

EECS 559 HW 1

HONOR CODE: : I affirm that I did not use prior-year homework solutions or any large language model (LLM) in solving or writing up this assignment.

QUESTION 1:

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

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

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:

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

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:

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

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 \mathbf{x}_+ and \mathbf{x}_- as follows:

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

We can define:

$$\mathbf{z} \triangleq \begin{bmatrix} \mathbf{x}_+ \\ \mathbf{x}_- \end{bmatrix}$$

And clearly:

$$\begin{aligned} \|\mathbf{x}\|_1 &= \sum_{i=0}^n |x_i| \\ &= \sum_{i=0}^n x_i \mathbf{1}_{\{x_i \geq 0\}} + \sum_{i=0}^n x_i \mathbf{1}_{\{x_i < 0\}} \\ &= \mathbf{1}_n^T \mathbf{x}_+ + \mathbf{1}_n^T \mathbf{x}_- \\ &= \begin{bmatrix} \mathbf{1}_n^T & \mathbf{1}_n^T \end{bmatrix} \mathbf{z} \end{aligned}$$

We can now recast the problem as:

$$\min_{\mathbf{z} \in \mathbb{R}^n} \mathbf{c}^T \mathbf{z} \quad \text{s.t.} \quad \mathbf{a}_i^T \mathbf{z} = \mathbf{b}_i \quad \forall i \in [n]$$

Where:

$$\mathbf{c} \triangleq \begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_n \end{bmatrix}, \quad \mathbf{a}_i \triangleq \mathbf{M}_{[i,:]}, \quad \mathbf{M} \triangleq [\mathbf{A} \quad -\mathbf{a}]$$

QUESTION 2: Consider the following unconstrained optimization problem:

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

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

- (a) Show that if \mathbf{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 \mathbf{x}_0 be a local minimizer of f , meaning that $\exists \epsilon > 0$ such that:

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

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

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

And by the convexity of f :

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

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

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

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

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 \mathbf{x}_0 is a local minimizer of f . Since strict convexity implies convexity, we know that \mathbf{x}_0 is also a global minimizer (from the previous argument).

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

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

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

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 *strictly convex* the local minimizer \mathbf{x}_0 is a *unique global* minimizer. \square

QUESTION 2: Continued

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

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

PROOF: (\Rightarrow) (By contradiction) Let \mathbf{x}_0 be a global minimum of f , we assume $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$. Since f is convex, it satisfies the first order condition, $\forall \mathbf{y} \in \mathbb{R}^n$:

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

We can specifically choose some arbitrary \mathbf{y} :

$$\mathbf{y}(t) = \mathbf{x}_0 + t \nabla f(\mathbf{x}_0) \quad \forall t \in \mathbb{R}$$

We apply this to the first order condition to get:

$$f(\mathbf{x}_0 - t \nabla f(\mathbf{x}_0)) \geq f(\mathbf{x}_0) + t \|\nabla f(\mathbf{x}_0)\|^2$$

$\exists t$ s.t.:

$$f(\mathbf{x}_0 - t \nabla f(\mathbf{x}_0)) \leq f(\mathbf{x}_0) + t \|\nabla f(\mathbf{x}_0)\|^2$$

Unless $\|\nabla f(\mathbf{x}_0)\|^2 = 0$ which is possible iff $\nabla f(\mathbf{x}_0) = \mathbf{0}$

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

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

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

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

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

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

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

Which implies:

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

Therefore g is a convex function.

- (b) $g(\mathbf{x}) \triangleq \max\{f_1(\mathbf{x}), \dots, f_n(\mathbf{x})\}$ is convex:

SOLUTION: Consider some $\beta \in [0, 1]$ and $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$.

$$\begin{aligned} f_i(\beta \mathbf{x}_1 + (1 - \beta) \mathbf{x}_2) &\leq \beta f_i(\mathbf{x}_1) + (1 - \beta) f_i(\mathbf{x}_2) \\ \max_i f_i(\beta \mathbf{x}_1 + (1 - \beta) \mathbf{x}_2) &\leq \max_i [\beta f_i(\mathbf{x}_1) + (1 - \beta) f_i(\mathbf{x}_2)] \end{aligned}$$

Notice that $\forall a_i, b_i \in \mathbb{R}$, $\max\{a_i + b_i\} \leq \max\{a_i\} + \max\{b_i\}$. Therefore:

$$\begin{aligned} \max_i f_i(\beta \mathbf{x}_1 + (1 - \beta) \mathbf{x}_2) &\leq \beta \max_i f_i(\mathbf{x}_1) + (1 - \beta) \max_i f_i(\mathbf{x}_2) \\ g(\beta \mathbf{x}_1 + (1 - \beta) \mathbf{x}_2) &\leq \beta g(\mathbf{x}_1) + (1 - \beta) g(\mathbf{x}_2) \end{aligned}$$

Therefore the convexity of f implies the convexity of g .

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

SOLUTION: Consider $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^m$, and $\alpha \in [0, 1]$

$$f(\mathbf{A}(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) + \mathbf{b}) = f(\alpha(\mathbf{Ax}_1 + \mathbf{b}) + (1 - \alpha)(\mathbf{Ax}_2 + \mathbf{b}))$$

By the convexity of f :

$$\begin{aligned} f(\mathbf{A}(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) + \mathbf{b}) &\leq \alpha f(\mathbf{Ax}_1 + \mathbf{b}) + (1 - \alpha) f(\mathbf{Ax}_2 + \mathbf{b}) \\ g(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) &\leq \alpha g(\mathbf{x}_1) + (1 - \alpha) g(\mathbf{x}_2) \end{aligned}$$

Therefore the convexity of f implies the convexity of g .

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

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

Where $\mathbf{x}_i \in \mathbb{R}^n$ 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:

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

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:

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

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 ($m = k \geq 2$):

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

W.l.o.g., we can express \mathbf{x}_k as follows:

$$\mathbf{x}_k = \alpha \mathbf{y}_k + (1 - \alpha) \mathbf{y}_{k+1} \quad \text{where } \alpha \in [0, 1]$$

Where \mathbf{y}_k and \mathbf{y}_{k+1} are arbitrary vectors in \mathbb{R}^n . We can rewrite (\star) as follows:

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

And by the convexity of f :

$$\lambda_k f(\alpha \mathbf{y}_k + (1 - \alpha) \mathbf{y}_{k+1}) \leq \lambda_k \alpha f(\mathbf{y}_k) + \lambda_k (1 - \alpha) f(\mathbf{y}_{k+1})$$

Which leaves us with:

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

And since $\sum_{i=1}^{k-1} \lambda_i + \lambda_k (\alpha + 1 - \alpha) = 1$, the $m = k$ case being true implies that the $m = k + 1$ case is also true. Therefore by mathematical induction, $\forall m \geq 2, \forall i \in [m], \forall \lambda_i \geq 0, \forall \mathbf{x}_i \in \mathbb{R}^n$:

$$f : \mathbb{R}^n \mapsto \mathbb{R} \text{ convex} \iff f\left(\sum_{i=1}^m \lambda_i \mathbf{x}_i\right) \leq \sum_{i=1}^m \lambda_i f(\mathbf{x}_i) \quad \text{where } \sum \lambda_i = 1$$

□

QUESTION 5:

- (a) Show that the sequence $x_k = 1 + \left(\frac{1}{2}\right)^{2^k}$ is Q-quadratically convergent to 1.

SOLUTION: We know that $\lim_{k \rightarrow \infty} x_k = 1$, consider:

$$\begin{aligned} \left\| 1 + \left(\frac{1}{2}\right)^{2^k} - 1 \right\|_2 &\leq \gamma \left\| 1 + \left(\frac{1}{2}\right)^{2^{k-1}} - 1 \right\|_2^2 \\ \left(\frac{1}{2}\right)^{2^k} &\leq \gamma \left[\left(\frac{1}{2}\right)^{2^{k-1}} \right]^2 \\ \left(\frac{1}{2}\right)^{2^k} &\leq \gamma \left(\frac{1}{2}\right)^{2^k} \end{aligned}$$

Which is true $\forall \gamma \leq 1, \forall k \geq 1$, therefore the sequence $\{x_k\}$ is Q-quadratically convergent.

- (b) Does the sequence $x_k = \frac{1}{k!}$ converge Q-superlinearly? Q-quadratically?

SOLUTION: We know that $\lim_{k \rightarrow \infty} x_k = 0$, consider:

$$\begin{aligned} \left\| \frac{1}{k!} \right\|_2 &\leq \gamma \left\| \frac{1}{(k-1)!} \right\|_2^p \\ \frac{1}{k!} &\leq \gamma \left[\frac{1}{(k-1)!} \right]^p \\ \frac{((k-1)!)^{p-1}}{k} &\leq \gamma \\ (p-1) \log((k-1)!) - \log(k) &\leq \log(\gamma) \\ (p-1) \sum_{i=1}^{k-1} \log(i) - \log(k) &\leq \log(\gamma) \end{aligned}$$

The sequence above is clearly only upper bounded if $p-1 \leq 0$, therefore $\{x_k\}$ is Q-convergent with $p \leq 1$. So the sequence is neither Q-quadratically nor Q-superlinearly convergent.

QUESTION 5: Continued

(c) Consider the sequence $\{x_k\}$ defined by

$$x_k = \begin{cases} \left(\frac{1}{4}\right)^{2^k} & k \text{ even} \\ \frac{x_{k-1}}{k} & k \text{ odd} \end{cases}$$

Is this sequence Q-superlinearly convergent? Q-quadratically convergent? R-quadratically convergent? Justify your answer.

SOLUTION: We know that $\lim_{k \rightarrow \infty} x_k = 0$, consider the case where k is odd:

$$\begin{aligned} \|x_k\|_2 &\leq \gamma \|x_{k-1}\|_2^p \\ \frac{x_{k-1}}{k} &\leq \gamma x_{k-1}^p \\ \frac{1}{k} \left(\frac{1}{4}\right)^{(1-p)2^k} &\leq \gamma \end{aligned}$$

There can only be such γ if $p \leq 1$, therefore $\{x_k\}$ is neither Q-quadratically nor Q-superlinearly convergent. Now consider the sequence $\{\rho_k\}$ defined as follows:

$$\rho_k = \left(\frac{1}{4}\right)^{2^k} \quad \forall k \in \mathbb{N}$$

Notice that:

$$\begin{aligned} \rho_k &\geq x_k \quad \forall k \in \mathbb{N} \\ \rho_k &\geq \|x_k - 0\|_2 \end{aligned}$$

We know from (a) that $\{\rho_k\}$ is Q-quadratically convergent, therefore $\{x_k\}$ is R-quadratically convergent.