实用泛函分析　吕和祥 王天明 大连理工大学出版社

具有不等式约束的极值问题

Definition 1.

设X是线性赋范空间，M和N是X的两个子集，如果存在有界线性泛函$x^* \in X^*$及常数c使得对所有的$x \in M, y \in N$,均有

$$\langle x, x^* \rangle \geqslant c \geqslant \langle y, x^* \rangle \tag{1}$$

则称泛函x^*分离集合M与N。

有界线性泛函x^*的几何意义是：

如果集合

$$H = \{x | x \in X, \langle x, x^* \rangle = c\}$$

是X的闭超平面，那么式(1)就表示超平面H分离集合M与N。

Theorem 2.

设M是线性赋范空间X中的凸集，点$x_0 \in X$不是M的内点，则存在一个闭超平面H，其分离x_0与M。

这个定理的意思是存在线性泛函$x^* \in X^*$,满足

$$\langle x, x^* \rangle > \langle x_0, x^* \rangle \tag{2}$$

定理说存在一点与凸集可分离，如果这样的点形成一个凸集，那么这两个凸集从直观上看也应该是可分离的。

下面讨论不等式约束泛函的极值问题。

设

M是线性赋范空间X中的凸集合，

$f(x), \varphi_1(x), \varphi_2(x), \cdots, \varphi_n(x)$是定义在M上的凸函数。

讨论极值问题：

$$\begin{array}{llll} \min: & f(x), & x \in M & \\ s.t. & \varphi_i(x) \leqslant 0, & i = 1, 2, \cdots, n, & x \in M \end{array} \tag{3}$$

这是不等式约束泛函的极值问题。

Theorem 3.

(Kuhn-Tucker鞍点定理)设$\tilde{x}$是不等式约束最小值问题(3)的解，那么存在$\tilde{\lambda} \geqslant 0$,使得

$$L(\tilde{x}, \lambda) \leqslant L(\tilde{x}, \tilde{\lambda}) \leqslant L(x, \tilde{\lambda}) \tag{4}$$

和

$$\langle \tilde{\lambda}, \varphi(\tilde{x}) \rangle \geqslant 0 \tag{5}$$

对于所有$\lambda \geqslant 0$和$x \in M$均成立，其中

$$\begin{aligned} L(x,\lambda) &= f(x) + \langle \varphi(x), \lambda \rangle \\ \varphi(x) &= (\varphi_1(x), \varphi_2(x), \cdots, \varphi_m(x)) \\ \lambda &= (\lambda_1, \lambda_2, \cdots, \lambda_m) \in R^m \end{aligned}$$

Proof.

首先指出，方程组

$$\begin{cases} f(x) - f(\tilde{x}) < 0 \\ \varphi(x) \leqslant 0 \end{cases}$$

在集合M中无解，因为约束最小值点$\tilde{x}$在M内。

对于每一个$x \in M$,定义集合V_x:

$$V_x = \left\{ \begin{pmatrix} v_0 \\ \boldsymbol{v} \end{pmatrix} | v_0 \in R, \boldsymbol{v} \in R^m, 0 \leqslant f(x) - f(\tilde{x}) < v_0, \varphi(x) < \boldsymbol{v} \right\}$$

定义

$$V = \bigcup_{x \in M} V_x$$

由假设，V不包括原点，集合V是凸的，$\begin{pmatrix} v_0^1 \\ \boldsymbol{v}^1 \end{pmatrix} \in V_{x_1}$，$\begin{pmatrix} v_0^2 \\ \boldsymbol{v}^2 \end{pmatrix} \in V_{x_2}$，意味着$\begin{pmatrix} v_0^\theta \\ \boldsymbol{v}^\theta \end{pmatrix} \in V_{x_\theta}$，$0 < \theta < 1$，其中

$$\begin{aligned} v_0^\theta &= \theta v_0^1 + (1-\theta) v_0^2 \\ \boldsymbol{v}^\theta &= \theta \boldsymbol{v}^1 + (1-\theta) \boldsymbol{v}^2 \\ x_\theta &= \theta x_1 + (1-\theta) x_2 \end{aligned}$$

对于不包含零点的凸集V，由定理2知，存在$\tilde{\lambda}_0, \tilde{\lambda}$满足

$$\tilde{\lambda}_0 v^0 + \langle \tilde{\lambda}, \boldsymbol{v} \rangle > \tilde{\lambda}_0 \cdot 0 + \langle \tilde{\lambda}, 0 \rangle = 0 \tag{6}$$

对于所有的$\begin{pmatrix} v^0 \\ \boldsymbol{v} \end{pmatrix} \in V, (\tilde{\lambda}_0, \tilde{\lambda}) \neq 0,$

在式(6)中，因为v^0和$\boldsymbol{v}$的每一个分量可以取任意大于零的数，为了保持这个不等式恒成立，那么必定有 $\tilde{\lambda}_0 \geqslant 0, \tilde{\lambda} \geqslant 0$。(由于$\boldsymbol{v}$的分量可能会小于0,因此，$\tilde{\lambda}_0 > 0$)，对于任意$\varepsilon > 0$，取

$$v^0 = f(x) - f(\tilde{x}) + \varepsilon, \boldsymbol{v} = \varphi(x) + \boldsymbol{\mu}$$

其中$\mu = \varepsilon(1,1,\cdots,1)^T$,则式(6)成为

$$\tilde{\lambda}_o[f(x) - f(\tilde{x})] + \tilde{\lambda}_0 \varepsilon + \langle \tilde{\lambda}, \varphi(x) \rangle + \varepsilon \sum_{i=1}^m \tilde{\lambda}_i > 0$$

因为ε是任意的，有

$$\tilde{\lambda}_0[f(x) - f(\tilde{x})] + \langle \tilde{\lambda}, \varphi(x) \rangle \geqslant 0 \tag{7}$$

由$\tilde{\lambda} \geqslant 0, \varphi(x) \leqslant 0$,得$\langle \tilde{\lambda}, \varphi(x) \rangle \leqslant 0$，对所有M中的$x$均成立。得

$$\langle \tilde{\lambda}, \varphi(\tilde{x}) \rangle \leqslant 0$$

在式(7)中令$x=\tilde{x}$,则得

$$\langle\tilde{\lambda},\varphi(\tilde{x})\rangle\geqslant 0$$

鉴于以上两个不等式，得

$$\langle\tilde{\lambda},\varphi(\tilde{x})\rangle=0 \tag{8}$$

由式(7)(8)得

$$\tilde{\lambda}_0 f(\tilde{x})+\langle\tilde{\lambda},\varphi(\tilde{x})\rangle\leqslant\tilde{\lambda}_0 f(x)+\langle\tilde{\lambda},\varphi(x)\rangle \tag{9}$$

这就是要证明的式(4)右端的不等式。

因为$\varphi(\tilde{x})\leqslant 0$，那么对于任意$\lambda\geqslant 0$，有$\langle\lambda,\varphi(\tilde{x})\rangle\leqslant 0$，所以

$$\tilde{\lambda}_0 f(\tilde{x})+\langle\lambda,\varphi(\tilde{x})\rangle\leqslant\tilde{\lambda}_0 f(\tilde{x})=\tilde{\lambda}_0 f(\tilde{x})+\langle\tilde{\lambda},\varphi(\tilde{x})\rangle \tag{10}$$

这是式(4)左端的不等式。

现证明$\tilde{\lambda}_0>0$。否则，如果$\tilde{\lambda}_0=0$,由式(8)(9)，有

$$0\leqslant\langle\tilde{\lambda},\varphi(x)\rangle$$

由已知条件$\varphi(x)<0$或$\langle\tilde{\lambda},\varphi(x)\rangle<0$,与上式矛盾，所以$\tilde{\lambda}_0>0$。

对于充分性，如果$\langle\tilde{\lambda},\tilde{x}\rangle,\tilde{\lambda}\geqslant 0,\tilde{x}\in M$满足鞍点定理式(4)(5),那么$\tilde{x}$是约束最小值问题的解。

实际上由不等式(10)得

$$\langle\lambda-\tilde{\lambda},\varphi(\tilde{x})\rangle\leqslant 0 \qquad \forall\lambda\geqslant 0 \tag{11}$$

由此推得$\varphi(\tilde{x})\leqslant 0$,再由$\tilde{\lambda}>0$得到$\langle\tilde{\lambda},\varphi(\tilde{x})\rangle\leqslant 0$，在式(11)中令$\lambda=0$，则得$\langle\tilde{\lambda},\varphi(\tilde{x})\rangle\geqslant 0$,由最后两式得

$$\langle\tilde{\lambda},\varphi(\tilde{x})\rangle=0$$

于是，由不等式(9)得

$$f(\tilde{x})\leqslant f(x)+\frac{1}{\tilde{\lambda}_0}\langle\tilde{\lambda},\varphi(x)\rangle$$

因为

$\langle\tilde{\lambda},\varphi(x)\rangle<0,\tilde{\lambda}_0>0,$

所以

$f(\tilde{x})\leqslant f(x),\forall x\in M,\varphi(x)<0,$

即$\tilde{x}$是约束问题的最小点。由前述的充分性论证发现，完全不需要凸性的要求。 □

Theorem 4.

(Kuhn-Tucker鞍点定理的微分形式)

令f和$\varphi_i,1\leqslant i\leqslant m$在一个包含元素$\tilde{x}$的开集上伽脱可微，如果$\tilde{x}$使得$f(x)$在约束条件

$$\varphi_i(x)\leqslant 0,1\leqslant i\leqslant m$$

下取最小值，那么存在非负拉格朗日乘子$\lambda_1,\lambda_2,\cdots,\lambda_m$满足下面方程

$$\begin{aligned}\nabla f(\tilde{x})+\sum_{i=1}^{m}\lambda_i\nabla\varphi_i(\tilde{x}) &= 0\\ \sum_{i=1}^{m}\lambda_i\varphi_i(\tilde{x}) &= 0 \qquad \lambda_i\geqslant 0,1\leqslant i\leqslant m\end{aligned}$$

与等式约束不同的是，不等式约束的拉格朗日乘子是大于或等于零的，即$\lambda\in R_+^m$。