1. Let A be an $m \times n$ matrix and $\mathbf{c} \in \mathbb{R}^n$. Prove that exactly one of the following two systems has a solution:

   - $A\mathbf{x} \leq \mathbf{0}, \mathbf{x} \geq 0, \mathbf{c}^T\mathbf{x} > 0$
   - $A^T\mathbf{y} \geq \mathbf{c}, \mathbf{y} \geq \mathbf{0}$

   Proof. We will prove this by a series of equivalences. Let us begin with the second system. It is equivalent to $\begin{bmatrix} A & -I \end{bmatrix}^T \begin{bmatrix} y \\ s \end{bmatrix} = c, y \geq 0, s \geq 0$. By Farkas lemma, this is equivalent to the system $\begin{bmatrix} A \\ -I \end{bmatrix} x \leq 0, c^T x > 0$ not having a solution. And finally, this last system is equivalent to $Ax \leq 0, c^T x > 0, x \geq 0$. Thus, either $Ax \leq 0, x \geq 0, c^T x > 0$ has a solution, or $A^T y \geq c, y \geq 0$ has a solution, but not both. □

2. (Soft thresholding)

   (a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be the Euclidean norm function $f(\mathbf{x}) = \|\mathbf{x}\|_2$. Prove that

   $$\partial f(\mathbf{x}) = \begin{cases} \{ \xi \in \mathbb{R}^n \mid \|\xi\|_2 \leq 1 \} & \text{if } \mathbf{x} = 0 \\ \frac{\mathbf{x}}{\|\mathbf{x}\|_2} & \text{if } \mathbf{x} \neq 0. \end{cases}$$

   Solution. Let us begin with the case $\bar{x} = 0$. If a vector p satisfies $\|p\| \leq 1$, then[1]

   $$\begin{aligned} p^T x &\leq \|x\| \\ p^T x - p^T \bar{x} + \|\bar{x}\| &\leq \|x\| \\ p^T(x - \|x\|) + f(\bar{x}) &\leq f(x) \end{aligned}$$

   So $p \in \partial f(0)$. Now suppose that $p \in \partial f(0)$. Then

   $$\begin{aligned} f(x) &\geq f(0) + p^T(x - 0) \\ \|x\| &\geq p^T x \end{aligned}$$

   for all $x \in \mathbb{R}^n$. Letting $x = \frac{p}{\|p\|}$. Then $\|p\| \geq \|p\|^2$ or $\|p\| \leq 1$. Thus,

   $$\partial f(0) = \{\xi \in \mathbb{R}^n \mid \|\xi\| \leq 1\}$$

   For the case $\bar{x} \neq 0$, if a vector p satisfies $\frac{x}{\|x\|} = p$, then

   $$\begin{aligned} \bar{x}^T x &\leq \|\bar{x}\|\|x\| \\ \|\bar{x}\| + \frac{\bar{x}^T x}{\|\bar{x}\|} - \frac{\bar{x}^T \bar{x}}{\|\bar{x}\|} &\leq \|x\| \\ \|\bar{x}\| + p^T x - p^T \bar{x} &\leq \|x\| \\ f(\bar{x}) + p^T(x - \bar{x}) &\leq f(x) \end{aligned}$$

[1] Sorry for the horrible formatting

So $p \in \partial f(\bar{x})$. On the other hand, suppose that $p \in \partial f(\bar{x})$ for $\bar{x} \neq 0$. Then

$$\|x\| \geq \|\bar{x}\| + p^T(x - \bar{x})$$

for any $x \in \mathbb{R}^n$. Letting $x = 2\bar{x}, x = p, x = 0$, we obtain $\|p\| \leq 1, \|p\| \geq 1$, and $p^T x = \|x\|$. Now note that[2]

$$\begin{aligned}\|x\| &= p^T x \\ &= \|p\|\|x\|\cos(\theta)\end{aligned}$$

where θ is the angle between x and p. This implies that

$$1 \leq \cos(\theta)$$

So $\cos(\theta) = 1$. That is, x and p are in the same direction, i.e. $p = \alpha x$ for some $\alpha \in \mathbb{R}$. But we also know that $\|x\| = p^T x = \alpha x^T x = \alpha\|x\|^2$. So $\alpha = \frac{1}{\|x\|}$ and thus $p = \frac{x}{\|x\|}$.

(b) Using the result of part (a), compute the optimal solution of

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_2,$$

where $\mathbf{y} \in \mathbb{R}^n$ and $\lambda > 0$ are given data. (Fact: If f_1, f_2 are convex functions on $\mathbb{R}^n$, then for any $\mathbf{x} \in \mathbb{R}^n$, $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.)

Solution. Let $f(x) = \frac{1}{2}\|x - y\|_2^2 + \lambda\|x\|_2$. Assume that $x \neq 0$; then we have that $\nabla f(x) = (x - y) + \frac{\lambda x}{\|x\|}$. In order for $x \neq 0$ to be an optimal solution, $\nabla f(\bar{x}) = 0$. We can rearrange this to get that $y = (1 + \frac{\lambda}{\|x\|})x$. That is, x is in the direction of y. Furthermore, taking norm, we get that $\|y\| = \|x\| + \lambda$ or $\|x\| = \|y\| - \lambda$ if $\|y\| \geq \lambda$. Thus, when $\|y\| \geq \lambda$, our optimal solution is

$$x = \frac{y}{\|y\|}(\|y\| - \lambda) = y\left(1 - \frac{\lambda}{\|y\|}\right)$$

If $\lambda > \|y\|$, then our assumption that $x \neq 0$ is false. Thus our optimal solution in this case is $x = 0$[3].

3. Suppose that S is a closed and convex subset of $\mathbb{R}^n$ with nonempty interior. Let $f : S \to \mathbb{R}$ be differentiable on $\text{int}(S)$. State if the following are true or false, justifying your answer. (Theorems proved in class can be used directly.)

(a) If f is convex on S, then $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x})$ for all $x \in S$ and $\bar{x} \in \text{int}(S)$.

Solution. Since S is a convex set, we know that $\int(S)$ is also convex. By direct result in class, this implies that $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x})$ for all $x \in \text{int}(S)$ and $\bar{x} \in \text{int}(S)$. All we need to do is show that this can be extended for points in the

[2] Collaborated with Blake Splitter for the geometric argument

[3] Sloppying handling of cases

boundary $S \setminus \text{int}\, S$. Let x be a point in the boundary of S. Then there exists a sequence $\{x_n\} \subset \text{int}(S)$ such that $\{x_n\}$ converges to x. For each x_n, we know that

$$f(x_n) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x_n - \bar{x})$$

Since f is differentiable on the interior, it is also continuous. Thus, taking the limit as $n \to \infty$ on each side gives

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$$

which is what we wanted to show.

(b) If $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$ for all $x \in S$ and $\bar{x} \in \text{int}(S)$, then f is convex on S.

Solution. This is not true in general. Consider a function $f : [0, 1] \times [0, 1] \to \mathbb{R}$. Let $f(x, y) = 1 - x^2$ if $y = 1$ and $f(x, y) = 0$ otherwise. The function is clearly non-convex over S (it is nonconvex over one edge of the boundary), but $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$ for any $x \in S, \bar{x} \in \text{int}(S)$ since the right hand side is always zero and the left hand side is nonnegative.

4. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$, given by $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$, where

$$A = \begin{bmatrix} 2 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & \theta \end{bmatrix}.$$

What is the Hessian of f? For what values of θ is f strictly convex?

Solution. From the calculus review notes, we know that the Hessian of a quadratic form with nonsymmetric matrix is $H(f) = A + A^T$. Thus,

$$H(f) = \begin{bmatrix} 4 & 3 & 4 \\ 3 & 6 & 3 \\ 4 & 3 & 2\theta \end{bmatrix}$$

We will show conditions in which the Hessian is positive by computing all leading principle minors. If all leading principle minors are positive, then we are good. We list the determinants below (and noting the fact that $4 > 0$)

- $\det H = 30\theta - 60 \implies \theta > 2$
- $M_{11} = 12\theta - 9 \implies \theta > \frac{9}{12}$

Thus, H is SPD, and $f(x)$ is strictly convex, whenever $\theta > 2$. [4]

5. Prove that the geometric mean function

$$f : \mathbb{R}^n_+ \to \mathbb{R}_+, \quad f(\mathbf{x}) = \left(\prod_{i=1}^{n} x_i \right)^{1/n}$$

is concave.

[4] Coincidentally, this is also the condition for H to be positive semidefinite

Proof. For simplicity, let $\Pi = (\prod_{i=1}^{n} x_i)^{\frac{1}{n}}$. We will prove that f is concave by showing its Hessian is negative semidefinite. Observe that the partial of f with respect to x_k is

$$\frac{\partial f(x)}{\partial x_k} = \frac{1}{n}\left(\frac{\Pi}{x_k}\right)$$

And thus through rigorous computation we can obtain that

$$\frac{\partial^2 f(x)}{\partial x_k^2} = \frac{-(n-1)\Pi}{n^2 x_k^2}$$

and

$$\frac{\partial^2 f(x)}{\partial x_k x_l} = \frac{\Pi}{n^2 x_k x_l}$$

Thus, setting $y_i = \frac{1}{x_i}$, we can write

$$z^T H z = \left(\sum_{i=1}^{n} y_i z_i\right)^2 - n\sum_{i=1}^{n} y_i^2 z_i^2$$

for any $v \in \mathbb{R}^n$. To be negative semidefinite, we need this to be less than zero. This follows from the convexity of x^2 (it has strictly positive second derivative). By convexity, we have that

$$f(\frac{1}{n}a_1 + \cdots + \frac{1}{n}a_n) \leq \frac{1}{n}[f(a_1) + \ldots f(a_n)]$$

and so letting $f(x) = x^2$ we get

$$\left(\sum_{i=1}^{n} \frac{1}{n}a_i\right)^2 \geq \frac{1}{n}\sum_{i=1}^{n} a_i^2$$

Rearrangement and letting $a_i = y_i z_i$ gives the desired inequality. So H is negative semidefinite and f is concave. $\square$