优化方法

2016-2017学年工科硕士课程

第1章　引论

1.0 数学基础知识

- n维向量, $a = (a_1, a_2, \cdots, a_n)^T$,

1.0 数学基础知识

- n维向量, $a = (a_1, a_2, \cdots, a_n)^T$,
- 内积, $a \cdot b = a^T b$,

1.0 数学基础知识

- n维向量, $a = (a_1, a_2, \cdots, a_n)^T$,
- 内积, $a \cdot b = a^T b$,
- 内点,边界点,闭集,开集,

1.0 数学基础知识

- n维向量, $a = (a_1, a_2, \cdots, a_n)^T$,
- 内积, $a \cdot b = a^T b$,
- 内点,边界点,闭集,开集,
- 多元函数$f : \mathbf{R}^n \to \mathbf{R}$, f的梯度

$$\mathrm{grad} f(x) = \nabla f(x) = (\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n})^T,$$

- n维向量，$a = (a_1, a_2, \cdots, a_n)^T$,
- 内积，$a \cdot b = a^T b$,
- 内点,边界点,闭集,开集,
- 多元函数$f : \mathbf{R}^n \to \mathbf{R}$, f的梯度

$$\mathrm{grad} f(x) = \nabla f(x) = (\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n})^T,$$

- f的海森阵(Hesse阵)

$$\nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{n \times n}.$$

方向导数和泰勒公式

- f在x处关于方向d的方向导数,

$$f^{'}(x,d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t} = \nabla f(x)^T d.$$

方向导数和泰勒公式

- f在x处关于方向d的方向导数，

$$f'(x,d)=\lim_{t\downarrow 0}\frac{f(x+td)-f(x)}{t}=\nabla f(x)^T d.$$

方向导数和泰勒公式

- f在x处关于方向d的方向导数,

$$f'(x,d)=\lim_{t\downarrow 0}\frac{f(x+td)-f(x)}{t}=\nabla f(x)^T d.$$

几何意义?

方向导数和泰勒公式

- f在x处关于方向d的方向导数，

$$f'(x,d) = \lim_{t\downarrow 0} \frac{f(x+td)-f(x)}{t} = \nabla f(x)^T d.$$

  几何意义？

- 多元函数的泰勒展开公式，

方向导数和泰勒公式

- f在x处关于方向d的方向导数,

$$f'(x,d) = \lim_{t \downarrow 0} \frac{f(x+td)-f(x)}{t} = \nabla f(x)^T d.$$

几何意义?

- 多元函数的泰勒展开公式,

方向导数和泰勒公式

- f在x处关于方向d的方向导数，

$$f^{'}(x, d) = \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t} = \nabla f(x)^T d.$$

几何意义？

- 多元函数的泰勒展开公式，

$$\begin{aligned} f(x) = \quad & f(x_0) + \nabla f(x_0)^T (x - x_0) \\ & + \frac{1}{2}(x - x^0)^T \nabla^2 f(x)(x - x_0) + o(\|x - x_0\|^2). \end{aligned}$$

方向导数和泰勒公式

- f在x处关于方向d的方向导数,

$$f'(x,d) = \lim_{t\downarrow 0} \frac{f(x+td)-f(x)}{t} = \nabla f(x)^T d.$$

几何意义?

- 多元函数的泰勒展开公式,

$$\begin{aligned} f(x) = \quad & f(x_0) + \nabla f(x_0)^T (x - x_0) \\ & + \frac{1}{2}(x - x^0)^T \nabla^2 f(x)(x - x_0) + o(\|x - x_0\|^2). \end{aligned}$$

- 点x_0的ε邻域

$$N(x_0,\varepsilon) = \{x \| \|x - x_0\| < \varepsilon\}.$$

其中向量模$\|x\| = [\sum_{i=1}^n x_i^2]^{1/2}$,称为欧氏范数. $\|x - x_0\|$表示两点之间的距离,称为欧氏距离.

利用最优化的理论和方法解决生产实际和自然科学中的具体问题,一般分为两个步骤:

(1) 建立数学模型.

(2) 进行数学加工和求解.

例1.1

某化工厂生产A,B,C,D四种产品,生产每种产品一吨所消耗的工时和产值如下表:

产品	A	B	C	D
工时(小时)	100	300	400	75
产值(千元)	1	5	10	0.5

要求全厂产值在1000万元以上,求当消耗的总工时最少时,该厂生产各种产品的数量.

例1.1

某化工厂生产A,B,C,D四种产品,生产每种产品一吨所消耗的工时和产值如下表:

产品	A	B	C	D
工时(小时)	100	300	400	75
产值(千元)	1	5	10	0.5

要求全厂产值在1000万元以上,求当消耗的总工时最少时,该厂生产各种产品的数量.

$$\begin{aligned} \min \quad & y = 100x_1 + 300x_2 + 400x_3 + 75x_4 \\ \text{s.t.} \quad & x_1 + 5x_2 + 10x_3 + 0.5x_4 \geq 10000, \\ & x_i \geq 0, \quad i = 1, 2, 3, 4. \end{aligned} \tag{1-1}$$

例1.7

设有B亿元资金可用于投资,有m个项目$A_1, A_2, \cdots, A_m$可供挑选. 若对项目A_i进行投资,需花费资金a_i亿元,可获益c_i亿元,试确定最佳的投资方案.

例1.7

设有B亿元资金可用于投资,有m个项目$A_1, A_2, \cdots, A_m$可供挑选. 若对项目A_i进行投资,需花费资金a_i亿元,可获益c_i亿元,试确定最佳的投资方案.

$$
\begin{aligned}
\min \quad & f_1(x) = \sum_{i=1}^{m} a_i x_i \\
\max \quad & f_2(x) = \sum_{i=1}^{m} c_i x_i \\
\text{s.t.} \quad & \sum_{i=1}^{m} a_i x_i \leq B, \\
& x_i \in I = \{0, 1\}, \quad i = 1, \cdots, m.
\end{aligned}
\tag{1-7}
$$

1.2.1 最优化问题的提法和基本概念

最优化问题可简写为

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \ i = 1, 2, \cdots, m, \\ & h_j(x) = 0, \ j = m+1, m+2, \cdots, p. \end{aligned} \tag{2-1}$$

令

$$R = \{x | g_i(x) \leq 0, i = 1, 2, \cdots, m; h_j(x) = 0, \ j = m+1, m+2, \cdots, p\},$$

称R为问题(2-1)的可行集或容许集,称$x \in R$ 为可行解或容许解.

最优解的定义

定义2.1

- 若有$x^* \in R$,使得$\forall x \in R$,均有$f(x^*) \leq f(x)$,则称x^*为问题(2-1)的(全局)最优解(点)或全局极小点.

最优解的定义

定义2.1

- 若有$x^* \in R$,使得$\forall x \in R$,均有$f(x^*) \leq f(x)$,则称x^*为问题(2-1)的(全局)最优解(点)或全局极小点.
- 若有$x^* \in R$,使得$\forall x \in R \cap N(x^*, \varepsilon)$,均有$f(x^*) \leq f(x)$,则称$x^*$为问题(2-1)的局部最优解(点)或局部极小点.

最优解的定义

定义2.1

- 若有$x^* \in R$,使得$\forall x \in R$,均有$f(x^*) \le f(x)$,则称x^*为问题(2-1)的(全局)最优解(点)或全局极小点.
- 若有$x^* \in R$,使得$\forall x \in R \cap N(x^*, \varepsilon)$,均有$f(x^*) \le f(x)$,则称$x^*$为问题(2-1)的局部最优解(点)或局部极小点.
- 严格全局极小点, $\forall x \in R$, $x \neq x^*$, $<$

最优解的定义

定义2.1

- 若有$x^* \in R$,使得$\forall x \in R$,均有$f(x^*) \leq f(x)$,则称x^*为问题(2-1)的(全局)最优解(点)或全局极小点.
- 若有$x^* \in R$,使得$\forall x \in R \cap N(x^*, \varepsilon)$,均有$f(x^*) \leq f(x)$,则称$x^*$为问题(2-1)的局部最优解(点)或局部极小点.
- 严格全局极小点, $\forall x \in R$, $x \neq x^*$, $<$
- 严格局部极小点, $\forall x \in R \cap N(x^*, \varepsilon)$, $x \neq x^*$, $<$

定理2.1

设$f(x), g_i(x), i = 1, 2, \cdots, m, h_j(x), j = m+1, \cdots, p$在$\mathbf{R}^n$上连续,则问题(2-1)的可行集$R$为闭集,它的全局最优集合也为闭集.

定理2.1

设$f(x), g_i(x), i = 1, 2, \cdots, m, h_j(x), j = m+1, \cdots, p$在$\mathbf{R}^n$上连续,则问题(2-1)的可行集$R$为闭集,它的全局最优集合也为闭集.

定理2.2(一阶必要条件)

设$f(x) \in C^1$定义在$\Omega \subset \mathbf{R}^n$上,$x^*$是$\Omega$ 的一个内点,若x^*为$f(x)$的一个极小点,则$\nabla f(x^*) = 0$.

满足条件$\nabla f(x^*) = 0$的点x^*称为$f(x)$的稳定点或驻点.

1.2.2 二维最优化问题的几何意义

研究如下的问题:

$$
\begin{aligned}
\min \quad & z = f(x_1, x_2) = (x_1 - 1.5)^2 + x_2^2 \\
\text{s.t.} \quad & g_1(x) = x_1^2 + x_2^2 - 1 \leq 0 \\
& g_2(x) = 2x_1 + x_2 - 1 \geq 0 \\
& g_3(x) = x_1 \geq 0 \\
& g_4(x) = x_2 \geq 0
\end{aligned}
\tag{2-3}
$$

研究如下的问题:

$$
\begin{aligned}
\min \quad & z = f(x_1, x_2) = (x_1 - 1.5)^2 + x_2^2 \\
\text{s.t.} \quad & g_1(x) = x_1^2 + x_2^2 - 1 \leq 0 \\
& g_2(x) = 2x_1 + x_2 - 1 \geq 0 \\
& g_3(x) = x_1 \geq 0 \\
& g_4(x) = x_2 \geq 0
\end{aligned} \tag{2-3}
$$

图解法，其求解步骤可概述如下:

(1) 画出问题的可行集R的图形.

(2) 作出目标函数的等值线族.

(3) 通过观察等值线族与R,确定使目标函数取得最小值的可行点x^*,即为所求的最优解.

1.2.2 最优化问题分类 I

(1) 线性规划 (Linear Programming,LP) 若$f(x), g_i(x), h_j(x)$都是x的线性函数.

(2) 二次规划 (Quadratic Programming,QP) 若$f(x)$是x的二次函数,$g_i(x), h_j(x)$都是的线性函数.

(3) 非线性规划 (Nonlinear Programming,NLP) 若$f(x), g_i(x), h_j(x)$中至少有一个是x的非线性函数. 一般分为两种:无约束非线性规划问题和约束非线性规划问题.

(4) 整数规划 (Integer Programming,IP) 若某些或全部设计变量取非负的整数值.
- 纯整数规划问题或全整数规划问题.
- 混合型整数规划问题.
- 0-1规划问题.

(5) 几何规划 (Geometric Programming) 若目标函数和约束函数都是设计变量的广义多项式, 形如$\phi(x) = \sum_{i=1}^{m} c_i \prod_{j=1}^{n} x_j^{r_{ij}}$, 其中$c_i, r_{ij}$可取任意实数值,$x_i > 0$.

(6) 多目标规划 (Multiobjective Programming) 若目标函数 $f(x) = (f_1(x), f_2(x), \cdots, f_p(x))^T, p \geq 2$.

(7) 其他 动态规划,不可微规划,参数规划,随机规划等.

1.3 凸集和凸函数

凸集和凸函数的理论一般称为凸分析,是最优化的理论基础,本节介绍凸集和凸函数最基本的知识.

1.3.1 凸集

定义3.1

设$\Omega \subset \mathbf{R}^n$,如果对于任意的点$x, y \in \Omega$,连接点$x, y$的线段上的一切点都在$\Omega$中,即对$0 \le \mu \le 1$的一切$\mu$,总有$\mu x + (1-\mu)y \in \Omega$,则称$\Omega$为一个凸集.

显然单点集和整个空间$\mathbf{R}^n$都是凸集,我们规定空集$\emptyset$为凸集.

凸集的例子

例3.1

满足线性规划问题的约束条件$Ax = b, x \geq 0$的一切点x所组成的集合R是一个凸集.

凸集的例子

例3.1

满足线性规划问题的约束条件$Ax = b, x \geq 0$的一切点x所组成的集合R是一个凸集.

例3.2

设$c \neq 0$是已给定的n维向量,b是已给常数,集合

$$H = \{x | x \in \mathbf{R}^n, c^T x = b\}$$

称为超平面,它也是一个凸集.

凸组合

定义3.2

设实数

$$\alpha_i \geq 0(i=1,2,\cdots,p), \sum_{i=1}^{p} \alpha_i = 1, x^i \in \mathbf{R}^n(i=1,2,\cdots,p),$$

称$x=\sum_{i=1}^{p} \alpha_i x^i$为点$x^1, x^2, \cdots, x^p$的一个凸组合.

凸组合

定义3.2

设实数

$$\alpha_i \geq 0(i = 1, 2, \cdots, p), \sum_{i=1}^{p} \alpha_i = 1, x^i \in \mathbf{R}^n(i = 1, 2, \cdots, p),$$

称$x = \sum_{i=1}^{p} \alpha_i x^i$为点$x^1, x^2, \cdots, x^p$的一个凸组合.

定理3.1

集合$\Omega \subset \mathbf{R}^n$为凸集的充要条件是:点$x^i \in \Omega, i = 1, 2, \cdots, p$的任意凸组合仍在$\Omega$中.

证明.

充分性显然. 必要性用数学归纳法证明. □

定理3.2

任意一组凸集的交集仍为凸集.

定理3.2

任意一组凸集的交集仍为凸集.

定义3.3

包含集合$\Omega \subset \mathbf{R}^n$的所有凸集的交集称为$\Omega$的凸包，记为$\mathrm{Co}(\Omega)$或$H(\Omega)$.

有定理3.2可知$\mathrm{Co}(\Omega)$为凸集,它实际上是$\mathbf{R}^n$中包含Ω的最小凸集.

定理3.2

任意一组凸集的交集仍为凸集.

定义3.3

包含集合$\Omega \subset \mathbf{R}^n$的所有凸集的交集称为$\Omega$的凸包，记为$\mathrm{Co}(\Omega)$或$H(\Omega)$.

有定理3.2可知$\mathrm{Co}(\Omega)$为凸集,它实际上是$\mathbf{R}^n$中包含Ω的最小凸集.

定义3.4

设$\Omega \subset \mathbf{R}^n$,如果对于任何$x \in \Omega$及所有的$\alpha \geq 0$,都有$\alpha x \in \Omega$,则称集合$\Omega$为一个锥. 一个同时为凸集的锥称为凸锥.

1.3.2 凸函数

定义3.5

设$f(x)$是定义在非空凸集$\Omega \subset \mathbf{R}^n$上的函数,若对任意$x, y \in \Omega$,不等式

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{3-1}$$

对于$0 \leq \lambda \leq 1$中的一切λ都成立,则称$f(x)$为Ω上的凸函数.

1.3.2 凸函数

定义3.5

设$f(x)$是定义在非空凸集$\Omega \subset \mathbf{R}^n$上的函数,若对任意$x, y \in \Omega$,不等式

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \tag{3-1}$$

对于$0 \le \lambda \le 1$中的一切λ都成立,则称$f(x)$为Ω上的凸函数.

- 严格凸函数, $x \ne y$, $\lambda \in (0,1)$, $<$
- 凹函数, $\ge$
- 严格凹函数, $x \ne y$, $\lambda \in (0,1)$, $>$

凸函数的性质

定理3.3

设$f_1(x), f_2(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则$f_1(x) + f_2(x)$在$\Omega$上也是凸的.

凸函数的性质

定理3.3

设$f_1(x), f_2(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则$f_1(x) + f_2(x)$在$\Omega$上也是凸的.

定理3.4

设$f(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则对任意常数$a \geq 0$,函数$af(x)$也是凸的.

凸函数的性质

定理3.3

设$f_1(x), f_2(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则$f_1(x) + f_2(x)$在$\Omega$上也是凸的.

定理3.4

设$f(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则对任意常数$a \geq 0$,函数$af(x)$也是凸的.

推论

设$f_1(x), f_2(x), \cdots, f_p(x)$是凸集$\Omega \subset \mathbf{R}^n$ 上的凸函数,
$\lambda_i \geq 0 (i = 1, 2, \cdots, p)$,则非负的线性组合$\sum_{i=1}^{p} \lambda_i f_i(x)$在$\Omega$上也是凸的.

定理3.5

设$f(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则对任一个实数$c$,水平集

$$\Omega_c = \{x | x \in \Omega, f(x) \leq c\}$$

是凸集.

水平集

定理3.5

设$f(x)$是凸集$\Omega \subset \mathbf{R}^n$上的凸函数,则对任一个实数$c$,水平集

$$\Omega_c = \{x | x \in \Omega, f(x) \leq c\}$$

是凸集.

推论

设$f_1(x), f_2(x), \cdots, f_p(x)$是凸集$\Omega \subset \mathbf{R}^n$ 上的凸函数,$c_i (i = 1, 2, \cdots, p)$为实常数,则$\Omega$中同时满足$f_1(x) \leq c_1, \cdots, f_p(x) \leq c_p$的点构成的集合$\Omega_c$为凸集.

凸函数的充要条件I

定理3.6

定义在凸集$\Omega \subset \mathbf{R}^n$上的可微函数$f(x)$为凸函数的充要条件是:对所有的$x, y \in \Omega$都有

$$f(y) \geq f(x) + (\nabla f(x))^T (y - x).$$

凸函数的充要条件I

定理3.6

定义在凸集$\Omega \subset \mathbf{R}^n$上的可微函数$f(x)$为凸函数的充要条件是:对所有的$x, y \in \Omega$都有

$$f(y) \geq f(x) + (\nabla f(x))^T (y - x).$$

证明.

主要利用一阶Taylor公式

$$\begin{aligned} f(\lambda y + (1-\lambda)x) &= f(x + \lambda(y - x)) \\ &= f(x) + \lambda(\nabla f(x))^T (y - x) + o(\lambda \|y - x\|) \end{aligned}$$

其中$\lambda \in (0, 1)$. □

凸函数的充要条件II

定理3.7

设$\Omega \subset \mathbf{R}^n$为含有内点的凸集,$f(x) \in C^2$,则$f(x)$在$\Omega$上为凸函数的充要条件是: $f(x)$的Hesse矩阵$F(x) = \nabla^2 f(x)$在整个Ω上是半正定的.

凸函数的充要条件II

定理3.7

设$\Omega \subset \mathbf{R}^n$为含有内点的凸集,$f(x) \in C^2$,则$f(x)$在$\Omega$上为凸函数的充要条件是：$f(x)$的Hesse矩阵$F(x) = \nabla^2 f(x)$在整个$\Omega$上是半正定的.

证明.

由定理3.6和二阶Taylor公式

$$\begin{aligned} f(y) = \quad & f(x) + (\nabla f(x))^T (y - x) \\ & + \frac{1}{2}(y - x)^T F(x + \alpha(y - x))(y - x) \end{aligned}$$

其中$\alpha \in (0, 1)$. □

严格凸函数的充分条件

定理3.8

设$\Omega \subset \mathbf{R}^n$为非空开凸集,$f(x) \in C^2$, 若$\nabla^2 f(x)$在$\Omega$上处处正定,
则$f(x)$在Ω上为严格凸函数.

$\nabla^2 f(x)$在Ω上处处正定是严格凸函数的充分非必要条件. 反例$f(x) = x^4$.

严格凸函数的充分条件

定理3.8

设$\Omega \subset \mathbf{R}^n$为非空开凸集,$f(x) \in C^2$, 若$\nabla^2 f(x)$在$\Omega$上处处正定,
则$f(x)$在Ω上为严格凸函数.

$\nabla^2 f(x)$在Ω上处处正定是严格凸函数的充分非必要条件. 反例$f(x) = x^4$.

例3.3

讨论$f(x) = 5x_1^2 - 6x_1x_2 + 5x_2^2$在凸集$\Omega \subset \mathbf{R}^n$ 上是否为严格凸函数.

严格凸函数的充分条件

定理3.8

设$\Omega \subset \mathbf{R}^n$为非空开凸集,$f(x) \in C^2$, 若$\nabla^2 f(x)$在$\Omega$上处处正定,
则$f(x)$在Ω上为严格凸函数.

$\nabla^2 f(x)$在Ω上处处正定是严格凸函数的充分非必要条件. 反例$f(x) = x^4$.

例3.3

讨论$f(x) = 5x_1^2 - 6x_1x_2 + 5x_2^2$在凸集$\Omega \subset \mathbf{R}^n$ 上是否为严格凸函数.

解:

$F(x) = \nabla^2 f(x) = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix}$中所有的顺序主子式为

$$10 > 0, \begin{vmatrix} 10 & -6 \\ -6 & 10 \end{vmatrix} = 64 > 0.$$

所以$F(x)$处处正定, 因此$f(x)$为严格凸函数.

例3.4

讨论$f(x)=\frac{1}{2}x^TAx+b^Tx+c$,其中$A$为对称阵,$b=(b_1,b_2,\cdots,b_n)^T$为给定向量,$c$为常数,则

(1) 当A为半正定时(记为$A\geq 0$), $f(x)$为凸函数.

(2) 当A为正定时(记为$A>0$), $f(x)$为严格凸函数.

例3.4

讨论$f(x)=\frac{1}{2}x^TAx+b^Tx+c$,其中$A$为对称阵,$b=(b_1,b_2,\cdots,b_n)^T$为给定向量,$c$为常数,则

(1) 当A为半正定时(记为$A\geq 0$), $f(x)$为凸函数.

(2) 当A为正定时(记为$A>0$), $f(x)$为严格凸函数.

解：

因为$F(x)=\nabla^2 f(x)=A$, 由定理3.7和定理3.8即得到结论。

1.3.3 凸规划

考虑如下的非线性规划问题:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0,\ \ i = 1, 2, \cdots, m, \\ & h_j(x) = 0,\ \ j = m+1, m+2, \cdots, p. \end{aligned} \tag{3-7}$$

因为

$$h_i(x) = 0 \Longleftrightarrow h_i(x) \leq 0, \text{且} - h_i(x) \leq 0.$$

因此,不失一般性,我们可以考虑仅包含不等式约束的优化问题:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0,\ \ i = 1, 2, \cdots, l. \end{aligned} \tag{3-8}$$

定义3.6

在问题(3-8)中,若$f(x), g_i(x)(i=1,2,\cdots,l)$均为可行集$R$上的凸函数,则称这样的问题为凸规划问题.

凸规划

定义3.6

在问题(3-8)中,若$f(x), g_i(x) (i = 1, 2, \cdots, l)$均为可行集$R$上的凸函数,则称这样的问题为凸规划问题.

定理3.9

凸规划问题(3-8)的可行集R是凸集.

定理3.10

对于凸规划问题(3-8), 目标函数$f(x)$的任一局部极小点都是$f(x)$在非空可行集R上的全局极小点.

定理3.10

对于凸规划问题(3-8),目标函数$f(x)$的任一局部极小点都是$f(x)$在非空可行集R上的全局极小点.

定理3.11

对于凸规划问题(3-8),若$f(x)$在非空可行集R上是严格凸函数,则问题(3-8)的全局极小点是唯一的.

拟凸函数

定义3.7

设$\Omega \subset \mathbf{R}^n$为非空凸集,$f : \Omega \to \mathbf{R}^1$,若$\forall x, y \in \Omega$,均有

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}, \forall \lambda \in (0,1)$$

成立,则称$f(x)$在Ω上为拟凸函数.

拟凸函数

定义3.7

设$\Omega \subset \mathbf{R}^n$为非空凸集,$f : \Omega \to \mathbf{R}^1$,若$\forall x, y \in \Omega$,均有

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}, \forall \lambda \in (0, 1)$$

成立,则称$f(x)$在Ω上为拟凸函数.

- 严格拟凸函数, $\forall x, y \in \Omega$, $f(x) \neq f(y)$, $<$
- 拟凹函数, $-f$拟凸
- 严格拟凹函数, $-f$严格拟凸

严格拟凸函数不一定是拟凸函数; 凸函数一定是拟凸函数.

伪凸函数

定义3.7

设$\Omega \subset \mathbf{R}^n$为非空开凸集,$f : \Omega \to \mathbf{R}^1$为可微函数,如果对于

$$\nabla f(x^1)^T(x^1 - x^2) \leq 0, \forall x^1, x^2 \in \Omega$$

必有$f(x^1) \leq f(x^2)$,则称$f(x)$在Ω上为伪凸函数.

伪凸函数

定义3.7

设$\Omega \subset \mathbf{R}^n$为非空开凸集,$f:\Omega \to \mathbf{R}^1$为可微函数,如果对于

$$\nabla f(x^1)^T(x^1 - x^2) \le 0, \forall x^1, x^2 \in \Omega$$

必有$f(x^1) \le f(x^2)$,则称$f(x)$在Ω上为伪凸函数.

- 严格伪凸函数, $\nabla f(x^1)^T(x^1 - x^2) \le 0, \forall x^1, x^2 \in \Omega,\ x^1 \ne x^2,\ <$
- 伪凹函数, $-f$伪凸
- 严格伪凹函数, $-f$严格伪凸

可微的凸函数是伪凸函数; 严格伪凸函数是伪凸函数; 可微的伪凸函数是严格拟凸函数, 也是拟凸函数.

广义凸规划

若问题(3-8)的可行集R是凸集，$f(x)$是R上的(严格)拟凸函数或(严格)伪凸函数，则称(3-8)为广义凸规划问题.

若问题(3-8)的可行集R是凸集，$f(x)$是R上的(严格)拟凸函数或(严格)伪凸函数，则称(3-8)为广义凸规划问题.

定理3.15

设(3-8)的可行集R是凸集，$f(x)$是R上的严格拟凸函数，则广义凸规划(3-8)的任一局部最优解x^*是全局最优解.

广义凸规划

若问题(3-8)的可行集R是凸集，$f(x)$是R上的(严格)拟凸函数或(严格)伪凸函数，则称(3-8)为广义凸规划问题.

定理3.15

设(3-8)的可行集R是凸集，$f(x)$是R上的严格拟凸函数，则广义凸规划(3-8)的任一局部最优解x^*是全局最优解.

定理3.16

设(3-8)的可行集R是开凸集，$f(x)$是R上的伪凸函数，$x^* \in R$，若$\nabla f(x^*) = 0$，则x^*是(3-8)的全局最优解.