Name: Hongda Li   AMATH 514 SPRING 2022 HW2

1 Problem 2.2

Proposition 1 (Linear Map Preserves Convexity)**.** Let $C \subseteq \mathbb{R}^n$ be a closed and convex set, let A be a $m \times n$ matrix. Show that the set $\{Ax | x \in C\}$ is again convex.

To start the proof, we introduce the following notations to make discussion better:

1.) Let $A[C]$ be the range of the function over the set C, denoting the set $\{Ax : x \in C\}$.

2.) Let A^{-1} denotes the pre-image of the linear operator. When applying to a vector it's the set $A^{-1}(y) := \{x \in \mathbb{R}^n : Ax = y\}$.

3.) Let $A^{-1}[X|C]$ be the pre-image of the operator on the set X intersecting with C, then: $A^{-1}[X|C] := \{x \in C : Ax \in X\}$, if $X = \{a\}$ is singleton, we denote $A^{-1}[a|C]$

4.) If we denote λC where C is a set, we are scaling all the elements in the set by the scalar λ, meaning that $\lambda C := \{\lambda x : x \in C\}$.

The objective of the proposition is to show that the set $A[C]$ is a convex set when the set C is convex.

Proof.

$$\begin{aligned}
\forall a, b \in A[C], a \neq b \qquad (1.0.1)\\
a \in A[C] \iff A[A^{-1}(a) \cap C] = a \qquad (1.0.2)\\
b \in A[C] \iff A[A^{-1}(b) \cap C] = b \qquad (1.0.3)
\end{aligned}$$

For any a, b in the image of the set C through A is the same as looking for the intersection of pre-image of a, b intersecting with C; by definition both $A^{-1}(b) \cap C, A^{-1}(b) \cap C$ is non-empty. Let's consider the convex combination of any 2 points in $A[C]$ we have:

$$\begin{aligned}
&\forall \lambda \in (0, 1) : &(1.0.4)\\
\lambda a + (1 - \lambda)b &= \lambda A[A^{-1}(a) \cap C] + (1 - \lambda)A[A^{-1}(b) \cap C] &(1.0.5)\\
&= A[\lambda A^{-1}(a)] \cap C + A[(1 - \lambda)A^{-1}(b) \cap C] &(1.0.6)\\
&= A[\underbrace{\lambda A^{-1}(a) \cap C + (1 - \lambda)A^{-1}(b) \cap C}_{\in C}] \quad \text{Convexity of C} &(1.0.7)\\
&\implies A[\lambda A^{-1}(a) \cap C + (1 - \lambda)A^{-1}(b) \cap C] \in A[C] &(1.0.8)
\end{aligned}$$

Using the property of the learning mapping, we can group together the sets of pre-images of a, b intersecting C, because each element of the pre-images are in the set C which is convex, then a convex combinations of any of its element is still in the set, therefore, any convex combination of a, b from $A[C]$ is still in the set $A[C]$. $\square$

Problem 2.4

Proposition 2. If $z \in \text{conv}(X)$, then there exists affinely indepedent vectors $\{x_1, \cdots, x_m\} \subseteq X$ such that x is in the convex hull of those vectors.

Before proving it, we need to envoke a lemma. We also introduce the notation $[n]$ to be the set of natural indices going from $1, \cdots, n$.

Lemma 1. Suppose that $x \in \text{conv}(\{x_i\}_{i=1}^n)$ and the set of vectors $\{x_i\}_{i=1}^n$ is Affinely Dependent, then $x \in \text{conv}(\{x_i\}_{i\in\mathcal{I}})$ where $\mathcal{I} \subsetneq [n]$. If x is in the convex hull of countably set of vectors and the set of vectors are Affline Linear Dependent, then it can be represented as a convex hull of a subset of those Affline Dep vectors such that the cardinality is strictly less.

Proof.

$$x \in \text{conv}(\{x_i\}_{i=1}^n) \tag{1.0.9}$$

$$\implies x = \sum_{i=1}^{n} \lambda_i \quad \text{s.t: } \lambda_i > 0 \quad \forall i \in [n] \tag{1.0.10}$$

If any of the λ_i is already zero, then we kick out those x_i out of the set and then go back to the top of the proof. Next, we consider the property of Affline Linear Depdent set, (Aff Dep) for short.

$$\{x_i\}_{i\in[n]} \text{ is Aff Dep} \tag{1.0.11}$$

$$\iff \exists u_j \neq 0, j \in [n], \langle \mathbf{1}, \vec{\mu} \rangle = 0 : \mathbf{0} = \sum_{i=1}^{n} \mu_i x_i \tag{1.0.12}$$

Firstly, choose a special u_j such that: u_j is not zero (it exists, asserted by the definition of Aff Dep), and $j \in \arg\max_{i\in[n]}(|\mu_i|)$. Fix the j and now consider the consequence:

$$\mathbf{0} = \frac{\lambda_j}{\mu_j}\left(\sum_{i=1}^{n} \mu_i x_i\right) \tag{1.0.13}$$

$$\mathbf{0} = \left(\sum_{i\neq j, i=1}^{m} \frac{\lambda_j \mu_i}{\mu_j} x_j\right) + \lambda_j x_j \tag{1.0.14}$$

$$x = \left(\sum_{i\neq j, i=1}^{n} \lambda_i x_i\right) + \lambda_j x_j \tag{1.0.15}$$

$$\implies x = \sum_{i=1, i\neq j}^{n} \left(\lambda_j - \frac{\lambda_j \mu_j}{\mu_i} x_i\right) \tag{1.0.16}$$

$$x = \sum_{i=1, i\neq j}^{n} \lambda_j \left(1 - \frac{\mu_i}{\mu_j} x_i\right) \tag{1.0.17}$$

We start with the definition of a Aff Dep, and then we multiply both size by a none zero scalar λ_j/μ_j. Then we pull out the j th term from the sum. The third line is from the definition of $x \in \text{conv}(\{x_i\}_{i\in[n]})$. Please obseve that $1 - \mu_i/\mu_j \geq 0 \; \forall \; i \in [n]$ will always be a non-negative because j is chosen such that $|\mu_j|$ is as large as possible. Using that fact, we know that $x \in \text{conv}(\{x_i\}_{i\in\mathcal{I}})$ where $\mathcal{I} \subsetneq [n]$, and $\mathcal{I}$ contains the indices that makes $1-\mu_i/\mu_j$. The 2 sets are not equal because by definition at least one of the coefficient for $1 - \frac{\mu_i}{\mu_j}$ is zero. $\square$

Using the lemma, we can prove the proposition inductively. Given any set X, we choose any countable subset such that $\{x_i\}_{i\in[n]}$ as a subset of X and $x \in \text{conv}(\{x_i\}_{i\in[n]})$, for any X by the definition of Convex Hull. Then, there are only 2 possible cases about the set $\{x_i\}_{i\in[n]}$:

1.) The set $\{x_i\}_{i\in[n]}$ is Aff Dep, then we can use the lemma and get a smaller set $\{x_i\}_{i\in\mathcal{I}}$ such that it's a strict subset of the former set, and $x \in \text{conv}(\{x_i\}_{i\in\mathcal{I}})$.

2.) If the set $\{x_i\}_{i\in[n]}$ is Aff InDep, then we are done.

Repeat the above process, each time we redefine $\{x_i\}_{i\in n_k} := \{x_i\}_{i\in\mathcal{I}}$. Then we have:

$$\{x_i\}_{i\in[n_k]} \subsetneq \{x_i\}_{i\in[n_{k-1}]} \subsetneq \cdots \subsetneq \{x_i\}_{i\in[n]} \tag{1.0.18}$$

Then the sequence of sets must terminates, and when it terminates it has to be the case that they are linear Afflinely Independent. More specifically, whenever the set is a singleton, containing only one element, then it has to be the case that the set is Afflinely Independent (This is trivial). Therefore, for all sets of $\{x_i\}_{i\in[n]}$ we started with as a subset of X, the inductive always terminates with an non-empty set.

Problem 2.5

To prove the theorem, we introduce 2 lemmas about covex sets, and set projections to simplify things.

Definition 1 (Set Projection)**.** We define $\text{proj}_Q(z)$ be the closest point in Q to the point z measured by the 2-norm. Mathmatically:

$$\text{proj}_Q(z) = \left\{ \|x^+ - z\|_2^2 : x^+ = \inf_{x\in Q} \|x - z\|_2^2 \right\} \tag{1.0.19}$$

Lemma 2 (Set Difference Prserves Convexity)**.** Let set Q_1, Q_2 and convex, then we define $Q_1 + Q_2 := \{x + y : x \in Q_1, y \in Q_2\}$.

Proof. Considering choosing any 2 points from the set $Q_1 + Q_2$; we can say that:

$$\exists \; q_1 \in Q_1, q_2 \in Q_2 : x := q_1 + q_2 \tag{1.0.20}$$
$$\exists \; q_3 \in Q_1, q_4 \in Q_2 : y := q_3 + q_4 \tag{1.0.21}$$

Let's consider the convex combinations of these 2 points:

$$\lambda x = \lambda(q_1 + q_2) \tag{1.0.22}$$
$$(1-\lambda)y = (1-\lambda)(q_3 + q_4) \tag{1.0.23}$$
$$\implies \lambda x + (1-\lambda)y = \lambda(q_1 + q_2) + (1-\lambda)(q_3 + q_4) \tag{1.0.24}$$
$$= \lambda q_1 + \lambda q_2 + (1-\lambda)q_3 + (1-\lambda)q_4 \tag{1.0.25}$$
$$\text{By convexity of } Q_1, Q_2; = \underbrace{\lambda q_1 + (1-\lambda)q_3}_{\in Q_1} + \underbrace{\lambda q_2 + (1-\lambda)q_4}_{\in Q_3} \tag{1.0.26}$$
$$\implies \lambda x + (1-\lambda)y \in Q_1 + Q_2 \tag{1.0.27}$$

The set $Q_1 + Q_2$ is still convex. □

Next, take notice that, if D is convex, then the set $-D := \{-y, y \in D\}$ is still going to be a convex set, this is trivial. Then $C - D$ can be interpreted as $C + (-D)$ and it will still be a convex set by the above lemma that we proved.

Lemma 3 (Obtuse Angle Theorem)**.** For any closed, convex, non-empty set Q in the finite Eulidean space, the projection (it's a singleton set) of any points on to the set to the point itself make an abstuse angle with all the other points in the set Q. Mathematically:

$$\forall y \exists z : \{z\} = \underset{Q}{\operatorname{proj}}(y) \tag{1.0.28}$$
$$\implies \langle y - z, x - z\rangle \le 0 \quad \forall x \in Q \tag{1.0.29}$$

Note: The statement is stronger than what we need to prove the problem, but it's stated here because I learned it in AMATH 516.

Proof. We considering connecting a line segment from the projection point z to another point $x \in Q$, then we take the derivative along that line segment.

$$x(t) := z + t(x - z)\ \forall t \in [0, 1] \tag{1.0.30}$$
$$\implies x(0) = z, x(1) = x \tag{1.0.31}$$
$$\varphi(t) := \frac{1}{2}\|y - x(t)\|_2^2 \tag{1.0.32}$$
$$\varphi(t) \ge \varphi(0) \tag{1.0.33}$$
$$Q \text{ Convex } \implies \varphi(t) \in Q\ \forall t \in [0, 1] \tag{1.0.34}$$
$$\implies \lim_{t \searrow 0} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(0) = \langle y - x(0), x'(t)\rangle \tag{1.0.35}$$
$$= \langle y - z, x - z\rangle \ge 0 \tag{1.0.36}$$
$$-\langle z - y, x - z\rangle \ge 0 \tag{1.0.37}$$

Take note that, $\varphi(t) \ge \varphi(0)$, which is how we get the inequality at the second last statement. In this statement we made use of the fact that the line segment is always in the set Q, and the monotone property of $\varphi(t)$ to get the proof work.

Next, we wish to use the minimizer property to show that the set of projections of the set Q onto the point y is unique:

$$
\begin{aligned}
& z, z' \in \underset{Q}{\operatorname{proj}}(y) && (1.0.38)\\
\Longrightarrow\ & \langle y - z, z - z'\rangle \le 0 \ \wedge\ \langle y - z', z - z'\rangle \le 0 && (1.0.39)\\
\Longrightarrow\ & \langle z - z', z - z'\rangle = \|z - z'\|_2^2 \le 0 && (1.0.40)
\end{aligned}
$$

The point $z' = z$ of they were to be the projection and at the same time both satisfying the obtuse angle property. □

1.1 Problem 2.5(i)

Proposition 3. Let $C, D \in \mathbb{R}^n$ such that they are both bounded, closed and convex and $C \cap D = \emptyset$ Then there exists a hyperplane separating elements from the 2 sets.

Proof. We begin the proof by defining the set $C - D := \{x - y : x \in C, y \in D\}$. Immediately obseve that if $C \cap D = \emptyset$, then:

$$
\mathbf{0} \neq a := x^+ - y^+ = \arg\min_{z} \{\|z\| : z \in (C - D)\} \tag{1.1.1}
$$

This minimizer exists and it's going to be unique. This is true by applying the uniqueness of set projection (Last part of the Obtuse angle lamma) together with **Lemma 2** ($C - D$ is also a convex set), in addition, we convince ourselves that the set $C - D$ is also closed and bounded, this is true by the fact that both C, D are closed and bounded, they are compact. So that the minimizer is at least in set $C - D$ and there exists $x^+ \in C, y^+ \in D$ such that $x^+ - y^+ = a$.
The minimizer a won't be $\mathbf{0}$ because $C \cup D = \emptyset$. From here, we make the claim that $y^+ \in C, x^+ \in D$ where $\|x^+\|, \|y^+\| \neq \infty$ are also unique because they satisfies:

$$
x^+ = \operatorname{proj}_C(y^+) \wedge y^+ = \operatorname{proj}_D(x^+) \tag{1.1.2}
$$

Especially if any of x^+, y^+ are not the projection onto the other set (or both), then there is room for improving the distance between $x^+ - y^+$, contradicting the fact that $x^+ - y^+$ is suppose to be the minmizer on the set $C - D$. They are also unique because C, D are convex. Next, we invoke the hyper plane separation theorem to separate the point $\mathbf{0}$ (the origin) with the convex set $C - D$, giving us:

$$
\begin{aligned}
& \langle a, x^+ - y^+\rangle > \delta > 0 = \langle a, 0\rangle && (1.1.3)\\
& \delta := \frac{1}{2}\|x^+ - y^+\|^2 = \frac{1}{2}\|a\|^2 && (1.1.4)\\
\Longrightarrow\ & \langle a, x^+\rangle > \delta + \langle a, y^+\rangle > \langle a, y^+\rangle && (1.1.5)
\end{aligned}
$$

The first line is true by using the hyperplane separation theorem to separate $\mathbf{0}$ with $C-D$, and $x^+ - y^+$ are in $C - D$, and then we simply just move the $-y^+$ around to show the separation.

Next, recall that the set C, D are also convex, and y^+, x^+ are points of projection of other points outside of the set, therefore, we use the Obtuse Angle Lemma of convex set:

$$\begin{aligned}
\forall x \in C : \langle x - x^+, y^+ - x^+ \rangle &\leq 0 && (1.1.6)\\
-\langle a, x - x^+ \rangle &\leq 0 && (1.1.7)\\
\langle a, x - x^+ \rangle &\geq 0 && (1.1.8)\\
\langle a, x \rangle &\geq \langle a, x^+ \rangle && (1.1.9)\\
\implies \langle a, x \rangle &\geq \delta + \langle a, y^+ \rangle && (1.1.10)
\end{aligned}$$

By a smilar token, we derive that:

$$\begin{aligned}
\forall y \in D : \langle y - y^+, x^+ - y^+ \rangle &\leq 0 && (1.1.11)\\
\langle a, y - y^+ \rangle &\leq 0 && (1.1.12)\\
\langle a, y \rangle &\leq \langle a, y^+ \rangle && (1.1.13)\\
\implies \delta + \langle a, y^+ \rangle &\geq \langle a, y \rangle && (1.1.14)
\end{aligned}$$

Therefore, the hyper plan that separating all the points in C, D is: $\{x : \langle a, x \rangle = \delta + \langle a, y^+ \rangle\}$ □

1.2 Problem 2.5 (ii)

Take notice that, the fact that x^+, y^+ exists and they are not just infinite is laying on the fact that the ste $C - D$ is compact. To make separation impossible, we simply consider sets that are not bounded, and having boundary that are approaching each other asymptoptically. For a counter example, consider the set in $\mathbb{R}^2$: $C := \{(x, y) : x = 0\}$ and the set: $D := \{(x, y) : y \geq \frac{1}{x}, x \geq 0\}$. In this case the sets asymptoptically approaches the vertical line at $x = 0$, which make it impossible to choose to closest points in the set C, D, both sets are unbounded but still closed and convex.

2 Problem 2.6(ii)

For this problem, I use the sub-matrix rank theorem and code to assist with looking for all the vertices of the polyhedra. This is a good choice if you give me a 30 by 3 matrix, I can still give you all the vertices in reasonably amount of time. Observe the following Proposition we proved during lecture time:

Proposition 4. Let A be an $m \times n$ matrix.

$$\begin{aligned}
&z \in \{x : Ax \leq b\} \text{ is vertex } \iff \text{rank}(A_{\mathcal{I},:}) = n && (2.0.1)\\
&\mathcal{I} = \{i : (A_{i,:})z = b\} && (2.0.2)
\end{aligned}$$

When coding it up using a computer, I check every possible combinations of rows by brute force and see if the determinant of the sub matrices are non-zero. If it's non zero, then we found a vertice trapped by those tight constraints. To do that, I simply wrote 2 recursive functions that generates all the subsets of $[m]$ with cardinality n, and then I use those sets (stored as a nested vectors of vectors) to index the row of my matrix to get all the possible sub-matrices. In the HW, the matrix define the polytope is 6×3. Giving us a maximum of 20 sub-matrices to check for. Here is the code we have:

```
using LinearAlgebra

"""
    List out all the combinations
"""
function Combinator!(
    s::Vector,
    start::Int, # offset for the array index.
    m::Int,     # number of elements to choose.
    accumulate::Vector,
    results::Vector{Vector}
)
    if m <= 1
        for e in start:(s|>length)
            push!(results, vcat(accumulate, e))
        end
    return end
    for II in start:((s|>length) + 1 - m)
        push!(accumulate, s[II])
        Combinator!(
            s, II + 1, m - 1, accumulate, results
        )
        accumulate |> pop!
    end

return end

function Combinator(s::Int, m::Int)
    s = 1:s |> collect
    v = Vector{Int}()
    r = Vector{Vector}()
    Combinator!(s, 1, m, v, r)
return r end


function VertexSearch(A::AbstractMatrix)
    TightConstraints = Vector{Vector{Int}}()
    Vertices = Vector{Vector{AbstractFloat}}()
    for Indices in Combinator(size(A, 1), size(A, 2))
        SubMatrix = A[Indices, :]
        x = SubMatrix\(b[Indices])
        if (SubMatrix |> det) != 0 && all(A*x .<= b)
            SubMatrix |> display
            println("This is in the Polytope")
            push!(Vertices, x)
        end
    end
    println("List of tight constraints are: ")
    TightConstraints |> display
    println("List of all vertices are")
    Vertices |> display

return end

A = [1 1 0; 0 1 1; 1 0 1; -2 -1 0; 0 -1 -2; -2 0 -1]
b = [2; 4; 3; 3; 3; 2]
VertexSearch(A)
```

Excuting the code produces the following results:

```
3×3 Matrix{Int64}:
1  1  0
0  1  1
1  0  1
This is in the Polytope
3×3 Matrix{Int64}:
    1  1   0
    0  1   1
-2  0  -1
This is in the Polytope
3×3 Matrix{Int64}:
1   1   0
1   0   1
0  -1  -2
This is in the Polytope
3×3 Matrix{Int64}:
    1   1   0
    0  -1  -2
-2   0  -1
This is in the Polytope
3×3 Matrix{Int64}:
    0   1  1
    1   0  1
-2  -1  0
```

```
This is in the Polytope
3×3 Matrix{Int64}:
  0   1   1
 -2  -1   0
 -2   0  -1
This is in the Polytope
3×3 Matrix{Int64}:
  1   0   1
 -2  -1   0
  0  -1  -2
This is in the Polytope
3×3 Matrix{Int64}:
 -2  -1   0
  0  -1  -2
 -2   0  -1
This is in the Polytope
List of tight constraints are:
Vector{Int64}[]
List of all vertices are
8-element Vector{Vector{AbstractFloat}}:
 [0.5, 1.5, 2.5]
 [-1.3333333333333333, 3.3333333333333335, 0.6666666666666666]
 [3.6666666666666665, -1.6666666666666665, -0.6666666666666666]
 [0.19999999999999996, 1.7999999999999998, -2.4]
 [-1.3333333333333335, -0.33333333333333304, 4.333333333333333]
 [-2.25, 1.5, 2.5]
 [1.5, -6.0, 1.5]
 [-0.6666666666666666, -1.6666666666666667, -0.6666666666666666]
```

There are 8 vertices, just like a cube.