Miscellenous Notes

University of Waterloo
Nicholas Pun

Contents

CO 463 - Convex Optimization

1.1 2020/01/08 - Introduction

Definition 1.1.1. Let $x, y \in \mathbb{R}^n$, $0 \leq \lambda \leq 1$, $z(\lambda) = \lambda x + (1-\lambda)y$ is a convex combination of x and y

This simple definition leads to many strong algebraic and topological results.

For this course, we will work in Euclidean space $\mathbb{E}^n$ with inner product $\langle x, y\rangle$ and norm $||x|| = \sqrt{\langle x, y\rangle}$.

On $\mathbb{R}^n$, we will use the familiar dot product $\langle x, y\rangle = x^\intercal y = \sum_{i=1}^n x_i y_i$ and norm $||x|| = \sqrt{\sum_{i=1}^n x_i^2}$.

Definition 1.1.2. $C \subseteq \mathbb{E}^n$ is convex set if $x, y \in C$ $0 \leq \lambda \leq 1 \Rightarrow \lambda x + (1-\lambda)y \in C$ (or equivalently, $x, y \in C \Rightarrow [x, y] \subseteq C$)

Note. We can write $z(\lambda)$ in the following ways:

$$\begin{aligned} z(\lambda) &= \lambda x + (1-\lambda)y \\ &= y + \lambda(x-y) \\ &= x + (1-\lambda)(y-x) \end{aligned}$$

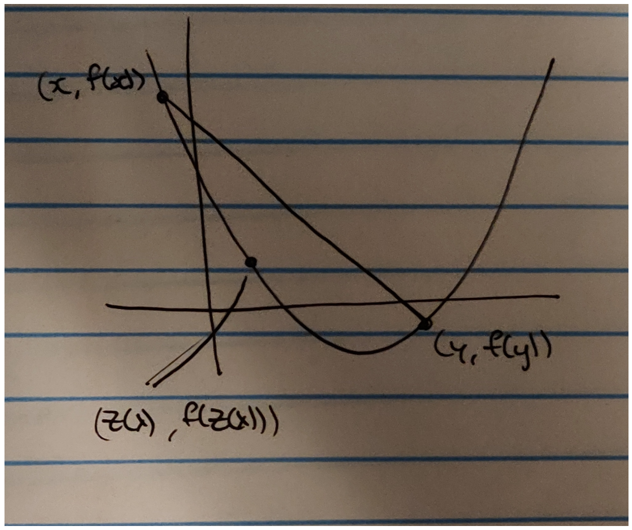
The second equation can be interpreted as: "Beginning at y and moving towards x". And, the third equation can be interpreted as: "Beginning at x and moving towards y"

Definition 1.1.3. Let $C \subseteq \mathbb{E}^n$, C convex set. Then $f : C \to \mathbb{R}$ is a convex function if:

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

$\forall x, y \in C,\ \forall 0 \leq \lambda \leq 1$

Example 1.1.1.



The line between $(x, f(x))$ and $(y, f(y))$ is called the secant line. For convex function, the graph lies below the secant line.

The region above the graph is called the epigraph, denoted $\operatorname{epi} f$. It is defined as:

$$\operatorname{epi} f = \{(r, x) \in \mathbb{R} \times \mathbb{E}^n \;:\; x \in C, f(x) \subseteq r\}$$

for $f : C \to \mathbb{R}$, C convex.

We will see that:

Theorem 1.1.1. f is a convex function if and only if $\operatorname{epi} f$ is a convex set

Suppose we have the following minization problem:

$$\begin{aligned} p^* = &\min f(x) \\ &\text{s.t } x \in C \end{aligned}$$

where $C \in \mathbb{E}^n$ convex, $f : C \to \mathbb{R}^n$ convex function.

We will show that: f convex $\Rightarrow$ (local min iff global min)

Applications:

- Linear Programming
- Convex Relaxation

1.2 2020/01/10 - Linear Algebra and Calculus Review

Definition 1.2.1. The closed ball with center $\overline{x} \in \mathbb{E}$ and radius δ is

$$B_\delta(\overline{x}) = \{x \;:\; ||x - \overline{x} \leq \delta\}$$

The unit ball is denoted $B(\overline{x})$

We say $S \subseteq \mathbb{E}$ is bounded if $\exists \gamma > 0$ such that $x \in S \Rightarrow ||x|| \leq \gamma$

Let $\mathbb{E} = M^{m \times n}$ ($m \times n$ matrices) with the trace inner product and the frobenius norm. If we have a mapping $L : \mathbb{E} \to \mathbb{F}$ between Euclidean spaces, this is called a linear transformation if:

$$L(ax + by) = aL(x) + bL(y)$$

$\forall a, b \in \mathbb{R},\ \forall x, y \in \mathbb{E}$

There are many examples of linear transformations:

Example 1.2.1. Matrix-Vector multiplication, i.e. Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be an $m \times n$ matrix. A is a linear transformation.

Example 1.2.2. Let $L : \mathcal{S}^n \to \mathbb{R}^n$ (where $\mathcal{S}$ is the set of $n \times n$ real symmetric matrices). $L(x) = diag(x)$ is a linear transformation.

Let $L : \mathbb{E} \to \mathbb{F}$ be a linear transformation, the adjoint of L is L^*, defined by $\langle L(x), y\rangle = \langle x, L^*(y)\rangle\ \forall x \in \mathbb{E},\ \forall y \in \mathbb{F}$

What is $diag^*$?: Let $A \in M^{n \times n}$,

$$diag(A) = \begin{pmatrix} A_{11} \\ \vdots \\ A_{nn} \end{pmatrix}$$

$$diag^*\left(\begin{pmatrix} A_{11} \\ \vdots \\ A_{nn} \end{pmatrix}\right) = \begin{bmatrix} A_{11} & & 0 \\ & \ddots & \\ 0 & & A_{nn} \end{bmatrix}$$

We can verify this by computing: $\langle diag(A), v\rangle = \langle A, diag^*(v)\rangle$

Definition 1.2.2. If $L : \mathbb{E} \to \mathbb{F}$ is a linear transformation, then L is called a linear operator.

Definition 1.2.3. If $L : \mathbb{E} \to \mathbb{F}$ and $L = L^*$, then L is a self-adjoint operator and $\mathbb{E} = \mathbb{F}$

1.2.1 Differentiation

Let $f : \mathbb{E} \to \mathbb{R}$ be a real-valued function.

We say f is $\mathcal{C}^1$ if its first partial derivatives are continuous. Similarly, we say f is $\mathcal{C}^2$ if the second partial derivatives are continuous.

WLOG, assume $A = A^\intercal$, and say we have the following quadratic function:

$$q(x) = \underbrace{\frac{1}{2}x^\intercal Ax}_{=\langle x, Ax\rangle} + \underbrace{b^\intercal x}_{=\langle b,x\rangle} + c$$

Then,

$$\begin{aligned} q(x+d) &= \frac{1}{2}(x+d)^\intercal A(x+d) + b^\intercal(x+d) + c \\ &= \frac{1}{2}x^\intercal Ax + b^\intercal x + c + (Ax)^\intercal + b^\intercal d + \frac{1}{2}d^\intercal Ad \\ &= q(x) + \langle Ax+b, d\rangle + \frac{1}{2}d^\intercal Ad \end{aligned}$$

This is a Taylor series! $\langle Ax+b, d\rangle$ is linear in d and $\frac{1}{2}d^\intercal Ad$ is in $o(||d||)$ if $A \neq 0$

$\nabla q(x) = Ax + b$ is the <u>gradient</u> of q at x, and in general, if $f(x+d) = f(x) + \langle v, d\rangle + o(||d||)$, then $v = \nabla f(x)$.

1.3 2020/01/15 - Review and Convex Sets

CO 739 - Information Theory and Applications