Homework 1

Sandy Auttelet

Due: January 16th at 11:59PM

Definition: A set $\Omega \subseteq \mathbb{R}^n$ is said to be convex if for every $x, y \in \Omega$ and all $0 \leq \lambda \leq 1, z = \lambda x + (1-\lambda)y \in \Omega$.

1. (Graded) Construct a formal proof of the following claim:
   Claim: The set $P = \{x \in \mathbb{R}^n | Ax = b, Cx \leq d\}$ is a convex set in $\mathbb{R}^n$, where A, b, C, d are appropriately sized real matrices/vectors.

   Proof:
   Let $z = \lambda x + (1-\lambda)y$ for $0 \leq \lambda \leq 1$ and some $x, y \in P$. Now we need to show that $Az = b$ and $Cz \leq d$ to show that $z \in P$. Suppose we start with the first condition to show that $Az = b$ then,

$$z = \lambda x + (1-\lambda)y \Rightarrow Az = A[\lambda x + (1-\lambda)y] = A\lambda x + A(1-\lambda)y$$

$$= \lambda(Ax) + (1-\lambda)(Ay) = \lambda b + (1-\lambda)b = b.$$

   Then the first condition is satisfied. Now we need to show that $Cz \leq d$ to confirm convexity by,

$$z = \lambda x + (1-\lambda)y \Rightarrow Cz = C[\lambda x + (1-\lambda)y] = C\lambda x + C(1-\lambda)y$$

$$= \lambda(Cx) + (1-\lambda)(Cy) \leq \lambda d + (1-\lambda)d = d.$$

   Then the second condition is satisfied and thus P is a convex set. $\square$

There is no need to read past this point as the grader as the following problems are incomplete and I am just doing these exercises for myself.

Definition: A function $f : \Omega \to \mathbb{R}$ is said to be convex if Ω is convex and for every $x, y \in \Omega$ and all $0 \leq \lambda \leq 1$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if $\forall x, y \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y).$$

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is affine if

$$f(x) = a_0 + \sum_{i=1}^{n} a_i x_i \text{ where } a_0, \ldots, a_n \text{ are scalars.}$$

1. (Not Graded) Exercise 1.1:
   Suppose that a function $f : \mathbb{R}^n \to \mathbb{R}$ is both concave and convex. Prove that f is an affine function.

   Answer:
   Suppose we have a function $f : \mathbb{R}^n \to \mathbb{R}$. The only way for a function to be both concave and convex is for strict equality such that $\forall x, y \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$,

   $$f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda)f(y).$$

   The above statement implies linearity of the function. We can then define $0 = a_0$, $\lambda = a_1$ and $(1 - \lambda) = a_2$ and so on for this sequence. I didn't quite get to the point that $f(x) = x$. Is that what I'm supposed to get? Keep working, maybe use the fact of linearity some how to prove your point....

   Therefore, a function that is concave and convex is also affine. In fact, affine functions are the only functions that are both concave and convex by page 15 of the book.

2. (Not Graded) Exercise 1.2:
   Suppose that $f_1, \ldots, f_m$ are convex functions from $\mathbb{R}^n \to \mathbb{R}$ and let

   $$f(x) = \sum_{i=1}^{m} f_i(x).$$

   Show that,

   (a) If each $f_i(x)$ is convex, so is $f(x)$.

   Answer:
   Suppose each $f_i(x)$ is convex and define $z_i = \lambda x_i + (1 - \lambda)y_i$. We need to show that the sum of convex functions is also convex. We will use the approach of induction to prove this claim. Suppose we start with the base case that $m = 1$,

   $$f_1(z_1) = f_1(\lambda x_1 + (1 - \lambda)y_1) \leq \lambda f_1(x_1) + (1 - \lambda)f_1(y_1)$$

   which implies of course that f is convex. Then suppose $m = 2$ such that

   $$f_1(z_1) = f_1(\lambda x_1 + (1 - \lambda)y_1) \leq \lambda f_1(x_1) + (1 - \lambda)f_1(y_1)$$

   and

   $$f_2(z_2) = f_2(\lambda x_2 + (1 - \lambda)y_2) \leq \lambda f_2(x_2) + (1 - \lambda)f_2(y_2).$$

   If we sum together both $f_1(z_1)$ and $f_2(z_2)$ then

$$
\begin{aligned}
f(z) = f_1(z_1) + f_2(z_2) &= f_1(\lambda x_1 + (1-\lambda)y_1) + f_2(\lambda x_2 + (1-\lambda)y_2) \\
&\le \lambda f_1(x_1) + (1-\lambda)f_1(y_1) + \lambda f_2(x_2) + (1-\lambda)f_2(y_2) \\
&= \lambda(f_1(x_1) + f_2(x_2)) + (1-\lambda)(f_1(y_1) + f_2(y_2)) = \lambda f(x) + (1-\lambda)f(y).
\end{aligned}
$$

so f is convex. Then we suppose f is convex for $m = k$ with some arbitrary $k \in \mathbb{N} > 2$. This means that

$$
\begin{aligned}
f(z) &= f_1(z_1) + f_2(z_2) + \cdots + f_k(z_k) \\
&= f_1(\lambda x_1 + (1-\lambda)y_1) + f_2(\lambda x_2 + (1-\lambda)y_2) + \cdots + f_k(\lambda x_k + (1-\lambda)y_k) \\
&\le \lambda f_1(x_1) + (1-\lambda)f_1(y_1) + \lambda f_2(x_2) + (1-\lambda)f_2(y_2) + \cdots + \lambda f_k(x_k) + (1-\lambda)f_k(y_k) \\
&= \lambda(f_1(x_1)+f_2(x_2)+\cdots+f_k(x_k))+(1-\lambda)(f_1(y_1)+f_2(y_2)+\cdots+f_k(y_k)) = \lambda f(x)+(1-\lambda)f(y).
\end{aligned}
$$

Suppose now that $m = k+1$ and since we know f is convex when $m = k$ we will use that fact here such that,

$$
\begin{aligned}
&f_1(\lambda x_1 + (1-\lambda)y_1) + f_2(\lambda x_2 + (1-\lambda)y_2) + \cdots + f_k(\lambda x_k + (1-\lambda)y_k) \\
&\le \lambda(f_1(x_1) + f_2(x_2) + \cdots + f_k(x_k)) + (1-\lambda)(f_1(y_1) + f_2(y_2) + \cdots + f_k(y_k))
\end{aligned}
$$

and

$$
f_{k+1}(\lambda x_{k+1} + (1-\lambda)y_{k+1}) \le \lambda f_{k+1}(x_{k+1}) + (1-\lambda)f_{k+1}(y_{k+1}).
$$

Summing these two inequalities together we get,

$$
\begin{aligned}
&f_1(\lambda x_1+(1-\lambda)y_1)+f_2(\lambda x_2+(1-\lambda)y_2)+\cdots+f_k(\lambda x_k+(1-\lambda)y_k)+f_{k+1}(\lambda x_{k+1}+(1-\lambda)y_{k+1}) \\
&\le \lambda(f_1(x_1) + f_2(x_2) + \cdots + f_k(x_k)) \\
&+(1-\lambda)(f_1(y_1) + f_2(y_2) + \cdots + f_k(y_k)) + \lambda f_{k+1}(x_{k+1}) + (1-\lambda)f_{k+1}(y_{k+1}) \\
&= \lambda(f_1(x_1)+f_2(x_2)+\cdots+f_k(x_k)+f_{k+1}(x_{k+1})+(1-\lambda)(f_1(y_1)+f_2(y_2)+\cdots+f_k(y_k)+f_{k+1}(y_{k+1})).
\end{aligned}
$$

To simplify by replacing f_i with f we get,

$$
f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y).
$$

Therefore, the sum of convex functions is also in itself a convex function. This concept makes sense considering convex functions are also linear transformations. Since linear transformations can be partitioned into linear combinations, the result falls out of that combination.

(b) If each f_i is piecewise linear and convex, so is f.

Answer:
The idea for this proof is to essentially approximate curvature by discretizing a curved surface. Using Figure 1.2 on page 17, you can visualize this result. Consider the below figure:

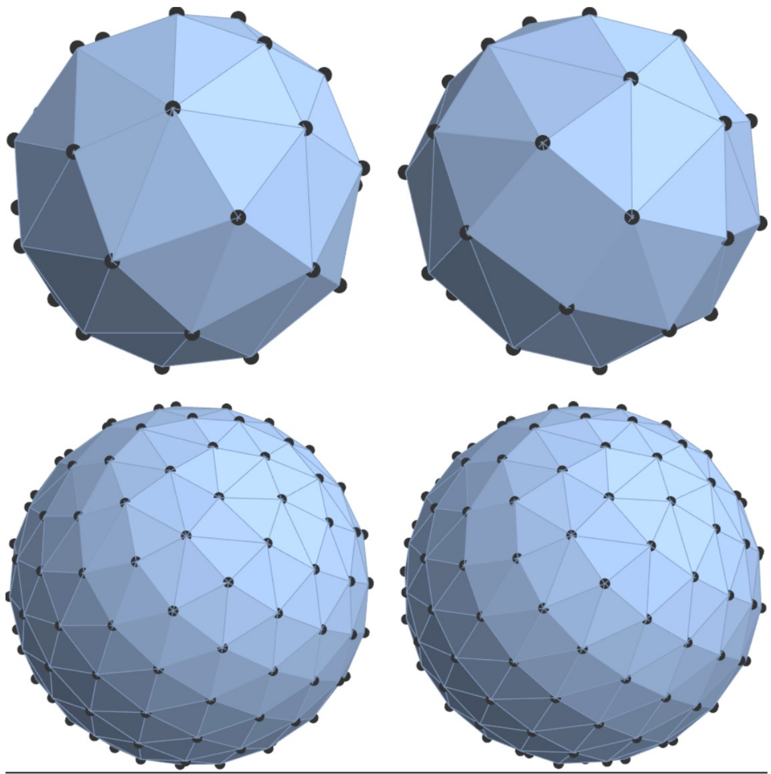

Figure 1: Convex Hull

Here, we can see if we take the limit as $i \to \infty$ for our f_i function and sum all f_i's together to get f, we can get a good approximation of a ball, i.e. the convex hull. We did this in our research, resulting in the normalization reorienting the basis for each point to approximate weights. This problem was in fact a minimization problem using the piecewise linear convex objective function, the flat norm. Explore this proof more with Kevin or Curtis to get a better idea of what it is you did since this is obviously the same problem, but the proof is not obvious yet.