condition on $\alpha \overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2$ being in S is such that,

$$f(\alpha \overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) = m \tag{6}$$

Which has already been shown. Thus $\mathcal{M}$ is a convex set.

(b) Show that if f is strictly convex, then $\mathcal{M}$ has only one element.

We will proceed by contradiction. Suppose that $\mathcal{M}$ has two or more elements. If this is the case then there exists a $\overrightarrow{\mathbf{v}}_1 \neq \overrightarrow{\mathbf{v}}_2$ with $\overrightarrow{\mathbf{v}}_1, \overrightarrow{\mathbf{v}}_2 \in \mathbb{R}^{\mathrm{n}}$ such that,

$$f(\overrightarrow{\mathbf{v}}_1) = f(\overrightarrow{\mathbf{v}}_2) = m \tag{7}$$

Since this is a condition of being in $\mathcal{M}$. Furthermore, we have the definition of strict convexity, for any $\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}} \in \mathbb{R}^{\mathrm{n}}$, and $\alpha \in (0,1)$, we have

$$f(\alpha \overrightarrow{\mathbf{x}} + (1-\alpha)\overrightarrow{\mathbf{y}}) < \alpha f(\overrightarrow{\mathbf{x}}) + (1-\alpha) f(\overrightarrow{\mathbf{y}}) \tag{8}$$

Therefore, we have,

$$f(\alpha \overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) < \alpha f(\overrightarrow{\mathbf{v}}_1) + (1-\alpha)\overrightarrow{\mathbf{v}}_2 \tag{9}$$

$$f(\alpha \overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2) < m \tag{10}$$

However, this cannot be the case, because then, $f(\alpha \overrightarrow{\mathbf{v}}_1 + (1-\alpha)\overrightarrow{\mathbf{v}}_2)$ would be less than the minimum of f. Contradiction. Therefore, there cannot exist more than one element in $\mathcal{M}$ if f is strictly convex.

2. **(2 points). Let $M \in \mathbb{R}^{n\times n}$ be a symmetric matrix, $\overrightarrow{\mathbf{b}} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. For $\overrightarrow{\mathbf{x}} \in \mathbb{R}^n$ we define**

$$f(x) = \overrightarrow{\mathbf{x}}^T M \overrightarrow{\mathbf{x}} + \langle \overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{b}} \rangle + c$$

f is called a quadratic function.

(a) **Compute the gradient $\nabla f(\overrightarrow{\mathbf{x}})$ and the Hessian $H_f(\overrightarrow{\mathbf{x}})$ at all $\overrightarrow{\mathbf{x}} \in \mathbb{R}^n$. Show that f is convex if and only if M is positive semi-definite.**

We can re-write $f(\overrightarrow{\mathbf{x}})$ as,

$$f(\overrightarrow{\mathbf{x}}) = \overrightarrow{\mathbf{x}}^T M \overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{x}}^T \overrightarrow{\mathbf{b}} + c \tag{11}$$

Then we have,

$$\begin{aligned} \nabla f(\overrightarrow{\mathbf{x}}) &= \nabla(\overrightarrow{\mathbf{x}}^T M \overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{x}}^T \overrightarrow{\mathbf{b}} + c) \\ &= \nabla(\overrightarrow{\mathbf{x}}^T M \overrightarrow{\mathbf{x}}) + \nabla(\overrightarrow{\mathbf{x}}^T \overrightarrow{\mathbf{b}}) \\ &= 2M\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{b}} \end{aligned} \tag{12}$$

And furthermore, we can calculate the Hessian by taking the Jabobian of the gradient,

$$H_f(\overrightarrow{\mathbf{x}}) = \mathbf{J}[\nabla f(\overrightarrow{\mathbf{x}})] = 2M \tag{13}$$

We also have the following proposition,

> *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice-differentiable function. We denote by H_f the Hessian matrix of f. Then f is convex if and only if for all $x \in \mathbb{R}^n$, $H_f(x)$ is positive semi-definite. [Prop. 1]*

Therefore, if we show that when M is positive semi-definite, it is equivalent to H_f being positive semi-definite, we will have proved the original statement. If M is positive semi-definite, we have,

$$\overrightarrow{\mathbf{x}}^T M \overrightarrow{\mathbf{x}} \geq 0 \tag{14}$$

$$\overrightarrow{\mathbf{x}}^T(2M)\overrightarrow{\mathbf{x}} \geq 2(0) = 0 \tag{15}$$

$$\overrightarrow{\mathbf{x}}^T H_f(\overrightarrow{\mathbf{x}})\overrightarrow{\mathbf{x}} \geq 0 \tag{16}$$

Therefore, when M is positive semi-definite, this is equivalent to $H_f(\overrightarrow{\mathbf{x}})$ being positive semi-definite, and we achieve the desired result: f is convex if and only if M is positive semi-definite.

(b) In this question, we assume M to be positive semi-definite. Show that f admits a minimizer if and only if $\overrightarrow{\mathbf{b}} \in Im(M)$.

The fact that f admits a minimizer is equivalent to saying that $\nabla f(\overrightarrow{\mathbf{x}}) = 0$ for some $\overrightarrow{\mathbf{x}}$. Equivalently, it must be the case that for some $\overrightarrow{\mathbf{x}}$

$$\nabla f(\overrightarrow{\mathbf{x}}) = 0 \tag{17}$$

$$2M\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{b}} = 0 \tag{18}$$

$$2M\overrightarrow{\mathbf{x}} = -\overrightarrow{\mathbf{b}} \tag{19}$$

$$M\overrightarrow{\mathbf{x}} = -\frac{\overrightarrow{\mathbf{b}}}{2} \tag{20}$$

This equation is equivalent to the notion that $\overrightarrow{\mathbf{b}}$ is in the image of M, since if it were not, there would be no $\overrightarrow{\mathbf{x}}$ that satisfies this equation. Since all of the statements used are equivalencies, we have that f admits a minimizer if and only if $\overrightarrow{\mathbf{b}} \in Im(M)$.

3. (3 points). We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex if there exists $\alpha > 0$ such that the function $\vec{\mathbf{x}} \to f(x) - \frac{\alpha}{2}||\vec{\mathbf{x}}||^2$ is convex. In other words, f is strongly convex if there exists $\alpha > 0$ and a convex function $g : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x) = g(x) + \tfrac{\alpha}{2}||\vec{\mathbf{x}}||^2$$

(a) Show that a strongly convex function is strictly convex. (Hint: start by showing that $\vec{\mathbf{x}} \to ||\vec{\mathbf{x}}||^2$ is strictly convex).

First, we define a function,

$$h(\vec{\mathbf{x}}) = \frac{\alpha}{2}||\vec{\mathbf{x}}||^2 = \frac{\alpha}{2}\langle \vec{\mathbf{x}}, \vec{\mathbf{x}} \rangle = \frac{\alpha}{2}\vec{\mathbf{x}}^T\vec{\mathbf{x}} \tag{21}$$

Where it is evident that we have,

$$\nabla h(x) = \alpha \vec{\mathbf{x}}^T \tag{22}$$

$$H_h(x) = \alpha \, Id_n \tag{23}$$

And furthermore, we have the following proposition,

If for all $x \in \mathbb{R}^n$, the Hessian $H_f(x)$ is positive definite, then f is strictly convex. [Prop. 2]

Additionally, if we are to show that $H_h(x)$ is positive definite, then it must be the case that,

$$\vec{\mathbf{x}}^T H_f(x) \vec{\mathbf{x}} > 0 \tag{24}$$

For all $\vec{\mathbf{x}} \in \mathbb{R}^n$ (with the exception of $\vec{\mathbf{x}} = \vec{\mathbf{0}}$). So then we have,

$$\vec{\mathbf{x}}^T(\alpha \, Id_n)\vec{\mathbf{x}} > 0 \tag{25}$$

$$\alpha \vec{\mathbf{x}}^T(Id_n \, Id_n)\vec{\mathbf{x}} > 0 \tag{26}$$

$$\alpha \vec{\mathbf{x}}^T(Id_n^T \, Id_n)\vec{\mathbf{x}} > 0 \tag{27}$$

$$\alpha(\overrightarrow{\mathbf{x}}^T Id_n^T)(Id_n \overrightarrow{\mathbf{x}}) > 0 \tag{28}$$

$$\alpha(Id_n \overrightarrow{\mathbf{x}})^T (Id_n \overrightarrow{\mathbf{x}}) > 0 \tag{29}$$

$$\alpha||Id_n \overrightarrow{\mathbf{x}}||^2 > 0 \tag{30}$$

$$\alpha||\overrightarrow{\mathbf{x}}||^2 > 0 \tag{31}$$

Which is obviously true for all $\overrightarrow{\mathbf{x}} \in \mathbb{R}^n$ and $\alpha > 0$ (with the exception of $\overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{0}}$). This shows that $h(x)$ is strictly convex.

We now have the following problem to solve. Given that $g(x)$ is a convex function and $h(x)$ is a strictly convex function, show that $f(x) = g(x) + h(x)$ is a strictly convex function. We can write the following for all $\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}} \in \mathbb{R}^n$, and $t \in (0, 1)$,

$$\begin{aligned} f(t\overrightarrow{\mathbf{x}} + (1-t)\overrightarrow{\mathbf{y}}) &= g(t\overrightarrow{\mathbf{x}} + (1-t)\overrightarrow{\mathbf{y}}) + h(t\overrightarrow{\mathbf{x}} + (1-t)\overrightarrow{\mathbf{y}}) \\ &\leq tg(\overrightarrow{\mathbf{x}}) + (1-t)g(\overrightarrow{\mathbf{y}}) + h(t\overrightarrow{\mathbf{x}} + (1-t)\overrightarrow{\mathbf{y}}) \\ &< tg(\overrightarrow{\mathbf{x}}) + (1-t)g(\overrightarrow{\mathbf{y}}) + th(\overrightarrow{\mathbf{x}}) + (1-t)h(\overrightarrow{\mathbf{y}}) \\ &= tf(\overrightarrow{\mathbf{x}}) + (1-t)f(\overrightarrow{\mathbf{y}}) \end{aligned} \tag{32}$$

Where the first and last lines follow directly from the definition of f, the second line follows from g being convex, and the third line follows from h being strictly convex. The inequality then forces,

$$f(t\overrightarrow{\mathbf{x}} + (1-t)\overrightarrow{\mathbf{y}}) < tf(\overrightarrow{\mathbf{x}}) + (1-t)f(\overrightarrow{\mathbf{y}}) \tag{33}$$

Which proves that f, a strongly convex function, is strictly convex.

(b) Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. Show that ϕ is strongly convex if and only if there exists $\alpha > 0$ such that for all $\overrightarrow{\mathbf{x}} \in \mathbb{R}^n$ the eigenvalues of $H_\phi(x)$ are greater or equal than α.

First we note that strong convexity on ϕ implies that,

$$\phi(\overrightarrow{\mathbf{x}}) = g(\overrightarrow{\mathbf{x}}) + \frac{\alpha}{2}||\overrightarrow{\mathbf{x}}||^2 \tag{34}$$

$$H_\phi(\overrightarrow{\mathbf{x}}) = H_g(\overrightarrow{\mathbf{x}}) + \alpha Id_n \tag{35}$$

Now, $\phi(\overrightarrow{\mathbf{x}})$ being strongly convex is equivalent to

$$\phi(\overrightarrow{\mathbf{x}}) - \frac{\alpha}{2}||\overrightarrow{\mathbf{x}}||^2 \tag{36}$$

Being convex, which furthermore, is equivalent to,

$$H_\phi(\overrightarrow{\mathbf{x}}) - \alpha Id_n \tag{37}$$

Being positive semi-definite for some choice of α (by [*Proposition 1*]). Then, taking the smallest eigenvalue of $H_\phi(\overrightarrow{\mathbf{x}})$, λ_{min} and its corresponding eigenvector, $\overrightarrow{\mathbf{v}}$, we can write the following expressions,

$$H_\phi(x)\overrightarrow{\mathbf{v}} = \lambda_{min}\overrightarrow{\mathbf{v}} \tag{38}$$

$$H_\phi(x)\overrightarrow{\mathbf{v}} - \alpha Id_n\overrightarrow{\mathbf{v}} = \lambda_{min}\overrightarrow{\mathbf{v}} - \alpha Id_n\overrightarrow{\mathbf{v}} \tag{39}$$

$$(H_\phi(x) - \alpha Id_n)\overrightarrow{\mathbf{v}} = (\lambda_{min} - \alpha)\overrightarrow{\mathbf{v}} \tag{40}$$

But since we know that when $H_\phi(\overrightarrow{\mathbf{x}}) - \alpha Id_n$ is convex, and therefore positive semi-definite, all of the eigenvalues associated with it are non-negative. Therefore, $\phi(\overrightarrow{\mathbf{x}})$ is strongly convex, if and only if all of the eigenvalues associated with $H_\phi(\overrightarrow{\mathbf{x}})$ are greater than or equal to α.

4. (3 points). Let $A \in \mathbb{R}^{n \times m}$ and $\overrightarrow{\mathbf{y}} \in \mathbb{R}^{n}$. For $\overrightarrow{\mathbf{x}} \in \mathbb{R}^{m}$ we define

$$f(x) = ||A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}}||^2$$

(a) Compute the gradient $\nabla f(x)$ and the Hessian $H_f(x)$ at all $\overrightarrow{\mathbf{x}} \in \mathbb{R}^{m}$. Show that f is convex.

First, we have that,

$$\begin{aligned}
\nabla f(x) &= \nabla\left(||A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}}||^2\right) = \nabla\left((A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}})^T(A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}})\right) \\
&= \nabla\left((\overrightarrow{\mathbf{x}}^T A^T - \overrightarrow{\mathbf{y}}^T)(A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}})\right) \\
&= \nabla\left(\overrightarrow{\mathbf{x}}^T A^T A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{x}}^T A^T\overrightarrow{\mathbf{y}} - \overrightarrow{\mathbf{y}}^T A\overrightarrow{\mathbf{x}} + \overrightarrow{\mathbf{y}}^T\overrightarrow{\mathbf{y}}\right) \\
&= \nabla\left(\overrightarrow{\mathbf{x}}^T A^T A\overrightarrow{\mathbf{x}}\right) - \nabla\left(\overrightarrow{\mathbf{x}}^T A^T\overrightarrow{\mathbf{y}} - \overrightarrow{\mathbf{y}}^T A\overrightarrow{\mathbf{x}}\right) \\
&= \nabla\left(\overrightarrow{\mathbf{x}}^T A^T A\overrightarrow{\mathbf{x}}\right) - A^T\overrightarrow{\mathbf{y}} - (\overrightarrow{\mathbf{y}}^T A)^T \\
&= \nabla\left((A\overrightarrow{\mathbf{x}})^T A\overrightarrow{\mathbf{x}}\right) - 2A^T\overrightarrow{\mathbf{y}} \\
&= 2A^T A\overrightarrow{\mathbf{x}} - 2A^T\overrightarrow{\mathbf{y}} \\
&= 2A^T(A\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}})
\end{aligned} \tag{41}$$

Which would mean that the Hessian, $H_f(x)$, is simply,

$$H_f(x) = \mathbf{J}[\nabla f(\overrightarrow{\mathbf{x}})] = 2A^T A \tag{42}$$

Now, from *Proposition 1*, we know that f is convex if $H_f(x)$ is positive semi-definite. For all $\overrightarrow{\mathbf{x}} \in \mathbb{R}^{n}$ we must show that,

$$\overrightarrow{\mathbf{x}}^T H_f(x)\overrightarrow{\mathbf{x}} \geq 0 \tag{43}$$

However, this is clearly true, since,

$$\overrightarrow{\mathbf{x}}^T H_f(x)\overrightarrow{\mathbf{x}} \geq 0 \tag{44}$$

$$2\overrightarrow{\mathbf{x}}^T(A^T A)\overrightarrow{\mathbf{x}} \geq 0 \tag{45}$$

$$2\overrightarrow{\mathbf{x}}^T(A^T A)\overrightarrow{\mathbf{x}} \geq 0 \tag{46}$$

$$2(\overrightarrow{\mathbf{x}}^T A^T)(A\overrightarrow{\mathbf{x}}) \geq 0 \tag{47}$$
$$2(A\overrightarrow{\mathbf{x}})^T(A\overrightarrow{\mathbf{x}}) \geq 0 \tag{48}$$
$$2||A\overrightarrow{\mathbf{x}}||^2 \geq 0 \tag{49}$$

Which is obviously true, since the norm of $A\overrightarrow{\mathbf{x}}$ will be non-negative for all values of $\overrightarrow{\mathbf{x}}$. Therefore, $H_f(x)$ is positive semi-definite and f is convex.

(b) Show that if $rank(A) < m$, then f is not strictly convex.

If $rank(A) < m$, then A is not full-rank, and $\mathcal{N}(A)$ is populated with at least one non-trivial $\overrightarrow{\mathbf{x}}$. In other words, there exists some $\overrightarrow{\mathbf{x}} \neq 0$ such that,

$$A\overrightarrow{\mathbf{v}} = \overrightarrow{\mathbf{0}} \tag{50}$$

Then we have that,

$$\begin{aligned} f(\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}) &= ||A(\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}) - \overrightarrow{\mathbf{y}}||^2 \\ &= ||A\overrightarrow{\mathbf{u}} + A\overrightarrow{\mathbf{v}} - \overrightarrow{\mathbf{y}}||^2 \\ &= ||A\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{y}}||^2 = f(\overrightarrow{\mathbf{u}}) \end{aligned} \tag{51}$$

Now, in order to show that f is strictly convex, it must be the case that for all $\overrightarrow{\mathbf{v}}, \overrightarrow{\mathbf{u}} \in \mathbb{R}^n$ and $t \in (0, 1)$.

$$f(t\overrightarrow{\mathbf{v}} + (1-t)\overrightarrow{\mathbf{u}}) < tf(\overrightarrow{\mathbf{v}}) + (1-t)f(\overrightarrow{\mathbf{u}}) \tag{52}$$

If we use $\overrightarrow{\mathbf{u}} = \overrightarrow{\mathbf{0}}$ and $\overrightarrow{\mathbf{v}} \in \mathcal{N}(A)$, we get,

$$||A(t\overrightarrow{\mathbf{v}} + (1-t)\overrightarrow{\mathbf{u}}) - \overrightarrow{\mathbf{y}}||^2 < t||A\overrightarrow{\mathbf{v}} - \overrightarrow{\mathbf{y}}||^2 + (1-t)||A\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{y}}||^2 \tag{53}$$
$$||tA\overrightarrow{\mathbf{v}} + (1-t)A\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{y}}||^2 < t||A\overrightarrow{\mathbf{v}} - \overrightarrow{\mathbf{y}}||^2 + (1-t)||A\overrightarrow{\mathbf{u}} - \overrightarrow{\mathbf{y}}||^2 \tag{54}$$
$$||-\overrightarrow{\mathbf{y}}||^2 < t||-\overrightarrow{\mathbf{y}}||^2 + (1-t)||-\overrightarrow{\mathbf{y}}||^2 \tag{55}$$
$$||-\overrightarrow{\mathbf{y}}||^2 < ||-\overrightarrow{\mathbf{y}}||^2 \tag{56}$$

This is clearly not true. So, if $rank(A) < m$, f is not strictly convex.

(c) Show that is $rank(A) = m$, then f is strongly convex (use the definition and results of *Problem 9.3*).

If we can show that,

$$H_f(\overrightarrow{\mathbf{x}}) - \alpha Id_n \tag{57}$$

$$2A^T A - \alpha Id_n \tag{58}$$

Is positive semi-definite, then we will have shown that f is strongly convex (*Problem 3b*). Since $rank(A) = m$, we know that $A^T A$ is positive-definite, since,

$$\overrightarrow{\mathbf{x}}^T A^T A \overrightarrow{\mathbf{x}} > 0 \tag{59}$$

$$(A\overrightarrow{\mathbf{x}})^T A \overrightarrow{\mathbf{x}} > 0 \tag{60}$$

$$||A\overrightarrow{\mathbf{x}}||^2 > 0 \tag{61}$$

Because $A\overrightarrow{\mathbf{x}} \neq 0$ for all $\overrightarrow{\mathbf{x}} \neq \overrightarrow{\mathbf{0}}$. Therefore, $2A^T A$ has all positive eigenvalues. Taking from *Problem 3b*, we observe the eigenvalue equation relating to the smallest eigenvalue of $2A^T A$ to be

$$(2A^T A - \alpha Id_n)\overrightarrow{\mathbf{v}} = (\lambda_{min} - \alpha)\overrightarrow{\mathbf{v}} \tag{62}$$

And if we choose α to be λ_{min}, then the smallest eigenvalue of $2A^T A - \alpha Id_n$ becomes zero. Hence, $2A^T A - \alpha Id_n$, is positive semi-definite, and f is strongly convex, given that $rank(A) = m$.