KKT条件:

problem: $\min f(x)$

$$\text{s.t. } h_i(x) = 0 \quad i = 1 \cdots m$$
$$g_j(x) \le 0 \quad j = 1 \cdots p.$$

$$L(x,\lambda,\mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x)$$

first order condition:

$$\nabla L(x^*,\lambda,\mu) = 0$$

$$\Leftrightarrow \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(x^*)$$

互补松弛:

$$\mu_j g_j(x^*) = 0 \Leftrightarrow j \notin A(x),\ g_j(x) < 0$$

then $\mu_j = 0$

KKT证明：(利用

原问题 $\Leftrightarrow$

$$\min f(x)$$

$$\text{s.t.}\quad h_i(x)=0 \quad i=1\cdots m \rightarrow \lambda_i$$

$$g_j(x)+z_j^2=0 \quad j=1\cdots p \rightarrow \mu_j$$

此时变量为 x, z.

由 Lagrange Theorm.

一阶条件证明

$$\begin{pmatrix}\nabla f(x)\\ 0\end{pmatrix}+\sum_{i=1}^{m}\lambda_i\begin{pmatrix}\nabla h_i(x)\\ 0\end{pmatrix}+\sum_{j=1}^{p}\mu_j\begin{pmatrix}\nabla g_j(x)\\ 0\\ 2z_j\\ \vdots\end{pmatrix}=0$$

算Z

$$\Leftrightarrow \begin{cases}\nabla f(x)+\sum_{i=1}^{m}\lambda_i\nabla h_i(x)+\sum_{j=1}^{p}\mu_j\nabla g_j(x)=0\\ \mu_j z_j=0\end{cases}$$

互补松弛来源，也与互补松弛等价

若 $z_j=0 \Leftrightarrow j\in A(x)$，$\mu$ 无所谓

若 $z_j\neq 0 \Rightarrow \mu_j=0$　$j\notin A(x)$

二阶条件：

$$y^T\left(\begin{pmatrix}\nabla^2 f(x) & 0\\ 0 & 0\end{pmatrix}+\sum_{i=1}^{m}\lambda_i\begin{pmatrix}\nabla^2 h_i(x) & 0\\ 0 & 0\end{pmatrix}+\sum_{j=1}^{p}\mu_j\begin{pmatrix}\nabla^2 g_j(x) & 0\\ 0 & \ddots & 2\end{pmatrix}\right)y = 0$$

for All y: $\begin{pmatrix}\nabla h(x)\\ 0\end{pmatrix}^T y \geq 0$ ①

$$\forall j=1\ldots p \quad \begin{pmatrix}\nabla g(x)\\ \vdots\\ z_j\end{pmatrix} y \geq 0 \quad ②$$

$\to z_j = 0$ when $j \in A(x)$

取 $y = \begin{pmatrix}0\\ \vdots\\ \vdots\end{pmatrix}$ → 下一块第 j 个

带 y 进 (*) 有 $2\mu_j \geq 0 \Rightarrow \mu_j \geq 0$

取 $y = \begin{pmatrix}w\\ 0\end{pmatrix}$ $\quad w \in S = \left\{ w \,\middle|\, \begin{array}{l}\nabla h_i(x)^T w = 0 \quad i=1\ldots m\\ \nabla g_j(x)^T w = 0 \quad j \in A(x)\end{array}\right\}$

y 满足 ① ②

带入(*)

$$W^T\left(\nabla^2 f(x)+\sum_{i=1}^{m}\nabla^2 h_i(x)\cdot\lambda_i+\sum_{j=1}^{p}\mu_j\nabla^2 g_j(x)\right)W \geq 0$$

证明:

eg. $\min: -\sum_{i=1}^{n}\log(x_i+Q_i)$

s.t. $x\geq 0$ $\sum_{i=1}^{n}x_i=1$.

解: $L=-\sum_{i=1}^{n}\log(x_i+Q_i)+\lambda\left(\sum_{i=1}^{n}x_i-1\right)-\mu^T x$

一阶条件:

① $\nabla L(x^*,\lambda,\mu)=\begin{pmatrix}-\frac{1}{x_i+Q_i}+\lambda-\mu_i\\ \vdots\\ -\frac{1}{x_n+Q_n}+\lambda-\mu_n\end{pmatrix}=0$

② $\mu_i\cdot x_i=0$

$\therefore\begin{cases}\mu_i\cdot x_i=0\\ \lambda-\mu_i-\frac{1}{x_i+Q_i}=0\end{cases}$

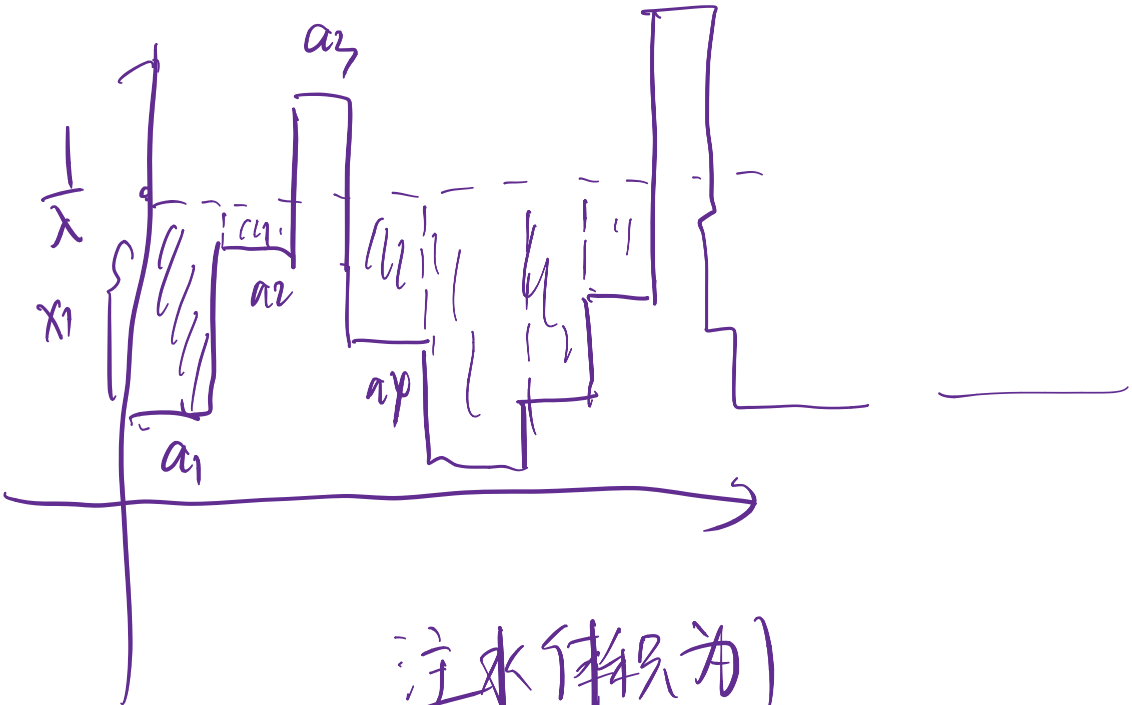
$\lambda(x_i+Q_i)-\mu_i(x_i+Q_i)=1$

$\Leftrightarrow\lambda=\frac{1+\mu_i Q_i}{x_i+Q_i}$ $\forall i$

$\mu_i\geq 0$

用互补松弛:

三种情况
$$\begin{cases} H_i > 0,\ x_i = 0 & \Rightarrow \lambda = H_i + \frac{1}{Q_i} > \frac{1}{Q_i} \Leftrightarrow Q_i > \frac{1}{\lambda} \\ H_i = 0 \quad x_i > 0 & \Rightarrow \lambda = \frac{1}{x_i + Q_i} < \frac{1}{Q_i} \Leftrightarrow Q_i < \frac{1}{\lambda} \\ H_i = 0 \quad x_i = 0 & \Rightarrow \lambda = \frac{1}{Q_i} \Leftrightarrow Q_i = \frac{1}{\lambda} \end{cases}$$



注水算法

KKT becomes sufficient when the problem is convex

f, g_j convex
h_i affine

KKT条件 不特指都是指一阶条件.

$$KKT \Leftrightarrow \left(\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*)\right)^T$$

$$\nabla f(x^*)^T (x - x^*) \leq f(x) - f(x^*)$$

$$\nabla h_i(x^*)(x - x^*) = 0$$

$$h_i(x) = a_i^T x + b_i$$

$$\nabla h_i(x) = a_i$$

$$\nabla h_i(x^*)^T (x - x^*) = a(x - x^*) = 0$$

$$\nabla g_j(x^*)^T (x - x^*) \leq g_j(x) - g_j(x^*) = g_j(x) \leq 0$$

Fritz John condition

- However, the following condition always holds: There are $\lambda_i,\ i=1,\ldots,m,$ and $\mu_j \geq 0,\ j=0,1,\ldots,r,$ such that

$$\mu_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{r} \mu_j \nabla g_j(x^*) = 0,$$

与KKT的差别

where $\mu_j g_j(x^*) = 0$ for $j=1,\ldots,r,$ and not all these multipliers are zero.

Other constraint qualifications

- There are several alternative conditions to guarantee the regularity of the constraints, thus ensuring the existence of Lagrangian multipliers.
- A famous constraint qualification is known as Mangasarian-Fromovitz constraint qualification:
- The vectors $\nabla h_i(x^*)$ $(i=1,\ldots,m)$ are linearly independent, and there exists d such that

$$(\nabla h_i(x^*))^{\mathrm{T}} d = 0,\ i=1,\ldots,m,$$

and

$$(\nabla g_j(x^*))^{\mathrm{T}} d < 0,\ \text{for all } j \in A(x^*).$$

- It is easy to see that with this condition, one must have $\mu_0 > 0$.

proof

上海财经大学
信息管理与工程学院
SCHOOL INSTITUTION OF MANAGEMENT & ENGINEERING

Suppose $\mu_0 = 0$

$$\left(\sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j=1}^{p} \mu_j \nabla g_j(x)\right)^T d = 0$$

$$\Leftrightarrow \underbrace{\sum_{i=1}^{m} \lambda_i \nabla h_i(x)^T d}_{=0} + \underbrace{\sum_{j=1}^{p} \mu_j \nabla g_j(x)^T d}_{>0} = 0$$

$\therefore \mu_j = 0$ for $j = 1 \ldots p$

与 μ_j 不都为0矛盾