CIS 621 Assignment 1

Steven Walton

January 23, 2019

1)

Let S be the set $S = x_1, \cdots, x_n$. Suppose that some subset s, where s be the subset $s \subseteq S$, where $\sum s = w$ for some $w \in \mathbb{Z}^+$

Proof

1) Prove that $L \in$ NP where L is this problem

This problem can be solved with the certificate and verifier method. That is we can guess a random subset s and test $\sum s$. The testing, verification, can be determined in polynomial time.

2) Reduction

We know the knapsack problem is NP-Complete. Knapsack can be reduced to S because this problem is a subset of knapsack. We search for the highest sum by summing over different subsets of the total items that can be placed into the knapsack. If we let the weight constraint become non-existent then the problem is the same, except we are looking for a specific value, not the largest.

2)

Given

$$a, x \in \mathbb{R}^2 : a = \begin{bmatrix} a_1 \\ a_1 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} b_1, b_2 \in \mathbb{R}$$

2a)

Prove

$$a^T x = \|a\| \|x\| \cos\theta$$

Proof
Let $c = a - x$, from definition of triangles

$$
\begin{aligned}
< c, c > &= < a - x, a - x > \text{ by definition} \\
&= < a, a > - < a, x > - < x, a > + < x, x > \text{ by expansion} \\
&= < a, a > - < a, x > - < a, x > + < x, x > \text{ communative property of real inner products} \\
&= < a, a > -2 < a, x > + < x, x > \text{ by addition} \\
&= < a, a > + < x, x > -2 < a, x >
\end{aligned}
$$

Compare to Law of Cosines

$$c^2 = a^2 + x^2 - 2ax\cos\theta$$

We note that $c^2 = \|c\|^2 =< c, c >$. We use the shorthand $c = \|c\|$

$$
\begin{aligned}
a^2 + x^2 - 2ax\cos\theta &= < a, a > + < x, x > -2 < a, x > \\
< a, a > + < x, x > -2ax\cos\theta &= \text{ by definition} \\
-2ax\cos\theta &= -2 < a, x > \text{ by reduction} \\
ax\cos\theta &= < a, x >
\end{aligned}
$$

Using our definition of inner products $< a, x >= a^T x$

$$
\begin{aligned}
\therefore a^T x &= < a, x > \\
&= ax\cos\theta \\
&= \|a\|\|x\|\cos\theta
\end{aligned}
$$

2b)

Let

$$a^T x = b_1, a^T x = b_2$$

Calculate the distance between b_1 and b_2
We will say that $\forall b, \frac{b}{\sqrt{\|a\|}}$ is the normalized signed distance from of the plane to the origin. We note that the distance is

$$d = \left|\frac{b_1}{\hat{a}} - \frac{b_2}{\hat{a}}\right|$$

Where $\hat{a}$ is the normalized vector that is orthogonal to the plane. Because $\hat{a}$ is the same for both planes, we can say

$$d = \frac{|b_1 - b_2|}{\hat{a}}$$

We could be done here but I think you want more. So let's set a point in the first plane. We'll set the point $x_1 = 0$, by redefining the coordinate system appropriately.

$$\begin{aligned} \therefore a^T x &= \begin{bmatrix} a_1 a_2 \end{bmatrix} \begin{bmatrix} 0 \\ x_2 \end{bmatrix} \\ &= a_2 x_2 = b_1 \end{aligned}$$

Similarly we get $b_2 = a_2 x_2$ Substituting this back in we get

$$d = \frac{|a_2 x_2 - a_2 x_2|}{\hat{a}}$$

For less confusion we will say

$$x_1 = \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} x_2 = \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix}$$

and let's say that $a^T x_1 = b_1; a^T x_2 = b_2$a
Then we can re-express the above equation with unique x solutions as

$$\begin{aligned} d &= \frac{|a_2 x_{12} - a_2 x_{22}|}{\hat{a}} \\ d &= \frac{a_2}{\hat{a}} |x_{12} - x_{22}| \end{aligned}$$

Where these represent two components to two different, and arbitrary, x vectors that satisfy the solution to the planes.

3)

Given $x_0, \cdots, x_k \in \mathbb{R}^n$, transform S and express as $S = \{x | Ax \leq b\}$
Where $S = \{x \in \mathbb{R}^n | \|x - x_0\| \leq \|x - x_i\|, i = 1, \cdots, k\}$
We can rewrite this as

$$
\begin{gathered}
S = \|x - x_0\| \leq \|x - x_i\| \\
< x - x_0, x - x_0 > \leq < x - x_i, x - x_i > \\
x^T x - 2x_0^T x + x_0^T x_0 \leq x^T x - 2x_i^T x + x_i^T x_i \\
2(xi - x_0)^T x \leq x_i^T - x_0^T x_0
\end{gathered}
$$

Which is a halfspace

$$\therefore S = \{x | Ax \preceq b\}$$

4)

Let f be twice differentiable and $dom(f)$ be convex $(\nabla f(x), \nabla f(y) \neq 0)$.
Prove f is convex iff $(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0$
Proof Left to right One:

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

Two:

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

If we combine One and Two get get

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

We left out where the $f(x)$'s and $f(y)$'s cancel.
Right to left We can introduce the $f(x)$'s and $f(y)$'s because they will cancel and we can always add 0 to an equation.

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

We can see that this equation is the same as the combination of One and Two.

5)

Prove the following are convex

5a)

$f(x) = \max\{f_1(x), \cdots, f_m(x)\}$ where $f_i(x), i = 1, \cdots, m$ are convex
We will rewrite into the shorthand $f(x) = \max_i f_i(x)$

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \max_i (\theta x_i + (1-\theta)y_i) \\ \max_i (\theta x_i + (1-\theta)y_i) &\leq \theta \max_i x_i + (1-\theta) \max_i y_i \\ f(\theta x + (1-\theta)y) &\leq \theta f(x) + (1-\theta) f(y) \end{aligned}$$

We can state the second line because the maximal value of the function will always be less than the values provided by the maximals of the variables independently.

5b)

$f(x) = \min_{y \in C} g(x, y)$ where $g(x, y)$ is convex for x and y and C is a convex set.
We can say that the $dom f = \{x | (x, y) \in dom f \text{ for some } y \in C\}$ Let $\epsilon > 0$. The $y_1, y_2 \in C$ s.t. $g(x_i, y_i) \leq f(x_i) + \epsilon$ If we restrict $\theta \in [0, 1]$

$$\begin{aligned} f(\theta x_1 + (1-\theta)x_2) &= \min_{y \in C} g(\theta x_1 + (1-\theta)x_2, y) \\ &\leq g(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \\ &\leq \theta g(x_1, y_1) + (1-\theta) g(x_2, y_2) \\ &\leq \theta f(x_1) + (1-\theta) f(x_2) + \epsilon \\ \therefore f(\theta x_1 + (1-\theta)x_2) &\leq \theta f(x_1) + (1-\theta) f(x_2) + \epsilon \end{aligned}$$

5c)

$f(x, y) = \frac{|x_1|^p + \cdots + |x_n|^p}{y^{p-1}}$ where $p > 1, x \in \mathbb{R}^n, y \in \mathbb{R}^+$
We will prove this by stating that the Hessian matrix is at least semi-positive definite

We will construct each derivative independently

$$f(x) = \frac{\sum |x|^p}{y^{p-1}}$$

$$\frac{\partial f}{\partial x} = p\frac{\sum |x|^{p-1}}{y^{p-1}}$$

$$\boxed{\frac{\partial^2 f}{\partial x^2} = (p^2 - p)\frac{\sum |x|^{p-2}}{y^{p-1}}}$$

$$\frac{\partial f}{\partial y} = (1-p)\frac{\sum |x|^p}{y^p}$$

$$\boxed{\frac{\partial^2 f}{\partial y^2} = (p^2 - p)\frac{\sum |x|^p}{y^{p+1}}}$$

$$\boxed{\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = (p - p^2)\frac{\sum |x|^{p-1}}{y^p}}$$

From here we can look at the determinate D

$$\begin{aligned} D &= (p^2 - p)\frac{\sum |x|^{p-2}}{y^{p-1}}(p^2 - p)\frac{\sum |x|^p}{y^{p+1}} + (p - p^2)^2 \left(\frac{\sum |x|^{p-1}}{y^p} \right) \\ &= (p^2 - p)^2 \left(\frac{\sum |x|^{p-2}}{y^{p-1}} + \frac{\sum |x|^{p-1}}{y^p} \right) \geq 0 \end{aligned}$$

We know that this determinate is greater than 0 by breaking down the components. The p values on the left have to be $\geq 0 \because$ it is quadratic.
$\because x \in \mathbb{R}^n$ and that we are taking absolute values, we know that $\sum |x|^k \geq 0 \; \forall k, x \in \mathbb{R}$
Lastly, we defined $y \in \mathbb{R}^+$, or that y is a positive real number.
$\therefore$ neither the numerator nor the denominator can be negative and thus the determinate cannot be less than 0.
$\because D \geq 0$ we can say the function is positive semi-definite and $\therefore$ convex

5d)

$$f(x_{11}, \cdots, x_{1n}, x_{21}, \cdots, x_{2n}, \cdots, x_{m1}, \cdots, x_{mn}) = \sum_{i=1}^{m} \sum_{j=1}^{n} ((x_{ij} + 1)\log(x_{ij} + 1) - x_{ij})$$

where $x_{ij} \in \mathbb{R}_{++}, i = 1, \cdots, m, j = 1, \cdots, n$
For any i and j we can see that the first derivative is

$$\frac{\partial f}{\partial x_{ij}} = \log(x_{ij} + 1)$$

We notice that all non-diagonal terms cancel out if we take another derivative. The second derivative is

$$\frac{\partial f^2}{\partial^2 x_{ij}} = \frac{1}{x_{ij} + 1} \delta_{ij}$$

Where $\delta_{ij} = 1$ iff $i = j$ else 0

Because we have the condition that $x_{ij} \in \mathbb{R}_{++} \forall i, j$ we can state that all the diagonals of the Hessian matrix are positive numbers and all non-diagonals are 0. Therefore the determinate of this matrix is ≥ 0 and $\therefore$ convex.

6)

Let $f(x) = \max\{|a^T x + b|, \log \frac{1}{c^T x+d}\}; a, c, x \in \mathbb{R}^n; b, d \in \mathbb{R}$

Is $f(x)$ convex?

While we know that the log function is concave, we know that the negative of a concave function is convex (and vise versa). Because the function is a fractional we can rewrite as $-\log c^T + d$, which we know is convex. We also know that the absolute value function is convex. In 5a we proved that a max function is convex if all the components are convex. $\therefore f(x)$ has to be convex.