最优化方法

授课教师：徐海涛

2024年秋

第五章 有约束极小化非线性规划

5.1 有约束极小化非线性规划问题

- 大部分工程优化问题都可以归为有约束极小化问题，即在一定的约束条件下对目标函数极小化。
- 在机械领域，常见的例子是在应力和可接受的变形约束下，合理的设计结构，使得重量最小。
- 第4章已针对线性约束问题讨论了重要的概念，如起作用约束、拉格朗日乘子、利用计算机求解方法及一些几何概念。这些概念同样适用于非线性优化问题。
- 本章所讨论的数值方法，针对的是非线性约束，但是很多方法在迭代过程中仍会调用线性规划求解方法
- 本章以两变量问题的图示化求解作为开始，讨论利用EXCEL规划求解功能和MATLAB进行求解的方法；接下来给出非线性规划的标准形式，讨论了最优性条件、相关的几何概念和凸性

5.1 有约束极小化非线性规划问题

有约束的优化问题都可以写为一般形式的非线性规划：

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leqslant 0 \quad i = 1, \cdots, m \\ & h_j(\mathbf{x}) = 0 \quad j = 1, \cdots, \ell \end{aligned} \tag{5.1}$$

其中，决策变量 $\mathbf{x} = [x_1, x_2, \cdots, x_n]^{\mathrm{T}}$ 是一个 n 维实数列向量，f 为目标函数或成本函数，g 表示不等式约束，h 表示等式约束。通常利用 $\mathbf{x}^0$ 表示起始点，$\mathbf{x}^*$ 表示最优点，$\mathbf{x}_k$ 表示（当前）第 k 步迭代的点。同时满足等式约束和不等式约束的点 $\mathbf{x}$，称为可行点。对于点 $\mathbf{x}$，如果 $g_i(\mathbf{x}) = 0$，则称不等式约束 $g_i(\mathbf{x}) \leqslant 0$ 是点 $\mathbf{x}$ 处的起作用约束；如果 $g_i(\mathbf{x}) \leqslant 0$，则 $g_i(\mathbf{x}) \leqslant 0$ 是点 $\mathbf{x}$ 处的不起作用约束。等式约束 $h_i(\mathbf{x}) = 0$ 是任意可行点的起作用约束。形式上，问题（5.1）可以表示为

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \Omega \end{aligned} \tag{5.2}$$

其中，Ω 为由所有约束条件组成的可行域，即 $\Omega = \{\mathbf{x} : \mathbf{g} \leqslant \mathbf{0}, \mathbf{h} = \mathbf{0}\}$。对于式（5.2）这种模型，可知对于点 $\mathbf{x}^* = [x_1^*, x_2^*, \cdots, x_n^*]^{\mathrm{T}}$，如果所有的 $\mathbf{x} \in \Omega$，都有 $f(\mathbf{x}^*) \leqslant f(\mathbf{x})$，则称 $\mathbf{x}^*$ 是全局极小点。

5.1 有约束极小化非线性规划问题（例题）

例 5.1　桁架问题

该示例来自于文献［Fox，1971］，演示了非线性规划在机械设计中的应用。这是一个两杆平面桁架，如图 E5.1 所示。

两个杆都是薄壁钢管，通过锁销连接，承受 $2P$ 大小的向下负载，如图 E5.1 所示。假定钢管的壁厚 t 不变，半跨的长度也固定为 B。请为钢管设计合适的平均直径 d 和桁架高度 H。

每个杆承受的应力为

$$\sigma = \frac{P}{\pi t}\frac{(B^2+H^2)^{1/2}}{Hd} \qquad \text{(a)}$$

应力不应该超过上界 σ_{all}，故有约束条件为

$$\frac{P}{\sigma_{all}\pi t}\frac{(B^2+H^2)^{1/2}}{Hd} - 1 \leqslant 0 \qquad \text{(b)}$$

上述约束条件进行了归一化处理，这一点对于数值求解方法非常重要，因为无论约束条件的实际大小是多少，归一化处理之后都可以将它们限定在区间［0，1］之内。

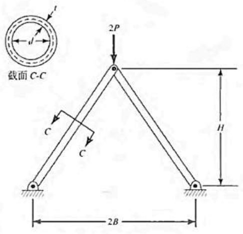

图 E5.1　两杆平面桁架

5.1 有约束极小化非线性规划问题（例题）

由于两个杆处于压缩状态，因此还应该注意避免弯曲。临界屈曲载荷的欧拉公式为

$$\sigma_{cr} = \frac{\pi^2 E(d^2+t^2)}{8(B^2+H^2)} \qquad \text{(c)}$$

利用式（a）和式（c），可将约束条件 $\sigma \leqslant \sigma_{cr}$ 写为归一化的形式：

$$\frac{8P(B^2+H^2)^{1.5}}{\pi^3 Et Hd(d^2+t^2)} - 1 \leqslant 0 \qquad \text{(d)}$$

显然，还必须满足 $H \geqslant 0$，$d \geqslant 0$。桁架重量，也就是目标函数为

$$f = 2\rho\pi dt(B^2+H^2)^{1/2} \qquad \text{(e)}$$

其中，ρ 为桁架材料的密度。这样一来，问题就成为在约束条件（b）和约束条件（d）下，使得目标函数 f 最小。得到了最优解 (H^*, d^*) 之后，还应该保证钢管是薄壁的，即横截面 $d/t \gg 1$。当 $\rho = 0.3\ \mathrm{lb/in^3}$，$P = 33\,000\ \mathrm{lb}$，$B = 30\ \mathrm{in}$，$t = 0.1\ \mathrm{in}$，$E = 30\times10^6\ \mathrm{psi}$，$\sigma_{all} = 100\,000$ 时，利用数值方法，求得最优解为 $H = 20\ \mathrm{in}$，$d = 1.9\ \mathrm{in}$。屈曲应力和屈服应力对应的约束都是起作用约束。如果修改问题的参数（如半跨距离 B 或屈服应力上限），最优解就会随之变化，相应的起作用约束也会不同。如果为决策变量增加上下界，或增加形变约束，按照当前的参数，可能不会存在可行解。

在设计过程中引入优化理论和技术，能够让设计工程师从繁杂的计算中解脱出来，使其将更多的精力投入到更有价值、更高层次的事情中。比如，设计工程师通过分析最优设计方案对应的起作用约束，判断是否能够通过更换材料、更改桁架拓扑结构或增加约束等方式来进一步降低成本。

5.2 两变量优化问题的图示化求解

两变量问题（$n=2$）可以用图示化的方式求解，1.5 节中例 1.12 已经演示过这样的求解过程。决策变量 $\mathbf{x} = [x_1, x_2]^T$，问题可写为 minimize$f(\mathbf{x})$，subject to $g_i(\mathbf{x}) \leqslant 0, i = 1, \cdots, m$。首先定义 $x_1 - x_2$ 平面，接下来对于每个 g_i，绘制 $g_i(x_1, x_2) = 0$ 对应的 $x_1 - x_2$ 关系曲线，只需要固定 x_1，然后找到对应的 x_2 即可。找出曲线对应的可行侧，即满足 $g_i(\mathbf{x}) < 0$ 的一侧。所有 m 条曲线对应的可行侧交集就是可行域。接下来绘制目标函数的等值线 $f(\mathbf{x}) = c$，令 c 不断变化，如 $c = 10, 20, 30$ 等，即可绘制出不同水平下的等值线。函数 f 的等值线是一组“平行”或不相交的曲线，位于可行域内的最小等值线就是函数的极小值，对应的点就是极小点。有时候，极小值水平下的等值线可能与某条约束方程曲线重合，在这种情况下，该约束条件曲线位于可行域的部分都是极小点（即有无数种最优设计方案）。

如果还存在等式约束，那么就必须在曲线 $h_i(\mathbf{x}) = 0$ 的交集上寻找最优解，且一定是位于由不等式约束 g_i 构成的可行域内。注意，对于二维问题，两个等式约束相交于一点，也就无所谓最优解了。

5.2 两变量优化问题的图示化求解

例 5.2　考虑前面给出的两变量问题，该问题的图示化求解方式如图 E5.2 所示。d 和 H 的最优解与数值方法得到的结果是一致的，参见例 5.1。首先，对 H 从 1 到 50 进行离散化取值，将这些离散化的 H 代入例 5.1 中的方程（b），可以直接利用解析方法求出对应的 d。利用 EXCEL 的“散点图”绘图功能可绘制出 H 和 d 之间的关系曲线，对应方程 $g_1 = 0$。以此类推，可绘制出 $g_2 = 0$ 的曲线，需要指出的是，对于 g_2，当 H 给定时，需要求解一个三次方程才能得到 d，这必须采用数值方法求解。

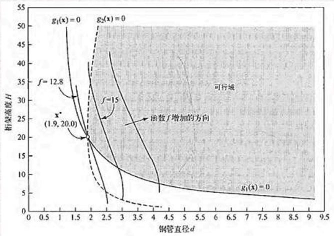

图 E5.2

5.3 利用EXCEL规划求解功能和MATLAB求解非线性优化问题

EXCEL规划求解功能

与第 4 章的 4.3 节中给出的求解过程相同，首先要在 EXCEL 表格 A1 – G7 中构造目标函数和约束条件，并给出变量的初始值和一些常数值。然后在【数据】菜单下的【分析】选项卡中选择【规划求解】，启动规划求解功能，求解定义好的优化问题。可按照如下格式构建问题：

高度 H	直径 d	目标函数 f	约束 1	约束 2
5	0.5	2.866441015	11.77891288	121.8370137

设置好各参数、约束条件和目标函数之后，就可以开展计算了。按照 EXCEL 规划求解功能的做法，需要在执行了确定目标、设定极大化还是极小化、通过更改可变单元格、添加遵守约束等操作之后，才能够开展计算。这些设置可参照 4.3 节进行。将 H 和 d 初始值分别设定为 $H=5$，$d=0.5$。

5.3 利用EXCEL规划求解功能和MATLAB求解非线性优化问题

EXCEL规划求解功能

选择【非线性 GRG】求解方法，可得如下求解结果：

目标单元格（最小值）

单元格	名称	初值	终值
\$C \$2	目标函数 f	2.866441015	12.81253417

可变单元格

单元格	名称	初值	终值	整数
\$A \$2	高度 H	5	20.23700348	约束
\$B \$2	直径 d	0.5	1.878339648	约束

5.3 利用EXCEL规划求解功能和MATLAB求解非线性优化问题

EXCEL规划求解功能

约束

单元格	名称	单元格值	公式	状态	型数值
\$D \$2	约束 1	3.43536E – 06	\$D \$2 < = 0	到达限制值	0
\$E \$2	约束 2	1.96753E – 05	\$E \$2 < = 0	到达限制值	0

灵敏度分析报告中还提供了最优解对应的拉格朗日乘子。

单元格	名称	终值	拉格朗日乘子
\$D \$2	约束 1	3.43536E – 06	– 5.614045636
\$E \$2	约束 2	1.96753E – 05	– 2.403980668

5.4 非线性优化问题的标准形式及转换方法

- 非线性优化问题必须转换为标准形式。本节将讨论标准形式转换过程遵循的4条原则

(i) maximize f <=> minimize $-f$

(ii) 把minimize: maximum $\{f_1, f_2, \cdots, f_p\}$ 的“极小极大化”问题改写为

minimize $f = x_{n+1}$

subject to $f_{i(x)} - x_{n+1} \le 0$, $i = 1, \ldots, p$

其中，x_{n+1}为新引入的变量。针对标准形式的求解方法都可以用于求解上述问题。

5.4 非线性优化问题的标准形式及转换方法

（iii）逐点约束。约束条件必须针对某个参数，如 $\theta\in[0,1]$ 的每个值都成立。可以采用将参数离散化的形式进行处理，如将 θ 离散化为参数 $\theta_i, i=1,\cdots,n_d$，其中，n_d 为离散点的总数量。这样的参数一般用 θ 表示。假定约束为

$$\psi(\mathbf{x},\theta)\leqslant 0,\ \theta\in[\theta_{\min},\theta_{\max}] \tag{a}$$

可改写为

$$\psi(\mathbf{x},\theta_i)\leqslant 0,\quad i=1,\cdots,n_d \tag{b}$$

类似这种逐点约束还有其他一些处理方式。但是，按照式（b）这种处理方式，并利用起作用约束集方法进行求解，是一种非常稳健的方法。本章最后的例 5.20 应用的就是这种方法。

5.4 非线性优化问题的标准形式及转换方法

（iv）回归分析一般用于构建方程来拟合一组数据，原则就是使拟合误差极小化，而误差一般采用不同类型的范数进行表示，比如，目标函数为

$$f=\text{minimize}\sum_{i=1}^{n}|\psi_i(\mathbf{x})| \tag{c}$$

其中，ψ_i 表示方程模型与实际数据之间的误差。在某些情况下，这种方式存在问题。比如，对于模型 $y=|x|$，式（c）所示的函数在 $\psi_i=0$ 对应的点处不可导。这可以利用如下方式进行处理：

$$\begin{aligned}&\text{minimize} && \sum_{i=1}^{n}(p_i+n_i)\\ &\text{subject to} && \psi_i=p_i-n_i\\ & && p_i\geqslant 0,\quad n_i\geqslant 0\end{aligned}$$

其中，p_i 和 $n_i, i=1,\cdots,n$ 分别表示正值误差和负值误差。

5.4 非线性优化问题的标准形式及转换方法

其中，ψ_i 表示方程模型与实际数据之间的误差。在某些情况下，这种方式存在问题。比如，对于模型 $y=|x|$，式（c）所示的函数在 $\psi_i=0$ 对应的点处不可导。这可以利用如下方式进行处理：

$$\begin{aligned}&\text{minimize} && \sum_{i=1}^{n}(p_i+n_i)\\ &\text{subject to} && \psi_i=p_i-n_i\\ & && p_i\geqslant 0,\quad n_i\geqslant 0\end{aligned}$$

其中，p_i 和 $n_i, i=1,\cdots,n$ 分别表示正值误差和负值误差。

最后，考虑如下目标函数：

$$\underset{\mathbf{x}}{\text{minimize}}\ \underset{i}{\text{maximize}}\ |\psi_i(\mathbf{x})|$$

其含义为使得最大的误差达到最小。可以按照如下方式进行处理：

$$\begin{aligned}&\text{minimize} && Z\\ &\text{subject to} && -Z\leqslant\psi_i(\mathbf{x})\leqslant Z\end{aligned}$$

这样一来，目标函数就是可导的，Z 是新引入的变量。随着 Z 的降低，误差的上界和下界同时缩小。这也是前面第（ii）情况中用到的思路。

5.5 最优性必要条件

- 若点 x^* 是函数 f 的一个无约束局部极小点，则 $\nabla f(x^*)=0$。
- 这是局部极小点的必要条件。这一条件的核心是“零斜率”，利用该条件，可以手工计算函数的极小点，还可作为数值算法的停止规则。

5.5 最优性必要条件

- 本节将讨论有约束问题式(5. 1) 或式(5. 2)的最优性必要条件。从图5.1中可以看出，当最优点位于可行域边界上时，对应的目标函数斜率并不是零。如果最优点位于可行域内部，即所有的约束方程都是不起作用的，则零斜率的条件仍然成立。

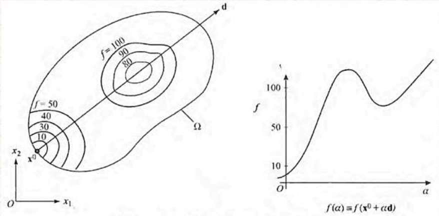

图 5.1　极小点位于边界上时，梯度不为零

5.5 最优性必要条件

- 接下来分两步给出最优性条件
- 首先，利用拉格朗日乘子法给出只包括等式约束的优化问题的最优性条件。该方法在机械领域中应用非常广泛，与变分法同时应用，能够得到很多问题的解析解，如最速降线问题在平面上寻找一条曲线，使得球体能够以最短的时间从高点A下降到低点B。

纯等式约束的优化问题——拉格朗日乘子法

满足等式约束优化问题最优性必要条件的点实际上是平稳点，可以是函数 f 的极小点、极大点或者鞍点（拐点）。这与无约束优化问题的最优性必要条件是类似的。包含 ℓ 个纯等式约束的优化问题可以写为

$$\text{minimize} \quad f(\mathbf{x}) \tag{5.3a}$$

$$\text{subject to} \quad h_j(\mathbf{x}) = 0, \quad j = 1, \cdots, \ell \tag{5.3b}$$

5.5 最优性必要条件（等式约束）

纯等式约束的优化问题——拉格朗日乘子法

满足等式约束优化问题最优性必要条件的点实际上是平稳点，可以是函数 f 的极小点、极大点或者鞍点（拐点）。这与无约束优化问题的最优性必要条件是类似的。包含 ℓ 个纯等式约束的优化问题可以写为

$$\text{minimize} \quad f(\mathbf{x}) \tag{5.3a}$$

$$\text{subject to} \quad h_j(\mathbf{x}) = 0, \quad j = 1, \cdots, \ell \tag{5.3b}$$

首先考虑从约束方程式（5.3b）中消除一个决策变量，比如将 x_n 用其他变量 x_i 进行表示，并将其代入到 f 中。这样一来，函数 f 的变量就只有 $n-1$ 个，以此类推，可以消除掉所有的约束，从而将有约束问题转换为无约束问题，就可以直接给出其最优性条件。这种方法是经过证明了的，在某些情况下可以使用。但是，绝大部分情况下，变量消除是一项非常繁琐的工作。而且，式（5.3b）中任一变量 $x_1, x_2, \cdots, x_n$ 都是平等的，因此，没有任何理由选定某个变量作为被消除的变量，而其他的变量就可以作为独立变量得以保留。拉格朗日设计了一种形式上非常美观的方法，可以在不用开展变量消除工作的前提下给出最优性条件。首

5.5 最优性必要条件（等式约束）

纯等式约束的优化问题——拉格朗日乘子法

构造拉格朗日函数

$$L(\mathbf{x}) = f(\mathbf{x}) + \sum_{j=1}^{\ell} \lambda_j h_j(\mathbf{x})$$

其中，λ_j 是一个标量，表示与约束条件 h_j 相对应的乘子。如果 $\mathbf{x}^*$ 是一个平稳点（极小点、极大点或者鞍点），则一定会有以下条件成立：

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{\ell} \lambda_j \frac{\partial h_j}{\partial x_i} = 0, \quad i = 1, \cdots, n \tag{5.5a}$$

$$h_j(x) = 0, \quad j = 1, \cdots, \ell \tag{5.5b}$$

式（5.5）也可以写为 $\nabla_{\mathbf{x}} L = 0$ 和 $\mathbf{h} = 0$。这就是有约束优化问题的最优性必要条件。利用该条件，可以获得候选的极小点，或者断言某个点 $\mathbf{x}^*$ 不是极小点。为了获得极小点 $\mathbf{x}^*$ 及其对应的拉格朗日乘子 λ，则应该求解方程式（5.5），得到所有的候选点，然后比较候选点对应的函数值，最小函数值对应的点就是极小点。利用最优性充分条件也可以从候选点中挑选出极小点，这是 5.6 节讨论的内容。

5.5 最优性必要条件（等式约束）

例 5.3 考虑优化问题：

$$\begin{aligned} \text{minimize} \quad & 2x_1 + x_2 \\ \text{subject to} \quad & x_1^2 + x_2^2 - 1 = 0 \end{aligned}$$

拉格朗日函数为 $L = 2x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 1)$，最优性条件为

$$\frac{\partial L}{\partial x_1} = 2 + 2\lambda x_1 = 0, \quad x_1 = -\frac{1}{\lambda}$$

$$\frac{\partial L}{\partial x_2} = 1 + 2\lambda x_2 = 0, \quad x_2 = -\frac{1}{2\lambda}$$

明理，精工，笃学，致远 21

5.5 最优性必要条件（等式约束）

将 x_1 和 x_2 的表达式代入到约束方程 $h = 0$ 中，可得 $\lambda = \pm\sqrt{5}/2$。代入到 x_1 和 x_2 中，可得

$$\lambda = \frac{\sqrt{5}}{2} : x_1^* = -\frac{2}{\sqrt{5}}, \quad x_2^* = -\frac{1}{\sqrt{5}} \Rightarrow \text{图E5.3中的点}A$$

$$\lambda = -\frac{\sqrt{5}}{2} : x_1^* = \frac{2}{\sqrt{5}}, \quad x_2^* = \frac{1}{\sqrt{5}} \Rightarrow \text{图E5.3中的点}B$$

显然，极小点为 A。正如前面提到的，必要条件只能给出两个候选的极小点。利用 5.6 节中讨论的充分条件可以在这些候选点中找出真正的极小点（见例 5.8）。

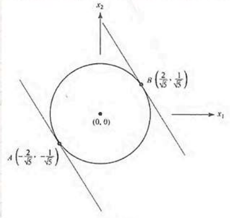

图 E5.3

明理，精工，笃学，致远 22

5.5 最优性必要条件（不等式约束）

不等式约束优化问题：下降和可行方向的概念以及最优性条件（KKT 条件）

$$\text{minimize} \quad f(\mathbf{x}) \tag{5.6a}$$

$$\text{subject to} \quad g_i(\mathbf{x}) \leqslant 0, \quad i = 1, \cdots, m \tag{5.6b}$$

的最优性必要条件。首先要给出下降方向和可行方向的定义，后面即将要讨论的数值方法中也要用到这些定义。所谓方向，指的是决策变量空间或 $x_1, x_2, \cdots, x_n$ 空间中的一个向量。只考虑连续可导函数，即函数的导数也是连续的。

明理，精工，笃学，致远 23

5.5 最优性必要条件（不等式约束）

下降方向

在点 $\mathbf{x}_k$ 处，如果方向 $\mathbf{d}$ 满足

$$\nabla f^{\mathrm{T}} \mathbf{d} < 0 \tag{5.7}$$

则 $\mathbf{d}$ 是一个下降方向，其中 ∇f 为函数 f 在 $\mathbf{x}_k$ 处的梯度。这意味着对于足够小的 $\alpha > 0$，不等式

$$f(\mathbf{x}_k + \alpha\mathbf{d}) < f(\mathbf{x}_k)$$

成立。由此可知，如果沿着方向 $\mathbf{d}$ 对决策变量进行调整，则可以使得函数值 f 下降。第 3 章中提到过，∇f 是函数值 f 增加的方向。因此，如果 $\mathbf{d}$ 指向的是 ∇f 的反方向所在的半空间，则 $\mathbf{d}$ 是一个下降方向，如图 5.2 所示。

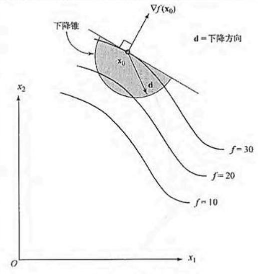

图 5.2 下降锥和下降方向

明理，精工，笃学，致远 24

5.5 最优性必要条件（不等式约束）

下降方向

例 5.4 函数为 $f=x_1x_2$，$\mathbf{x}^0=[1,3]^{\mathrm{T}}$，$\mathbf{d}=[1,1]^{\mathrm{T}}$，试问 $\mathbf{d}$ 是否是下降方向？

解：$\nabla f(\mathbf{x}^0)=[3,1]^{\mathrm{T}}$，故 $(\nabla f(\mathbf{x}^0))^{\mathrm{T}}\mathbf{d}=4>0$，可知 $\mathbf{d}$ 不是下降方向。

5.5 最优性必要条件（不等式约束）

可行方向

接下来考虑点 $\mathbf{x}_k$ 处的可行方向，$\mathbf{x}_k\in\Omega$ 为优化问题式（5.6b）的可行域。对于方向 $\mathbf{d}$，如果存在一个 $\bar{\alpha}>0$，使得在所有的 α，$0\leqslant\alpha\leqslant\bar{\alpha}>0$ 下，都有 $(\mathbf{x}_k+\alpha\mathbf{d})\in\Omega$，则称 $\mathbf{d}$ 是点 $\mathbf{x}_k$ 处的可行方向。考虑到 $g_i(\mathbf{x})$ 是可导函数，可以给出一个更为实用的下降方向定义。在可行点 $\mathbf{x}_k$ 处，假定有一个或多个不等式约束是起作用约束，即有一个或多个 $g_i=0$。定义起作用约束的下标集 I：

$$I=\{i:g_i(x_k)=0,\quad i=1,\cdots,m\} \tag{5.8}$$

5.5 最优性必要条件（不等式约束）

可行方向

正则点

对于可行点 $\mathbf{x}_k$，如果梯度向量 $\nabla g_i(\mathbf{x}_k)$，$i\in I$ 线性无关，则 $\mathbf{x}_k$ 是一个正则点。

如果 $\mathbf{x}_k$ 是一个正则点，则方向 $\mathbf{d}$ 满足

$$\nabla g_i^{\mathrm{T}}\mathbf{d}<\mathbf{0}, \qquad i\in I \tag{5.9}$$

时，$\mathbf{d}$ 是一个可行方向。这意味着对于足够小的 $\alpha>0$，$(\mathbf{x}_k+\alpha\mathbf{d})$ 也会是可行点，或 $g_i(\mathbf{x}_k+\alpha\mathbf{d})<0$，$i=1,\cdots,m$ 成立。实际上，对于每个 $i\in I$，式（5.9）定义了一个半空间，所有半空间的交集构成一个可行锥，$\mathbf{d}$ 就位于这个可行锥中，如图 5.3 所示。

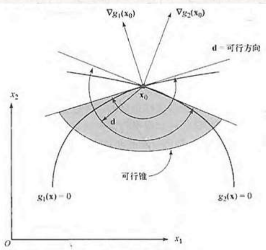

图 5.3 可行锥和可行方向

5.5 最优性必要条件（不等式约束）

可行方向

例 5.5 考虑优化问题：

$$\text{minimize}\quad f=-(x_1+x_2)$$

$$g_1\equiv x_1^2+4x_2^2-1\leqslant 0,\quad g_2\equiv -x_1\leqslant 0,\quad g_3\equiv -x_2\leqslant 0$$

初始点为 $\mathbf{x}_0=\left[\frac{1}{\sqrt{5}},\frac{1}{\sqrt{5}}\right]^{\mathrm{T}}$。试分别判断方向向量 $\mathbf{d}_1=[1,0]^{\mathrm{T}}$，$\mathbf{d}_2=[1,-0.5]^{\mathrm{T}}$，$\mathbf{d}_3=[0,-1]^{\mathrm{T}}$ 是否是初始点处的下降方向？是否是可行方向？是否是可行下降方向？

解：有 $\nabla f(\mathbf{x}_0)=[-1,-1]^{\mathrm{T}}$，利用式（5.7）可知，$\mathbf{d}_1$ 和 $\mathbf{d}_2$ 是下降方向，$\mathbf{d}_3$ 不是下降方向。

初始点处的起作用约束下标集为 $I=\{1\}$，只有 g_1 是 $\mathbf{x}_0$ 处的起作用约束。有 $\nabla g_1(\mathbf{x}_0)=\frac{1}{\sqrt{5}}[2,8]^{\mathrm{T}}$，利用式（5.9）可判断出 $\mathbf{d}_1$ 不是可行方向，$\mathbf{d}_2$ 和 $\mathbf{d}_3$ 是可行方向。

由此可知，只有 $\mathbf{d}_2$ 是可行下降方向，如图 E5.5 所示。

5.5 最优性必要条件（不等式约束）

初始点处的起作用约束下标集为 $I=\{1\}$，只有 g_1 是 $\mathbf{x}_0$ 处的起作用约束。有 $\nabla g_1(\mathbf{x}_0)=\frac{1}{\sqrt{5}}[2,8]^{\mathrm{T}}$，利用式（5.9）可判断出 $\mathbf{d}_1$ 不是可行方向，$\mathbf{d}_2$ 和 $\mathbf{d}_3$ 是可行方向。

由此可知，只有 $\mathbf{d}_2$ 是可行下降方向，如图 E5.5 所示。

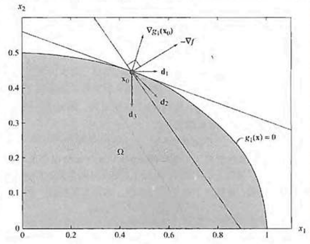

图 E5.5

明理，精工，笃学，致远

本节结束，谢谢

Thank you for listening

课程名称：最优化方法

明理，精工，笃学，致远