凸函数的世界

凸分析和凸优化

枫聆

2021 年 6 月 24 日

目录

数学优化问题

Definition 1.1. 数学优化问题或者说优化问题可以写成如下形式

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le b_i,\ i = 1, 2, \cdots, m.
\end{array}
$$

其中向量 $x = (x_1, \cdots, x_n) \in \mathbf{R}^n$ 称为问题的优化变量，函数 $f_0 \colon \mathbf{R}^n \to \mathbb{R}$ 称为目标函数，函数 $f_i \colon \mathbf{R}^n \to \mathbb{R}$ 被称为约束函数，常数 b_i 称为约束上限或者约束边界.

Definition 1.2. 那些满足约束的向量 z，即使得上述不等式成立的向量，它们构成一个解集 Z. 这个解集中使得 $f_0(z)$ 最小的那些 x^* 称为当前优化问题的最优解，即

$$\forall z \in Z, f_i(z) \le b_i,\ i = 1, 2, \cdots, m \text{ and } f(x^*) \le f(z).$$

基本概念

仿射集 (affine set)

Definition 2.1. ($\mathbf{R}^n$(n 维实数向量空间) 上直线的定义) 对任意两个 $\mathbf{R}^n$ 中不同两个 $\mathbf{x}$ 和 $\mathbf{y}$，形如

$$\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) = (1-\lambda)\mathbf{x} + \lambda\mathbf{y},\ \lambda \in \mathbf{R}$$

的点集被称为过 $\mathbf{x}$ 和 $\mathbf{y}$ 的直线.

Definition 2.2. 对于 $\mathbf{R}^n$ 中的子集 M，如果对于任意的 $\mathbf{x}, \mathbf{y} \in M$ 和 $\lambda \in \mathbf{R}$ 都有 $(1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in M$，则称 M 为 $\mathbf{R}^n$ 中的仿射集(affine set). 相关书与仿射集同义的名词有仿射流形(affine manifold)，仿射变量(affine variety)，线性变量(linear variety) 或者flat(平坦的).

Theorem 2.3. $\mathbf{R}^n$ 上的所有子空间都是仿射集. 反过来含 $\mathbf{0}$ 的仿射集都是子空间.

证明. 由向量空间子空间的定义，给定任意的子空间 V，在scalar multiplication和vector addition下封闭的，自然也满足仿射集的定义.

相反，若给定一个仿射集 M，有 $\mathbf{0} \in M$，因此

$$\forall \mathbf{x} \in M, \lambda x = (1-\lambda)\mathbf{0} + \lambda\mathbf{x} \in M,$$

这就证明了scalar multiplication. 对任意的 $\forall \mathbf{x}, \mathbf{y} \in M$，有

$$\mathbf{x} + \mathbf{y} = 2(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}) \in M.$$

即证明了vector addition. □

Definition 2.4. 对于 $M \subseteq \mathbf{R}^n$ 及 $\mathbf{a} \in M$. M 关于 $\mathbf{a}$ 的平移(translate) 定义为

$$M + \mathbf{a} = \{\, \mathbf{x} + \mathbf{a} \mid \mathbf{x} \in M \,\}.$$

Definition 2.5. 给定仿射集 M 和仿射集 L，若可以找到一个 $\mathbf{a}$ 使得满足关系

$$M = L + \mathbf{a},$$

则称 M 和 L 平行.

Proposition 2.6. 仿射集的平移仍为仿射集.

证明. 给定仿射集 M 和它的平移 $M + \mathbf{a}$，取任意的 $\mathbf{x}, \mathbf{y} \in M$，那么

$$(1-\lambda)(\mathbf{x} + \mathbf{a}) + \lambda(\mathbf{y} + \mathbf{a}) = (1-\lambda)\mathbf{x} + \lambda\mathbf{y} + \mathbf{a} \in M + \mathbf{a}.$$

□

Proposition 2.7. 仿射集的平行是一种等价关系.

证明. 自反性，取 $\mathbf{a} \in M$ 即得；对称性，若 $M = L + \mathbf{a}$，则 $L = M + (-\mathbf{a})$；传递性, 若 $M = L + \mathbf{a}, L = L_1 + \mathbf{b}$，则 $M = L_1 + (\mathbf{a} + \mathbf{b})$. □

Theorem 2.8. 每个非空仿射集 M 一定平行于唯一的子空间 L. 这个 L 由

$$L = M - M = \{ \mathbf{x} - \mathbf{y} \mid \mathbf{x}, \mathbf{y} \in M \}.$$

给定.

证明. 先证唯一，假设 M 同时平行于两个子空间 L_1, L_2，我们前面已经证明平行是一个等价关系，因此存在 $\mathbf{a}$ 使得 $L_1 = L_2 + \mathbf{a}$，由于 $\mathbf{0} \in L_1$，所以 $-\mathbf{a} \in L_2$，所以 $\mathbf{a} \in L_2$, 那么 $L_1 = L_2$(相当于往子空间里面在扔一个它自己的元素进去，并不会改变这个子空间).

现在我们要找到这样一个子空间, 根据前面的定理知道含 $|0$ 的仿射集才是子空间, 很明显使得 $\mathbf{a} \in M$, 那么 $\mathbf{0} \in M - \mathbf{a}$. 这样的 $\mathbf{a}$ 是任意的, 并且所有仿射 $M + \mathbf{a}$ 都是相同的, 所以这样所以 $L = \sum_{a \in M} M - a = M - M$. □

Definition 2.9. 与非空仿射集平行的子空间的维数称为这个仿射集的维数. 维数为 0，1 和 2 分别称为点，线，面. $\mathbf{R}^n$ 中 $(n-1)$ 维仿射集被称为超平面(hyperplane).

Definition 2.10. 给定 $\mathbf{R}^n$ 上子空间 L，对于任意 $\mathbf{y} \in L$，使得关系

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0$$

成立的 $x \in \mathbf{R}^n$ 构成的集合称为 L 的正交补(orthogonal complement)，记为 $L^\perp$. 显然 $L^\perp$ 也是一个子空间，且满足

$$\dim L + \dim L^\perp = n.$$

Definition 2.11. (建立在正交性理论上平行于超平面仿射集表示方法) 若一维子空间是有当个单个非零向量 $\mathbf{b}$ 构成，那么其对应的正交 $n-1$ 维子空间可以表示为 $\{ \mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} \perp \mathbf{b} \}$. 自然地，与其平行的仿射集为

$$\begin{aligned} \{ \mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} \perp \mathbf{b} \} + \mathbf{a} \quad &= \{ \mathbf{x} + \mathbf{a} \mid \langle \mathbf{x}, \mathbf{b} \rangle = 0 \text{ and } \mathbf{x} \in \mathbf{R}^n \} \\ &= \{ \mathbf{y} \in \mathbf{R}^n \mid \langle \mathbf{y} - \mathbf{a}, \mathbf{b} \rangle = 0 \} \\ &= \{ \mathbf{y} \in \mathbf{R}^n \mid \langle \mathbf{y}, \mathbf{b} \rangle = \beta \} \end{aligned},$$

其中 $\beta = \langle \mathbf{a}, \mathbf{b} \rangle$.

Theorem 2.12. 给定 $\beta \in \mathbf{R}$ 以及非零 $\mathbf{b} \in \mathbf{R}$，集合

$$H = \{ \mathbf{x} \in \mathbf{R}^n \mid \langle \mathbf{x}, \mathbf{b} \rangle = \beta \}$$

为 $\mathbf{R}^n$ 的超平面.

证明. □