运筹学通论I

胡晓东

应用数学研究所

中国科学院数学与系统科学研究院

Http://www.amt.ac.cn/member/huxiaodong/

3. 最优性条件 - 无约束

定理 20. 设函数 $f(x)$ 在点 x^* 处可微。若存在一个向量 d 满足 $\nabla f(x^*)d < 0$，则存在一个实数 $\delta > 0$ 使得，对任意$\lambda \in (0, \delta)$，有 $f(x^*+\lambda d) \le f(x^*)$，即 d 是函数 $f(x)$ 在点 x^* 处的一个下降方向。

证明. (练习) 因为函数 $f(x)$ 在点 x^* 处可微，所以有

$$f(x^*+\lambda d) = f(x^*) + \lambda \nabla f(x^*)\, d + \lambda\, \|d\|\, \alpha(x^*;\lambda d),$$

其中当$\lambda \to 0$ 时，$\alpha(x^*;\lambda d) \to 0$。等式两边同除$\lambda$，即可得证。

定理 21 (一阶必要条件) 设函数 $f(x)$ 在点 x^* 处可微。若 x^*是一个局部最小解，则 $\nabla f(x^*) = 0$。

证明. (练习) 若 $\nabla f(x^*) \neq 0$，则取 $d = -\nabla f(x^*)$。再由**定理 20**，即可得证。

3. 最优性条件 - 无约束（续一）

定理 22 (二阶必要条件) 设 $f(x)$ 在点 x^* 处二阶可微。若 x^* 是一个局部最小解，则 $\nabla f(x^*)=0$，且海森阵 $H(x^*)$ 是半正定的。

证明. (**练习***) 考虑任意一个方向向量 d。因为 $f(x)$ 在点x^* 处二阶可微，所以有

$$f(x^*+\lambda d) = f(x^*) + \lambda\nabla f(x^*)d + \lambda^2 d^T H(x^*)d/2 + \lambda^2\,\|d\|^2\alpha(x^*;\lambda d),$$

其中当$\lambda \to 0$ 时，$\alpha(x^*;\lambda d) \to 0$。再由 **定理 21** 可得$\nabla f(x^*)=0$。将上述等式两边同除 λ^2，可得

$$\frac{f(x^*+\lambda d) - f(x^*)}{\lambda^2} = d^T H(x^*)d/2 + \|d\|^2\alpha(x^*;\lambda d)。$$

又因为 x^* 是一个局部最小解，所以当λ充分小时，有

$$f(x^*+\lambda d) \geq f(x^*) 。 \quad \ldots\ldots$$

3. 最优性条件 - 无约束（续二）

定理 23 (充分条件) 设 $f(x)$ 在点x^* 处二阶可微。 若 $\nabla f(x^*)=0$ 且海森阵 $H(x^*)$ 是正定的，则 x^* 是一个局部最小解。

证明. (练习*) 对任意向量 x，都有

$$f(x) = f(x^*) + \nabla f(x^*) (x - x^*)$$
$$+ (x - x^*)^T H(x^*) (x - x^*)/2 + \|x - x^*\|^2 \alpha(x^*; x - x^*),$$

其中当$x \to x^*$ 时，$\alpha(x^*; x - x^*) \to 0$。

采用反证法。假设 x^* 不是一个局部最小解，则存在一个序列$\{x_k\} \to x^*$ 使得对每个 k 都有 $f(x_k) < f(x^*)$。现定义一个序列 $(x_k - x^*)/\| x_k - x^*\| = d_k$。则当 $k \to \infty$ 时，$d_k \to d$，其中$\|d\|=1$。又因为$\nabla f(x^*) = 0$，所以有 $d^T H(x^*) d \le 0$，与 $H(x^*)$ 是正定矩阵的假设相矛盾。

3. 最优性条件 - 有约束

定义 12. 设 S 是 E^n 上的一个非空闭集。则 S 在点 x^* 处的**可行方向锥** $D(x^*)$ 定义为

$$D(x^*) \equiv \{d \mid d \neq 0 \text{ 且} \exists\, \delta > 0\text{，使得 } \forall \lambda \in (0, \delta)\text{，} x^* + \lambda d \in S\}\text{。}$$

每一个向量 $d \in D(x^*)$ 都称为点 x^* 处的可行方向。

由**定义 5** 可知，在点 x^* 处沿着方向 $d \in D(x^*)$ 做一个微小移动，即可到达一个可行点（解）。另外，由**定理 20** 可知，若 $\nabla f(x^*) d < 0$，则 d 是点 x^* 处的一个改进方向。下述定理说明，若 x^* 是一个局部最小解，且 $\nabla f(x^*) d < 0$，则 $d \notin D(x^*)$。

3. 最优性条件 - 有约束(续一)

定理 24. (必要条件) 考虑求最小值问题 $\min\{f(x) \mid x\in S\}$，其中 S 是 E^n上的一个非空集合。设函数 $f(x)$ 在点 x^* 处可微。若 x^* 是一个局部最小解，则 $F(x^*) \cap D=\varnothing$，其中$D$ 是 S 在点 x^* 处的可行方向锥，$F(x^*)=\{d \mid \nabla f(x^*)d < 0\}$。

证明 (练习*) 采用反证法。假设存在一个向量 $d\in F(x^*) \cap D$。则由 **定理 20** 和局部最小解的定义，可得如下矛盾：

存在某个 $\delta_1 > 0$, 使得对任意 $\lambda\in(0, \delta_1)$ 都有 $f(x^*+\lambda d) \le f(x^*)$

而且

存在某个 $\delta_2 > 0$, 使得对任意 $\lambda\in(0, \delta_2)$ 都有 $x^* + \lambda d\in S$。与 x^* 是一个局部最小解的假设矛盾。

3. 最优性条件 - 不等式约束

定理 25. (必要条件) 考虑求最小值问题 $\min\{ f(x) | g_i(x) \le 0, i=1,2,\ldots, m$ 且 $x \in S \}$，其中 S 是 E^n 上的一个非空开集合。设 x^* 是一个可行解，并令 $I=\{i \mid g_i(x^*)=0\}$。另设函数 $f(x)$ 和 $g_i(x)$，$i \in I$，在点 x^* 处都是可微的，且 $g_i(x)$，$i \notin I$，在点 x^* 处是连续的。若 x^* 是一个局部最小解，则 $F(x^*) \cap G(x^*)=\varnothing$，其中 $F(x^*) =\{d \mid \nabla f(x^*)d < 0 \}$，$G(x^*)=\{ d \mid \nabla g_i(x^*)d < 0$ 且 $i \in I \}$。

证明. (练习*) 设 $d \in G(x^*)$。则存在一个实数 $\delta_1 > 0$ 使得对任意 $\lambda \in (0, \delta_1)$ 都有 $x^* + \lambda d \in S$。此外，还存在一个实数 $\delta_2 > 0$，使得对任意 $\lambda \in (0, \delta_2)$ 且 $i \notin I$，都有 $g_i(x^*+\lambda d) < 0$。因而，存在一个实数 $\delta_3 > 0$，使得对任意 $\lambda \in (0, \delta_3)$ 且 $i \in I$ 都有 $g_i(x^*+\lambda d) < g_i(x^*)$。因此 $d \in D$，其中 D 是点 x^* 处的可行方向锥。 由此可得 $G(x^*) \subset D$。再由**定理 24** 即得 ……

3. 最优性条件 - 不等式约束(续一)

Min $(x-3)^2 + (y-2)^2$

s.t. $x^2 + y^2 \le 5$

$x + y \le 3$

$x,\ y \ge 0.$

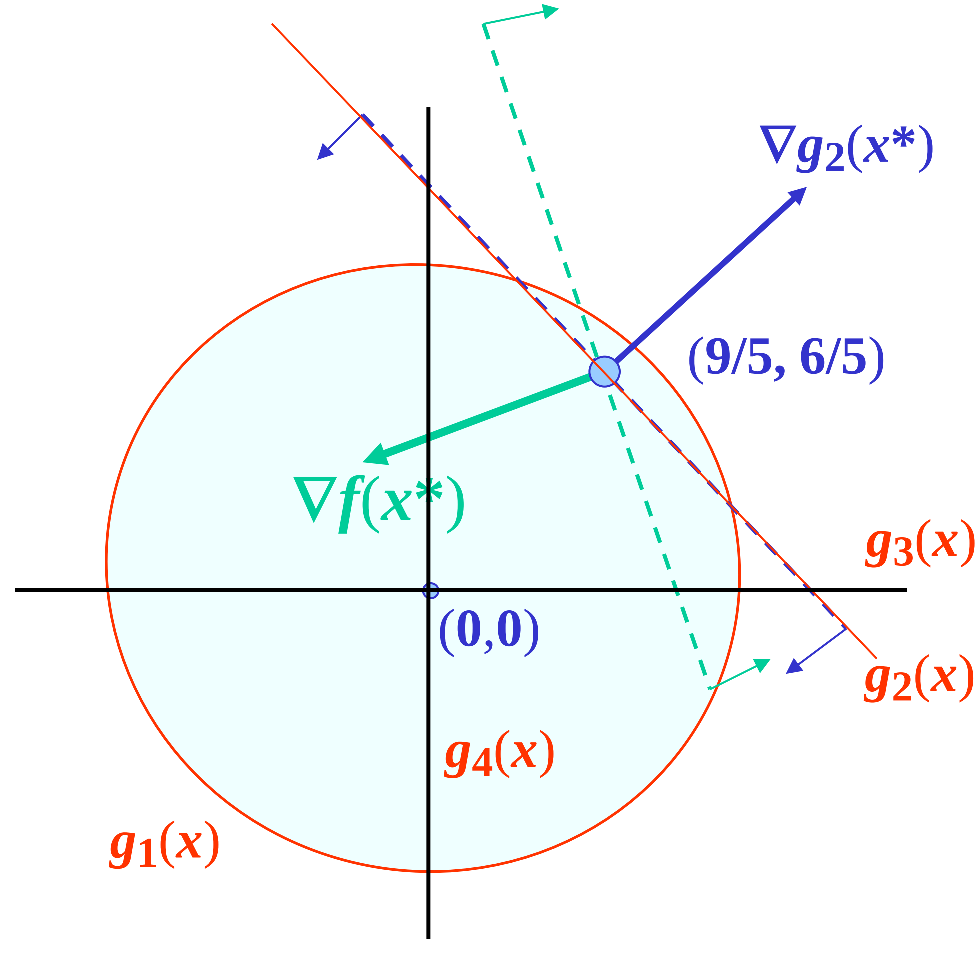
$g_1(x) = x^2 + y^2 - 5,$

$g_2(x) = x + y - 3,$

$g_3(x) = -x,$

$g_4(x) = -y;$

$\nabla f(x) = [2(x-3), 2(y-2)]$

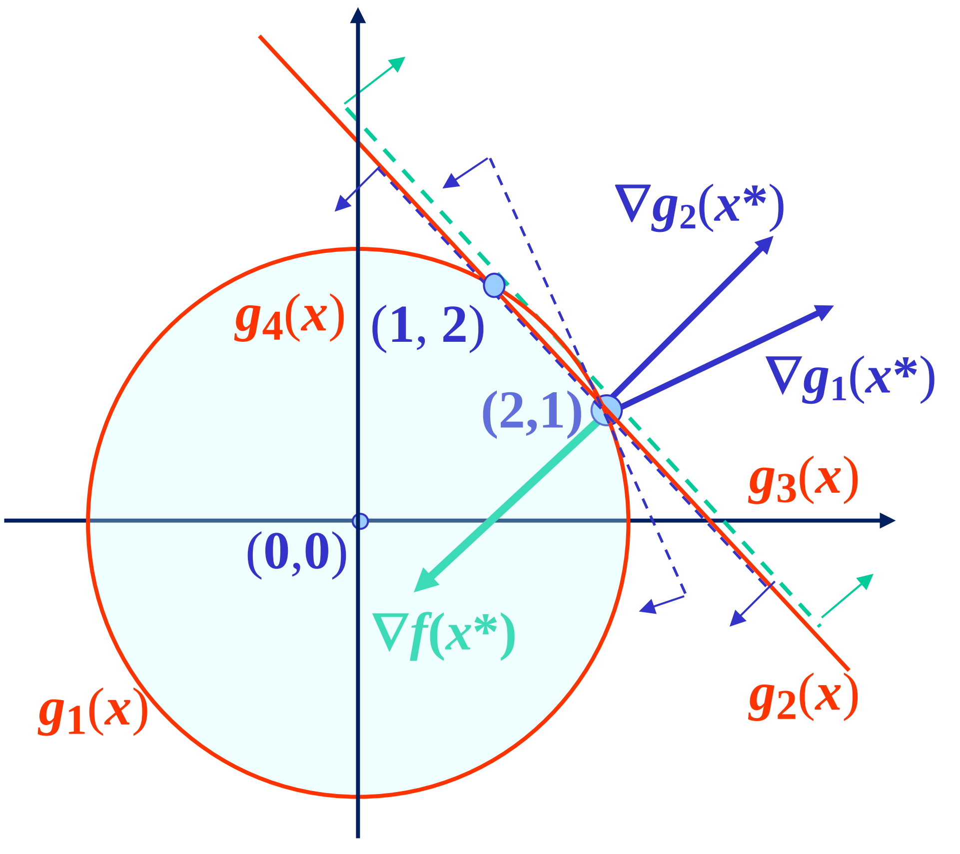


在非最优解处

$F(x^*) \cap G(x^*) \neq \varnothing$

3. 最优性条件 - 不等式约束(续二)

$$\text{Min}\quad (x-3)^2 + (y-2)^2$$
$$\text{s.t.}\quad x^2 + y^2 \le 5$$
$$x + y \le 3$$
$$x,\ y \ge 0.$$



$x^*=(2,1)$

$\nabla g_1(x^*)=(4,\ 2)$,

$\nabla g_2(x^*)=(1,\ 1)$,

$\nabla f(x^*)=(-2,-2)$.

x^* 是一个最优解

3. 最优性条件 - 不等式约束(续三)

定理 26. (**Fritz John 条件**, 1948) 考虑求最小值问题

$$\min\{\, f(x) \mid g_i(x) \le 0,\ i=1,2,\ldots,m,\ \text{且}\ x\in S\},$$

其中S 是 E^n上的一个非空凸集。 设 x^* 是一个可行解，且令 $I=\{\, i \mid g_i(x^*)=0\}$。 此外，假定 $f(x)$ 和 $g_i(x)$，$i\in I$，在点 x^* 处可微，且 $g_i(x)$，$i\notin I$，在 x^* 处连续。若 x^* 是一个局部最小解，则存在实数 u_0 和 u_i，$i\in I$，使得

$$u_0\nabla f(x^*) + \sum_{i\in I} u_i\nabla g_i(x^*)=0, \qquad u_0,\ u_i \ge 0,\ i\in I\ \text{且}\ (u_0, u_I)\neq 0。$$

此外，若 $g_i(x)$，$i\notin I$，也在点 x^* 处可微，则

$$u_0\nabla f(x^*) + \sum_{i=1}^{i=m} u_i\nabla g_i(x^*)=0,$$ 且对每一个指标 i 都有

$$u_i\, g_i(x^*)=0, \quad (u_0, u)\neq 0。$$

3. 最优性条件 - 不等式约束(续四)

证明. 由**定理 25** 可知，不存在一个向量 d 使得 $\nabla f(x^*)d < 0$ 且对每一个 $i \in I$，都有 $\nabla g_i(x^*)d < 0$。现令 A 为 $\nabla f(x^*)$ 和 $\nabla g_i(x^*)$，$i \in I$，作为行向量的矩阵。因而 $Ad < 0$ 没有解。再由**Gordan 定理**，可知存在一个非零向量 $p \geq 0$ 使得 $A^T p = 0$。再将向量 p 的分量分别记为 u_0 和 u_i，$i \in I$，即可得证定理的第一部分。

定理的第二部分也可证得，只要令 $u_i = 0$ 当 $i \notin I$ 时。

练习. 证明**Gordan 定理**

设 A 是一个 $m \times n$ 的矩阵，则以下两个线性系统恰有一个有解

系统1 $Ax < 0$， x 是一个 n 维向量；

系统2 $A^T y = 0$ 且 $y \geq 0$， y 是一个非0的 m 维向量。

提示： 若**系统1**无解，则 $S_1 \cap S_2 = \varnothing$，其中 $S_1 = \{z = Ax \mid x \in E^n\}$，$S_2 = \{z \in E^m \mid z < 0\}$。

3. 最优性条件 - 不等式约束(续五)

在**Fritz John 条件**中，实数 u_0 和每一个u_i 通常称为**拉格朗日乘子**。每一个条件 $u_i g_i(x^*) = 0$ 称为**互补松弛条件**。如果相应的不等式不是严格的，$g_i(x) < 0$，那么一定有$u_i = 0$。同样地，只有当相应的不等式是严格的时候，才可能有 $u_i > 0$。

如果**拉格朗日乘子** u_0 等于0，那么 **Fritz John 条件**就与目标函数 $f(x)$ 的梯度无关。这个条件仅仅说明了，存在所有严格约束条件函数的梯度的非负且非平凡的线性和，其值为**0**。因此，这个条件对于寻找问题的最优解没有实际帮助。

人们提出了许多条件，以确保**拉格朗日乘子** $u_0 > 0$, 它们通常称为**约束规格**。特别地，**Karush** [1939，硕士论文，芝加哥大学数学系]，**Kuhn** 和 **Tucker** [1951，普林斯顿大学] 独立的给出了最优解的必要条件，它们实际上就是 **Fritz John 条件**再加上性质 $u_0 > 0$。

3. 最优性条件 - 不等式约束(续六)

定理 27. (**K-K-T 必要条件**) 考虑求最小值问题 $\min\{f(x) \mid g_i(x) \le 0,\ i=1,2,\ldots,m$ 且 $x \in S\}$，其中 S 是 E^n 上的一个非空开集。设 x^* 是一个可行解，另令 $I = \{i \mid g_i(x^*)=0\}$。此外，设函数 $f(x)$ 和 $g_i(x)$，$i \in I$，在点 x^* 处是可微的，且 $g_i(x)$，$i \notin I$，在点 x^* 处是连续的，∇g_i，$i \in I$，是线性无关的。若 x^* 是一个局部最小解，则存在实数 u_i，$i \in I$，使得

$$\nabla f(x^*) + \sum_{i \in I} u_i \nabla g_i(x^*) = 0, \text{ 且 } u_i \ge 0,\quad i \in I。$$

此外，若函数 $g_i(x)$，$i \notin I$，在点 x^*处也是可微的，则

$$\nabla f(x^*) + \sum_{i=1}^{i=m} u_i \nabla g_i(x^*) = 0,$$

$$u_i g_i(x^*) = 0,\ u_i \ge 0,\quad i=1,\ldots,m。$$

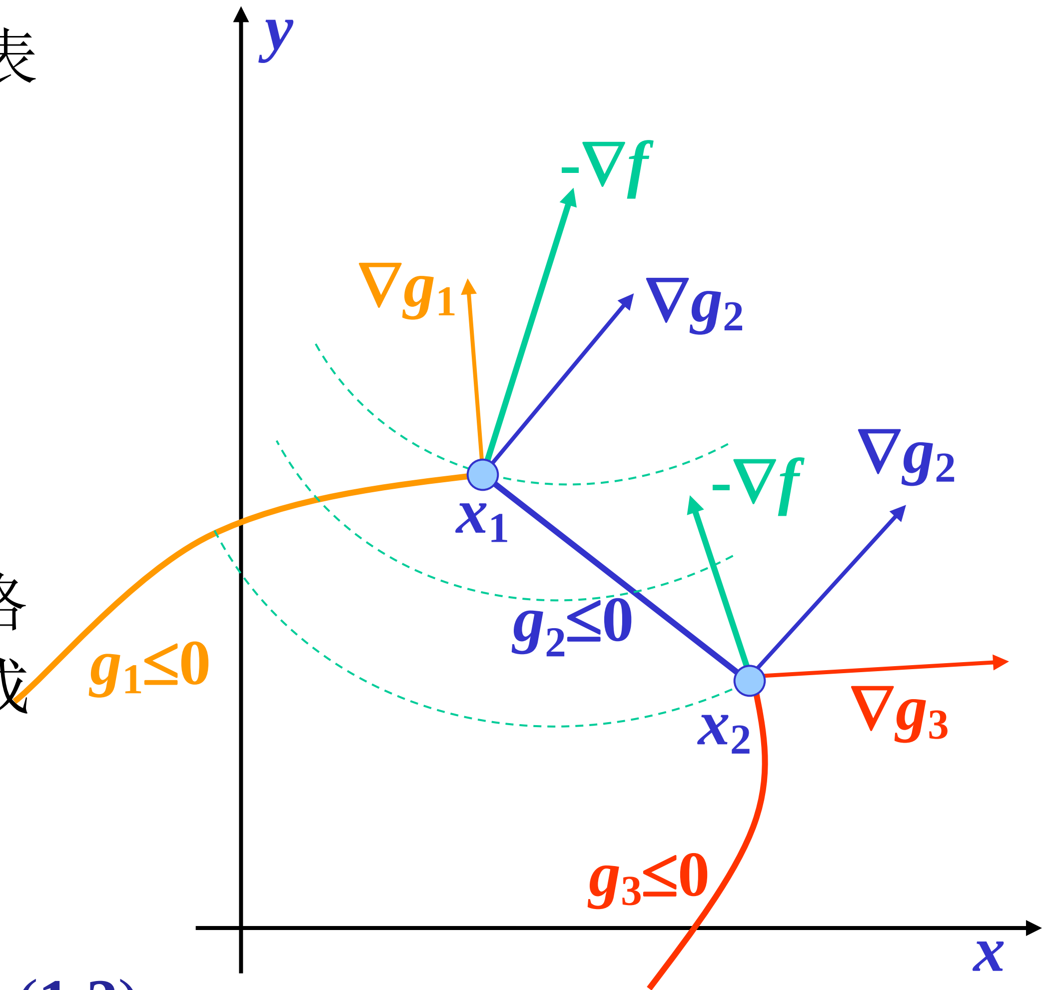
证明 (练习) 可由**定理 26** 直接推出。

3. 最优性条件 - 不等式约束(续七)

K-K-T 条件可以借助向量表述如下：

$$\nabla f(x^*) + u^T \nabla g(x^*) = 0,$$
$$u^T g(x^*) = 0,$$
$$u_i \geq 0。$$

亦即，$-\nabla f(x^*)$ 属于取严格约束相应函数的梯度所张成的锥。



练习题：前例中，**(2,1)** 和 **(1,2)** 是否都满足**K-K-T** 条件?

3. 最优性条件 - 不等式约束(续八)

定理 28. (**K-K-T 充分条件**) 考虑求最小值问题 $\min\{ f(x) \mid g_i(x) \le 0,\ i=1,\ldots,m$，且$x\in S\}$，其中 S 是E^n上的一个非空开集。设 x^* 是一个可行解，另令 $I=\{i \mid g_i(x^*)=0\}$。此外，假设 $f(x)$ 是伪凸函数，而 $g_i(x)$，$i\in I$，是拟凸函数且在点 x^* 处可微。若 **K-K-T** 条件在点 x^* 处成立，则 x^* 是一个全局最优解。

证明. (**练习****) 由函数 g_i 的拟凸性可知，对每个 $i\in I$ 都有

$$g_i(x^*+\lambda(x-x^*)) = g_i(\lambda x + (1-\lambda)x^*)$$
$$\le \max\{ g_i(x), g_i(x^*)\} = g_i(x^*)。$$

因而根据**定理19**，可得 $\nabla g_i(x^*)(x-x^*) \le 0$。将不等式乘以 u_i 并取遍 I 中的所有指标 i 相加，即得 $\sum u_i \nabla g_i(x^*)(x-x^*) \le 0$。再根据 K-K-T 条件，可得 $\nabla f(x^*)(x-x^*) \ge 0$。最后根据函数 $f(x)$ 在点 x^* 处是伪凸的(**定义 10**)，即得 $f(x^*) \le f(x)$。

定理29. (K-K-T 必要条件)

考虑求最小值问题 $\min\{ f(x) \mid g_i(x) \le 0,\ i=1,\ldots,m;\ h_j(x)=0,\ j=1,2,\ldots,k;\ x\in S\}$，其中 S 是一个开集合。设 x^* 是一个可行解，$I=\{i \mid g_i(x^*)=0\}$。假设目标函数 $f(x)$ 和约束函数 g_i，$i\in I$ 在点 x^*处可微，约束函数 g_i，$i\notin I$，在点 x^* 处连续，h_j 在点 x^* 处连续可微。再假设 ∇g_i，$i\in I$，和 ∇h_j 是线性无关的。若 x^* 是一个局部最小解，则存在实数 u_i，$i\in I$ 和 v_j，$j=1,2,\ldots,k$ 使得

$$\nabla f(x^*) + \sum_{i\in I} u_i \nabla g_i(x^*) + \sum_{j} v_j \nabla h_j(x^*) = 0,\ u_i \ge 0,\ i\in I。$$

若函数 g_i，$i\notin I$，也在点 x^* 处可微，则

$$\nabla f(x^*) + \sum_{i=1}^{i=m} u_i \nabla g_i(x^*) + \sum_{j=1}^{j=k} v_j \nabla h_j(x^*) = 0,$$

$$u_i\, g_i(x^*) = 0,\quad u_i \ge 0,\quad i=1,\ldots,m.$$

3. 最优性条件 - 等式与不等式约束(续一)

定理 30. (**K-K-T 充分条件**) 考虑求最小值问题

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } g_i(x) \le 0,\ i = 1,2, \ldots, m, \\ &\qquad h_j(x) = 0,\ i = 1,2, \ldots, k, \\ &\qquad\qquad x \in S, \end{aligned}$$

其中 S 是 E^n 上一个非空开集。

设 x^* 是一个可行解，$I = \{i \mid g_i(x^*)=0\}$。假设在点 x^* 处满足 **K-K-T 条件**。另令 $J=\{ j \mid v_j > 0\}$，$L=\{l \mid v_l <0\}$。若在点 x^* 处

$f(x)$ 是伪凸的,

g_i，$i \in I$，是拟凸的，

$h_j(x)$，$j \in J$，是拟凸的，且 $h(x)$，$l \in L$，是拟凹的，

则 x^* 是一个全局最优解。

3. 对偶问题

原始问题

$$\begin{aligned} &\text{Min} \quad f(x) \\ &\text{s.t.} \quad g_i(x) \le 0, \quad i = 1,2,\ldots,m; \\ &\qquad\ \ h_j(x) = 0, \quad j = 1,2,\ldots,k, \\ &\qquad\qquad x \in S. \end{aligned}$$

拉格朗日对偶问题

$$\begin{aligned} &\text{Max} \ \theta(u,v) = \inf\left\{ f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{j=1}^{k} v_j h_j(x) \mid x \in S \right\} \\ &\text{s.t.} \qquad\quad u \ge 0 \end{aligned}$$

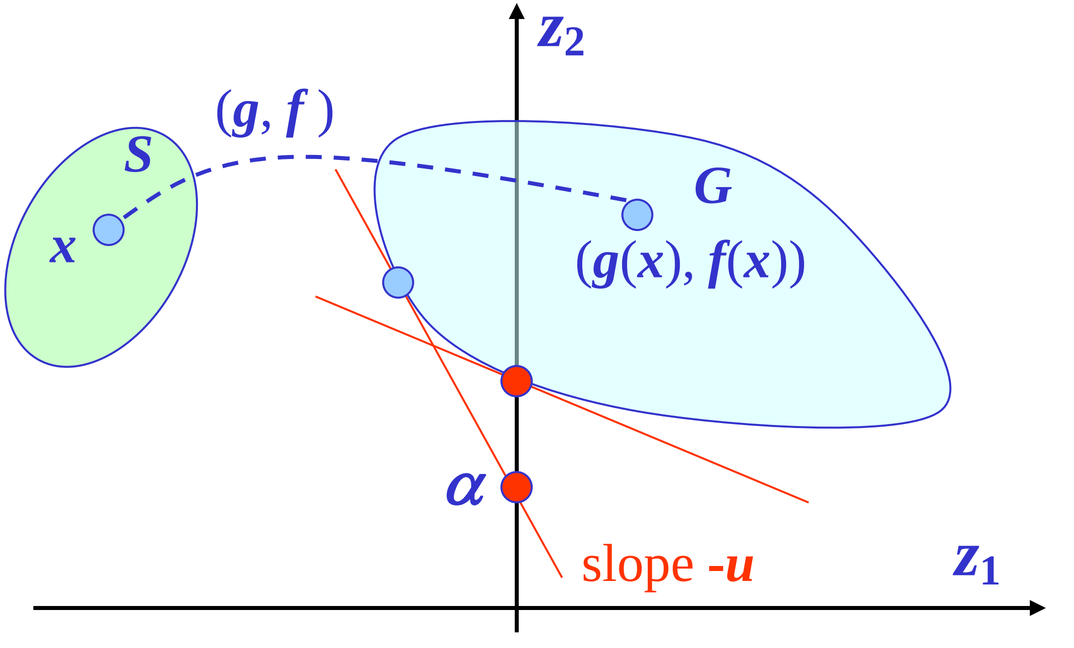
注意，在求最大值问题的目标函数 $\theta(u, v)$ 的表达式中，借助拉格朗日乘子 u_i 和 v_j，不等式和等式约束条件都融入到了目标函数中了。

思考题： 由拉格朗日对偶问题是否可推出线性规划对偶问题？

3. 对偶问题（续一）

为了给出对偶问题直观的几何上的解释，考虑仅带有不等式约束的原始问题。在(z_1, z_2)-平面上，我们用G 表示集合$\{(z_1, z_2) \mid z_1 = g(x), z_2 = f(x), \ x \in S\}$。**原始问题**是求$G$ 中位于z_2-轴左侧的坐标最小的一个点，而**对偶问题**是求支撑平面的一个斜率使其与 z_2-轴的交点坐标最大。

注意，$z_2 + u\, z_1 = \alpha$ 是一个直线方程，其斜率为 $-u$ 且在 z_2-轴上的截距为α 。



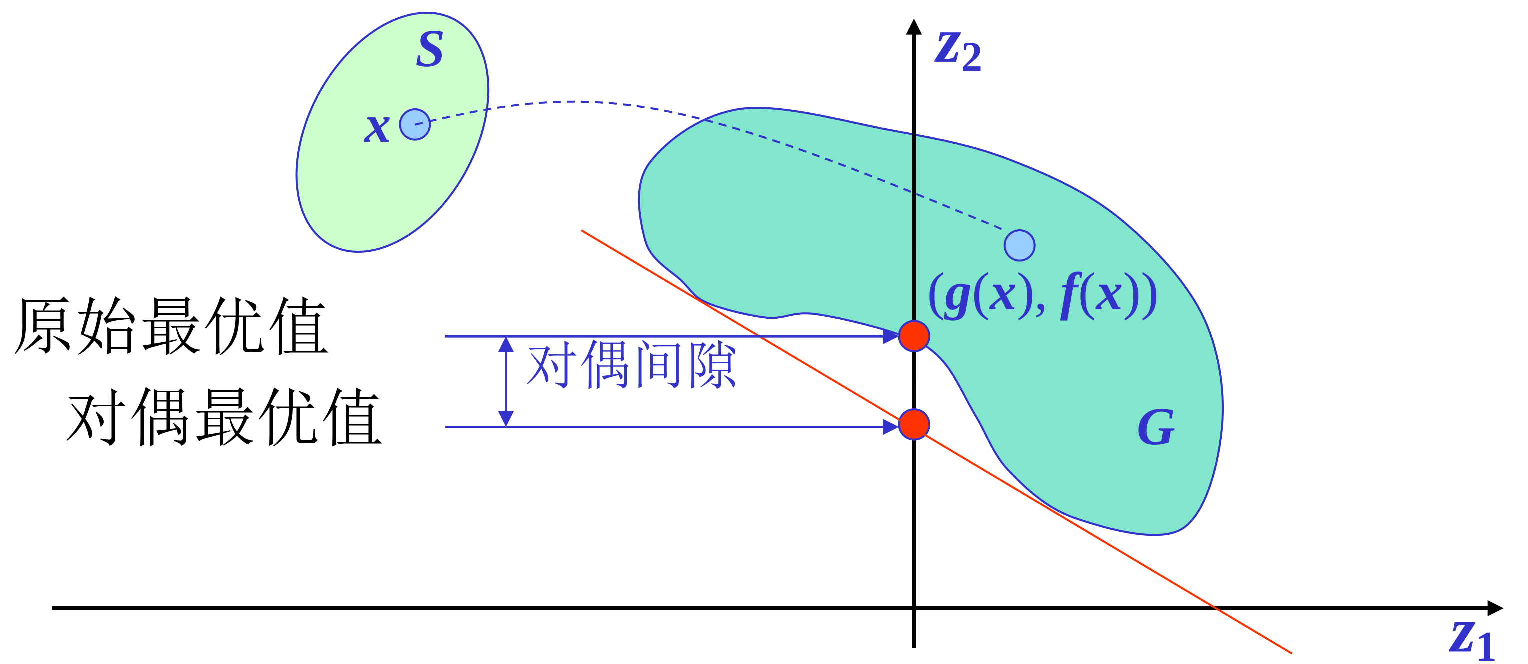
3. 对偶定理

定理 31 (弱对偶定理) 设 x 是原始问题的一个可行解，(u, v) 是对偶问题的一个可行解。则有 $f(x) \geq \theta(u, v)$。

证明 (练习). 由 $\theta(u,v)$ 的定义，可得 $u \geq 0$, $g_i(x) \leq 0$, $h_j(x) = 0$，

则有 $\theta(u,v) = \inf\{ f(x) + \Sigma\, u_i g_i(x) + \Sigma\, v_j h_j(x) \mid x \in S \}$

$$\leq f(x) + \Sigma\, u_i g_i(x) + \Sigma v_j h_j(x)$$



3. 对偶定理（续一）

推论 1 $\inf\{f(x)|\ g_i(x)\le 0,\ h_j(x)=0,\ x\in S\} \ge \sup\{\theta(u, v)\ |\ u \ge 0\}$。

推论 2 若 $f(x^*)\le\theta(u^*, v^*)$, 其中 $u\ge 0$， $g_i(x^*)\le 0,\ h_j(x^*)=0,\ x^*\in S$, 则 x^* 和 (u^*, v^*) 分别是原始问题和对偶问题的最优解。

推论 3 若 $\inf\{\ f(x)\ |\ g_i(x) \le 0,\ h_j(x) = 0,\ x\in S\ \}= -\infty$， 则当$u \ge 0$时， 有 $\theta(u, v) = -\infty$。

推论 4 若 $\sup\{\theta(u, v)\ |\ u \ge 0\} = \infty$， 则原始问题不存在可行解。

我们用向量函数表示法，$g(x): E^n\rightarrow E^m$， $h(x): E^n \rightarrow E^k$, 其中向量的第 i 个分量分别是 $g_i(x)$ 和 $h_i(x)$。

3. 对偶定理（续二）

引理 6 设 S 是 E^n 上的一个非空凸集， $\alpha(x)$ 和 $g(x)$ 是两个凸函数，令 $h(x)=Ax-b$。若下述系统 1 无解 x，则下述系统 2 有解 (u_0, u, v)。反之亦成立只要 $u_0>0$。

系统 1: $\alpha(x)<0,\ g(x)\le 0,\ h(x)=0,$ 存在某个 $x\in S$.

系统 2: $u_0\alpha(x)+u\,g(x)+v\,h(x)\ge 0,$ 对于所有 $x\in S$,

$$(u_0, u)\ge 0,\ (u_0, u, v)\ne 0$$

证明. (练习)** 假定**系统1** 不存在解。我们考虑集合

$$\Psi\equiv\{(p, q, r)\mid p>\alpha(x),\ q\ge g(x),\ r=h(x),\ \text{存在某个 } x\in S\}。$$

注意， Ψ 是一个凸集且 $0\notin\Psi$。 由**分离定理**可知，存在一个非0向量(u_0, u, v) 使得对每个 $(p, q, r)\in \mathrm{cl}\,\Psi$，都有$u_0p+uq+vr\ge 0$。又因 p 和 q 可取得任意地大，故有 $u_0\ge 0$， $u\ge 0$。再注意到，对于每个 $x\in S$， 都有 $(p, q, r)=(\alpha(x), g(x), h(x))\in \mathrm{cl}\,\Psi$ 。因而**系统2** 有解。

3. 对偶定理（续三）

证明 (续) 证明逆命题。假设**系统2**有解 (u_0, u, v)，其中$u_0 \geq 0$，$u \geq 0$，且对每个 $x \in S$，都满足$u_0\alpha(x) + ug(x) + vh(x) \geq 0$。现设存在某个 $x \in S$ 满足 $g(x) \leq 0, h(x)=0$。由上述不等式，可得$u_0\alpha(x) \geq 0$。而因为假设 $u_0 > 0$，$\alpha(x) \geq 0$，所以**系统1** 无解。

定理 32 (强对偶定理)

设 S 是E^n上的一个非空凸集合，$f(x)$ 和 $g(x)$ 都是凸函数，令$h(x) = Ax - b$。假设存在某个 $x^* \in S$ 满足 $g(x^*) < 0, h(x^*)=0$，且$0 \in \text{int}\, H(S)$，其中$H(S)=\{h(x) \mid x \in S\}$。则

$$\inf\{ f(x) \mid x \in S, g(x) \leq 0, h(x) = 0 \} = \sup\{ \theta(u, v) \mid u \geq 0 \}。$$

而且，若下确界 inf 是有限的，则$\sup\{\theta(u,v) \mid u \geq 0\}$ 在 (u^*, v^*) 处达到，其中$u^* \geq 0$。若 inf 在点 x^* 处达到，则 $u^*g(x^*)=0$。

3. 对偶定理（续四）

证明. (练习**) 令 $\gamma \equiv \inf\{ f(x) \mid x\in S, g(x) \le 0, h(x)=0 \}$。若 $\gamma=-\infty$，则 $\sup\{\theta(u, v) \mid u\ge 0\} = -\infty$。当 γ 是有限时，考虑系统 $f(x)-\gamma<0$, $g(x) \le 0$, $h(x) = 0$, $x\in S$。若它不存在解，则存在一个非 0 向量 (u_0, u, v)，其中 $(u_0, u) \ge 0$ 且对所有 $x\in S$，都满足

$$u_0 (f(x) - \gamma) + ug(x) + vh(x) \ge 0。$$

我们首先证明 $u_0 > 0$。应用反证法，假设 $u_0 = 0$。则有 $ug(x^*) + vh(x^*) \ge 0$，由此可得 $u=0$。因而对于所有 $x\in S$ 都有 $vh(x) \ge 0$。然而，$0\in \operatorname{int} H(S)$，故可选取 $x\in S$ 使得 $h(x)= -\lambda v$ 且 $\lambda > 0$。因此，$vh(x) = -\lambda \|v\|^2 \ge 0$，产生矛盾。又因为 $u_0 > 0$，所以可假设对所有 $x\in S$，都有 $f(x) + u^*g(x) + v^*h(x) \ge \gamma$。再由**定理31**，即得 $\theta(u^*, v^*) =\gamma$。

若下确界 inf 在点 x^* 处达到，则 $f(x^*)=\gamma$, $x^*\in S$, $g(x^*) \le 0$, $h(x^*)=0$。因而有 $ug(x^*)=0$。

3. 鞍点定理

定理 33 (鞍点定理)
设 S 是 E^n 上的一个非空凸集，$f(x): E^n \to E^1$，$g(x): E^n \to E^m$，$h(x): E^n \to E^k$。假定存在一个 $x^* \in S$ 和 (u^*, v^*)，其中$u^* \geq 0$ 且

$$\phi(x^*, u, v) \leq \phi(x^*, u^*, v^*) \leq \phi(x, u^*, v^*) \ \forall\, x \in S,\quad (u, v),\quad u \geq 0,$$

其中$\phi(x, u, v) \equiv f(x) + ug(x) + vh(x)$。则 x^* 和 (u^*, v^*) 分别是原始问题和对偶问题的最优解。

反之，设 S，$f(x)$和 $g(x)$ 都是凸的，且 $h(x)$ 是一个仿射函数。另设 $0 \in \text{int}\, H(S)$。若 x^* 是原始问题的一个最优解，则存在一个 (u^*, v^*)，其中$u^* \geq 0$ ，且满足上述不等式。

3. 鞍点定理（续一）

证明 (练习)** 因为对所有 $u \ge 0$ 和所有 $v \in E^k$，都有

$$f(x^*)+ug(x^*)+vh(x^*)=\phi(x^*,u,v)\le\phi(x^*,u^*,v^*)= f(x^*)+u^*g(x^*)+v^*h(x^*)$$

所以有 $g(x^*) \le 0$, $h(x^*)=0$， x^* 是原始问题的一个可行解。在上述不等式中，令 $u = 0$， $v = v^*$， 即得 $u^*g(x^*)\ge 0$。又因$u^*\ge 0$ 且 $g(x^*) \le 0$， 故有 $u^*g(x^*) = 0$。注意

$$\begin{aligned} f(x^*) &= f(x^*) + u^*g(x^*) + v^*h(x^*) \\ &= \phi(x^*, u^*, v^*) \\ &\le \phi(x,\ u^*, v^*) \quad \text{(由定理的假设)} \\ &= f(x) + u^*g(x) + v^*h(x). \end{aligned}$$

因此有 $f(x^*) \le \theta(u, v) = \inf\{ f(x) + ug(x) + vh(x) \mid x \in S \}$， 这是因为 对所有 $x \in S$， 上述不等式都成立。再由**推论2**可知， x^* 和 $u^* \ge 0$ 分别是原始问题和对偶问题的最优解。

3. 鞍点定理（续二）

证明 (续) 反之，设 x^* 是原始问题的一个最优解。则又定理 **32** 可知，存在一个 (u^*,v^*)，其中$u^*\geq 0$，满足 $f(x^*) = \theta(u^*,v^*)$ 和 $u^*g(x^*) = 0$。再由函数 $\theta(u, v)$ 的定义可得

$$
\begin{aligned}
f(x^*) &= f(x^*) + u^*g(x^*) + v^*h(x^*) = \phi(x^*, u^*,v^*) \\
&= \theta(u^*,v^*) = \inf \{ f(x) + u^*g(x) + v^*h(x) \mid x\in S \} \\
&\leq f(x) + u^*g(x) + v^*h(x)，\text{对任意 } x\in S \\
&= \phi(x, u^*,v^*),
\end{aligned}
$$

即得第二个不等式。

第一个不等式成立是因为 $u \geq 0$ 和 $g(x^*) \leq 0$，有

$\phi(x^*, u, v) = f(x^*) + ug(x^*) + vh(x^*) \leq f(x^*) = \phi(x^*, u^*,v^*)$。

在某些凸性的假设下，**K-K-T** 条件中的拉格朗日乘子同样也是鞍点判定条件中的系数。