切锥

定义 (切锥)

给定可行域$\mathcal{X}$ 及$x \in \mathcal{X}$, 若存在序列$\{z_k\}_{k=1}^{\infty} \subset \mathcal{X}$, $\lim_{k\to\infty} z_k = x$ 以及正标量序列$\{t_k\}_{k=1}^{\infty}$, $t_k \to 0$ 满足

$$\lim_{k\to\infty} \frac{z_k - x}{t_k} = d$$

则称向量d 为$\mathcal{X}$ 在点x 处的一个切向量. 所有点x 处的切向量构成的集合称为切锥, 用$T_{\mathcal{X}}(x)$ 表示.

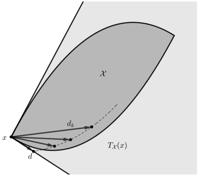

Figure: 不等式约束

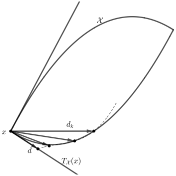

Figure: 等式约束

一般优化问题:

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) \\
\text{s.t.} \quad & c_i(x) \le 0,\ i \in \mathcal{I} \\
& c_i(x) = 0,\ i \in \mathcal{E}
\end{aligned}
$$

定理 (几何最优性条件)

假设可行点x^* 是上述问题的一个局部极小点.如果$f(x)$和$c_i(x)$, $i \in \mathcal{I} \cup \mathcal{E}$ 在点x^* 处是可微的, 那么

$$d^{\mathrm{T}} \nabla f(x^*) \ge 0, \quad \forall d \in T_{\mathcal{X}}(x^*)$$

等价于

$$T_{\mathcal{X}}(x^*) \cap \{d \mid \nabla f(x^*)^{\mathrm{T}} d < 0\} = \emptyset.$$

证明

- 若$T_{\mathcal{X}}(x^*) \cap \{d \mid \nabla f(x^*)^{\mathrm{T}} d < 0\} \neq \emptyset$, 取$d \in T_{\mathcal{X}}(x^*)$ 且$f(x^*)^{\mathrm{T}} d < 0$.
- 存在$\{t_k\}_{k=1}^{\infty}$ 和$\{d_k\}_{k=1}^{\infty}$ 使得$x^* + t_k d_k \in \mathcal{X}$, 其中$t_k \to 0$ 且$d_k \to d$.
- 由于$\nabla f(x^*)^{\mathrm{T}} d < 0$, 对于充分大的$k$, 我们有

$$
\begin{aligned}
f(x^* + t_k d_k) &= f(x^*) + t_k \nabla f(x^*)^{\mathrm{T}} d_k + o(t_k) \\
&= f(x^*) + t_k \nabla f(x^*)^{\mathrm{T}} d + t_k \nabla f(x^*)^{\mathrm{T}} (d_k - d) + o(t_k) \\
&= f(x^*) + t_k \nabla f(x^*)^{\mathrm{T}} d + o(t_k) \\
&< f(x^*)
\end{aligned}
$$

这与x^* 的局部极小性矛盾.

线性化可行锥

定义 (线性化可行锥)

对于可行点$x \in \mathcal{X}$, 定义该点的积极集$\mathcal{A}(x) = \mathcal{E} \cup \{i \in \mathcal{I} \ : \ c_i(x) = 0\}$,
点x 处的**线性化可行方向锥**定义为

$$\mathcal{F}(x) = \left\{ d \;\middle|\; \begin{array}{l} d^{\mathrm{T}} \nabla c_i(x) = 0, \ \forall\, i \in \mathcal{E}, \\ d^{\mathrm{T}} \nabla c_i(x) \le 0, \ \forall\, i \in \mathcal{A}(x) \cap \mathcal{I} \end{array} \right\}$$

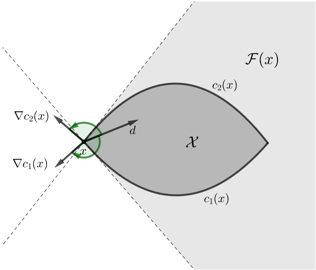

Figure: $\mathbb{R}^2$ 上的约束集合和线性化可行方向锥

定理 (线性化可行锥包含切锥)

设$c_i(x), i \in \mathcal{E} \cup \mathcal{I}$ 一阶连续可微, 则对任意可行点x 有$T_{\mathcal{X}}(x) \subseteq \mathcal{F}(x)$.

证明:不妨设积极集$\mathcal{A}(x) = \mathcal{E} \cup \mathcal{I}$, 设$d \in T_{\mathcal{X}}(x)$, 由定义

$$\lim_{k\to\infty} t_k = 0, \quad \lim_{k\to\infty} \frac{z_k - x}{t_k} = d \Leftrightarrow z_k = x + t_k d + e_k$$

其中残量e_k 满足$\|e_k\| = o(t_k)$.对$i \in \mathcal{E}$, 根据泰勒展开,

$$\begin{aligned}
0 &= \frac{1}{t_k} c_i(z_k) \\
&= \frac{1}{t_k}\left(c_i(x) + \nabla c_i(x)^{\mathrm{T}}(t_k d + e_k) + o(t_k)\right) \\
&= \nabla c_i(x)^{\mathrm{T}} d + \frac{\nabla c_i(x)^{\mathrm{T}} e_k}{t_k} + o(1).
\end{aligned}$$

注意到$\frac{\|e_k\|}{t_k} \to 0$, 令$k \to \infty$ 即可得到

$$\nabla c_i(x)^{\mathrm{T}} d = 0, \quad i \in \mathcal{E}.$$

线性化可行锥包含切锥

同理，对$i \in \mathcal{I}$，根据泰勒展开，

$$\begin{aligned} 0 &\geq \frac{1}{t_k} c_i(z_k) \\ &= \frac{1}{t_k}\left(c_i(x) + \nabla c_i(x)^{\mathrm{T}}(t_k d + e_k) + o(t_k)\right) \\ &= \nabla c_i(x)^{\mathrm{T}} d + \frac{\nabla c_i(x)^{\mathrm{T}} e_k}{t_k} + o(1). \end{aligned}$$

注意到$c_i(x) = 0,\ i \in \mathcal{I}$，因此我们有

$$\nabla c_i(x)^{\mathrm{T}} d \leq 0, \quad i \in \mathcal{I}$$

结合以上两点，最终可得到$T_{\mathcal{X}}(x) \subseteq \mathcal{F}(x)$

反例:切锥未必包含线性化可行锥

$$\begin{aligned} \min_{x\in\mathbb{R}} \quad & f(x) = x \\ \text{s.t.} \quad & c(x) = -x + 3 \le 0 \end{aligned}$$

- 则$T_{\mathcal{X}}(3) = \{d \mid d \ge 0\}, \mathcal{F}(3) = \{d : d \ge 0\}$, 于是$T_{\mathcal{X}}(3) = \mathcal{F}(3)$
- 将问题的约束变形为

$$c(x) = (-x+3)^3 \le 0$$

因为可行域不变, 故点$x^* = 3$ 处, 切锥$T_{\mathcal{X}}(x^*) = \{d : d \ge 0\}$ 不变

- 由$c'(x^*) = -3(x^* - 3)^2 = 0$ 知线性化可行锥$\mathcal{F}(x^*) = \{d \mid d \in \mathbb{R}\}$
- 此时, $\mathcal{F}(x^*) \supset T_{\mathcal{X}}(x^*)$ (严格包含)

约束品性的引入

- 线性化可行方向锥$\mathcal{F}(x)$ 受可行域$\mathcal{X}$ 代数表示方式的影响
- 切锥$T_{\mathcal{X}}(x)$ 仅由可行域$\mathcal{X}$ 决定
- 线性可行化方向锥容易计算, 但不能反映可行域$\mathcal{X}$ 的本质特征
- 切锥能反映可行域$\mathcal{X}$ 的本质特征, 但不容易计算
- 引入约束品性来沟通两者, 确保最优点x^* 处$T_{\mathcal{X}}(x^*) = \mathcal{F}(x^*)$, 从而可以用$\mathcal{F}(x)$ 取代$T_{\mathcal{X}}(x)$

线性无关约束品性

定义 (线性无关约束品性)

给定可行点x 及相应的积极集$\mathcal{A}(x)$.如果积极集对应的约束函数的梯度,即$\nabla c_i(x)$, $i \in \mathcal{A}(x)$, 是线性无关的, 则称**线性无关约束品性**(LICQ)在点x 处成立.

定理

给定任意可行点$x \in \mathcal{X}$, 若在该点LICQ 成立, 则有$T_{\mathcal{X}}(x) = \mathcal{F}(x)$.

- **证明**:不失一般性, 我们假设积极集$\mathcal{A}(x) = \mathcal{E} \cup \mathcal{I}$.记矩阵

$$A(x) = [\nabla c_i(x)]^{\mathrm{T}}_{i \in \mathcal{I} \cup \mathcal{E}}.$$

- 假设集合$\mathcal{A}(x)$ 的元素个数为m, 那么矩阵$A(x) \in \mathbb{R}^{m \times n}$ 并且$\mathrm{rank}(A) = m$.
- 令矩阵$Z \in \mathbb{R}^{n \times (n-m)}$ 为$A(x)$ 的零空间的基矩阵, 则Z 满足

$$\mathrm{rank}(Z) = n - m, \quad A(x)Z = 0.$$

线性无关约束品性

- 令$d \in \mathcal{F}(x)$ 为任意线性化可行方向，给定$\lim_{k\to\infty}\ t_k = 0$ 的正标量$\{t_k\}_{k=1}^{\infty}$，定义映射$R : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$:

$$R(z,t) = \begin{bmatrix} c(z) - tA(x)d \\ Z^{\mathrm{T}}(z - x - td) \end{bmatrix},$$

其中$c(z)$ 为向量值函数，其第i 个分量为$c_i(z)$.

- 由$\mathcal{A}(x) = \mathcal{E} \cup \mathcal{I}$ 知

$$R(x,0) = \begin{bmatrix} c(x) \\ Z^{\mathrm{T}}(x - x) \end{bmatrix} = 0, \quad \frac{\partial R(x,0)}{\partial z} = \begin{bmatrix} A(x) \\ Z^{\mathrm{T}} \end{bmatrix}.$$

- 根据Z 的构造，雅可比矩阵$\frac{\partial R(x,0)}{\partial z}$ 是非奇异的.因此，由隐函数定理，对任意充分小的t_k，都存在唯一的z_k，使得$R(z_k, t_k) = 0$.
- 由于$R(z_k, t_k) = 0$，故$c_i(z_k) = t_k \nabla c_i(x)^{\mathrm{T}} d$，$i \in \mathcal{I} \cup \mathcal{E}$.根据线性化可行方向$d$ 的定义，$c_i(z_k) \geq 0$，$i \in \mathcal{I}$，$c_i(z_k) = 0$，$i \in \mathcal{E}$，即z_k 为可行点.

线性无关约束品性

- 由泰勒展开得

$$0 = R(z_k, t_k) = \begin{bmatrix} c(z_k) - t_k A(x) d \\ Z^{\mathrm{T}}(z_k - x - t_k d) \end{bmatrix}$$
$$= \begin{bmatrix} A(x)(z_k - x) + e_k - t_k A(x) d \\ Z^{\mathrm{T}}(z_k - x - t_k d) \end{bmatrix}$$
$$= \begin{bmatrix} A(x) \\ Z^{\mathrm{T}} \end{bmatrix} (z_k - x - t_k d) + \begin{bmatrix} e_k \\ 0 \end{bmatrix}.$$

其中残量e_k 满足$\|e_k\| = o(t_k)$.

- 两边同时作用$\begin{bmatrix} A(x)^{\mathrm{T}} & Z \end{bmatrix}^{-\mathrm{T}}$ 并除以t_k, 则有

$$\frac{z_k - x}{t_k} = d + \begin{bmatrix} A(x) \\ Z^{\mathrm{T}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{e_k}{t_k} \\ 0 \end{bmatrix} \Rightarrow \lim_{k\to\infty} \frac{z_k - x}{t_k} = d,$$

即$d \in T_{\mathcal{X}}(x)$.故$\mathcal{F}(x) \subseteq T_{\mathcal{X}}(x)$.又$T_{\mathcal{X}}(x) \subseteq \mathcal{F}(x)$, 则两集合相同.

Mangasarian–Fromovitz约束品性(MFCQ)

定义 (Mangasarian–Fromovitz约束品性)

给定可行点x 及相应的积极集$\mathcal{A}(x)$.如果存在一个向量$w \in \mathbb{R}^n$, 使得

$$\begin{aligned} \nabla c_i(x)^{\mathrm{T}} w < 0, &\quad \forall i \in \mathcal{A}(x) \cap \mathcal{I}, \\ \nabla c_i(x)^{\mathrm{T}} w = 0, &\quad \forall i \in \mathcal{E}, \end{aligned}$$

并且等式约束对应的梯度集$\{\nabla c_i(x), \ i \in \mathcal{E}\}$ 是线性无关的, 则称点x 处*MFCQ*成立.

- Mangasarian–Fromovitz约束品性是LICQ的的一个常用推广, 简称为MFCQ.
- LICQ可以推出MFCQ, 但是反过来不成立.在MFCQ 成立的情况下, 我们也可以证明$T_{\mathcal{X}}(x) = \mathcal{F}(x)$.

线性约束品性

- 另外一个用来保证$T_{\mathcal{X}}(x) = \mathcal{F}(x)$ 的约束品性是线性约束品性.

定义 (线性约束品性)

若所有的约束函数$c_i(x)$, $i \in \mathcal{I} \cup \mathcal{E}$ 都是线性的, 则称线性约束品性成立.

- 当线性约束品性成立时, 也有$T_{\mathcal{X}}(x) = \mathcal{F}(x)$.
- 因此对只含线性约束的优化问题, 例如线性规划、二次规划, 很自然地有$T_{\mathcal{X}}(x) = \mathcal{F}(x), \forall\, x$.我们无需再关注约束函数的梯度是否线性无关.
- 一般来说, 线性约束品性和LICQ之间没有互相包含的关系.

Karush-Kuhn-Tucker(KKT)条件:引入

- 回顾几何最优性条件:

$$x^* \text{局部极小} \Leftrightarrow T_{\mathcal{X}}(x^*) \cap \{d \mid \nabla f(x^*)^{\mathrm{T}} d < 0\} = \varnothing.$$

- $T_{\mathcal{X}}(x^*) = \mathcal{F}(x^*)$ 时(约束品性成立), 上述条件变为

$$\left\{ d \;\middle|\; \begin{array}{l} d^{\mathrm{T}} \nabla f(x^*) < 0, \\ d^{\mathrm{T}} \nabla c_i(x^*) = 0, \ i \in \mathcal{E}, \\ d^{\mathrm{T}} \nabla c_i(x^*) \le 0, \ i \in \mathcal{A}(x^*) \cap \mathcal{I} \end{array} \right\} = \varnothing.$$

- 上式依然难以验证, 但可使用Farkas引理进行化简.

定理 (Farkas引理)

设p和q为两个非负整数, 给定$\mathbb{R}^n$中的向量$\{a_i\}_{i=1}^p, \{b_i\}_{i=1}^q$和$c$.则满足:

$$\begin{aligned} d^{\mathrm{T}} a_i &= 0, \ i = 1, 2, \cdots, p, \\ d^{\mathrm{T}} b_i &\geq 0, \ i = 1, 2, \cdots, q, \\ d^{\mathrm{T}} c &< 0 \end{aligned}$$

的d不存在当且仅当存在$\{\lambda_i\}_{i=1}^p$和$\mu_i \geq 0, \ i = 1, 2, \cdots, q$, 使得

$$c = \sum_{i=1}^{p} \lambda_i a_i + \sum_{i=1}^{q} \mu_i b_i.$$

证明:仅证必要性, 若这样的λ_i和μ_i不存在, 定义集合

$$S = \{z \mid z = \sum_{i=1}^{p} \lambda_i a_i + \sum_{i=1}^{q} \mu_i b_i, \ \lambda_i \in \mathbb{R}, \mu_i \geq 0\}.$$

- S 是一个闭凸锥.因为$c \notin S$, 由凸集的严格分离超平面定理可知:存在$d \in \mathbb{R}^n, \alpha \in \mathbb{R}$ 使得

$$d^{\mathrm{T}}c < \alpha < d^{\mathrm{T}}z, \quad \forall z \in S.$$

- 因为$0 \in S$, 所以

$$d^{\mathrm{T}}c < \alpha < d^{\mathrm{T}}0 = 0.$$

- 任取b_i 与$t \geq 0$, 有$tb_i \in S$, 故$\alpha < td^{\mathrm{T}}b_i$. 由$t$ 的任意性知$d^{\mathrm{T}}b_i \geq 0$.
- 同理, 任取$t \in \mathbb{R}$ 与a_i, 有$ta_i \in S$, 故$td^{\mathrm{T}}a_i < \alpha$.由$t$ 的任意性知$d^{\mathrm{T}}a_i = 0$.
- 综上, 此时的d 为不等式系统的解.

从Farkas引理到KKT条件

- 由Farkas引理，取$a_i = \nabla c_i(x^*), i \in \mathcal{E}, b_i = \nabla c_i(x^*), i \in \mathcal{A}(x^*) \cap \mathcal{I}$ 以及$c = -\nabla f(x^*)$，则$T_{\mathcal{X}}(x^*) = \mathcal{F}(x^*)$ 时几何最优性条件等价于：

$$-\nabla f(x^*) = \sum_{i\in\mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i\in\mathcal{A}(x^*)\cap\mathcal{I}} \lambda_i^* \nabla c_i(x^*),$$

其中$\lambda_i^* \in \mathbb{R}, i \in \mathcal{E}$，$\lambda_i^* \geq 0, i \in \mathcal{A}(x^*) \cap \mathcal{I}$.

- 如果补充定义$\lambda_i^* = 0,\ i \in \mathcal{I} \backslash \mathcal{A}(x^*)$，那么

$$-\nabla f(x^*) = \sum_{i\in\mathcal{I}\cup\mathcal{E}} \lambda_i^* \nabla c_i(x^*),$$

这恰好对应于拉格朗日函数关于x 的一阶最优性条件.

- **互补松弛条件**:对于任意的$i \in \mathcal{I}$，我们注意到

$$\lambda_i^* c_i(x^*) = 0.$$

这说明$i \in \mathcal{A}(x^*) \cap \mathcal{I}$ 时乘子$\lambda_i^* = 0$ 或$c_i(x^*) = 0$ 至少出现一种，当两种情况恰好只有一种满足时，我们也称严格互补松弛条件.

KKT条件:总结

假设x^* 是一般优化问题

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & c_i(x) \leq 0,\ i \in \mathcal{I} \\ & c_i(x) = 0,\ i \in \mathcal{E} \end{aligned}$$

的一个局部最优点.如果

$$T_{\mathcal{X}}(x^*) = \mathcal{F}(x^*)$$

成立, 那么存在拉格朗日乘子λ_i^* 使得如下条件成立:

$$\begin{aligned} \textbf{稳定性条件} \quad & \nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i\in\mathcal{I}\cup\mathcal{E}} \lambda_i^* \nabla c_i(x^*) = 0, \\ \textbf{原始可行性条件} \quad & c_i(x^*) = 0,\ \forall i \in \mathcal{E}, \\ \textbf{原始可行性条件} \quad & c_i(x^*) \leq 0,\ \forall i \in \mathcal{I}, \\ \textbf{对偶可行性条件} \quad & \lambda_i^* \geq 0,\ \forall i \in \mathcal{I}, \\ \textbf{互补松弛条件} \quad & \lambda_i^* c_i(x^*) = 0,\ \forall i \in \mathcal{I}. \end{aligned}$$

KKT条件:注记

- 称满足KKT条件的变量对(x^*, λ^*) 为KKT对.
- 称x^* 为KKT点.
- 如果局部最优点x^* 处$T_{\mathcal{X}}(x^*) \neq \mathcal{F}(x^*)$, 那么$x^*$ 不一定是KKT点.
- KKT条件只是必要的, KKT点不一定是局部最优点.

依旧考虑一般优化问题:

$$\min_{x\in\mathbb{R}^n} f(x), \qquad \text{s.t. } c_i(x) \le 0,\ i \in \mathcal{I};\ c_i(x) = 0,\ i \in \mathcal{E}.$$

若x^* 是满足KKT条件的点, 假设$T_{\mathcal{X}}(x^*) = \mathcal{F}(x^*)$, 则$\forall d \in \mathcal{F}(x^*)$,

$$d^{\mathrm{T}}\nabla f(x^*) = -\sum_{i\in\mathcal{E}} \underbrace{\lambda_i^* d^{\mathrm{T}}\nabla c_i(x^*)}_{=0} - \sum_{i\in\mathcal{A}(x^*)\cap\mathcal{I}} \underbrace{\lambda_i^* d^{\mathrm{T}}\nabla c_i(x^*)}_{\le 0} \ge 0,$$

此时一阶条件无法判断x^* 是否是最优值点.

- 若$d^{\mathrm{T}}\nabla f(x^*) = 0$, 则需要利用二阶信息来进一步判断在其可行邻域内的目标函数值.
- 拉格朗日函数在这些方向上的曲率即可用来判断x^* 的最优性.
- 首先引入临界锥来精确刻画这些方向.

临界锥

定义 (临界锥)

设(x^*, λ^*) 是满足*KKT*条件的*KKT*对, 定义临界锥为

$$\mathcal{C}(x^*, \lambda^*) = \{d \in \mathcal{F}(x^*) \mid \nabla c_i(x^*)^{\mathrm{T}} d = 0,\ \forall i \in \mathcal{A}(x^*) \cap \mathcal{I} \text{ 且 } \lambda_i^* > 0\},$$

其中$\mathcal{F}(x^*)$ 为点x^* 处的线性化可行方向锥.

- 临界锥是线性化可行方向锥$\mathcal{F}(x^*)$ 的子集.
- 沿着临界锥中的方向进行优化, 所有等式约束和$\lambda_i^* > 0$ 对应的不等式约束(此时这些不等式均取等)都会尽量保持不变.
- 当$d \in \mathcal{C}(x^*, \lambda^*)$ 时, $\forall i \in \mathcal{E} \cup \mathcal{I}$ 有$\lambda_i^* \nabla c_i(x^*)^{\mathrm{T}} d = 0$, 故
$$d^{\mathrm{T}} \nabla f(x^*) = \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) = 0.$$
- 临界锥定义了依据一阶导数不能判断是否为下降或上升方向的线性化可行方向, 必须使用高阶导数信息加以判断.

二阶最优性条件

考虑一般优化问题:

$$\min_{x\in\mathbb{R}^n} f(x), \qquad \text{s.t.} \;\; c_i(x) \le 0,\; i\in\mathcal{I};\; c_i(x)=0,\; i\in\mathcal{E}.$$

定理 (二阶最优性条件)

必要性:假设x^* 是问题的一个局部最优解,并且$T_{\mathcal{X}}(x^*)=\mathcal{F}(x^*)$ 成立.令λ^* 为相应的拉格朗日乘子,即(x^*,λ^*) 满足*KKT*条件,那么

$$d^{\mathrm{T}}\nabla^2_{xx}L(x^*,\lambda^*)d \ge 0, \quad \forall d\in\mathcal{C}(x^*,\lambda^*).$$

充分性:假设在可行点x^* 处,存在一个拉格朗日乘子λ^*,使得(x^*,λ^*) 满足*KKT*条件.如果

$$d^{\mathrm{T}}\nabla^2_{xx}L(x^*,\lambda^*)d > 0, \quad \forall d\in\mathcal{C}(x^*,\lambda^*),\; d\neq 0,$$

那么x^* 为问题的一个严格局部极小解.

回顾无约束优化问题的二阶最优性条件:

- 问题: $\min_{x\in\mathbb{R}^n} f(x)$.
- 必要条件:若x^* 是f 的一个局部极小点, 则$\nabla f(x^*) = 0, \nabla^2 f(x^*) \succeq 0$.
- 充分条件:若$\nabla f(x^*) = 0, \nabla^2 f(x^*) \succ 0$, 则$x^*$ 是f 的一个局部极小点.

约束优化问题的二阶最优性条件也要求某种"正定性", 但只需要考虑临界锥$\mathcal{C}(x^*, \lambda^*)$ 中的向量而无需考虑全空间的向量.

有些教材中将其称为"投影半正定性".

$$\min \quad x_1^2 + x_2^2, \qquad \text{s.t.} \quad \frac{x_1^2}{4} + x_2^2 - 1 = 0,$$

其拉格朗日函数为

$$L(x, \lambda) = x_1^2 + x_2^2 + \lambda(\frac{x_1^2}{4} + x_2^2 - 1).$$

该问题可行域在任意一点$x = (x_1, x_2)^{\mathrm{T}}$ 处的线性化可行方向锥为

$$\mathcal{F}(x) = \{(d_1, d_2) \mid \frac{x_1}{4} d_1 + x_2 d_2 = 0\}.$$

因为只有一个等式约束且其对应函数的梯度非零, 故有LICQ成立, 于是$\mathcal{F}(x) = T_{\mathcal{X}}(x)$.若$(x, \lambda)$ 为KKT对, 由于无不等式约束,
故$\mathcal{C}(x, \lambda) = \mathcal{F}(x)$.
可以计算出其4 个KKT对

$$(x^{\mathrm{T}}, \lambda) = (2, 0, -4), \quad (-2, 0, -4), \quad (0, 1, -1) \quad 和 \quad (0, -1, -1).$$

考虑第一个KKT对$y=(2,0,-4)^{\mathrm{T}}$, 计算可得

$$\nabla_{xx}^2 L(y)=\begin{bmatrix}0 & 0\\ 0 & -6\end{bmatrix},\quad \mathcal{C}(y)=\{(d_1,d_2)\mid d_1=0\}.$$

取$d=(0,1)$, 则

$$d^{\mathrm{T}}\nabla_{xx}^2 L(y)d=-6<0,$$

因此y 不是局部最优点.类似地, 对第三个KKT对$z=(0,1,-1)$,

$$\nabla_{xx}^2 L(z)=\begin{bmatrix}\frac{3}{2} & 0\\ 0 & 0\end{bmatrix},\quad \mathcal{C}(z)=\{(d_1,d_2)\mid d_2=0\}.$$

对于任意的$d=(d_1,0)$ 且$d_1\neq 0$,

$$d^{\mathrm{T}}\nabla_{xx}^2 L(z)d=\frac{3}{2}d_1^2>0.$$

因此, z 为一个严格局部最优点.