EECS 559 HW 1

QUESTION 1:

(a) Given $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{y} \in \mathbb{R}^m$, show that the basis pursuit problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{x}\|_1 \quad \textbf{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$$

is a convex problem, and recast it as a *linear program* in the standard form.

SOLUTION: Firstly, we can show that this is a convex optimization problem by showing that both the objective function and constraints are convex. We start by examining the objective function:

$$\begin{aligned}\|\boldsymbol{x}\|_1 &= \sum_i |x_i| \\ &= \sum_i \max\{x_i, -x_i\}\end{aligned}$$

We know that x_i and $-x_i$ are linear functions, therefore convex, thus the pointwise maximum of these two functions is convex. Since the objective function is a nonnegative sum of these pointwise maxima, it is also convex. The constraint:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$$

is linear, therefore convex. which means that this *is a convex optimization problem.*

We can now recast this as a linear program in the standard form. We start by defining $\boldsymbol{x}_+$ and $\boldsymbol{x}_-$ as follows:

$$\begin{aligned}\boldsymbol{x}_+ &\triangleq \sum_i x_i \mathbb{1}_{\{x_i\geq 0\}} \boldsymbol{e}_i \\ \boldsymbol{x}_- &\triangleq \sum_i -x_i \mathbb{1}_{\{x_i< 0\}} \boldsymbol{e}_i\end{aligned}$$

Where $\boldsymbol{e}_i$ is the i^{th} cannonical basis vector of $\mathbb{R}^n$. Noitce that:

$$\boldsymbol{x} = \boldsymbol{x}_+ - \boldsymbol{x}_-$$

QUESTION 2: Consider the following unconstrained optimization problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$$

Where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex function.

(a) Show that if $\boldsymbol{x}_0$ is a local minimizer of f, then it is also a global solution (need not to be unique). Moreover, if f is *strictly* convex, then show that the solution is unique (if exists).

PROOF: (By contradiction) Let $\boldsymbol{x}_0$ be a local minimizer of f, meaning that $\exists \epsilon > 0$ such that:

$$f(\boldsymbol{x}_0) \leq f(\boldsymbol{x}) \ \ \forall \boldsymbol{x} \in \mathcal{B}(\boldsymbol{x}_0, \epsilon) \triangleq \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}_0 - \boldsymbol{x}\|_2 \leq \epsilon\}$$

Assume $\exists \boldsymbol{x}_* \in \mathbb{R}^n$ such that $f(\boldsymbol{x}_*) < f(\boldsymbol{x}_0)$. $\forall \epsilon$ sufficiently small, $\exists \alpha \in (0, 1)$ such that:

$$\boldsymbol{y} \triangleq \alpha \boldsymbol{x}_0 + (1 - \alpha)\boldsymbol{x}_* \in \mathcal{B}(\boldsymbol{x}_0, \epsilon)$$

And by the convexity of f:

$$f(\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}_0) + (1 - \alpha) f(\boldsymbol{x}_*)$$

Since $f(\boldsymbol{x}_*) < f(\boldsymbol{x}_0)$, we can write:

$$f(\boldsymbol{y}) < \alpha f(\boldsymbol{x}_0) + (1 - \alpha) f(\boldsymbol{x}_0) = f(\boldsymbol{x}_0)$$

Which is when we arrive at our contradiction. Since for any arbitrary $\epsilon > 0$ there is an element $\boldsymbol{y} \in \mathcal{B}(\boldsymbol{x}_0, \epsilon)$ such that $f(\boldsymbol{x}_0) > f(\boldsymbol{y})$, our assumption is false. Therefore, any local minimum is a global minimum. □

Let's now assume that f is *strictly convex*. Assuming it exists, we can show the uniqueness of the global minimizer as follows.

PROOF: (By contradiction) Let $\boldsymbol{x}_0$ is a local minimizer of f. Since strict complexity implies convexity, we know that $\boldsymbol{x}_0$ is also a global minimizer (from the previous argument).

$$f(\boldsymbol{x}_0) \leq f(\boldsymbol{x}) \ \ \forall \boldsymbol{x} \in \mathbb{R}^n$$

Assume $\exists \boldsymbol{x}_* \in \mathbb{R}^n$ such that $f(\boldsymbol{x}_0) = f(\boldsymbol{x}_*)$. By strict convexity of f, $\forall \alpha \in (0, 1)$:

$$f(\alpha \boldsymbol{x}_0 + (1 - \alpha)\boldsymbol{x}_*) < \alpha f(\boldsymbol{x}_0) + (1 - \alpha) f(\boldsymbol{x}_*) = f(\boldsymbol{x}_0)$$

Therefore there is an element in the domain that achieves a lower value than our global minimizer, which is our contradiction. Therefore our assumption is wrong. Given that f is *strongly convex* the local minimizer $\boldsymbol{x}_0$ is a *unique global* minimizer. □

QUESTION 2: Continued

(b) Furthermore, suppose $f \in \mathcal{C}^1$ (i.e., continuously differentiable). Then show that $\boldsymbol{x}_0$ is a *global* minimizer of f iff:

$$\nabla f(\boldsymbol{x}_0) = \mathbf{0}$$

PROOF: ($\Rightarrow$) Let $\boldsymbol{x}_0$ be a global minimum of f. Since f is convex, it satisfies the first order condition, $\forall \boldsymbol{y} \in \mathbb{R}^n$:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}_0) + (\boldsymbol{y} - \boldsymbol{x}_0)^T \nabla f(\boldsymbol{x}_0)$$

TODO: FINISH FORWARD

PROOF: ($\Leftarrow$) Assume $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$. Since f is convex, it satisfies the first order condition, $\forall \boldsymbol{y} \in \mathbb{R}^n$:

$$\begin{aligned} f(\boldsymbol{y}) &\geq f(\boldsymbol{x}_0) + (\boldsymbol{y} - \boldsymbol{x}_0)^T \nabla f(\boldsymbol{x}_0) \\ &= f(\boldsymbol{x}_0) + (\boldsymbol{y} - \boldsymbol{x}_0)^T \mathbf{0} \\ &= f(\boldsymbol{x}_0) \end{aligned}$$

So $\forall \boldsymbol{y} \in \mathbb{R}^n$, $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$ implies that $\boldsymbol{x}_0$ is a global minimum of f.

QUESTION 3: Given $f(\boldsymbol{x}), f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})$ are convex on $\mathbb{R}^m$, show that:

(a) If $\alpha_i \geq 0 \;\; \forall i \in [n]$, then $g(\boldsymbol{x}) \triangleq \sum \alpha_i f_i(\boldsymbol{x})$ is a convex function.

Consider some $\beta \in [0,1]$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. By the confexity of f_i:

$$\beta f_i(\boldsymbol{x}) + (1-\beta) f_i(\boldsymbol{y}) \geq f_i(\beta \boldsymbol{x} + (1-\beta)\boldsymbol{y})$$

Which implies:

$$\begin{aligned}
\sum \alpha_i \left[\beta f_i(\boldsymbol{x}) + (1-\beta) f_i(\boldsymbol{y})\right] &\geq \sum \alpha_i f_i(\beta \boldsymbol{x} + (1-\beta)\boldsymbol{y}) \\
\beta \sum \alpha_i f_i(\boldsymbol{x}) + (1-\beta) \sum \alpha_i f_i(\boldsymbol{y}) &\geq g(\beta \boldsymbol{x} + (1-\beta)\boldsymbol{y}) \\
\beta g(\boldsymbol{x}) + (1-\beta) g(\boldsymbol{y}) &\geq g(\beta \boldsymbol{x} + (1-\beta)\boldsymbol{y}) \;\; \forall \beta \in [0,1] \;\; \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n
\end{aligned}$$

Therefore g is a convex function.

(b) $g(\boldsymbol{x}) \triangleq \max\{f_1(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x})\}$ is convex:

Consider some $\beta \in [0,1]$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. We will construct two sets:

$$\begin{aligned}
A &\triangleq \{\beta f_i(\boldsymbol{x}) + (1-\beta) f_i(\boldsymbol{y}) | i \in [n]\} \\
B &\triangleq \{f_i(\beta \boldsymbol{x} + (1-\beta)\boldsymbol{y}) | i \in [n]\}
\end{aligned}$$

By the convexity of f_i, each element in B is upper bounded by at least one element in A which means that:

$$\max A \geq \max B$$

TODO: FINISH THIS ONE

(c) Given $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{b} \in \mathbb{R}^n$, $g(\boldsymbol{x}) \triangleq f(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$ is convex.

QUESTION 4: Show that $f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex iff for all integers $m \geq 2$:

$$f\left(\sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{m} \lambda_i f(\boldsymbol{x}_i) \tag{$\star$}$$

Where $x_i \in \mathbb{R}^n$ and $\lambda_i \geq 0 \ \forall i \in [m]$ satisfying $\sum \lambda_i = 1$.

PROOF: (by induction) Consider the base case ($m = 2$). The expression above boils down to the following:

$$f(\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2) \leq \lambda_1 f(\boldsymbol{x}_1) + \lambda_2 f(\boldsymbol{x}_2)$$

Notice that $\lambda_1 + \lambda_2 = 1$ is equivalent to $\lambda_2 = 1 - \lambda_1$. And since $\lambda_2 \geq 0$ must be true, we know that $\lambda_1 \in [0, 1]$. After substituting we get the definition of convexity:

$$f(\lambda_1 \boldsymbol{x}_1 + (1 - \lambda_1)\boldsymbol{x}_2) \leq \lambda_1 f(\boldsymbol{x}_1) + (1 - \lambda_1) f(\boldsymbol{x}_2)$$

So the base case ($m = 2$) the inequality ($\star$) is true iff f is convex.

Now let's assume ($\star$) is true for some fixed value ($m = k \geq 2$):

$$f\left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{k} \lambda_i f(\boldsymbol{x}_i)$$

W.l.o.g., we can express $\boldsymbol{x}_k$ as the sum of two more vectors, and λ_k as the sum of two more non-negative scalars.

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{y}_k + \boldsymbol{y}_{k+1} \\ \lambda_k &= \tau_k + \tau_{k+1} \end{aligned}$$

Notice that the sum of all of the scalars still holds ($\sum_{i=1}^{k-1} \lambda_i + \tau_k + \tau_{k+1} = 1$). Making this substitution into ($\star$) yields:

$$f\left(\sum_{i=1}^{k-1} \lambda_i \boldsymbol{x}_i + \tau_k \boldsymbol{y}_k + \tau_{k+1} \boldsymbol{y}_{k+1}\right) \leq \sum_{i=1}^{k-1} \lambda_i f(\boldsymbol{x}_i) + (\tau_k + \tau_{k+1}) f(\boldsymbol{y}_k + \boldsymbol{y}_{k+1})$$

QUESTION 5: