上海大学

SHANGHAI UNIVERSITY

非线性规划

主要内容

- 基本概念
- 凸函数
- 线性搜索算法
- 无约束极值问题
- 约束极值问题

上海大学 SHANGHAI UNIVERSITY

约束优化问题的最优性条件

$$\begin{cases} \min \quad f(x) \\ s.t. \quad g_i(x) \geq 0, \quad i = 1, 2, \cdots, m ---- \text{不等式约束} \\ \qquad\quad h_j(x) = 0, \quad j = 1, 2, \cdots, l. ---- \text{等式约束} \\ \qquad\quad x \in E^n \end{cases}$$

$$S = \left\{ x \middle| \begin{array}{ll} g_i(x) \geq 0, & i = 1, 2, \cdots, m \\ h_j(x) = 0, & j = 1, 2, \cdots, l \end{array} \right\} - \text{可行集或可行域}$$

约束优化问题的最优性条件

定义： 对$\min\limits_{x\in E^n} f(x)$，设$\bar{x}\in E^n$是任给一点，

$d\neq 0$，若存在$\delta>0$，使得对任意的$\lambda\in(0,\delta)$，

有$f(\bar{x}+\lambda d)<f(\bar{x})$，则称$d$为$f(x)$在点$\bar{x}$处的

下降方向(descent direction)。

$F_0=\left\{d \,\middle|\, \nabla f(\bar{x})^T d<0\right\}$，称为点$\bar{x}$处的下降方向集。

约束优化问题的最优性条件

定　义：　设集合$S \subset E^n, \bar{x} \in S, d$为非零向量，若存在数$\delta > 0$,使得对任意$\lambda \in (0, \delta)$，都有

$$\bar{x} + \lambda d \in S$$

则称d为集合S在$\bar{x}$的可行方向(feasible direction)。

$$D = \left\{ d \,\middle|\, d \neq 0, \bar{x} \in clS, \exists \delta > 0, \forall \lambda \in (0, \delta), \text{有} \bar{x} + \lambda d \in S \right\}$$

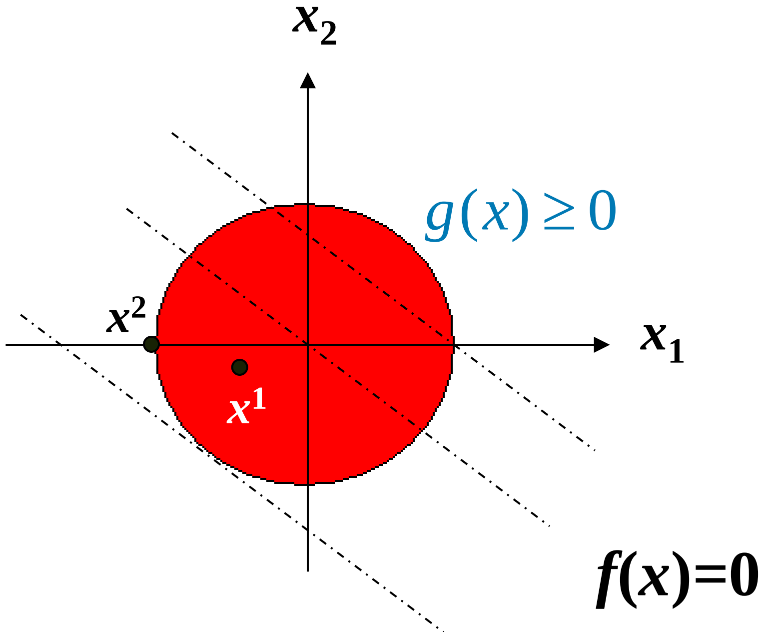
是$\bar{x}$处的可行方向锥。

约束优化问题的最优性条件

例：考虑如下约束优化问题：

$$\min\ f(x) = x_1 + x_2$$

$$s.t.\quad g(x) = 1 - x_1^2 - x_2^2 \geq 0$$



对于任意内点x^1，可行方向锥$D = R^2$.

对于边界点$x^2 = (-1, 0)^T$,可行方向锥

$$D = \{d \in R^2 \mid d_1 > 0\}.$$

约束优化问题的最优性条件

定理1（几何最优性条件）：考虑问题

$$\begin{aligned} &min \quad f(x) \\ &s.t. \quad\ \ x \in S \end{aligned}$$

设S是E^n的非空集合，$\bar{x} \in S$，$f(x)$在$\bar{x}$处可微，若$\bar{x}$是局部最优解，则$F_0 \cap D = \varnothing$。

证明：设存在$d \in F_0 \cap D$,则$d \in F_0, d \in D$.

$\because d \in F_0, \therefore \exists \delta_1 > 0$,对$\forall \lambda \in (0, \delta_1)$,有$f(\bar{x} + \lambda d) < f(\bar{x})$;

$\because d \in D, \therefore \exists \delta_2 > 0$,对$\forall \lambda \in (0, \delta_2)$,有$\bar{x} + \lambda d \in S$.

取$\delta = \min(\delta_1, \delta_2)$,则当$\lambda \in (0, \delta)$,有

$\bar{x} + \lambda d \in S$且$f(\bar{x} + \lambda d) < f(\bar{x})$，与$\bar{x}$为局部最优解矛盾。

不等式约束优化问题

一阶最优性条件

$$(1)\quad \begin{cases} \min & f(x) \\ s.t. & g_i(x) \geq 0, i=1,\cdots,m. \end{cases}$$

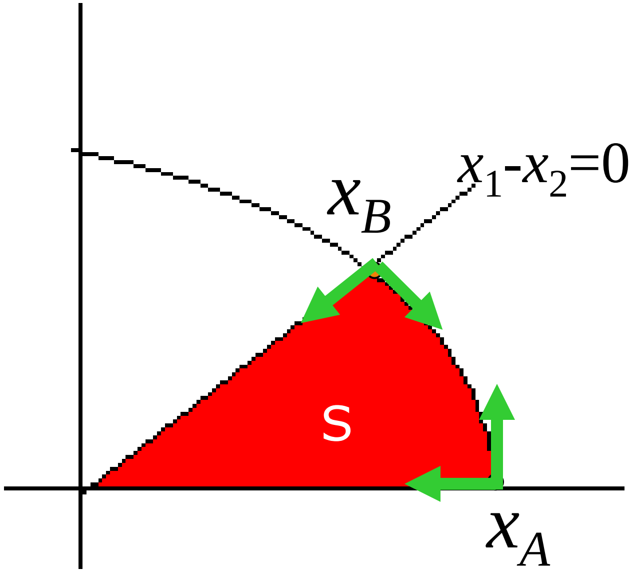
可行域 $S=\{x \mid g_i(x)\geq 0, i=1,\cdots,m\}$

定义　若问题(1)的一个可行点$\bar{x}$(即$\bar{x}\in S$)使某个不等式约束$g_i(x)\geq 0$变成等式，即$g_i(\bar{x})=0$，则该不等式约束称为关于可行点$\bar{x}$的起作用约束（或等式约束）；否则，若$\bar{x}$使得某个$g_i(\bar{x})>0$，则该不等式约束称为关于可行点$\bar{x}$ 的不起作用约束（或松约束）。

记　$I=\{i \mid g_i(\bar{x})=0, \bar{x}\in S\}$

例：约束

$$\begin{cases} g_1(x) = -x_1^2 - x_2^2 + 1 \geq 0 \\ g_2(x) = x_1 - x_2 \geq 0 \\ g_3(x) = x_1 \geq 0 \\ g_4(x) = \quad x_2 \geq 0 \end{cases}$$



$$x_A = (1,0)^T,\ x_B = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)^T$$

$$G_0 = \left\{ d \,\middle|\, \nabla g_i(\bar{x})^T d > 0, i \in I \right\}$$

称G_0为S在点$\bar{x}$处的局部约束方向锥（或内方向锥）。

不等式约束优化问题

$$G_0=\left\{d \,\middle|\, \nabla g_i(\bar{x})^T d>0, i\in I\right\}$$

称G_0为S在点$\bar{x}$处的局部约束方向锥（或内方向锥）。

证：对于起作用约束：$g_i(\bar{x})=0$;存在$\delta>0$,

使得任意$\lambda\in$（0，δ），有$g_i(\bar{x}+\lambda d)\geq g_i(\bar{x})=0$,

从而$\lim\limits_{\lambda\to 0}\dfrac{g_i(\bar{x}+\lambda d)-g_i(\bar{x})}{\lambda}=\nabla g_i(\bar{x})^T d\geq 0.$

由泰勒公式，$g_i(\bar{x}+\lambda d)=g_i(\bar{x})+\lambda\nabla g_i(\bar{x})^T d+o(\lambda)$,

当λ足够小时，只要$\nabla g_i(\bar{x})^T d>0$,就有$g_i(\bar{x}+\lambda d)\geq 0.$

不等式约束优化问题

定理2：设$\bar{x} \in S$，$f(x)$和$g_i(x)(i \in I)$在$\bar{x}$处可微，$g_i(x)(i \notin I)$在$\bar{x}$处连续，如果$\bar{x}$是问题(1)的局部最优解，则$F_0 \cap G_0 = \varnothing$。

证明：由定理**1**，在$\bar{x}$处，有$F_{\mathbf{0}} \cap D = \varnothing$.

设$d \in G_{\mathbf{0}}$，则$\nabla g_i(\bar{x})^T d > \mathbf{0}$，$i \in I$；令$\tilde{g}_i(x) = -g_i(x)$，$i \in I$.

则$\nabla \tilde{g}_i(\bar{x})^T d < \mathbf{0}$，$i \in I$;

由引理，存在$\delta_{\mathbf{1}} > \mathbf{0}$,当$\lambda \in (\mathbf{0}, \delta_{\mathbf{1}})$时，有 $\tilde{g}_i(\bar{x} + \lambda d) < \tilde{g}_i(\bar{x})$，$i \in I$

即 $g_i(\bar{x} + \lambda d) > g_i(\bar{x}) = \mathbf{0}$，$i \in I$.

当$i \notin I$时，$g_i(\bar{x}) > \mathbf{0}$,$\because g_i(\bar{x})(i \notin I)$在$\bar{x}$连续,

$\therefore$存在$\delta_{\mathbf{2}} > \mathbf{0}$,当$\lambda \in (\mathbf{0}, \delta_{\mathbf{2}})$时，有$g_i(\bar{x} + \lambda d) > \mathbf{0}, i \notin I$.

令$\delta = \min(\delta_{\mathbf{1}}, \delta_{\mathbf{2}})$，则当$\lambda \in (\mathbf{0}, \delta)$时，有$g_i(\bar{x} + \lambda d) > \mathbf{0}, \forall i$,

$\Rightarrow \bar{x} + \lambda d \in S \quad \Rightarrow d \in D \Rightarrow \quad G_{\mathbf{0}} \subseteq D \quad \Rightarrow F_{\mathbf{0}} \cap G_{\mathbf{0}} = \varnothing$

不等式约束优化问题

Farkas引理： 设A为$m\times n$矩阵，c为n维列向量，则 $Ax\le 0$，$c^T x>0$有解的充要条件是$A^T y=c$，$y\ge 0$无解。

证明:“$\Rightarrow$”（反证法）假设存在$y\ge 0$，使得$A^T y=c$

得 $y^T A=c^T$

设$\bar{x}$为$Ax\le 0, c^T x>0$的一个解,

则有$A\bar{x}\le 0, c^T\bar{x}>0$

$$\Rightarrow \quad y^T A\bar{x}=c^T\bar{x}>0 \qquad (1)$$

$\because \quad y\ge 0,$但$A\bar{x}\le 0$

$\therefore \quad y^T A\bar{x}\le 0$ 与（1）矛盾。

不等式约束优化问题

“$\Leftarrow$” 设$A^T y = c,\ y \geq 0$无解，令

$$S = \{z \mid z = A^T y,\ y \geq 0\}, 则 c \notin S$$

可以证明S为闭凸集，由点与凸集强可分离定理，

$\exists x \neq 0,\ \varepsilon > 0,$ 使得对$\forall z \in S,$ 有$x^T c \geq \varepsilon + x^T z$

$\because \varepsilon > 0, \quad \therefore \quad x^T c > x^T z$

$\Rightarrow c^T x > z^T x = y^T Ax$

即对任意的$y \geq 0$，有$c^T x > y^T Ax$ （1）

令$y = 0$，得$c^T x > 0$

$\because c^T x$为一定数，y的分量可取任意大，

$\therefore$由（1），必有$Ax \leq 0.$

即非零向量x是$Ax \leq 0$，$c^T x > 0$的解。

不等式约束优化问题

Gordan定理： 设A为$m\times n$矩阵，那么$Ax<0$有解的充要条件是不存在非零向量$y\geq 0$，使得$A^T y=0$。

证明：

"$\Rightarrow$" 设存在$\bar{x}$，使得$A\bar{x}<0$

若存在非零向量$y\geq 0$, 使得$A^T y=0$

则有$y^T A=0$, $\Rightarrow y^T A\bar{x}=0$

$\because A\bar{x}<0$

$\therefore y$的各分量不可能为非负数, 与$y\geq 0$矛盾.

不等式约束优化问题

"$\Leftarrow$"（证等价命题）即若$Ax<0$无解，则存在非零向量$y\geq 0$，使得$A^T y=0$.

设$Ax<0$无解，令$S_1=\{z\,|\,z=Ax,\ x\in E^n\}$　$S_2=\{z\,|\,z<0\}$

$\because Ax<0$无解，$\therefore S_1\bigcap S_2=\varnothing$,由凸集分离定理知，

存在非零向量y，使得对$\forall x\in E^n,\forall z\in S_2$，有

$$y^T Ax\geq y^T z \qquad (1)$$

特别地，当$x=0$时，有$y^T z\leq 0$。$\because z<0$，它的分量可取任意负数，$\therefore y\geq 0$

在（1）中令$z\to 0$，则对$\forall x\in E^n$,有

$$y^T Ax\geq 0 \qquad (2)$$

令$x=-A^T y$，代入（2），得$-y^T AA^T y\geq 0$，即$-\left\|A^T y\right\|\geq 0$，$\therefore\ A^T y=0$.

因此，存在非零向量$y\geq 0$,使得$A^T y=0$.

不等式约束优化问题

定理3(*Fritz John*条件) 设$\bar{x}\in S, I=\{i \mid g_i(\bar{x})=0\}$,

$f(x), g_i(x)(i\in I)$在$\bar{x}$处可微，$g_i(x)(i\notin I)$在$\bar{x}$处连续，若$\bar{x}$是问题(1)的局部最优解，则存在不全为零的数$w_0, w_i(i\in I)$,使得

$$\begin{cases} w_0\nabla f(\bar{x})-\sum_{i\in I} w_i\nabla g_i(\bar{x})=0 \\ w_0, \ w_i\geq 0, \quad i\in I. \end{cases}$$

$\bar{x}$ 称为**Fritz John 点(即满足Fritz John条件的点).**

不等式约束优化问题

证明：由定理2，在点$\bar{x}$, $F_0 \cap G_0 = \varnothing$,即不等式 $\begin{cases} \nabla f(\bar{x})^T d < 0 \\ -\nabla g_i(\bar{x})^T d < 0, \quad i \in I \end{cases}$ 无解。

设$I = \{i_1, i_2, \cdots, i_s\}$, 令$A = \begin{pmatrix} \nabla f(\bar{x})^T \\ -\nabla g_{i_1}(\bar{x})^T \\ \vdots \\ -\nabla g_{i_s}(\bar{x})^T \end{pmatrix}$

有$Ad < 0$无解,由$Gordan$定理,存在$w = \left(w_0, w_{i_1}, \cdots, w_{i_s}\right)^T \geq 0, w \neq 0$,使得

$$A^T w = 0, \text{即} \left(\nabla f(\bar{x}), -\nabla g_{i_1}(\bar{x}), \cdots, -\nabla g_{i_s}(\bar{x})\right) \begin{pmatrix} w_0 \\ w_{i_1} \\ \vdots \\ w_{i_s} \end{pmatrix} = 0$$

$$\Rightarrow w_0 \nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0.$$

不等式约束优化问题

定理3'(*Fritz John*条件) 设$\bar{x} \in S, I = \{i \mid g_i(\bar{x}) = 0\}$, $f(x), g_i(x)$在$\bar{x}$处可微，若$\bar{x}$是问题(1)的局部最优解，则存在不全为零的数$w_0, w_1, \cdots, w_m$, 使得

$$\begin{cases} w_0 \nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) = 0 \\ w_i g_i(\bar{x}) = 0, \quad i = 1, 2, \cdots, m \\ w_0, \ w_i \geq 0, \quad i = 1, 2, \cdots, m. \end{cases}$$

互补松弛条件

不等式约束优化问题

问题

$$\begin{cases} \min \quad f(x) \\ s.t. \quad g_i(x) \leq \mathbf{0}, i = \mathbf{1}, \mathbf{2}, \cdots, m \end{cases}$$

的*Fritz John*条件为

$$\begin{cases} w_{\mathbf{0}} \nabla f(\bar{x}) + \sum_{i \in I} w_i \nabla g_i(\bar{x}) = \mathbf{0} \\ w_{\mathbf{0}},\ w_i \geq \mathbf{0}, \quad i \in I. \end{cases}$$

不等式约束优化问题

例：设非线性规划问题

$$\begin{cases} \min \ f(x) = (x_1-3)^2 + (x_2-2)^2 \\ s.t. \quad g_1(x) = -x_1^2 - x_2^2 + 5 \geq 0 \\ \qquad g_2(x) = -x_1 - 2x_2 + 4 \geq 0 \\ \qquad g_3(x) = x_1 \geq 0 \\ \qquad g_4(x) = x_2 \geq 0 \end{cases}$$

$$\begin{cases} w_0 \nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) = 0 \\ w_i g_i(\bar{x}) = 0, \quad i = 1, 2, \cdots, m \\ w_0, \ w_i \geq 0, \quad i = 1, 2, \cdots, m. \end{cases}$$

判别点$x^{(1)} = (2,1)^T$和$x^{(2)} = (0,0)^T$是否是$Fritz\ John$点？

解：$\nabla f(x) = \left(2(x_1-3), 2(x_2-2)\right)^T$

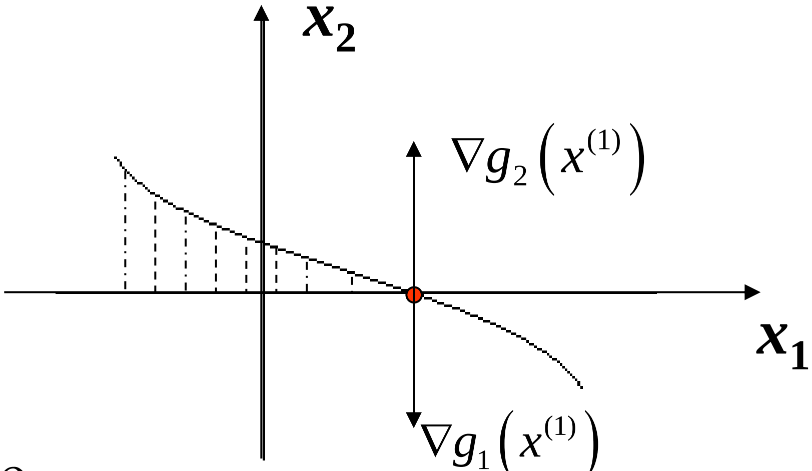
$$\nabla g_1(x) = (-2x_1, -2x_2)^T, \nabla g_2(x) = (-1, -2)^T$$

$$\nabla g_3(x) = (1, 0)^T, \quad \nabla g_4(x) = (0, 1)^T$$

不等式约束优化问题

例2. 设非线性规划问题：

$$\begin{cases} \min -x_1 \\ s.t. \quad g_1(x) = -x_2 + (1-x_1)^3 \geq 0 \\ \qquad\ \ g_2(x) = x_2 \geq 0 \end{cases}$$



判别点$x^{(1)} = (1,0)^T$是否是*Fritz John*点？

解：$\because \nabla f(x) = (-1,0)^T$，$\nabla g_1(x) = (-3(1-x_1)^2, -1)^T, \nabla g_2(x) = (0,1)^T$

在点$x^{(1)} = (1,0)^T$处, $I = \{1,2\}$，$\nabla f(x^{(1)}) = (-1,0)^T, \nabla g_1(x^{(1)}) = (0,-1)^T$,

$\nabla g_2(x) = (0,1)^T$，设有$w_0\begin{pmatrix}-1\\0\end{pmatrix} - w_1\begin{pmatrix}0\\-1\end{pmatrix} - w_2\begin{pmatrix}0\\1\end{pmatrix} = 0$

$\Rightarrow w_0 = 0$, 取$w_1 = w_2 > 0 \Rightarrow x^{(1)} = (1,0)^T$是*Fritz John*点。

不等式约束优化问题

例：$\min\ f(x)=(x_1-1)^2+(x_2-1)^2$

$$s.t.\quad g_1(x)=(1-x_1-x_2)^3\geq 0$$

$$g_2(x)=x_1\geq 0$$

$$g_3(x)=x_2\geq 0$$

最优解为

$$x^*=\left(\frac{1}{2},\frac{1}{2}\right)^T.$$

$\because$ 直线$x_1+x_2=1$上所有可行点$\bar{x}$使$\nabla g_1(\bar{x})=0$,

$\therefore$ 取$w_0=0, w_1=a>0$，$w_2=w_3=0$，总有

$$w_0\nabla f(\bar{x})-w_1\nabla g_1(\bar{x})-w_2\nabla g_2(\bar{x})-w_3\nabla g_3(\bar{x})=0$$

说明在直线$x_1+x_2=1$上每个可行点$\bar{x}$都是*Fritz John*点，但除x^*外，都不是最优解。

*Fritz John*条件只是局部最优点的必要条件。

不等式约束优化问题

定理2. 考虑问题$\begin{cases} \min & f(x) \\ s.t. & g_i(x) \geq 0, i = 1, \cdots, m \end{cases}$

设$\bar{x} \in S$, $f, g_i (i \in I)$在$\bar{x}$处可微，$g_i (i \notin I)$在$\bar{x}$连续，$\{\nabla g_i(\bar{x}) \mid i \in I\}$线性无关，若$\bar{x}$是局部最优解，则存在非负数$w_i, i \in I$，使得

$$\nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0.$$

不等式约束优化问题

证明：由定理1，存在不全为零的非负数

$w_0, w_i', i \in I$，使得

$$w_0 \nabla f(\bar{x}) - \sum_{i \in I} w_i' \nabla g_i(\bar{x}) = 0.$$

显然$w_0 \neq 0$，否则$\nabla g_i(\bar{x})(i \in I)$线性相关，矛盾。

于是，令$w_i = \dfrac{w_i'}{w_0} \geq 0 \ (i \in I)$，得

$$\nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0.$$

不等式约束优化问题

- **Karush-Kuhn-Tucker(KKT或KT)条件**

定理2'. 考虑问题$\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \ge 0, i = 1, \cdots, m \end{cases}$

设$\bar{x} \in S$, f, g_i在$\bar{x}$可微，$\{\nabla g_i(\bar{x}) \mid i \in I\}$线性无关，

若$\bar{x}$是局部最优解，则存在数$w_i, i = 1, 2, \cdots, m$, 使得

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) = 0$$

$$w_i g_i(\bar{x}) = 0 \quad i = 1, 2, \cdots, m$$

$$w_i \ge 0 \quad i = 1, 2, \cdots, m.$$

互补松弛条件

不等式约束优化问题

例：给定非线性规划问题
$$\begin{cases} \min \ f(x) = (x_1 - 1)^2 + x_2 \\ s.t. \quad g_1(x) = -x_1 - x_2 + 2 \geq 0 \\ \qquad g_2(x) = x_2 \geq 0 \end{cases}$$

求满足KKT条件的点。

解：$\nabla f(x) = (2(x_1 - 1), 1)^T, \nabla g_1(x) = (-1, -1)^T, \nabla g_2(x) = (0, 1)^T$

设x为满足KKT条件的点，则有

$$\begin{cases} \nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) = 0 \\ w_i g_i(\bar{x}) = 0, \quad i = 1, 2, \cdots, m \\ w_i \geq 0, \quad i = 1, 2, \cdots, m. \end{cases}$$

$$\begin{cases} \begin{pmatrix} 2(x_1 - 1) \\ 1 \end{pmatrix} - w_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} - w_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ w_1(-x_1 - x_2 + 2) = 0 \\ w_2 x_2 = 0, \quad -x_1 - x_2 + 2 \geq 0 \\ w_1, w_2 \geq 0, \quad x_2 \geq 0 \end{cases}$$

$\Rightarrow x_1 = 1, x_2 = 0, w_1 = 0, w_2 = 1,$ 即KKT点为$(1, 0)^T$.

不等式约束优化问题

例:求下列非线性规划问题的KKT点.

$$\min \ f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$$

$$s.t. \quad g_1(x) = -x_1^2 - x_2^2 + 5 \geq 0$$

$$g_2(x) = -3x_1 - x_2 + 6 \geq 0$$

$$\nabla f(x) = \begin{pmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{pmatrix}$$

$$\nabla g_1(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix} \quad \nabla g_2(x) = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$$

SHANGHAI UNIVERSITY 上海大学

不等式约束优化问题

设x为满足KKT条件的点，则有

$$
\begin{cases}
4x_1 + 2x_2 - 10 + 2w_1x_1 + 3w_2 = 0 \\
2x_1 + 2x_2 - 10 + 2w_1x_2 + w_2 = 0 \\
w_1(-x_1^2 - x_2^2 + 5) = 0 \\
w_2(-3x_1 - x_2 + 6) = 0 \\
-x_1^2 - x_2^2 + 5 \geq 0 \\
-3x_1 - x_2 + 6 \geq 0 \\
w_1, w_2 \geq 0
\end{cases}
$$

凸规划

$$
\begin{cases}
\min & f(x) \\
s.t. & g_i(x) \ge 0, i = 1,2,\cdots,m \\
 & h_j(x) = 0,\ j = 1,2,\cdots,l.
\end{cases}
$$

其中是$f(x)$凸函数，$g_i(x)$是凹函数，$h_j(x)$是线性函数。

凸规划

定理3.(一阶充分条件)

设问题$\begin{cases}\min\ f(x)\\ s.t.\quad g_i(x)\geq 0\ \ i=1,2,\cdots,m\end{cases}$

中，f是凸函数，$g_i(i=1,2,\cdots,m)$是凹函数，S为可行域，$\bar{x}\in S$，$I=\{i\mid g_i(\bar{x})=0\}$。$f$和$g_i(i\in I)$在点$\bar{x}$可微，$g_i(i\notin I)$在点$\bar{x}$连续，且在$\bar{x}$处$KKT$条件成立，则$\bar{x}$为整体极小点。

凸规划

证明：显然S为凸集,

$\because \bar{x} \in S, f$为凸函数且在$\bar{x}$可微, $\therefore$ 对$\forall x \in S$

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) -----(1)$$

又点$\bar{x}$处KKT条件成立,所以存在$w_i (i \in I), w_i \geq 0$

使得 $\nabla f(\bar{x}) = \sum_{i \in I} w_i \nabla g_i(\bar{x}) -------(2)$

代入(1)得 $f(x) \geq f(\bar{x}) + \sum_{i \in I} w_i \nabla g_i(\bar{x})^T (x - \bar{x}) ---(3)$

$\because g_i$是凹函数，$\therefore$ 当$i \in I$有

$$g_i(x) \leq g_i(\bar{x}) + \nabla g_i(\bar{x})^T (x - \bar{x})$$

$$\Rightarrow \nabla g_i(\bar{x})^T (x - \bar{x}) \geq g_i(x) - g_i(\bar{x}) = g_i(x) \geq 0 --(4)$$

将(4)代入(3),得 $f(x) \geq f(\bar{x}) \Rightarrow \bar{x}$是整体最优解.

一般约束问题

$$
(NP)\begin{cases} \min & f(x) \\ s.t. & g_i(x) \ge 0 \quad i=1,2,\cdots,m \\ & h_j(x)=0 \quad j=1,2,\cdots,l \end{cases}
$$

$$
\begin{cases} \min & f(x) \\ s.t. & g(x) \ge 0 \\ & h(x)=0 \end{cases} \qquad g(x)=\begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{bmatrix}, h(x)=\begin{bmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_l(x) \end{bmatrix}
$$

一般约束问题

定理1(*Fritz John*条件) 设$\bar{x}\in S, I=\{i \mid g_i(\bar{x})=0\}$,
$f(x), g_i(x)(i\in I)$在$\bar{x}$处可微，$g_i(x)(i\notin I)$在$\bar{x}$处连续，$h_j(j=1,2,\cdots,l)$在$\bar{x}$处连续可微，若$\bar{x}$是问题(NP)的局部最优解，则存在不全为零的数w_0，$w_i(i\in I)$和$v_j(j=1,2,\cdots,l)$，使得

$$\begin{cases} w_0\nabla f(\bar{x})-\sum_{i\in I} w_i\nabla g_i(\bar{x})-\sum_{j=1}^{l} v_j\nabla h_j(\bar{x})=0 \\ w_0,\ w_i\ge 0,\quad i\in I. \end{cases}$$

$$h_i(x)=0 \to \begin{cases} h_i(x)\ge 0 \\ -h_i(x)\ge 0 \end{cases}$$

一般约束问题

定理1'(*Fritz John*条件) 设$\bar{x}\in S, I=\{i\mid g_i(\bar{x})=0\}$，$f(x), g_i(x)$在$\bar{x}$处可微，$h_j(j=1,2,\cdots,l)$在$\bar{x}$处连续可微，若$\bar{x}$是问题$(NP)$的局部最优解，则存在不全为零的数$w_0, w_i(i\in I)$和$v_j(j=1,2,\cdots,l)$，使得

$$\begin{cases} w_0\nabla f(\bar{x})-\sum_{i=1}^{m}w_i\nabla g_i(\bar{x})-\sum_{j=1}^{l}v_j\nabla h_j(\bar{x})=0 \\ w_i g_i(\bar{x})=0,\quad i=1,2,\cdots,m \\ w_0,\ w_i\geq 0,\quad i=1,2,\cdots,m. \end{cases}$$

一般约束问题

定理2(KKT必要条件）考虑问题
$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0, i=1,\cdots,m \\ \qquad\ \ h_j(x)=0, \ j=1,\cdots,l \end{cases}$$

设$\bar{x}$为可行点,$I=\{i \mid g_i(\bar{x})=0\}$.$f,g_i(i\in I)$在$\bar{x}$处可微，$g_i(i\notin I)$在$\bar{x}$连续，$h_j(j=1,\cdots,l)$在$\bar{x}$连续可微，向量集

$$\{\nabla g_i(\bar{x}),\ \nabla h_j(\bar{x}) \mid i\in I, j=1,\cdots,l\}$$

线性无关，若$\bar{x}$是局部最优解，则存在数$w_i, i\in I$和$v_j(j=1,\cdots,l)$，使得

$$\nabla f(\bar{x})-\sum_{i\in I} w_i\nabla g_i(\bar{x})-\sum_{j=1}^{l} v_j\nabla h_j(\bar{x})=0.$$

$$w_i \geq 0 \quad (i\in I).$$

一般约束问题

定理2'(*KKT*必要条件）考虑问题$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0, i = 1,\cdots,m \\ \qquad\quad h_j(x) = 0,\ j = 1,\cdots,l \end{cases}$

设$\bar{x}$为可行点,$I = \{i \mid g_i(\bar{x}) = 0\}$. f, g_i在$\bar{x}$处可微，$h_j(j = 1,\cdots,l)$在$\bar{x}$连续可微，向量集$\{\nabla g_i(\bar{x}),\ \nabla h_j(\bar{x}) \mid i \in I, j = 1,\cdots,l\}$线性无关，若$\bar{x}$是局部最优解，则存在数$w_i, i \in I$和$v_j(j = 1,\cdots,l)$，使得

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{x}) = 0$$

$$w_i g_i(\bar{x}) = 0, i = 1,2,\cdots,m$$

$$w_i \geq 0 \quad i = 1,2,\cdots,m.$$

一般约束问题

定义广义的**Lagrange**函数:

$$L(x,w,v)=f(x)-\sum_{i=1}^{m}w_i g_i(x)-\sum_{j=1}^{l}v_j h_j(x)$$

$$=f(x)-w^T g(x)-v^T h(x)$$

其中$w=(w_1,w_2,\cdots,w_m)^T$

$v=(v_1,v_2,\cdots,v_l)^T$

乘子向量

$g(x)=(g_1(x),\cdots,g_m(x))^T$

$h(x)=(h_1(x),\cdots,h_l(x))^T.$

一般约束问题

定理2'(*KKT*必要条件）考虑问题 $\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0, i=1,\cdots,m \\ \qquad\ \ h_j(x)=0, \ j=1,\cdots,l \end{cases}$

设$\bar{x}$为可行点,$I=\{i \mid g_i(\bar{x})=0\}$. f, g_i在$\bar{x}$处可微，$h_j(j=1,\cdots,l)$在$\bar{x}$连续可微， 向量集

$$\{\nabla g_i(\bar{x}), \ \nabla h_j(\bar{x}) \mid i \in I, j=1,\cdots,l\}$$

线性无关，若$\bar{x}$是局部最优解，则存在乘子向量$\bar{w} \geq 0, \bar{v}$,

使得

$$\nabla_x L(\bar{x}, \bar{w}, \bar{v}) = 0$$

一般约束问题

一般情形的一阶必要条件(**KKT**必要条件)可表示为**:**

$$
\begin{cases}
\nabla_x L(x,w,v)=0 \\
w_i g_i(x)=0,\ i=1,2,\cdots,m \\
g_i(x)\ge 0,\ i=1,2,\cdots,m \\
h_j(x)=0,\ j=1,2,\cdots,l \\
w_i\ge 0,\ i=1,2,\cdots,m
\end{cases}
$$

一般约束问题-凸规划

定理3.(一阶充分条件)

设问题$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \ge 0 \ \ i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0 \ \ j=1,2,\cdots,l \end{cases}$

中，f是凸函数，$g_i(i=1,2,\cdots,m)$是凹函数，$h_j(j=1,2,\cdots,l)$是线性函数，S为可行域，$\bar{x}\in S$，$I=\{i \mid g_i(\bar{x})=0\}$。$f$和$g_i(i\in I)$在点$\bar{x}$可微，$h_j(j=1,2,\cdots,l)$在点$\bar{x}$连续，$g_i(i\notin I)$在点$\bar{x}$连续，且在$\bar{x}$处$KKT$条件成立，则$\bar{x}$为整体极小点。

一般约束问题-凸规划

证明：显然S为凸集，

$\because \bar{x} \in S, f$为凸函数且在$\bar{x}$可微，$\therefore$对$\forall x \in S$

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$$

而 $\nabla f(\bar{x}) = \sum_{i \in I} w_i \nabla g_i(\bar{x}) + \sum_{j=1}^{l} v_j \nabla h_j(\bar{x})$

$$\therefore f(x) \geq f(\bar{x}) + \sum_{i \in I} w_i \nabla g_i(\bar{x})^T (x - \bar{x}) + \sum_{j=1}^{l} v_j \nabla h_j(\bar{x})^T (x - \bar{x})$$

$\because g_i$是凹函数，$\therefore$当$i \in I$有$g_i(x) \leq g_i(\bar{x}) + \nabla g_i(\bar{x})^T (x - \bar{x})$

$\Rightarrow \nabla g_i(\bar{x})^T (x - \bar{x}) \geq g_i(x) - g_i(\bar{x}) = g_i(x) \geq 0$

$\because h_j$为线性函数$\therefore h_j(x) = h_j(\bar{x}) + \nabla h_j(\bar{x})^T (x - \bar{x})$

$\Rightarrow \nabla h_j(\bar{x})^T (x - \bar{x}) = 0 \Rightarrow f(x) \geq f(\bar{x}) \Rightarrow \bar{x}$是整体最优解.

一般约束问题-凸规划

推论1：设(NP)是线性约束的凸规划，

则$\bar{x}\in S$是整体最优解$\Leftrightarrow\bar{x}$是KKT点。

推论2：问题$\begin{cases}\min\ cx \\ s.t.\quad Ax\geq b, \\ \qquad\quad x\geq 0\end{cases}$则$x^*$是最优解

$$\Leftrightarrow \text{存在} w\in R^m, v\in R^n, \text{使得}\begin{cases} Ax^*\geq b \\ x^*\geq 0 \\ c-w^TA-v^T=0 \\ w^T(Ax^*-b)=0 \\ v^Tx^*=0 \\ w,v\geq 0\end{cases}$$

一般约束问题-凸规划

求解下列线性规划问题：

$$\begin{cases} \min -2x_1 + x_2 \\ s.t. \quad x_1 + x_2 + x_3 \ge 4 \\ \qquad x_1 + 2x_2 + 2x_3 \le 6 \\ \qquad x_1, x_2, x_3 \ge 0 \end{cases}$$

$$\Rightarrow \begin{cases} \min -2x_1 + x_2 \\ s.t. \quad x_1 + x_2 + x_3 - 4 \ge 0 \\ \qquad -x_1 - 2x_2 - 2x_3 + 6 \ge 0 \\ \qquad x_1, x_2, x_3 \ge 0 \end{cases}$$

一般约束问题-凸规划

由**KKT**条件,得

$$
\begin{cases}
-2-w_1+w_2-w_3=0 & (1)\\
1-w_1+2w_2-w_4=0 & (2)\\
-w_1+2w_2-w_5=0 & (3)\\
w_1(x_1+x_2+x_3-4)=0 & (4)\\
w_2(-x_1-2x_2-2x_3+6)=0 & (5)\\
w_3x_1=0 \quad w_4x_2=0 \quad w_5x_3=0 & (6)\\
w_i\ge 0, x_i\ge 0 & (7)\\
x_1+x_2+x_3-4\ge 0 & (8)\\
-x_1-2x_2-2x_3+6\ge 0 & (9)
\end{cases}
$$

$$
\begin{cases}
Ax^*\ge b\\
x^*\ge 0\\
c-w^TA-v^T=0\\
w^T(Ax^*-b)=0\\
v^Tx^*=0\\
w,v\ge 0
\end{cases}
$$

得到KKT点$(6,0,0)^T$.

$(6,0,0)^T$为整体最优解。

一般约束问题-凸规划

例：用**KKT**条件解下列问题：

$$\begin{cases} \min \ f = (x_1 - 1)^2 + (x_2 - 1)^2 \\ s.t. \quad -x_1 - x_2 + 2 \ge 0 \\ \qquad\quad x_2 - x_1 = 1 \\ \qquad\quad x_1 \ge 0 \\ \qquad\quad x_2 \ge 0 \end{cases}$$

$$\nabla f(x) = \big(2(x_1 - 1), 2(x_2 - 1)\big)^T \quad \nabla g_1(x) = (-1, -1)^T$$

$$\nabla g_2(x) = (1, 0)^T \quad \nabla g_3(x) = (0, 1)^T \quad \nabla h(x) = (-1, 1)^T$$

一般约束问题-凸规划

KKT条件为：

$$\begin{cases} 2(x_1-1)-w_1(-1)-w_2-v(-1)=0 \\ 2(x_2-2)-w_1(-1)-w_3-v=0 \\ -x_1-x_2+2\geq 0 \\ x_2-x_1=1 \\ x_1,x_2\geq 0 \\ w_1(-x_1-x_2+2)=0 \\ w_2x_1=0 \\ w_3x_2=0 \\ w_1,w_2,w_3\geq 0 \end{cases}$$

$x^*=\left(\dfrac{1}{2},\dfrac{3}{2}\right)^T$ 为KKT点.

因为$f(x)$为凸函数，$g_i(x)$，$h(x)$为线性函数，所以本问题为凸规划问题，

$\Rightarrow x^*$为全局最优解

$$f_{\min}=\frac{1}{2}$$

一般约束问题

例：

$$\min\ f(x)=x_2$$
$$s.t.\quad g(x)=x_1^2+x_2\ge 0$$

KKT点应满足方程组

$$\begin{cases}\begin{pmatrix}0\\1\end{pmatrix}-w\begin{pmatrix}2x_1\\1\end{pmatrix}=0\\ w(x_1^2+x_2)=0\\ x_1^2+x_2\ge 0\\ w\ge 0\end{cases}\quad\Longrightarrow\quad \begin{array}{l}w=1\\ x_1=x_2=0\end{array}$$

$x^*=(0,0)^T$ 不是极小点。

一般约束问题

例：

$$\min\ x_1^2 - 3x_2 - x_2^2$$
$$s.t.\quad x_2 = 0$$
$$\text{最优解} x^* = (0,0)^T$$

Lagrange函数为

$$L(x,v) = x_1^2 - 3x_2 - x_2^2 - vx_2 = x_1^2 - (v+3)x_2 - x_2^2$$

$$\nabla L(x,v) = \begin{pmatrix} 2x_1 \\ -(v+3) - 2x_2 \end{pmatrix}, \quad \nabla^2 L(x,v) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

注意：**Lagrange**函数的**Hessian**矩阵不定不能说明该函数有没有极值点。

例：给定非线性规划问题
$$\begin{cases} \min \ f(x)=(x_1-1)^2+x_2 \\ s.t. \quad g_1(x)=-x_1-x_2+2\ge 0 \\ \qquad\quad g_2(x)=x_2\ge 0 \end{cases}$$

求满足KKT条件的点。

解：$\nabla f(x)=(2(x_1-1),1)^T, \nabla g_1(x)=(-1,-1)^T, \nabla g_2(x)=(0,1)^T$

设x为满足KTT条件的点，则有

$$\begin{cases} \begin{pmatrix} 2(x_1-1) \\ 1 \end{pmatrix} - w_1\begin{pmatrix} -1 \\ -1 \end{pmatrix} - w_2\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ w_1(-x_1-x_2+2)=0 \\ w_2x_2=0, \quad -x_1-x_2+2\ge 0 \\ w_1, w_2\ge 0, \quad x_2\ge 0 \end{cases}$$

$\Rightarrow x_1=1, x_2=0, w_1=0, w_2=1$

即KKT点为$(1,0)^T$.

一般约束问题

KKT体现一阶信息，当一阶信息不够用，

有必要考虑二阶信息

可行点 $\bar{x}$ 是 KKT 点，$(\bar{x},\bar{w},\bar{v})$ 满足 KKT 条件，

$$\nabla_x L(\bar{x},\bar{w},\bar{v}) = 0$$

$$\nabla f(\bar{x}) - \sum_{i \in I} \bar{w}_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} \bar{v}_j \nabla h_j(\bar{x}) = 0$$

一般约束问题

定理（二阶充分条件）：设f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$是二次连续可微函数，$\bar{x}$为可行解，若存在$\bar{w},\bar{v}$，使$(\bar{x},\bar{w},\bar{v})$满足KKT条件且矩阵$\nabla^2 L_x(\bar{x},\bar{w},\bar{v})$在子空间$G$上是正定的，则$\bar{x}$是严格局部极小点。

其中 $$G=\left\{ d\neq 0 \;\middle|\; \begin{array}{l} \nabla g_i(\bar{x})^T d=0, i\in I(\bar{x})\text{且}\bar{w}_i>0 \\ \nabla g_i(\bar{x})^T d\geq 0, i\in I(\bar{x})\text{且}\bar{w}_i=0 \\ \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l \end{array} \right\}.$$

Lagrange对偶问题

$$\min \ f(x)$$
$$s.t. \quad g_i(x) \geq 0, i = 1, \cdots, m$$
$$h_j(x) = 0, j = 1, \cdots, l \qquad (1)$$
$$x \in D$$

集约束

定义(1)的对偶问题：

$$\max \ \theta(w, v)$$
$$s.t. \quad w \geq 0 \qquad (2)$$

其中$\theta(w,v) = \inf\left\{ f(x) - \sum_{i=1}^{m} w_i g_i(x) - \sum_{j=1}^{l} v_j h_j(x) \middle| x \in D \right\}$

若上式不存在有限下界时，令$\theta(w,v) = -\infty$.

$\theta(w,v)$称为$Lagrange$对偶函数。

Lagrange对偶问题

$$\max\ \theta(w,v)$$

$$s.t.\quad w \geq 0$$

其中$\theta(w,v)=\inf\left\{ f(x)-\sum_{i=1}^{m} w_i g_i(x)-\sum_{j=1}^{l} v_j h_j(x) \middle| x\in D \right\}$

$$L(x,w,v)=f(x)-\sum_{i=1}^{m} w_i g_i(x)-\sum_{j=1}^{l} v_j h_j(x)$$

Lagrange函数

对于任意的$x\in D$，$Lagrange$函数$L(x,w,v)$是w,v的线性函数，于是对偶函数$\theta(w,v)$作为线性函数的逐点下确界，必然是一个凹函数，所以，对偶问题是一个凸规划问题。

例：考虑线性规划问题

$$\begin{aligned} \min\ & cx \\ s.t.\quad & A_1 x \geq b_1 \\ & A_2 x = b_2 \\ & x \geq 0 \end{aligned}$$

若取集合约束$D=\{x|x\geq 0\}$，则该线性规划问题的Lagrange函数为

$$\theta(w,v) = \inf\left\{cx - w^T(A_1x - b_1) - v^T(A_2x - b_2) \mid x \in D\right\}$$

$$= \inf\{(c - w^T A_1 - v^T A_2)x + w^T b_1 + v^T b_2 \mid x \in D\}$$

$$= \begin{cases} w^T b_1 + v^T b_2 & \text{若} c - w^T A_1 - v^T A_2 \geq 0 \\ -\infty & \text{若} c - w^T A_1 - v^T A_2 \ngeq 0. \end{cases}$$

线性规划的对偶问题为：

$$\begin{aligned} \max\ & w^T b_1 + v^T b_2 \\ s.t.\quad & w^T A_1 + v^T A_2 \leq c \\ & w \geq 0 \end{aligned}$$

求下列非线性规划问题的对偶问题:

$$\begin{cases} \min\ x_1^2 + x_2^2 \\ s.t.\quad x_1 + x_2 - 4 \ge 0 \\ \qquad\quad x_1, x_2 \ge 0 \end{cases}$$

解1: 把变量的非负限制作为集约束，即

$$x \in D = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| x_1 \ge 0, x_2 \ge 0 \right\},$$

则 $\theta(w) = \inf\left\{x_1^2 + x_2^2 - w(x_1 + x_2 - 4) \mid x \in D\right\}.$

$$\theta(w) = \inf\left\{x_1^2 + x_2^2 - w(x_1 + x_2 - 4) \mid x \in D\right\}$$
$$= \inf\left\{x_1^2 - wx_1 + x_2^2 - wx_2 + 4w \mid x_1 \geq 0, x_2 \geq 0\right\}$$
$$= \inf\left\{x_1^2 - wx_1 \mid x_1 \geq 0\right\} + \inf\left\{x_2^2 - wx_2 \mid x_2 \geq 0\right\} + 4w$$

当$w \geq 0$时,

$$\inf\left\{x_1^2 - wx_1 \mid x_1 \geq 0\right\} = \left(\frac{w}{2}\right)^2 - w \times \frac{w}{2} = -\frac{w^2}{4}.$$

$$\inf\left\{x_2^2 - wx_2 \mid x_2 \geq 0\right\} = \left(\frac{w}{2}\right)^2 - w \times \frac{w}{2} = -\frac{w^2}{4}.$$

$$\therefore \theta(w) = -\frac{w^2}{4} - \frac{w^2}{4} + 4w = -\frac{w^2}{2} + 4w.$$

对偶问题为:

$$\boxed{\begin{cases} \max -\dfrac{w^2}{2} + 4w \\ s.t. \quad w \geq 0 \end{cases}}$$

对偶定理

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g(x) \geq 0 \\ & h(x) = 0 \\ & x \in D \end{aligned}$$

$$\begin{aligned} \max \quad & \theta(w, v) \\ s.t. \quad & w \geq 0 \end{aligned}$$

$$g(x) = \left(g_1(x), \cdots, g_m(x)\right)^T$$

$$h(x) = \left(h_1(x), \cdots, h_l(x)\right)^T$$

$$\theta(w, v) = \inf \left\{ f(x) - w^T g(x) - v^T h(x) \mid x \in D \right\}$$

定理1(弱对偶定理)

设x和(w,v)分别是原问题和对偶问题的可行解，则

$$f(x) \ge \theta(w,v)。$$

推论1：对于原问题和对偶问题，必有

$$\inf\left\{f(x) \mid g(x) \ge 0, h(x)=0, x \in D\right\} \ge \sup\left\{\theta(w,v) \mid w \ge 0\right\}.$$

推论2：若$f(\bar{x}) \le \theta(\bar{w},\bar{v})$，其中$\bar{x}$为原问题的可行解，$\bar{w} \ge 0$，则$\bar{x}$和$(\bar{w},\bar{v})$分别是原问题和对偶问题的最优解。

推论3：若$\inf\left\{f(x) \mid g(x) \ge 0, h(x)=0, x \in D\right\} = -\infty$，则对$\forall w \ge 0$，有$\theta(w,v) = -\infty$。

推论4：如果$\sup\left\{\theta(w,v) \mid w \ge 0\right\} = +\infty$，则原问题没有可行解。

对偶定理

$$\inf\{f(x)\mid g(x)\geq 0, h(x)=0, x\in D\}\underset{\text{记}}{=} f_{\min}$$

$$\sup\{\theta(w,v)\mid w\geq 0\}\underset{\text{记}}{=}\theta_{\max}$$

对偶间隙(dual gap):

$$\delta = f_{\min}-\theta_{\max}\geq 0$$

问题: $\delta = 0$成立的条件.

dual gap, 凸优化的时候等于0，非凸优化也有可能等于0。

Lagrange乘子的意义

$$(NP)\begin{cases}\min & f(x) \\ s.t. & g_i(x)\geq 0 \quad i=1,2,\cdots,m \\ & h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}$$

设(NP)的局部最优解为x^*,相应的*Lagrange*乘子为(w^*,v^*), $w^*\geq 0$.

Lagrange乘子的意义

对约束的右端项进行扰动

扰动问题

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq \varepsilon_i \quad i=1,2,\cdots,m \\ \qquad\quad h_j(x) = \lambda_j \quad j=1,2,\cdots,l \end{cases}$$

令$\varepsilon = (\varepsilon_1,\cdots,\varepsilon_m), \lambda = (\lambda_1,\cdots,\lambda_l)$

设扰动问题的局部最优解为$x^*(\varepsilon,\lambda)$,相应的$Lagrange$乘子为$(w^*(\varepsilon), v^*(\lambda))$，则当$(\varepsilon,\lambda)=(0,0)$时，有$x^*(0,0) = x^*$,

$(w^*(0), v^*(0)) = (w^*, v^*)$.

Lagrange乘子的意义

只有一个等式约束

$$(NP)\begin{cases}\min\ f(x) \\ s.t.\quad h(x)=0\end{cases}$$

设局部最优解为x^*,相应的乘子为v^*.

$$\text{扰动问题}\begin{cases}\min\ f(x) \\ s.t.\quad h(x)=\lambda\end{cases}$$

设局部最优解为$x^*(\lambda)$,相应的乘子为$v^*(\lambda)$.

$$\frac{d}{d\lambda}f\left(x^*(\lambda)\right)\Big|_{\lambda=0}=\nabla_x f\left(x^*(\lambda)\right)^T\frac{d}{d\lambda}x^*(\lambda)\Big|_{\lambda=0}$$

$$=\nabla_x f\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}.$$

Lagrange乘子的意义

由扰动问题的约束条件，得到

$$h\left(x^*(\lambda)\right)=\lambda$$

$$\therefore 1=\frac{d}{d\lambda}h\left(x^*(\lambda)\right)\Big|_{\lambda=0}=\nabla_x h\left(x^*(\lambda)\right)^T\frac{d}{d\lambda}x^*(\lambda)\Big|_{\lambda=0}$$

$$=\nabla_x h\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}.$$

由**KKT**条件 $\nabla f\left(x^*\right)-v^*\nabla h\left(x^*\right)=0$

得

$$\frac{d}{d\lambda}f\left(x^*(\lambda)\right)\Big|_{\lambda=0}=\nabla_x f\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}$$

$$=v^*\nabla_x h\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}=v^*.$$

Lagrange乘子的意义

只有一个不等式约束 $(NP)\begin{cases}\min & f(x) \\ s.t. & g(x)\geq 0\end{cases}$

设局部最优解为x^*,相应的乘子为w^*.

扰动问题 $\begin{cases}\min & f(x) \\ s.t. & g(x)\geq \varepsilon\end{cases}$

设局部最优解为$x^*(\varepsilon)$,相应的乘子为$w^*(\varepsilon)$.

并假设$x^*(0)=x^*, w^*(0)=w^*$.

分两种情况讨论

Lagrange乘子的意义

(1) $g\left(x^*\right)=0$,即$g(x)\ge 0$在x^*处是起作用约束.

当$|\varepsilon|$很小时，可以假设有$g\left(x^*(\varepsilon)\right)=\varepsilon$

即$g(x)$在$x^*(\varepsilon)$处为起作用约束，所以有

$$\frac{d}{d\varepsilon} f\left(x^*(\varepsilon)\right)\Big|_{\varepsilon=0}=w^*.$$

(2) $g\left(x^*\right)>0$,即$g(x)\ge 0$在x^*处是不起作用约束.

此时，x^*是无约束问题$\min f(x)$的局部最优解，

因此当$|\varepsilon|$很小时，x^*也是扰动问题的局部最优解，所以有

$$\frac{d}{d\varepsilon} f\left(x^*(\varepsilon)\right)\Big|_{\varepsilon=0}=0=w^*.$$

Lagrange乘子的意义

$$
(NP)\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \quad i = 1, 2, \cdots, m \\ \qquad\ \ h_j(x) = 0 \quad j = 1, 2, \cdots, l \end{cases}
$$

定理：设$f(x), g_i(x), h_j(x)$具有连续的二阶偏导数，x^*是(NP)的局部最优解，(w^*, v^*)是相应的$Lagrange$乘子向量. 假设$x^*(\lambda, \varepsilon)$是扰动问题的局部最优解，$\left(w^*(\lambda), v^*(\varepsilon)\right)$是相应的乘子向量，则有

$$
\nabla_\lambda f\left(x^*(\lambda)\right)\Big|_{\lambda=0} = w^*
$$

$$
\nabla_\varepsilon f\left(x^*(\varepsilon)\right)\Big|_{\varepsilon=0} = v^*.
$$

$$
\boldsymbol{z^* = \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i}
$$

影子价格

- **定义**：在一对 P 和 D 中，若 P 的某个约束条件的右端项常数b_i（第 i 种资源的拥有量）增加一个单位时，所引起目标函数最优值z* 的改变量称为第 i 种资源的**影子价格**，其值等于D问题中对偶变量y_i*。

- **1. 影子价格的数学分析：**

$$P: \max\ z = c^T x \qquad D: \min w = y^T b$$

$$\begin{cases} Ax \le b \\ x \ge 0 \end{cases} \qquad \begin{cases} A^T y \ge c \\ y \ge 0 \end{cases}$$

- **由对偶问题的基本性质可得：** $z^* = \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$

影子价格

- **2．影子价格的经济意义**
- **1）影子价格是一种边际价格**

$$z^* = \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$$

- 在其它条件不变的情况下，单位资源数量的变化所引起的目标函数最优值的变化。即对偶变量y_i就是第 i 种资源的影子价格。即：

$$\frac{\partial Z^*}{\partial b_i} = y_i^* \quad (i = 1,2\cdots m)$$

Lagrange乘子的意义

定义广义的**Lagrange**函数:

$$L(x,w,v)= f(x)-\sum_{i=1}^{m} w_i g_i(x)-\sum_{j=1}^{l} v_j h_j(x)$$

$$= f(x)-w^T g(x)-v^T h(x)$$

其中$w=(w_1,w_2,\cdots,w_m)^T$

$v=(v_1,v_2,\cdots,v_l)^T$

乘子向量

$g(x)=(g_1(x),\cdots,g_m(x))^T$

$h(x)=(h_1(x),\cdots,h_l(x))^T.$

Lagrange乘子的意义

例：某企业预算以2千元作为广告费，根据以往的经验，若以x_1千元作广播广告，x_2千元作报纸广告，销售金额为

$$-2x_1^2-10x_2^2-8x_1x_2+18x_1+34x_2\left(千元\right)$$

试问：

(1) 如何分配2千元广告费？

(2) 广告费预算作微小改变的影响如何？

Lagrange乘子的意义

解：最优化问题为

$$\min\ 2x_1^2+10x_2^2+8x_1x_2-18x_1-34x_2$$

$$s.t.\quad x_1+x_2-2=0,\qquad x_1\geq 0, x_2\geq 0$$

相应的KKT条件为

$$\begin{cases} 4x_1+8x_2-18-w_1-v=0 \\ 20x_2+8x_1-34-w_2-v=0 \\ w_1x_1=0,\quad w_2x_2=0 \\ x_1+x_2-2=0 \\ x_1,x_2,w_1,w_2\geq 0 \end{cases}$$

$$\boxed{\begin{aligned} &KKT\text{点为}x^*=(1,1)^T \\ &w_1^*=w_2^*=0 \\ &v=-6 \end{aligned}}$$

Lagrange乘子的意义

广告费作微小改动，考虑扰动问题

$$\min\ f(x) = 2x_1^2 + 10x_2^2 + 8x_1x_2 - 18x_1 - 34x_2$$

$$s.t.\quad x_1 + x_2 - 2 = \varepsilon$$

$$x_1 \ge 0, x_2 \ge 0$$

有 $$\frac{df\left(x^*(\varepsilon)\right)}{d\varepsilon}\Big|_{\varepsilon=0} = v^* = -6$$

当ε增加时，$f\left(x^*(\varepsilon)\right)$下降，即$-f\left(x^*(\varepsilon)\right)$上升，即当广告费增加后，销售金额也随着增加，而且销售金额的增加大约是广告费的6倍，可见适当增加广告费的预算是有利的。

罚函数法

借助罚函数把约束优化问题转化为无约束优化问题，进而用无约束最优化方法进行求解。

$$\text{罚函数法}\begin{cases}\text{外点罚函数法}\\\text{内点罚函数法}\\\text{乘子罚函数法}\end{cases}$$

Sequential unconstrained minimization technique序列无约束最小化技术

外点罚函数法

外点罚函数法：通过对不可行的迭代点施加惩罚，并随着迭代过程中对不可行性增大惩罚，迫使迭代点逐步向可行域靠近。一旦迭代点成为可行点。则这个迭代点就是原问题的最优解。

$$(A)\quad \begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \geq 0\ i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0\ \ j=1,2,\cdots,l \end{cases}$$

其中$f(x), g_i(x)(i=1,2,\cdots,m), h_j(x)(j=1,2,\cdots,l)$在$E^n$上连续。

$$S=\{x \mid g_i(x)\geq 0(i=1,2,\cdots,m), h_j(x)=0(j=1,2,\cdots,l)\}$$

外点罚函数法

引入罚项 $$p(x)=\sum_{i=1}^{m}\varphi[g_i(x)]+\sum_{j=1}^{l}\psi\left[h_j(x)\right]$$

其中$\varphi(y),\psi(y)$是连续函数，且满足

$$\begin{cases}\varphi(y)=0 & \text{当} y\geq 0\\ \varphi(y)>0 & \text{当} y<0\end{cases}$$

$$\begin{cases}\psi(y)=0 & \text{当} y=0\\ \psi(y)>0 & \text{当} y\neq 0\end{cases}$$

函数ϕ和ψ的典型取法：

$$\phi\left[g_i(x)\right]=\left[\max\left\{0,-g_i(x)\right\}\right]^{\alpha}\quad \psi\left[h_j(x)\right]=\left|h_j(x)\right|^{\beta}$$

其中$\alpha\geq 1,\beta\geq 1$均为给定常数，通常取$\alpha=\beta=2$。

外点罚函数法

$$
\begin{cases}
\min \ f(x) \\
s.t. \quad g_i(x) \geq 0 \ i = 1,2,\cdots,m \\
\qquad\quad h_j(x) = 0 \ j = 1,2,\cdots,l
\end{cases}
$$

$$\Longleftrightarrow \quad \min F(x,\sigma) = f(x) + \sigma p(x)$$

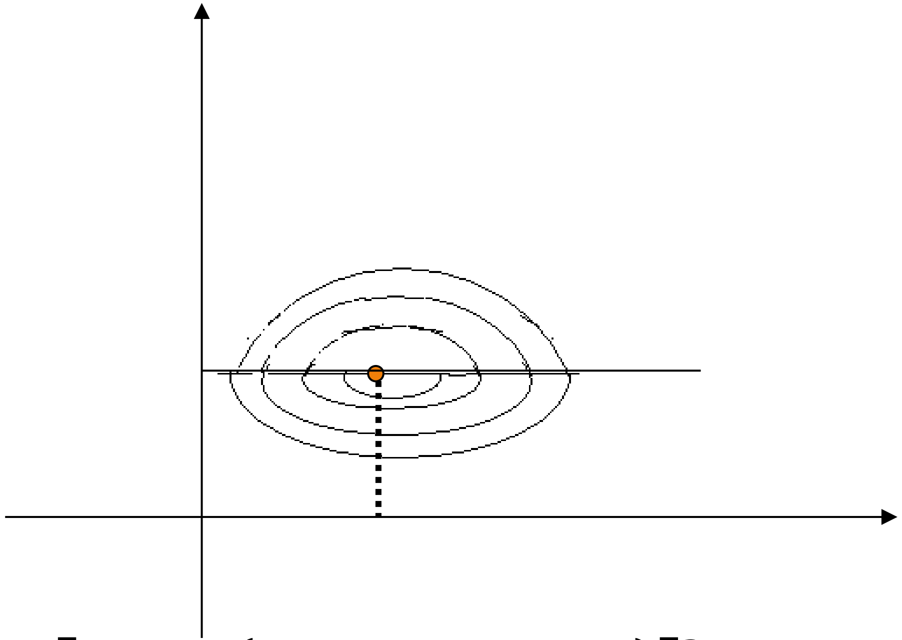
$$\min F(x,\sigma) = f(x) + \sigma\left[\sum_{i=1}^{m}\varphi[g_i(x)] + \sum_{j=1}^{l}\psi[h_j(x)]\right]$$

其中σ是很大的正数。$\bar{x}_\sigma \to x^*$ （当$\sigma \to +\infty$）

外点罚函数法

$$\min \ f(x) = (x_1 - 1)^2 + x_2^2$$

$$s.t. \quad g(x) = x_2 - 1 \ge 0$$



解：定义罚函数

$$F(x,\sigma) = (x_1 - 1)^2 + x_2^2 + \sigma\left[\max\{0, -(x_2 - 1)\}\right]^2$$

$$= \begin{cases} (x_1 - 1)^2 + x_2^2 & x_2 \ge 1 \\ (x_1 - 1)^2 + x_2^2 + \sigma(x_2 - 1)^2 & x_2 < 1 \end{cases}$$

外点罚函数法

$$\frac{\partial F}{\partial x_1} = 2(x_1 - 1) \quad \frac{\partial F}{\partial x_2} = \begin{cases} 2x_2 & x_2 \ge 1 \\ 2x_2 + 2\sigma(x_2 - 1) & x_2 < 1 \end{cases}$$

令 $\frac{\partial F}{\partial x_1} = 0 \quad \frac{\partial F}{\partial x_2} = 0$ 得

$$\overline{x}_\sigma = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \dfrac{\sigma}{1+\sigma} \end{pmatrix} \to \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\sigma \to +\infty)$$

外点罚函数法

$$\min\ x_1 + x_2$$

$$s.t.\quad x_1 - x_2^2 = 0$$

定义罚函数：$F(x,\sigma) = x_1 + x_2 + \sigma(x_1 - x_2^2)^2$

$$\nabla F(x,\sigma) = \begin{pmatrix} 1 + 2\sigma(x_1 - x_2^2) \\ 1 + 2\sigma(x_1 - x_2^2)(-2x_2) \end{pmatrix}$$

令$\nabla F(x,\sigma) = 0$,得

$$\bar{x}_\sigma = \left(\frac{1}{4} - \frac{1}{\sigma}, -\frac{1}{2}\right)^T \to \left(\frac{1}{4}, -\frac{1}{2}\right)^T (\sigma \to \infty)$$

外点罚函数法

$$
\begin{aligned}
\min\ \ & f(x) = x_1 + x_2 \\
s.t.\ \ & g_1(x) = -x_1^2 + x_2 \ge 0 \\
& g_2(x) = x_1 \ge 0
\end{aligned}
$$

解：定义罚函数

$$F(x,\sigma) = x_1 + x_2 + \sigma\left[\max\left\{0, x_1^2 - x_2\right\}\right]^2 + \sigma\left[\max\left\{0, -x_1\right\}\right]^2$$

$$\frac{\partial F}{\partial x_1} = 1 + 2\sigma\left[2\max\left\{0, x_1^2 - x_2\right\}x_1 + \max\left\{0, -x_1\right\}(-1)\right]$$

$$\frac{\partial F}{\partial x_2} = 1 + 2\sigma\max\left\{0, x_1^2 - x_2\right\}(-1)$$

外点罚函数法

令 $\dfrac{\partial F}{\partial x_1}=0 \quad \dfrac{\partial F}{\partial x_2}=0$ 得

$$\begin{cases} 1+4\sigma x_1(x_1^2-x_2)+2\sigma x_1=0 \\ 1-2\sigma(x_1^2-x_2)=0 \end{cases}$$

$$\Rightarrow \quad x_1=-\frac{1}{2+2\sigma} \quad x_2=\frac{1}{(2+2\sigma)^2}-\frac{1}{2\sigma}$$

当$\sigma\to+\infty$时，有$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\to\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

外点罚函数法

步骤:

1. 给定初始点$x^{(0)}$，初始罚因子$\sigma_1 > 0(\sigma_1 = 1)$，放大系数$c > 1$，允许误差$\varepsilon > 0$，置$k = 1$。

2. 以$x^{(k-1)}$为初始点，求解无约束 问题

$$\min \ f(x) + \sigma_k p(x)$$

设其极小点为$x^{(k)}$。

3. 若$\sigma_k p(x^{(k)}) < \varepsilon$，则停止计算，得到点$x^{(k)}$；否则，令$\sigma_{k+1} = c\sigma_k$，置$k := k+1$，返回2。

外点罚函数法

例:用外点法求解

$$\min\ (x-1)^2$$

$$s.t.\quad x-2\geq 0$$

解：取$x^{(0)}=0, \sigma_1=1$，令

$$p(x)=\left[\max\{0,-x+2\}\right]^2 \quad 则$$

$$p(x)=\begin{cases}0 & x\geq 2\\ (-x+2)^2 & x<2\end{cases}$$

外点罚函数法

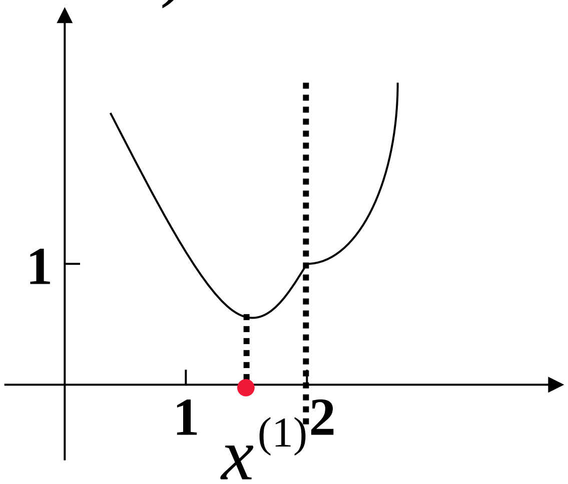
第一次迭代

求解无约束最优化问题:

$$\min F(x,\sigma_1)=(x-1)^2+1\times p(x)$$

$$其中F(x,\sigma_1)=\begin{cases}(x-1)^2 & x\ge 2\\ (x-1)^2+1\times(-x+2)^2 & x<2\end{cases}$$

解得: $x^{(1)}=\dfrac{3}{2}$



$$令\sigma_2=10\times\sigma_1=10$$

外点罚函数法

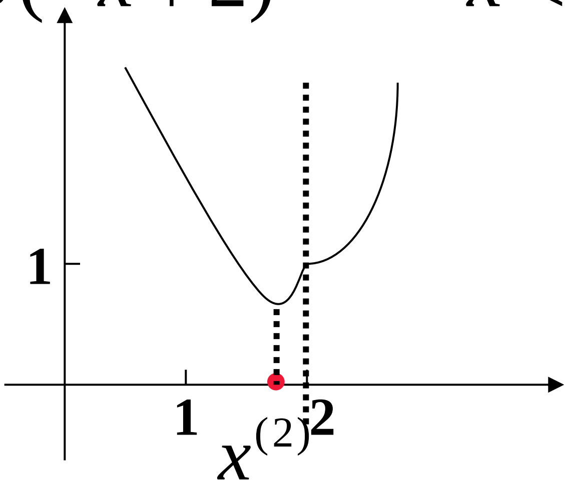
第二次迭代

求解无约束最优化问题：

$$\min F(x,\sigma_1)=(x-1)^2+10\times p(x)$$

其中$F(x,\sigma_1)=\begin{cases}(x-1)^2 & x\geq 2\\ (x-1)^2+10\times(-x+2)^2 & x<2\end{cases}$

解得： $x^{(1)}=\dfrac{21}{11}$



令$\sigma_3=10\times\sigma_2=100$

外点罚函数法

第三次迭代

求解无约束最优化问题:

$$\min\ F(x,\sigma_1)=(x-1)^2+100\times p(x)$$

其中$F(x,\sigma_1)=\begin{cases}(x-1)^2 & x\ge 2\\ (x-1)^2+100\times(-x+2)^2 & x<2\end{cases}$

解得： $x^{(2)}=\dfrac{201}{101}$

以此类推，得序列：

$$\frac{3}{2},\frac{21}{11},\frac{201}{101},\frac{2001}{1001},\cdots \qquad x^*=2$$

外点罚函数法

引理1 对于由外点法所产生的序列$\{x^{(k)}\}$，总有

(1) $F(x^{(k+1)},\sigma_{k+1}) \geq F(x^{(k)},\sigma_k)$

(2) $p(x^{(k+1)}) \leq p(x^{(k)})$

(3) $f(x^{(k+1)}) \geq f(x^{(k)})$

证明:(1)由$F(x,\sigma)=f(x)+\sigma p(x)$和$\sigma_{k+1}>\sigma_k$知

$$F(x^{(k+1)},\sigma_{k+1}) = f(x^{(k+1)})+\sigma_{k+1}p(x^{(k+1)})$$

$$\geq f(x^{(k+1)})+\sigma_k p(x^{(k+1)}) = F(x^{(k+1)},\sigma_k)$$

$\because x^{(k)}$是$F(x,\sigma_k)$的极小点,$\therefore$ 对$\forall x$,有$F(x,\sigma_k)\geq F(x^{(k)},\sigma_k)$

$$\Rightarrow F(x^{(k+1)},\sigma_k) \geq F(x^{(k)},\sigma_k)$$

$$\Rightarrow F(x^{(k+1)},\sigma_{k+1}) \geq F(x^{(k)},\sigma_k)$$

外点罚函数法

(2) $x^{(k)}$和$x^{(k+1)}$分别使$F(x,\sigma_k),F(x,\sigma_{k+1})$取极小

$$\therefore f(x^{(k+1)})+\sigma_k p(x^{(k+1)}) \geq f(x^{(k)})+\sigma_k p(x^{(k)}) \quad (*)$$

$$f(x^{(k)})+\sigma_{k+1} p(x^{(k)}) \geq f(x^{(k+1)})+\sigma_{k+1} p(x^{(k+1)})$$

$$\sigma_k p(x^{(k+1)})+\sigma_{k+1} p(x^{(k)}) \geq \sigma_k p(x^{(k)})+\sigma_{k+1} p(x^{(k+1)})$$

$$\Rightarrow \left(\sigma_{k+1}-\sigma_k\right) p(x^{(k)}) \geq \left(\sigma_{k+1}-\sigma_k\right) p(x^{(k+1)})$$

$$\Rightarrow p(x^{(k)}) \geq p(x^{(k+1)})$$

$$\boxed{\begin{array}{l}(2)\ p(x^{(k+1)}) \leq p(x^{(k)}) \\ (3)\ f(x^{(k+1)}) \geq f(x^{(k)})\end{array}}$$

(3) 由(*),得

$$f(x^{(k+1)})-f(x^{(k)}) \geq \sigma_k\left(p(x^{(k)})-p(x^{(k+1)})\right)$$

$$\geq 0$$

外点罚函数法

引理2 设x^*是问题(A)的一个最优解,则对$\forall k$,有

$$f(x^*) \geq F(x^{(k)},\sigma_k) \geq f(x^{(k)}).$$

证明:因为x^*是问题(A)的最优解,所以有

$$p(x^*) = 0$$

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \\ \qquad\quad h_j(x) = 0 \end{cases}$$

$\because x^{(k)}$是$F(x,\sigma_k)$的极小点

$$\therefore f(x^*) = F(x^*,\sigma_k) \geq F(x^{(k)},\sigma_k)$$

又$\because \sigma_k p(x^{(k)}) \geq 0$

$$\therefore F(x^{(k)},\sigma_k) = f(x^{(k)}) + \sigma_k p(x^{(k)}) \geq f(x^{(k)})$$

外点罚函数法

外点罚函数法的一个重要特点：

函数$F(x,\sigma)$是在整个空间E^n内进行优化,初始点可任意选择,且外点法也可用于非凸规划的最优化.

缺点：

1. 惩罚项$\sigma p(x)$的二阶偏导数一般不存在;

2. 外点法的中间结果不是可行解,不能作为近似解;

3. 当点$x^{(k)}$接近最优解时,罚因子σ_k很大.可能使罚函数性质变坏,使搜索产生极大困难

外点罚函数法

$$\min\ x_1^2 + x_2^2$$
$$s.t.\quad x_1 + 1 = 0$$

其罚函数为

$$F(x,\sigma) = x_1^2 + x_2^2 + \sigma(x_1+1)^2.$$

*Hessan*矩阵为

$$\nabla_x^2 F(x,\sigma) = \begin{pmatrix} 2+\sigma & 0 \\ 0 & 2 \end{pmatrix}$$

$$\text{条件数}=\frac{\text{最大特征值}}{\text{最小特征值}} = \frac{2+\sigma}{2} \to \infty\ (\sigma \to \infty)$$

外点罚函数法

$$\begin{cases} \min -x^4 \\ s.t. \quad x = 0 \end{cases} \qquad x^* = 0$$

若取$p(x)=x^2$，则$F(x,\sigma)=-x^4+\sigma x^2$，没有最小点.

若取$p(x)=x^8$，则$F(x,\sigma)=-x^4+\sigma x^8$，有极小点

$$x_\sigma = \left(\frac{1}{2\sigma}\right)^{\frac{1}{4}} \to 0$$

内点罚函数法

基本思想: 迭代总是从内点出发,并保持在可行域内部进行搜索.

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \ \ i = 1, 2, \cdots, m \end{cases}$$

其中$f(x), g_i(x)(i=1,2,\cdots,m)$是连续函数。

$$S = \{x \mid g_i(x) \geq 0, i = 1,2,\cdots,m, x \in E^n\}$$

$$\operatorname{int} S = \{x \mid g_i(x) > 0, i = 1,2,\cdots,m, x \in E^n\}$$

内点罚函数法

障碍因子

障碍函数 $G(x,r) = f(x) + rB(x)$

其中r是很小的正数，$B(x)$定义在可行域内部，它满足两个条件：

(1) $B(x)$是连续函数；

(2) 当点x趋向可行域边界时，$B(x) \to +\infty$。

两种最重要的形式：

倒数障碍函数

$$B(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)}$$

对数障碍函数

$$B(x) = -\sum_{i=1}^{m} \ln g_i(x)$$

内点罚函数法

两种障碍函数的比较

$$\min -x$$

$$s.t. \quad -x-1\geq 0$$

取 $B(x)=\dfrac{1}{-x-1}$,则内罚函数为

$$G(x,r)=-x-\frac{r}{x+1}$$

令 $\dfrac{\partial G}{\partial x}=-1+\dfrac{r}{(x+1)^2}=0$

得到 $\bar{x}=\bar{x}(r)=-1-\sqrt{r}$

当$r\to 0$时，有$\bar{x}(r)\to -1=x^*.$

内点罚函数法

两种障碍函数的比较

$$\min -x$$

$$s.t. \quad -x-1 \geq 0$$

取 $B(x) = -\ln(-x-1)$，则内罚函数为

$$G(x,r) = -x - r\ln(-x-1)$$

令 $\dfrac{\partial G}{\partial x} = -1 - \dfrac{r}{x+1} = 0$

得到 $\bar{x} = \bar{x}(r) = -1 - r$

当$r \to 0$时，有$\bar{x}(r) \to -1 = x^*$.

内点罚函数法

例:考虑约束优化问题
$$\min \frac{x}{2}$$
$$s.t. \quad x \geq 1$$

该问题的对数障碍函数为

$$G(x,r)=\frac{x}{2}-r\ln(x-1)$$

$G(x,r)$的最小点为:

$$x_r = 1+2r \to 1 (r \to 0)$$

$$G(x_r,r)=\frac{1}{2}+r-r\ln 2r \to \frac{1}{2}(r \to 0)$$

内点罚函数法

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \ \ i=1,2,\cdots,m \end{cases}$$

$$\Rightarrow \quad \begin{cases} \min G(x,r) = f(x) + rB(x) \\ s.t. \quad x \in \operatorname{int} S \end{cases}$$

$$\begin{cases} \min G(x,r_k) = f(x) + r_k B(x) \\ s.t. \quad x \in \operatorname{int} S \end{cases}$$

其中$\{r_k\}$为严格单调减且趋于0的障碍因子数列。

内点罚函数法

步骤:

1. 给定初始点$x^{(0)} \in \operatorname{int} S$，允许误差$\varepsilon > 0$，初始参数$r_1$,缩小系数$\beta \in (0,1)$,置$k = 1$。

2. 以$x^{(k-1)}$为初始点，求解下列问题

$$\begin{aligned} \min \quad & f(x) + r_k B(x) \\ s.t. \quad & x \in \operatorname{int} S \end{aligned}$$

设其极小点为$x^{(k)}$。

3. 若$r_k B(x^{(k)}) < \varepsilon$，则停止计算，得到点$x^{(k)}$；否则，令$r_{k+1} = \beta r_k$，置$k := k+1$，返回2。

内点罚函数法

例:用内点罚函数法求解下列问题

$$\min\ x_1 + x_2$$

$$s.t.\quad -x_1^2 + x_2 \geq 0$$

$$x_1 \geq 0$$

解：定义障碍函数

$$G(x,r) = x_1 + x_2 - r_k \ln\left(-x_1^2 + x_2\right) - r_k \ln x_1$$

令 $\dfrac{\partial G}{\partial x_1} = 1 - \dfrac{-2x_1 r_k}{-x_1^2 + x_2} - \dfrac{r_k}{x_1} = 0,\ \dfrac{\partial G}{\partial x_2} = 1 - \dfrac{r_k}{-x_1^2 + x_2} = 0$

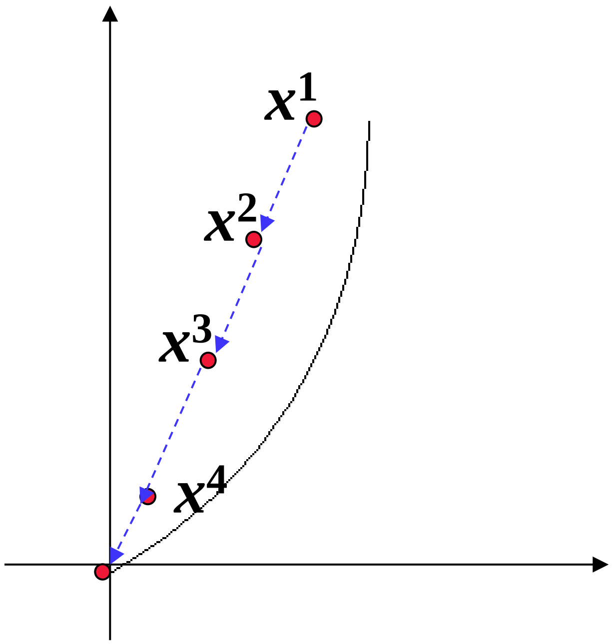
$$\Rightarrow x_1 = \frac{1}{4}\left(-1 + \sqrt{1+8r_k}\right),\ x_2 = \frac{3r_k}{2} - \frac{1}{8}\left(-1 + \sqrt{1+8r_k}\right)$$

当$r_k \to 0$时，$x^* = (0,0)^T$

内点罚函数法

$$x_1=\frac{1}{4}\left(-1+\sqrt{1+8r_k}\right),\quad x_2=\frac{3r}{2}-\frac{1}{8}\left(-1+\sqrt{1+8r_k}\right)$$

	r_k	$x_1(r_k)$	$x_2(r_k)$
1	1	0.5	1.25
2	0.5	0.309	0.595
3	0.25	0.183	0.283
4	0.1	0.085	0.107
5	0.0001	0	0



内点罚函数法

求初始内点的迭代步骤

1. 任取$x^{(0)} \in E^n, r_0 > 0$(如取$r_0 = 1$),置$k := 0$。

2. 令 $S_k = \{i \mid g_i(x^{(k)}) \le 0, 1 \le i \le m\}$,

$$T_k = \{i \mid g_i(x^{(k)}) > 0, 1 \le i \le m\}.$$

3. 若$S_k = \varnothing$，停止计算；否则，转4。

4. 构造函数

$$\widetilde{P}(x, r_k) = -\sum_{i \in S_k} g_i(x) + r_k \sum_{i \in T_k} \frac{1}{g_i(x)} \quad (r_k > 0)$$

记 $\widetilde{R}_k = \{x \mid g_i(x) > 0 \quad i \in T_k\}$

内点罚函数法

5. 以$x^{(k)}$为初始点，在$\tilde{R}_k$域内，求障碍函数$\tilde{P}(x, r_k)$的极小点:

$$\min \ \tilde{P}(x, r_k)$$
$$s.t. \quad x \in \tilde{R}_k$$

得$x^{(k+1)}$，转6。

6. 令$0 < r_{k+1} < r_k \left(如取 r_{k+1} = \frac{1}{10} r_k \right)$, 置$k := k+1$,

转2。

内点罚函数法

内点罚函数法优点

迭代总在可行域内进行，每一个中间结果都是可行解，可以作为近似解。

内点罚函数法缺点

选取初始可行点较困难，且只适用于含不等式约束的非线性规划问题。

混合罚函数法

$$
(A)\quad \begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \ i = 1,2,\cdots,m \\ \qquad\quad h_j(x) = 0 \ j = 1,2,\cdots,l \end{cases}
$$

引入增广目标函数

$$
P(x,r) = f(x) - r\sum_{i\in I_1} \ln g_i(x)
$$

$$
+\frac{1}{r}\left[\sum_{i\in I_2}(\min(0,-g_i(x))^2 + \sum_{j=1}^{l} h_j(x)^2\right]
$$

其中$I_1 = \{i \mid g_i(x) > 0, i = 1,\cdots,m\}, I_2 = \{i \mid g_i(x) \leq 0, i = 1,\cdots,m\}$

拉格朗日乘子法

考虑等式约束优化问题

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & c_i(x)=0,\ i=1,\cdots,m. \end{aligned} \tag{6.1}$$

其中$f(x)$和$c_i(x)$连续可微的，$m<n$。

设x^*是局部极小点且是正则点，那么由KKT条件，存在乘子向量
$\lambda^*=(\lambda_1^*,\cdots,\lambda_m^*)^T$, 使得

$$\nabla f(x^*)-\sum_{i=1}^{m}\lambda_i^*\nabla c_i(x^*)=0, \tag{6.2}$$

$$c_i(x^*)=0,\ i=1,\cdots,m. \tag{6.3}$$

拉格朗日乘子法

算法 6.1

步 *1.* 构造拉格朗日函数

$$L(x,\lambda) = f(x) - \sum_{i=1}^{m} \lambda_i c_i(x).$$

步 *2.* 解*KKT*条件方程组*(6.2)-(6.3)*得到(x^*, λ^*).

拉格朗日乘子法

例：

$$\min \quad f(x) = (x_1 - \frac{13}{3})^2 + (x_2 - \frac{1}{2})^2 - x_3 \tag{6.4}$$

$$\text{s.t.} \quad c_1(x) = x_1 + \frac{5}{3}x_2 - 10 = 0, \tag{6.5}$$

$$c_2(x) = (x_2 - 2)^2 + x_3 - 4 = 0. \tag{6.6}$$

解 1. 求约束梯度

$$\nabla c_1(x) = \begin{bmatrix} 1 \\ \frac{5}{3} \\ 0 \end{bmatrix}, \nabla c_2(x) = \begin{bmatrix} 0 \\ 2x_2 - 4 \\ 1 \end{bmatrix}.$$

2. 构造拉格朗日函数

$$\begin{aligned} L(x, \lambda) \quad = \quad & (x_1 - \frac{13}{3})^2 + (x_2 - \frac{1}{2})^2 - x_3 \\ & -\lambda_1(x_1 + \frac{5}{3}x_2 - 10) - \lambda_2((x_2 - 2)^2 + x_3 - 4). \end{aligned}$$

拉格朗日乘子法

3. 解KKT条件方程组:

$$
\begin{aligned}
\frac{\partial L}{\partial x_1} &= 2(x_1 - \frac{13}{3}) - \lambda_1 = 0, \\
\frac{\partial L}{\partial x_2} &= 2(x_2 - \frac{1}{2}) - \frac{5}{3}\lambda_1 - 2\lambda_2(x_2 - 2) = 0, \\
\frac{\partial L}{\partial x_3} &= -1 - \lambda_2 = 0, \\
\frac{\partial L}{\partial \lambda_1} &= -(x_1 + \frac{5}{3}x_2 - 10) = 0, \\
\frac{\partial L}{\partial \lambda_2} &= -((x_2 - 2)^2 + x_3 - 4) = 0.
\end{aligned}
$$

解得

$$
x^* = \begin{pmatrix} 8\frac{11}{18} \\ \frac{5}{6} \\ 2\frac{23}{36} \end{pmatrix}, \quad \lambda^* = \begin{pmatrix} \frac{77}{9} \\ -1 \end{pmatrix}. \quad \square
$$

增广拉格朗日乘子法

罚函数法的主要缺点是要求罚参数$\sigma \to +\infty$时，才能得到约束优化问题的解。 在罚函数的基础上提出的增广Lagrange乘子法（又称乘子罚函数法）可以克服这个缺点。

考虑等式约束优化问题

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) = 0, i \in \{1, 2, \cdots, m\}. \end{aligned} \tag{11.1}$$

增广拉格朗日函数为

$$P(x, \lambda, \sigma) = f(x) - \sum_{i=1}^{m} \lambda_i c_i(x) + \frac{1}{2} \sum_{i=1}^{m} \sigma_i c_i^2(x), \tag{11.2}$$

这里λ_i为乘子，σ_i为罚因子。显然， 增广拉格朗日函数是由拉格朗日函数再加上一个惩罚项$\frac{1}{2}\sum_{i=1}^{m} \sigma_i c_i^2(x)$构成的.

增广拉格朗日乘子法

给定$\lambda^{(k)}$, $\sigma^{(k)}$, 令x_{k+1}为无约束优化问题

$$\min_x P(x,\ \lambda^{(k)},\ \sigma^{(k)}) \tag{11.3}$$

的解。于是，我们立即有

$$\begin{aligned}
& \nabla_x P(x_{k+1},\ \lambda^{(k)},\ \sigma^{(k)}) \\
= \ & \nabla f(x_{k+1}) - \sum_{i=1}^{m} \lambda_i^{(k)} \nabla c_i(x_{k+1}) + \sum_{i=1}^{m} \sigma_i^{(k)} c_i(x_{k+1}) \nabla c_i(x_{k+1}) \\
= \ & \nabla f(x_{k+1}) - \sum_{i=1}^{m} [\lambda_i^{(k)} - \sigma_i^{(k)} c_i(x_{k+1})] \nabla c_i(x_{k+1}) \qquad (11.4) \\
= \ & 0.
\end{aligned}$$

与在x^*处的一阶必要条件相比，我们取

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} - \sigma_i^{(k)} c_i(x_{k+1}),\ i = 1, \cdots, m \tag{11.5}$$

做为下一次迭代的Lagrange乘子。

增广拉格朗日乘子法

这样，我们有

$$\nabla f(x_{k+1}) - \sum_{i=1}^{m} \lambda_i^{(k+1)} \nabla c_i(x_{k+1}) = 0.$$

在算法中，当

$$|c_i(x_{k+1})| \leq \frac{1}{4}|c_i(x_k)| \qquad (11.6)$$

不满足时，我们就扩大相应的罚因子，令

$$\sigma_i^{(k+1)} = 10\sigma_i^{(k)}.$$

11.6式说明：迭代点的极限满足等式约束，在迭代过程中，等式约束是不满足的。但是还是希望约束满足，所以当约束的满足情况没有变好时，就会增大惩罚因子。

增广拉格朗日乘子法

步 1. 给出初始点$x_1 \in R^n$, $\lambda^{(1)} \in R^m$, $\sigma_i^{(1)} > 0, i = 1, \cdots, m$, $\varepsilon \geq 0$, $k := 1$.

步 2. 求解(11.3)得到x_{k+1}. 如果$||c(x_{k+1})||_\infty \leq \varepsilon$, 则停。

步 3. 对$i = 1, \cdots, m$, 令

$$\sigma_i^{(k+1)} = \begin{cases} \sigma_i^{(k)}, & \text{如果(11.6)成立;} \\ \max(10\sigma_i^{(k)},\ k^2), & \text{否则.} \end{cases}$$

步 4. 由(11.5)计算乘子$\lambda_i^{(k+1)}$, $i = 1, \cdots, m$. $k := k + 1$; 转步2.

谢　　谢！