Convex Analysis

Mark Dean

Lecture Notes for Fall 2014 PhD Class - Brown University

1 Lecture 1

1.1 Introduction

We now move onto a discussion of convex sets, and the related subject of convex function. As we will see in a minute, the basic idea of convex sets is that they contain their own line segments: if I take any two points in the set, and draw a line between them, then all the points along that line are in the set. Convex sets are extremely important for a number of purposes. Perhaps from our point of view, the most useful is their role in optimization (this may ring a few bells - hopefully it will ring more as we go along)

1.2 Convex Sets

We begin by defining a convex set. As I said in the introduction, the key idea is that if I take any two points in a convex set and 'walk' from one point to another in a straight line, then I will not leave the set. Of course, I need to formalize this notion.

Definition 1 *Let V be a linear space. A subset $S \subset V$ is convex if*

$$\lambda x + (1-\lambda)y \in S \ \forall \ x, y \in S, \ \lambda \in (0,1)$$

We sometimes call $\lambda x + (1-\lambda)y$ a 'line segment'. As you can see, it is effectively a weighted average of the two points x and y. Note that, in order to define the idea of a convex set we need a notion of addition and scalar multiplication - the two properties that define a linear space. This is one of the reasons we spent so long discussing linear spaces at the start of the course.

We can extend the notion of a linear segment to more than two points in a set in the following way:

Definition 2 *A convex combination of a set S is a vector*

$$s = \lambda_1 s_1 + \lambda_2 s_2 + ... + \lambda_n s_n$$

where $n \in \mathbb{N}$, $s_i \in S \ \forall \ i$, $\lambda_i \in \mathbb{R}_+ \ \forall \ i$ and $\sum_{i=1}^{n} \lambda_i = 1$

Note how close this is to the idea of the span of a set S. The crucial difference is that we demand $\sum_{i=1}^{n} \lambda_i = 1$, and so (in some sense) we can only project inward from a set of points, rather than outward, as is allowed in the concept of a span. Note that, for convenience, we will use $P_n = \{\lambda \in \mathbb{R}^n | \lambda_i \geq 0 \forall i, \sum_{i=1}^{n} \lambda_i = 1\}$

The idea of a convex combination allows for an alternative characterization of a convex set

Lemma 1 *A set $S \subset M$ is convex if and only if it contains all convex combinations of S*

Proof. *The fact that a set that contains all its convex combinations is convex is trivial. We prove that a convex set contains all its convex combinations we prove by induction on k, the number of vectors used to form the convex combination. The fact that it is true for $k = 1$ (and 2) is trivial, so now assume it is true for k and we need to prove that it is true for $k + 1$. Let*

$$s = \lambda_1 s_1 + \lambda_2 s_2 + ... + \lambda_{k+1} s_{k+1}$$

be a convex combination of elements in S. Note that

$$\bar{s} = \frac{\lambda_2}{\sum_{i=2}^{k+1} \lambda_i} s_2 + ... + \frac{\lambda_{k+1}}{\sum_{i=2}^{k+1} \lambda_i} s_{k+1}$$

is a convex combination of k elements in S. By induction, $\bar{s} \in S$. But then

$$\begin{aligned} s &= \lambda_1 s_1 + \lambda_2 s_2 + ... + \lambda_{k+1} s_{k+1} \\ &= \lambda_1 s_1 + \sum_{i=2}^{k+1} \lambda_i \left(\frac{\lambda_2}{\sum_{i=2}^{k+1} \lambda_i} s_2 + ... + \frac{\lambda_{k+1}}{\sum_{i=2}^{k+1} \lambda_i} s_{k+1} \right) \\ &= \lambda_1 s_1 + \sum_{i=2}^{k+1} \lambda_i \bar{s} \\ &= \lambda_1 s_1 + (1 - \lambda_1) \bar{s} \end{aligned}$$

so, as S is convex, $s \in S$. ■

A useful structure that we are going to come back to time and again is the **convex hull** of a set. This is the smallest convex set that contains the set

Definition 3 *The convex hull of a set S is defined as*

$$co(S) = \cap \{C | C \text{ is convex and } S \subset C\}$$

While it is a useful property, it can be difficult to identify the convex hull of a particular set. A useful theorem in this regard is that the convex hull is equal to the set of convex combinations of elements is that set.

Theorem 1 *For any set S, $co(S) = K(S)$, where $K(S) = \{x | s \text{ is a convex combination of } S\}$*

Proof. *We first show that $K(S)$ is a subset of $co(S)$. Let $s = \lambda_1 s_1 + \lambda_2 s_2 + ... + \lambda_n s_n$ for some collection $\{s_i\}_{i=1}^n$ of vectors in S. Let C be any convex set that contains S. By lemma 1, we know that $s \in C$. Thus $K(S) \subset C$, and so $K(S) \subset co(S)$*

Now we need to show that $co(S)$ is a subset of $K(S)$. All we need to show is that $K(S)$ is convex. To see this, take x, y such that

$$\begin{aligned} x &= \sum_{i=1}^{m} \alpha_i x_i \\ y &= \sum_{i=1}^{n} \beta_i y_i \end{aligned}$$

where $\sum_{i=1}^{m}\alpha_i = 1$, $\sum_{i=1}^{n}\beta_i = 1$ *and* $x_i, y_i \in S\ \forall\ i$. *Then*

$$\begin{aligned} & \mu x + (1-\mu)y \\ = \ & \mu\sum_{i=1}^{m}\alpha_i x_i + (1-\mu)\sum_{i=1}^{n}\beta_i y_i \\ = \ & \sum_{i=1}^{m}\mu\alpha_i x_i + \sum_{i=1}^{n}(1-\mu)\beta_i y_i \end{aligned}$$

but, this is a convex combination of S, *as* $x_i, y_i \in S\ \forall\ i$ *and*

$$\begin{aligned} & \sum_{i=1}^{m}\mu\alpha_i + \sum_{i=1}^{n}(1-\mu)\beta_i \\ = \ & \mu\sum_{i=1}^{m}\alpha_i + (1-\mu)\sum_{i=1}^{n}\beta_i \\ = \ & \mu + (1-\mu) \\ = \ & 1 \end{aligned}$$

■

A useful extension of this proved by Caratheodory is that, in $\mathbb{R}^n$ we can generate the convex hull by taking convex combinations using at most $n+1$ vectors.

Theorem 2 *For any set* $S \in \mathbb{R}^n$, $co(S) = \left\{\sum_{i=1}^{n+1}\lambda_i x_i | x_i \in S \forall i, \lambda \in P_{n+1}\right\}$

Proof. *We know that*

$$\left\{\sum_{i=1}^{n+1}\lambda_i x_i | x_i \in S \forall i, \lambda \in P_{n+1}\right\} \subset K(S)$$

, and so

$$\left\{\sum_{i=1}^{n+1}\lambda_i x_i | x_i \in S \forall i, \lambda \in P_{n+1}\right\} \subset co(S)$$

thus, all we need to show is that $co(S) \subset \left\{\sum_{i=1}^{n+1}\lambda_i x_i | x_i \in S \forall i, \lambda \in P_{n+1}\right\}$.

Let $x \in co(S) = K(S)$. *Then there exists an* $y_1, ... y_m \in S$ *and* $\lambda \in P_m$ *such that*

$$x = \sum_{i=1}^{m}\lambda_i y_i$$

We only need to worry about the case where $m > n+1$. *In this case, the set* $\{y_1 - y_m, ..., y_{m-1} - y_m\}$ *is linearly dependent, so there is some* $\beta \neq 0$ *such that*

$$\sum_{i=1}^{m-1}\beta_i\left(y_i - y_m\right) = 0$$

Let $\beta_m = -\sum_{i=1}^{m-1} \beta_i$*, then*

$$\begin{aligned} \sum_{i=1}^{m} \beta_i y_i &= 0 \\ \sum_{i=1}^{m} \beta_i &= 0 \end{aligned}$$

But, as

$$\begin{aligned} x &= \sum_{i=1}^{m} \lambda_i y_i \\ &\Rightarrow \quad x = \sum_{i=1}^{m} \lambda_i y_i - t \sum_{i=1}^{m} \beta_i y_i \\ &= \sum_{i=1}^{m} (\lambda_i - t\beta_i)\, y_i \end{aligned}$$

Let $\bar{t} = \min\left\{\frac{\lambda_i}{\beta_i} | \beta_i > 0\right\} := \frac{\lambda_j}{\beta_j}$ *and* $\alpha_i = \lambda_i - \bar{t}\beta_i$

Notice that

$$\begin{aligned} \lambda_i - \bar{t}\beta_i &= \lambda_i - \frac{\lambda_j}{\beta_j}\beta_i \geq 0 \\ as\ \frac{\lambda_j}{\beta_j} &\leq \frac{\lambda i}{\beta i} \end{aligned}$$

and

$$\sum_{i=1}^{m} \alpha_i = \sum_{i=1}^{m} \lambda_i - \bar{t}\beta_i = \sum_{i=1}^{m} \lambda_i = 1$$

Also,

$$\alpha_j = \lambda_j - \frac{\lambda_j}{\beta_j}\beta_j = 0$$

so

$$x = \sum_{\substack{i=1 \\ i \neq j}}^{m} \alpha_i y_i$$

We can therefore discard one of the vectors in $y_i, ...y_m$*. Iterating on this procedure we can get down to* $n+1$ *vectors* ■