Advanced Linear Algebra Week 8 Day 1

2018/11/05 – Jonathan Hayase, updated by Prof. Weiqing Gu

1 Convexity

1.1 Convexity and optimization

Motivation: Consider optimization $\min f(x)$ such that $x \in \Omega$.

We're interested in global minima. We want to make a statement about global solutions.

Definition: A set $\Omega \subseteq \mathbb{R}^n$ is convex if for $x, y \in \Omega$, all $z = \lambda x + (1-\lambda)y \in \Omega$ for $\lambda \in [0,1]$, and z is called a convex combination.

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain dom is a convex set and for all $x, y \in$ dom for any $\lambda \in [0,1]$ we have

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

For $n = 1$, you could visualizes this as the function always lying below its secants.

Theorem: Consider an optimization problem $\min f(x)$ such that $x \in \Omega$ where f is a convex function and Ω is a convex set. Then any local minimum is a global minimum.

Proof: Let $\overline{x}$ be a local minimum. Therefore, $\overline{x} \in \Omega$ and there exists $\epsilon > 0$ such that $f(\overline{x}) \leq f(x)$ for all $x \in B(\overline{x}, \epsilon)$. Suppose for proof by contradiction, that there exists a $z \in \Omega$ with $f(z) < f(\overline{x})$. Since Ω is convex, $\lambda\overline{x} + (1-\lambda)z \in \Omega$ and since f is convex,

$$\begin{aligned} f(\lambda\overline{x} + (1-\lambda)z) &\leq \lambda f(\overline{x}) + (1-\lambda)f(z) \\ &< f(\overline{x}) + (1-\lambda)f(\overline{x}) \\ &= f(x). \end{aligned}$$

Let $\lambda = 1$, then $f(\overline{x}) < f(\overline{x})$. Thus we have attained our contradiction.

1.2 Midpoint convexity

Definition: A set $\Omega \subseteq \mathbb{R}^n$ is midpoint convex if for $\forall x, y \in \Omega$, the midpoint between x and y is also in Ω,

$$x, y \in \Omega \implies \frac{x+y}{2} \in \Omega$$

Theorem: A closed midpoint convex set is convex.

1.3 Examples of convex sets

1. Hyperplanes $\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}^T\boldsymbol{x} = b\right\}$ where $\boldsymbol{a} \neq 0 \in \mathbb{R}^n$ is the normal vector and $b \in \mathbb{R}$.

2. Half space $\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}^T\boldsymbol{x} \leq b\right\}$ where $\boldsymbol{a} \neq 0 \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

3. Solid ellipsoids $\left\{\boldsymbol{x} \in \mathbb{R}^n \mid (\boldsymbol{x}-\boldsymbol{x}_c)^T P(\boldsymbol{x}-\boldsymbol{x}_c) \leq r\right\}$ where $x_c \in \mathbb{R}^n$, $r \in \mathbb{R}$, and $P \succ 0$.
4. Set of symmetric, positive definite, matrices, $S_+^{n\times n} = \left\{P \in S^{n\times n} \mid P \succeq 0\right\}$. Proof: Let $\lambda \in [0,1]$. Suppose $y \in \mathbb{R}^n$, then

$$y^T(\lambda A + (1-\lambda)B)y = \lambda y^T A y + (1-\lambda) y^T B y \geq 0$$

so $\lambda A + (1-\lambda)B \succeq 0$, as desired.

Say in $\mathbb{R}^3$, $S_+^{2\times 2} = \left\{(x,y,z) \mid \left[\begin{smallmatrix} x & y \\ y & z \end{smallmatrix}\right] \succeq 0\right\}$.

$$\det \begin{bmatrix} x & y \\ y & z \end{bmatrix} = xz - y^2 > 0$$

1.4 Operations

If Ω_1 and Ω_2 are convex sets, $\Omega_1 \cap \Omega_2$ is convex, but $\Omega_1 \cup \Omega_2$ is not always convex.

Proof: Pick $x, y \in \Omega_1 \cap \Omega_2$. Then $\lambda x + (1-\lambda)y \in \Omega_1$, since $x, y \in \Omega_1$ and $\lambda x + (1-\lambda)y \in \Omega_2$ since $x, y \in \Omega_2$.

Application: A polyhedron can be written as $\left\{x | Ax \leq b\right\}$ where $A_{m\times m}$ and $b_{m\times 1}$.

Example:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \end{bmatrix}$$

2 Relations between convex functions and convex sets

2.1 Epigraphs

The epigraph epi(f) of a function $f : \mathbb{R}^n \to \mathbb{R}$ is a subset of $\mathbb{R}^{n+1}$ defined by

$$\text{epi} = \left\{(x,y) \mid x \in \text{dom}(f), f(x) \leq t\right\}$$

Recall: level curves $f(\boldsymbol{x}) = t$. If the function is convex, then the region bounded by its level curves is also convex.

Theorem: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if its epigraph is convex.

2.2 Convex Hulls

Given $x_1 \ldots, x_m \in \mathbb{R}^n$, a point of the form $\lambda_1 x_1 + \cdots + \lambda_m x_m$ where $\sum_{i=1}^m \lambda_i = 1$ is called a convex combination.

Lemma: Let $S \subseteq \mathbb{R}^n$ is convex iff it contains every convex combination of its points.

Definition: The convex hull of a set $S \subseteq RR^n$ denoted by conv(S) is the set of all convex combinations of the points in S.

$$\text{conv}(s) = \left\{\sum_{i=1}^m \lambda_i x_i \;\middle|\; x_i \in S, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1\right\}$$

3 Carathéodory's theorem

Consider a set $S \subseteq \mathbb{R}^d$. Then every point in $\text{conv}(S)$ can be written as a convex combination of $d+1$ points in S.

Prof. Gu came up with a solution (?) using some algebra.

Proof: Let $x \in \text{conv}(S)$. Then $x = \alpha_1 y_1 + \cdots + \alpha_m y_m$ for $y_i \in S$, $\alpha_i \geq 0$, $\sum \alpha_i = 1$. If $m < d$, we are done.

Suppose $m > d+1$, then we'll give another representation of x using $m-1$ points. (Rest later)

4 Separating Hyperplane Theorem

Theorem: Let C and D be two convex sets in $\mathbb{R}^n$ that do not intersect ($C \cap D = \emptyset$). Then there exists an $a \neq 0 \in \mathbb{R}^n$, $b \in \mathbb{R}$ such that the plane $\left\{x \in \mathbb{R}^n \mid a^T x = b\right\}$ separates the points of C and D. Written another way, $a^t x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$.

Remark: Neither of the inequalities can be made strict. Counterexample: suppose C is the negative real numbers and D is the nonnegative real numbers.

5 Hahn-Banach separation Theorem

Let $K \in \{\mathbb{R}, \mathbb{C}\}$ and V is a topological vector space over K. If A, B are convex and $A \cap B = \emptyset$, then if A open, then

1. There exists a continuous linear map from $\lambda : V \to K$ such that

$$\Re(\lambda(a)) < t \leq \Re(\lambda(b))$$

   for all $a \in A, b \in B$.

2. If V is locally compact, B is closed, then there exists a continuous linear map $\lambda : V \to K$ and $s, t \in \mathbb{R}$ such that $\Re(\lambda(a)) < t < s < \Re(\lambda(b))$ for $a \in A$ and $b \in B$.