最优化方法 Operations Research

约束优化问题的最优性条件

第六讲

问题的提出

例:某公司经营两种设备。第一种设备每件售价为30元，第二种设备每件售价为450元。且知，售出第一、二种设备分别需时为每件约0.5小时和（$2+0.25x_2$）小时，其中x_2为第二种设备售出数量。公司的总营业时间为800小时。

求：公司为获取最大营业额（销售额）的最优营业计划

[解]设公司应经营销售第一、二种设备数额分别为x_1件和x_2件，则有

$$\max\text{：} f(X)=30x_1+450x_2$$

$$s.t.\ 0.5x_1+2x_2+0.25x_2^2\le 800$$

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

$$\text{非线性规划}\begin{cases}\text{无约束优化} & \min\limits_{x\in E^n} \ f(x) \\ \text{约束优化} & \end{cases}$$

约束优化问题数学模型：

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

约束优化问题的最优性条件

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

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

------------可行集或可行域

思路：

几何------ 代数

直观------ 抽象

简单------ 复杂

？？？
Start with some basic concepts/knowledge

定义: 设集合$S \subset E^n$, $\bar{x} \in clS$, 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)。

$\bar{x}$处的可行方向锥(集)：

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

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)。

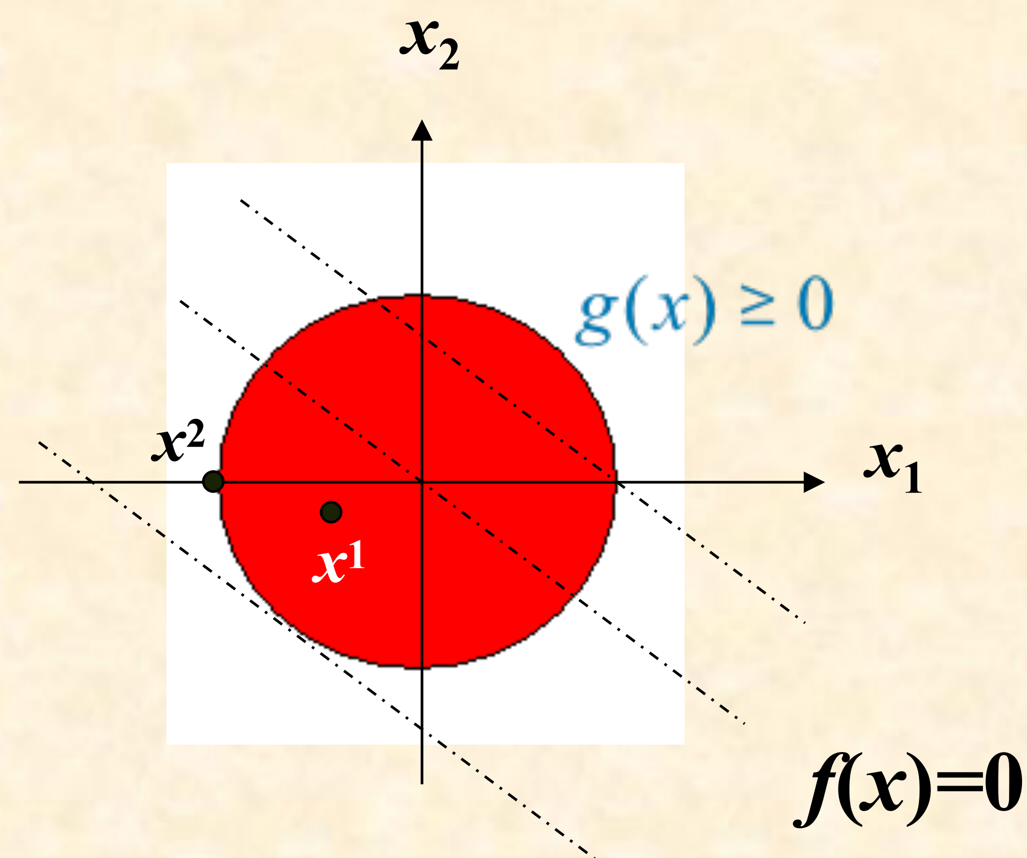
点 $\bar{x}$ 处的下降方向集:

$$F_0=\left\{d\,\middle|\,\nabla f(\bar{x})^T d<0\right\}$$

例：考虑如下约束优化问题：

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

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



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

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

一般约束问题的最优性条件

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

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

定理1（几何最优性条件）：考虑问题

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

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：设存在的$d \in F_0 \cap D$,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。

定义： 设$\bar{x}$为可行点，不等式约束中在$\bar{x}$起作用约束下标集记为I，如果向量组

$$\left\{\nabla g_i(\bar{x}), \nabla h_j(\bar{x}) \mid i \in I, j = 1, 2, \cdots, l\right\}$$

线性无关，则称$\bar{x}$为约束$g(x) \geq 0$和$h(x) = 0$的正则点。

定义： 称点集$\{x = x(t) \mid t_0 \leq t \leq t_1\}$为曲面$S = \{x \mid h(x) = 0\}$上的一条曲线，如果对所有$t \in [t_0, t_1]$均有$h(x(t)) = 0$.

如果导数$x'(t) = \dfrac{\mathrm{d}x(t)}{\mathrm{d}t}$存在，则称曲线是可微的.

曲面S上在点x处所有可微曲线的切向量组成的集合，称为曲面S在点x的切平面，记为$T(x)$.

定义子空间

$$H := \left\{ d \mid \nabla h(x)^T d = 0 \right\}$$

其中$\nabla h(x) = \left(\nabla h_1(x), \nabla h_2(x), \cdots, \nabla h_l(x) \right)^T$

结论：

若向量$d \in T(\bar{x})$，则有

$$d \in H = \left\{ d \mid \nabla h(\bar{x})^T d = 0 \right\}$$

定理：设$\bar{x}$是曲面$S=\{x\mid h(x)=0\}$上的一个正则点，则

$$T(\bar{x})=H=\{d\mid \nabla h(\bar{x})^T d=0\}.$$

证明：设$d\in H$,考虑非线性方程组$h(\bar{x}+td+\nabla h(\bar{x})y)=0$

其中$t\in R^1, y\in R^l$.该方程组有解$(y,t)=(0,0)$.

在$t=0$处，h关于y的$Jacobi$矩阵为$\nabla h(\bar{x})^T\nabla h(\bar{x})$

由隐函数定理，在$t=0$的邻域，存在连续可微函数

$y=y(t)(y(0)=0)$,使$h(\bar{x}+td+\nabla h(\bar{x})y(t))=0$成立

令$x(t)=\bar{x}+td+\nabla h(\bar{x})y(t)$,则$x(t)$为曲面$S$上过$x(0)$的一条曲线。在点$x(0)=\bar{x}$，切向量为

$$x'(0)=d+\nabla h(\bar{x})y'(0)$$

又 $\nabla h(\bar{x})^T d+\nabla h(\bar{x})^T\nabla h(\bar{x})y'(0)=0$

而$d\in H\Rightarrow \nabla h(\bar{x})^T d=0$, $\because$ $\bar{x}$是正则点, $\therefore$ $\nabla h(\bar{x})^T\nabla h(\bar{x})$可逆,

$\Rightarrow y'(0)=0\Rightarrow x'(0)=d\Rightarrow d\in T(\bar{x})$.

定理：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，$h_j(j=1,2,\cdots,l)$在$\bar{x}$处连续可微，且$\bar{x}$是$S=\{x\mid h(x)=0\}$的正则点。如果$\bar{x}$是问题(NP)的局部最优解，则在$\bar{x}$处，有

$$F_0\cap G_0\cap H_0=\varnothing,$$

其中 $$F_0=\left\{d\mid \nabla f(\bar{x})^T d<0\right\},$$

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

$$H_0=\left\{d\mid \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l\right\}.$$

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

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

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

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

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

设$\bar{x}$为可行点,$I=\{i \mid g_i(\bar{x})=0\}$.$f, g_i(i\in I)$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

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

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

定理2'(KKT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,$I=\{i\,|\,g_i(\bar{x})=0\}$.$f,g_i$在$\bar{x}$处可微，$h_j(j=1,\cdots,l)$在$\bar{x}$连续可微，向量集

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

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

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

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

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

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

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

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

向量形式

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

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

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

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

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

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

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$\bar{w} \ge 0, \bar{v}$,

使得

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

北京交通大學

一般情形的一阶必要条件(KKT必要条件)可表示为:

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

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

$\Leftrightarrow$ 存在$w \in R^m, v \in R^n$，使得

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

求下列线性规划问题的KKT点:

1.
$$\begin{cases} \min -2x_1 + x_2 \\ \text{s.t.} \quad x_1 + x_2 + x_3 \geq 4 \\ \qquad x_1 + 2x_2 + 2x_3 \leq 6 \\ \qquad x_1, x_2, x_3 \geq 0 \end{cases}$$

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

由KKT条件,得

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

得到KKT点$(6, 0, 0)^T$.

$(6,0,0)^T$为整体最优解。

$$2.\ \begin{cases} \min\ f(x) = x_2 \\ s.t.\quad g(x) = x_1^2 + x_2 \geq 0 \end{cases}$$

KKT点应满足方程组

$$\begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} - w \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix} = 0 \\ w(x_1^2 + x_2) = 0 \\ x_1^2 + x_2 \geq 0 \\ w \geq 0 \end{cases} \Rightarrow \begin{matrix} w = 1 \\ x_1 = x_2 = 0 \end{matrix}$$

$x^* = (0,0)^T$ 不是极小点。

例3：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$(1,0)^T$.

Lagrange乘子的意义

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

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

对约束的右端项进行扰动

扰动问题

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

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

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

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,$(w^*(\lambda),v^*(\varepsilon))$是相应的乘子向量，则有

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

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

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

$$-2x_1^2-10x_2^2-8x_1x_2+18x_1+34x_2\,(\text{千元})$$

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？

解： 最优化问题为

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

$$\begin{aligned} s.t. \quad & x_1 + x_2 - 2 = 0 \\ & x_1 \ge 0, x_2 \ge 0 \end{aligned}$$

相应的KKT条件为

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

KKT点为$x^* = (1,1)^T$

$w_1^* = w_2^* = 0$

$v = -6$

广告费作微小改动，考虑扰动问题

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

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

$$x_1 \geq 0, x_2 \geq 0$$

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

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。

Homework

写出下列问题的KKT点

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

$$\begin{cases} \min \left(x_1 - \dfrac{4}{9} \right)^2 + (x_2 - 2)^2 \\ s.t. \quad x_2 - x_1^2 \geq 0 \\ \qquad x_1 + x_2 \leq 6 \\ \qquad x_1, x_2 \geq 0 \end{cases}$$