Operation Research Notes

顾言午

2020 年 12 月 25 日

Contents

1 Introduction

运筹学 (Operations Research, OR) 是从二十世纪三四十年代 (即二战期间) 发展起来的一门应用性很强的学科。它的研究对象是人类对各种资源的运用及筹划活动，如战争、后勤、生产规划、经营管理、资本运作、工程设计等。研究内容就是资源筹划活动中各种问题的模型化及其数学方法。

解决实际问题的第一步往往是定性分析，具体的说

- 要有一个初步概念，即主要的决策是什么?
- 在选取决策方案时，有效性的度量是什么?
- 在对各方案进行比较时，这些度量之间可能要做哪些权衡?

定量分析对管理决策过程有着重要意义...

定量分析方法的标准步骤

- 表达问题 (Definition of the problem)
- 建立模型 (Construction of the model)
- 分析求解 (Solution of the model)
- 检验和改进 (Validation of the model)
- 解的实施 (Implementation of the solution)

表达问题:

列出问题的各种要素，它们包括一些可控的变量 (或称决策变量)，不可控变量 (参数)，各变量的约束条件以及为确定最佳解决方案拟采用的目标度量。

建立模型:

即把问题中的可控变量，参数，约束及目标的关系用一定的数学模型刻画出来。

分析求解:

分析模型并用各种数学方法和手段来求解模型，进而确定解对模型的技术条件的灵敏度。

检验和改进:

将模型的解应用到实际问题中，检验解的合理性和有效性，可能产生的问题和修改模型。

解的实施:

模型的解应用于实际问题，并最终解决实际问题。

建立模型: 实际问题 $\rightarrow$ 反映实际问题的数学模型 (如: 微分方程模型，统计模型，最优化模型等)。

各种优化模型及其求解的数学方法构成了运筹学的大部分内容，这些代表性模型可供选择但并不能完全适用于所有实际问题。

一般的运筹学参考书不会着重叙述建立模型的过程，这并不是说建模不重要; 相反，建模在任何时候任何场合都是极其重要的。

在很多实际应用问题中，从数学上看都是非适定的 (ill-posed), 即解不唯一。对于这样的实际问题，人们往往通过制定相应的目标准则，然后从众多的解中选出在一定条件下最好的解。

这些正是最优化理论与方法所研究的内容，本节对最优化的基本概念作一些简要的介绍，并给出最优化建模方法的相关知识。

最优化讨论的是为找出实际问题的最优解决策略而采取的模型化及方法，其过程是:

• 先把待解决的问题用最优化形式描述为在给定的约束条件下找出使某个目标函数达到最大 (小) 的解;

• 然后再采用数学上严密的算法来求解。

在这里我们强调最优化方法作为数学工具在现实中被广泛应用的事实，大多数有代表性的最优化算法都已有程序库和软件包。但是，有效利用这些成果是以有待解决的问题已被模型化成最优化问题的形式为前提的，而这一过程并非简单。

这个模型化的过程要求: 在收集客观数据的同时要有敏锐的洞察力和综合能力; 兼蓄有针对性的领域知识和技能。

最优化方法在生产规划、经济管理、工业工程、交通运输、国防等重要领域中有广泛的应用，并已受到政府部门、科研机构和产业界的高度重视。

例如: 选址问题，是运筹学和计算几何的一个交叉分支。
选址问题关注的是设施的最佳位置以减少诸多社会成本 (如交通、时间等), 同时要考虑类似 “避免扰民或居民区附近放置危险物品” 等因素。

最优化 (optimization) 一般是指在某种条件下作出最好的决策，或者是从众多的方案中选出最好的。这种问题经常用下面的数学模型描述:

在给定的约束条件 (constraint) 下，找出一个决策变量 (decision variable) 的值，使得被称为目标函数 (objective function) 的表达愿望尺度的函数值达到最小或最大。

一般说来决策变量有多个，因此用 n 维向量 $x = (x_1, ..., x_n)^T$ 来表示，于是把问题写成如下形式

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & x \in \mathcal{S} \subset \mathbb{R}^n \end{aligned} \tag{1}$$

在此，目标函数 f 是定义在包含 $\mathcal{S}$ 的适当集合上的实值函数。进一步，$\mathcal{S}$ 是该问题变量 x 的可取值之集合，称为问题的可行域 (feasible region).

如果可行域 $\mathcal{S}=\mathbb{R}$, 则称为无约束最优化问题:

$$\min_{x\in\mathbb{R}^n} \quad f(x) \tag{2}$$

带约束最优化问题通常可写为:

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & c_i(x)=0, i\in\mathcal{E} \\ & c_i(x)\geq 0, i\in\mathcal{I} \end{aligned} \tag{3}$$

这里 $c_i(x)$ 是约束函数，$\mathcal{E}$ 和 $\mathcal{I}$ 分别是等式约束的指标集和不等式约束的指标集。

当目标函数和约束函数均为线性函数时，问题称为线性规划 (Linear Programming);

当目标函数与约束函数中至少有一个是变量 x 的非线性函数时,问题称为非线性规划 (Nonlinear Programming)。

此外，根据决策变量、目标函数和要求的不同，最优化还被分为整数规划、动态规划、网络规划、非光滑规划、随机规划、多目标规划等。

满足约束条件 $x\in\mathcal{S}$ 的 x 称为问题的可行解 (feasible solution), 如果可行解 $x^*\in\mathcal{S}$ 进一步满足

$$f(x^*)\leq f(x), \forall x\in\mathcal{S} \tag{4}$$

则称 x^* 为问题 (1) 的全局最优解 (global optimal solution). 另外，在包含可行解 $x^*\in\mathcal{S}$ 的适当邻域 $U(x^*)$ 里，成立

$$f(x^*)\leq\ f(x), \forall x\in\mathcal{S}\cap U(x^*). \tag{5}$$

此时称 x^* 为问题 (1) 的局部最优解 (local optimal solution).

不少问题的目标函数或约束条件可能很复杂，要找出全局最优解非常困难，这时我们的目标就会是求出局部最优解。

最优性条件:

问题的最优解所满足的必要或者充分条件。它为最优化问题求解算法的设计、分析提供必不可少的理论基础。

无约束问题的极值条件

一阶必要条件: 设目标函数 $f(x)$ 在点 $\bar{x}$ 处可微, 若 $\bar{x}$ 是局部极小点, 则 $\nabla f(\bar{x})=0$.

二阶必要条件: 设目标函数 $f(x)$ 在点 $\bar{x}$ 处二次可微, 若 $\bar{x}$ 是局部极小点, 则 $\nabla f(\bar{x})=0$, 并且 Hesse 矩阵 $\nabla^2 f(\bar{x})\geq 0$.

二阶充分条件: 设目标函数 $f(x)$ 在点 $\bar{x}$ 处二次可微, 若 $\nabla f(\bar{x})=0$ 且 $\nabla^2 f(\bar{x})>0$, 则 $\bar{x}$ 是局部极小点.

充要条件: 设 $f(x)$ 是定义在 $\mathbb{R}^n$ 上的可微凸函数, 则 $\bar{x}$ 是整体极小点 (全局最优解) 的充要条件是 $\nabla f(\bar{x})=0$.

约束问题的最优性条件–Kuhn-Tucker 必要条件:

设 $\bar{x}$ 为约束问题 (3) 的可行点, 记 $\mathcal{I}_e = \{i \in \mathcal{I} | c_i(\bar{x}) = 0\}$,$f$ 和 $c_i, i \in \mathcal{I}_e$ 在点 $\bar{x}$ 可微,$c_i, i \notin \mathcal{I} \setminus \mathcal{I}_e$ 在点 $\bar{x}$ 连续, $c_i(i \in \mathcal{E})$ 在点 $\bar{x}$ 连续可微, 向量集 $\{\nabla c_i(\bar{x}) | i \in \mathcal{E} \cup \mathcal{I}_e\}$ 线性无关. 如果 $\bar{x}$ 是局部最优解, 则存在数 $\lambda_i \geq 0$ 和 μ_j 使得

$$\nabla f(\bar{x}) - \sum_{i \in \mathcal{I}_e} \lambda_i \nabla c_i(\bar{x}) - \sum_{\mathcal{E}} \mu_j \nabla c_i(\bar{x}) = 0 \tag{6}$$

定义 Lagrange 函数

$$L(x, \lambda, \mu) = f(x) - \sum_{i \in \mathcal{I}} \lambda_i c_i(x) - \sum_{i \in \mathcal{E}} \mu_i c_i(x).$$

若 $\bar{x}$ 为问题局部最优解, 则存在乘子向量 $\bar{\lambda} \geq 0, \bar{\mu}$ 使得 $\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$.

此时, 一阶必要条件可表达为

$$(K-T) \begin{cases} \nabla_x L(x, \lambda, \mu) = 0 \\ c_i(x) = 0, i \in \mathcal{E} \\ c_i(x) \geq 0, i \in \mathcal{I} \\ \lambda_i c_i(x) = 0, i \in \mathcal{I} \\ \lambda_i \geq 0, i \in \mathcal{I} \end{cases} \tag{7}$$

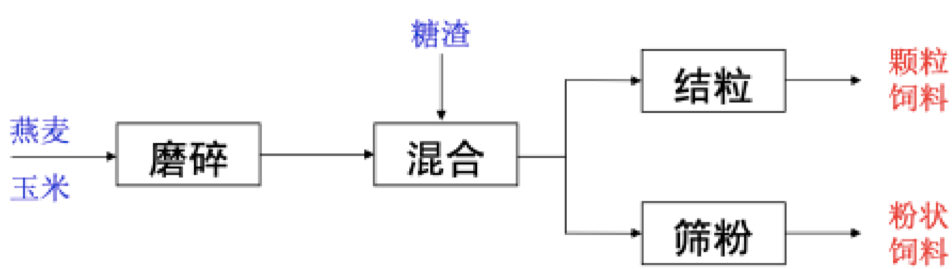
2 Linear Programming

线性规划模型

$$\begin{aligned} min \quad & c^T x \\ s.t. \quad & A_e x = b_e \\ & Ax \leq b. \end{aligned} \tag{8}$$

Example.

某饲料公司要生产两种类型的动物饲料: 粉状饲料和颗粒饲料. 生产这些饲料所需的原料由燕麦、玉米和糖渣组成. 首先需要将这些原料 (除糖渣外) 磨碎, 然后将所有原料混合形成饲料半成品. 在最后一个生产工序中, 将半成品制成颗粒状或粉末状, 从而得到最终的饲料产品.



饲料产品都要满足一些营养成分需求. 表 1 列出了各种原料的蛋白质, 脂肪和纤维含量百分比, 以及最终产品的相应含量要求.

原料	蛋白质	脂肪	纤维素
燕麦	13.6	7.1	7.0
玉米	4.1	2.4	3.7
糖渣	5.0	0.3	25.0
含量要求	≥ 9.5	≥ 2.0	≤ 6.0

表 1: 营养成分含量百分比

各种原料的可用量也有限制. 表 2 给出了各种原料的可用量以及对应的价格.

原料	可用量 (千克)	价格 (千克)
燕麦	11900	1.3
玉米	23500	1.7
糖渣	750	1.2

表 2: 原料可用量与价格

表 3 列出了各道工序的成本.

磨碎	混合	结粒	筛粉
2.5	0.5	4.2	1.7

表 3: 加工成本

如果每天需求量为 9 吨颗粒饲料,12 吨粉状饲料, 则各种原料应分别使用多少进行混合才能使得总成本最低?

在这里, 燕麦、玉米、糖渣的使用量为决策变量, 并分别设为 x_1 千克,x_2 千克,x_3 千克.

饲料配制最优化模型:

$$\begin{aligned}
min\quad & (1.3x_1 + 1.7x_2 + 1.2x_3) + 2.5(x_1 + x_2) + 0.5(x_1 + x_2 + x_3) \\
s.t.\quad & 13.6x_1 + 4.1x_2 + 5.0x_3 \geq 9.5(x_1 + x_2 + x_3), \\
& 7.1x_1 + 2.4x_2 + 0.3x_3 \geq 2.0(x_1 + x_2 + x_3), \\
& 7.0x_1 + 3.7x_2 + 25.0x_3 \leq 6.0(x_1 + x_2 + x_3), \\
& x_1 + x_2 + x_3 \geq 9000 + 12000, \\
& 0 \leq x_1 \leq 11900, \\
& 0 \leq x_2 \leq 23500, \\
& 0 \leq x_3 \leq 750.
\end{aligned} \tag{9}$$

生产出需求的 9 吨颗粒饲料和 12 吨粉状饲料, 最低成本为 150868 元. 详细的结果如下所示:

原料	燕麦 (千克)	玉米 (千克)	糖渣 (千克)
使用量	11897	8678.9	424.47

表 4: 饲料配制的最佳方案

某农场在现有土地上种植不同农作物的生产计划安排定性分析: 合理分配土地资源, 在约束条件及一定假设下获得好收成.

农作物	土地	用工	粪肥	化肥	利润
A	1	450	35	350	1500
B	1	600	25	400	1200
C	1	900	30	300	1300
总资源	100	63000	3300	33000	

表 5: 农作物生产相关数据

分析并建立模型

设农作物 A, B, C 的种植面积分别为 x_1, x_2, x_3 公顷, 则总利润是

$$1500x_1 + 1200x_2 + 1800x_3$$

资源约束:

土地限制 $x_1 + x_2 + x_3 \leq 100$,

劳力限制 $450x_1 + 600x_2 + 900x_3 \leq 63000$,

粪肥限制 $35x_1 + 25x_2 + 30x_3 \leq 3300$,

化肥限制 $350x_1 + 400x_2 + 300x_3 \leq 33000$,

此外, 实际种植面积的非负性限制, 即 $x_1, x_2, x_3 \geq 0$.

综上所述, 得到农业生产计划安排问题的数学模型 (线性规划) 为:

$$\begin{aligned} max \quad & 1500x_1 + 1200x_2 + 1800x_3 \\ s.t. \quad & x_1 + x_2 + x_3 \leq 100, \\ & 450x_1 + 600x_2 + 900x_3 \leq 63000, \\ & 35x_1 + 25x_2 + 30x_3 \leq 3300, \\ & 350x_1 + 400x_2 + 300x_3 \leq 33000, \\ & x_1, x_2, x_3 \geq 0. \end{aligned} \tag{11}$$

2.1 Description of Linear Programming

线性规划的一般形式:

$$
\begin{aligned}
min(max) \quad & z = c_1x_1 + ... + c_nx_n \\
s.t. \quad & a_{11}x_1 + ... + a_{1n}x_n \leq (=, \geq) b_1 \\
& \vdots \\
& a_{m1}x_1 + ... + a_{mn}x_n \leq (=, \geq) b_m
\end{aligned} \tag{12}
$$

线性规划问题总可以写成如下标准形式:

$$
(LP) \begin{cases} min & \sum_{j=1}^{n} c_jx_j \\ s.t. & \sum_{j=1}^{n} a_{ij}x_j = b_i, i = 1, ..., m \\ & x_j \geq 0, j = 1, ..., n \end{cases} \tag{13}
$$

或者用矩阵表示为:

$$
(LP) \begin{cases} min & c^Tx \\ s.t. & Ax = b \\ & x \geq 0 \end{cases} \tag{14}
$$

其中矩阵 $A \in \mathbb{R}^{m \times n}$,$c$ 是 n 维列向量,b 是 m 维列向量. 为了后面单纯形方法的计算需要, 我们还可以假设 $b \geq 0$.

可能的标准化步骤有:

1. 目标函数 $max\ f(x) \to min\ -f(x)$
2. 不等式约束的等式化 (引入松弛变量或者剩余变量)
3. 自由变量的非负化 $x_j = x_j' - x_j'', x_j', x_j'' \geq 0$
4. 变量上下界的处理

2.2 Conclusions of Linear Programming

结论 1:

在线性规划中, 约束条件均为线性等式及线性不等式, 所以可行域 $\mathcal{S}$ 是凸集.

极点. 设 $\mathcal{S}$ 是非空凸集, 若 $x = \lambda x^{(1)} + (1-\lambda)x^{(2)}, \lambda \in (0,1), x^{(1)}, x^{(2)} \in \mathcal{S}$, 必有 $x = x^{(1)} = x^{(2)}$, 则称 x 是凸集 $\mathcal{S}$ 的极点.

结论 2:

对于有界闭凸集, 任何一点都能表示成极点的凸组合.

结论 2 对无界集并不成立, 为此需引入极方向的概念.

设 $S \subset \mathbb{R}^n$ 是闭凸集 (可能无界),d 为非零向量. 如果对 $\mathcal{S}$ 中的每个 x 都有射线 $\{x + \lambda d | \lambda \geq 0\} \subset \mathcal{S}$, 则称向量 d 为 $\mathcal{S}$ 的方向.

又设 $d^{(1)}$ 和 $d^{(1)}$ 是 $\mathcal{S}$ 的两个方向, 若对任何正数 λ 有 $d^{(1)} \neq \lambda d^{(2)}$, 则称 $d^{(1)}$ 和 $d^{(2)}$ 是两个不同的方向.

极方向: 若 $\mathcal{S}$ 的方向 d 不能表示成该集合的两个不同方向的正的线性组合, 则称 d 为 $\mathcal{S}$ 的极方向.

多面集表示定理:

设 $\mathcal{S}=\{x|Ax=b,x\geq 0\}$ 为非空多面集, 则有

(1) 极点集非空, 且存在有限个极点 $x^{(1)},...,x^{(k)}$;

(2) 极方向集合为空集的充要条件是 $\mathcal{S}$ 有界;

(3) 若 $\mathcal{S}$ 无界, 则存在有限个极方向 $d^{(1)},...,d^{(l)}$;

(4) $x\in\mathcal{S}$ 的充要条件是

$$x=\sum_{i=1}^{k}\lambda_i x^{(i)}+\sum_{j=1}^{l}\mu_j d^{(j)}$$

其中 $\lambda_i\geq 0, i=1,...,k, \sum_{i=1}^{k}\lambda_i=1, \mu_j\geq 0, j=1,...,l.$

Pf.

(1). 令 y 为 $\mathcal{S}$ 中 0 分量最多的一点, 即

$$\#\{i|y_i=0\}\geq\max_{x\in\mathcal{S}}\#\{j|x_j=0\}.$$

下证明 y 为 $\mathcal{S}$ 的极点.

若不然, 存在 $x^{(1)}\neq x^{(2)}\in\mathcal{S},\lambda\in(0,1)$ 使得 $\lambda x^{(1)}+(1-\lambda)x^{(2)}=y$.

由于 $x^{(1)},x^{(2)}\geq 0$, 对 i 使得 $y_i=0$, 必有 $x_i^{(l)}=0, l=1,2$. 令 $\alpha=x^{(2)}-x^{(1)}$, 易见 $A\alpha=0$. 令 r 满足

$$|\frac{y_r}{\alpha_r}|=min\{\ |\frac{y_j}{\alpha_j}|\ |\ \alpha_j\neq 0\}$$

则 $y'=y-\frac{y_r}{\alpha_r}\alpha\in\mathcal{S}$, 且 $y'_r=0$; 这与 y 的定义矛盾. 因此 y 是极点. 这就证明了极点集非空.

由极点集–可行基解集等价定理, 极点与可行基解一一对应, 而可行基解的数量不多于 $\begin{pmatrix} n \\ m \end{pmatrix}$ 因此极点集有限.

(2)(3). 考察

$$\mathcal{S}_0\{x\mid Ax=0,x\geq 0,\sum_{i=1}^{n}x_i=1\}$$

及标准化后的方向集

$$\mathcal{D}=\{\frac{d}{\|d\|_1}\mid d\text{为 }\mathcal{S}\text{ 的方向}\}$$

可以看出 $\mathcal{S}_0=\mathcal{D}$, 因此 $\mathcal{D}$ 是多面集. 下证 $\mathcal{S}$ 标准化后的极方向与 $\mathcal{D}$ 的极点一一对应.

设 d 为 $\mathcal{D}$ 的极点, 并假设 $d=\lambda_1 d^{(1)}+\lambda_2 d^{(2)},\lambda_1,\lambda_2>0$. 因为

$$d=\lambda_1\|d^{(1)}\|_1\frac{d^{(1)}}{\|d^{(1)}\|_1}+\lambda_2\|d(2)\|_1\frac{d^{(2)}}{\|d^{(2)}\|_1}$$

我们不妨直接假设 $\|d^{(j)}\|_1=1, j=1,2$. 于是

$$1=\|d\|_1=\lambda_1\|d^{(1)}\|_1+\lambda_2\|d^{(2)}\|_1=\lambda_1+\lambda_2$$

即 d 是 $\mathcal{D}$ 上两点的凸组合. 由 d 为极点知 $d^{(1)} = d^{(2)} = d$, 因此 d 是 $\mathcal{S}$ 的极方向.

设 d 为 $\mathcal{S}$ 的极方向, 并且 $\|d\|_1 = 1$. 若有 $d^{(1)}, d^{(2)} \in \mathcal{D}$ 及 $\lambda \in (0,1)$, 使得 $d = \lambda d^{(1)} + (1-\lambda)d^{(2)}$, 因为这是两个方向的正线性组合, 因此有 $d = d^{(1)} = d^{(2)}$, 故 d 是 $\mathcal{D}$ 的极点.

若 $\mathcal{S}_0$ 为空集, 则任意 x 满足 $Ax = 0$ 且 $\|x\|_1 = 1$ 都具有负分量. 令

$$\delta = -max\{\sum_{i:x_i<0} x_i \mid Ax = 0, \|x\|_1 = 1\}$$

因为 $\{x \mid Ax = 0, \|x\|_1\}$ 是一个有界闭集, 因此最大值可以取到并严格大于 $0.\mathcal{S}$ 非空, 设 $y \in \mathcal{S}$, 则

$$S \subseteq y + \{\lambda x \mid Ax = 0, \|x\|_1 = 1, \lambda \in [0, \frac{\|y\|_1}{\delta}]\}$$

即 $\mathcal{S}$ 有界.

因此,$\mathcal{S}$ 无界则有 $\mathcal{S}_0$ 非空. 由 (1) 知 $\mathcal{S}_0$ 有极点, 因此 $\mathcal{S}$ 有极方向. 进一步, 极点个数有限即得极方向个数有限. 若极方向集合非空, 则 $\mathcal{S}$ 显然是无界的.

(4) 充分性显然.

必要性: 对 $\mathcal{S}$ 中元素的正分量个数进行归纳.

由 (1) 知正分量个数最少的一点 (零分量最多的一点) 为极点.

假设正分量个数小于 t 的点都具有 (1) 表示. 设 x 的正分量个数等于 t, 不妨设

$$x = (x_1, ..., x_t, 0, ..., 0)$$

若 $x_1, ..., x_t$ 对应 A 中的列向量 $a_1, ..., a_t$ 线性无关, 则由 $rank(A) = m$, 可以再选 $m-t$ 列构成 A 的一个可逆子矩阵 $B = (a_1, ..., a_m)$, 知 x 是可行基解, 亦为极点.

若 $a_1, ..., a_t$ 线性相关, 不妨设 $a_{s+1}, ..., a_t$ 为它们的最大线性无关组, 则 a_1 具有表示

$$a_1 = \mu_{s+1}a_{s+1} + ... + \mu_t a_t$$

在 A 的其余列中选择 $m + s-t$ 列构成可逆阵 B, 不妨记为

$$B = (a_{s+1}, ..., a_k, a_{k+1}, ..., a_{s+m})$$

令 $d_N = (1, 0, ..., 0), d_B = -B^{-1}Nd_N$, 构成的 d 满足 $Ad = 0$. 于是

$$Ad = a_1 + d_2a_2 + ... + d_na_n = a_1 + d_{s+1}a_{s+1} + ... + d_{s+m}a_{s+m} = 0$$

即

$$a_1 = -d_{s+1}a_{s+1} - ... - d_{s+m}a_{s+m}$$

代入上上式得

$$(\mu_{s+1} + d_{s+1})a_{s+1} + ... + \mu_t + d_t)a_t + d_{t+1}a_{t+1} + ... + d_{s+m}a_{s+m} = 0$$

故 $d_{t+1} = ... = d_{s+m} = 0$

- 若 $d \geq 0$, 令

$$\theta = \min_{1\leq i\leq t}\{\frac{x_i}{d_i} \mid d_i > 0\} = \frac{x_l}{d_l} > 0$$

并令 $x' = x - \theta D$, 则 x' 为可行解,$x'_l = 0$. 故 x' 前 k 个分量中至少有一个为 0, 而后 $n-t$ 个分量等于 0, 因此由归纳假设,

$$x' = \sum_{i=1}^{k} \lambda_i x^{(i)} + \sum_{j=1}^{l} \mu_j d^{(j)}$$

而 $\frac{d}{\|d\|_1}$ 为 $\mathcal{S}_0$ 上的一点, 由结论 2, 有界多面集内任意一点可以表示为极点的凸组合, 故 θd 可以表示为极方向的正线性组合. 故 $x = x' + \theta d$ 满足假设.

- 若 d 中有分量小于 0, 则令

$$\theta_1 = \min_{1\le i\le t}\{\frac{x_i}{d_i} \mid d_i > 0\}$$

$$\theta_2 = \min_{1\le i\le t}\{\frac{x_i}{-d_i} \mid d_i < 0\}$$

并令 $x' = x - \theta_1 d, x'' = x + \theta_2 d$, 则 x', x'' 的正分量个数都小于 k, 由归纳假设

$$x' = \sum_{i=1}^{k} \lambda_i^{(1)} x^{(i)} + \sum_{j=1}^{l} \mu_j^{(1)} d^{(j)}$$

$$x'' = \sum_{i=1}^{k} \lambda_i^{(2)} x^{(i)} + \sum_{j=1}^{l} \mu_j^{(2)} d^{(j)}$$

又由

$$x = \frac{\theta_2}{\theta_1+\theta_2} x' + \frac{\theta_1}{\theta_1+\theta_2} x''$$

得

$$x = \sum_{i=1}^{k} \lambda'_i x^{(i)} + \sum_{j=1}^{l} \mu'_j d^{(j)}$$

其中

$$\lambda_i = \frac{\theta_2 \lambda_i^{(1)} + \theta_1 \lambda_i^{(2)}}{\theta_1+\theta_2} \ge 0$$

$$\sum_i \lambda'_i = \frac{\theta_2}{\theta_1+\theta_2}\sum_i \lambda_i^{(1)} + \frac{\theta_1}{\theta_1+\theta_2}\sum_i \lambda_i^{(2)} = 1$$

$$\mu'_i = \frac{\theta_2 \mu_i^{(1)} + \theta_1 \mu_i^{(2)}}{\theta_1+\theta_2} \ge 0$$

即对 x 假设成立.

由数学归纳法原理, 证毕.

考虑标准形式

$$(LP)\begin{cases} min & c^T x \\ s.t. & Ax = b \\ & x \ge 0. \end{cases} \tag{15}$$

设可行域 $\mathcal{S} = \{x \mid Ax = b, x \ge 0\}$ 的极点为 $x^{(1)}, ..., x^{(k)}$, 极方向为 $d^{(1)}, ..., d^{(l)}$. 根据表示定理, 任何可行点 x 可表示为

$$x = \sum_{i=1}^{k} \lambda_i x^{(i)} + \sum_{j=1}^{l} \mu_j d^{(j)}.$$

将其代入标准型中得到:

$$\begin{aligned} min \quad & c^Tx = \sum_{i=1}^{k}\lambda_i(c^Tx^{(i)}) + \sum_{j=1}^{l}\mu_j(c^Td^{(j)}) \\ s.t. \quad & \lambda_i \geq 0, i=1,...,k, \sum_{i=1}^{k}\lambda_i \geq 0, j=1,...,l \end{aligned} \tag{16}$$

最优解的约定

可行域为空 $\Rightarrow$ 无解

可行域有界 $\Rightarrow$ 唯一解或者无穷多解

可行域无界 $\Rightarrow$ 唯一解或者无穷多解或者 $-\infty$

我们把"唯一解"和"无穷多解"称为模型存在最优解, 而把"无界解 $-\infty$"归入不存在最优解的情形.

结论 3:

标准形式线性规划模型的可行域非空, 存在有界最优解当且仅当所有 $c^Td^{(j)}$ 为非负数 (其中 $d^{(j)}$ 为可行域的极方向). 进一步有, 若存在有界最优解, 则目标函数的最优值一定能在某个极点 (最优极点) 处达到.

设线性规划标准形式

$$(LP)\begin{cases} min & c^Tx \\ s.t. & Ax=b \\ & x \geq 0. \end{cases} \tag{17}$$

假设 $A=(B,N)$, 其中 B 是 m 阶可逆矩阵 (不失一般性). 同时记 $x=(x_B^T,x_N^T)$, 其中 x_B 的分量与 B 中的列对应,x_N 的分量与 N 中的列对应.

这样 $Ax=b$ 即可写成

$$(B,N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} = b$$

即 $Bx_B+Nx_N=b \Rightarrow x_B=B^{-1}b–B^{-1}Nx_N$.

基解/基矩阵:

在上式中,x_N 的分量就是线性方程组 $Ax=b$ 的自由变量. 特别地令 $x_N=0$, 则得到解

$$x=\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

称为方程组的一个基解, 对应的 B 称为基矩阵.x_B 的各分量称为基变量,x_N 的各分量称为非基变量.

可行基解/可行基矩阵:

又若 $B^{-1}b \geq 0$, 则称 $x=\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ 为 (LP) 的可行基解, 相应的称 B 为可行基矩阵, $x_B=\begin{pmatrix} x_{B_1} \\ \vdots \\ x_{B_m} \end{pmatrix}$ 为一组可行基变量.

极点集–可行基解集等价定理

设线性规划标准型 (LP) 的可行域 $\mathcal{S}=\{x \mid Ax=b, x\geq 0\}$, 且 $rank(A_{m\times n})=m$, 则可行域 $\mathcal{S}$ 的极点集与 (LP) 的可行基解集是等价的.

Pf.

可行基解集 $\Rightarrow$ 极点集

令 $x=\begin{pmatrix} x_B \\ x_N \end{pmatrix}$ 为可行基解, 即 $x_N=0$,B 可逆. 设 $x^{(1)}, x^{(2)}\in\mathcal{S}$,$\lambda\in(0,1)$, 使得 $\lambda x^{(1)}+(1-\lambda x^{(2)}=x$. 由于 $x^{(1)}, x^{(2)}\geq 0$, 而 $x_N=0$, 故 $x_N^{(1)}=x_N^{(2)}=0$, 于是 $x^{(1)}+N=B^{-1}=x_B^{(2)}=x_B$, 即 $x^{(1)}=x^{(2)}=x$, 这说明 x 是一个极点.

极点集 $\Rightarrow$ 可行基解集

令 x 为一个极点. 令 $x_{p_1}, x_{p_2}, ..., x_{p_k}$ 为 x 中非零分量, 对应 A 中 $a_{p_1}, ..., a_{p_k}$ 列.

$$\text{断言}: a_{p_1}, ..., a_{p_k}\text{线性无关}.$$

若 $\exists\lambda_1, ..., \lambda_k$ 不全为 0 使得 $\sum_{i=1}^k \lambda_i a_{p_i}=0$, 令 $x_{p_i}^{(1)}=x_{p_i}+\delta\lambda_i, x_{p_i}^{(2)}=x_{p_i}-\delta\lambda_i, i=1, ..., k$,$x^{(1)}, x^{(2)}$ 其余分量均为 0, 那么由于 $x_{p_i}>0$, 取 $|\delta|$ 足够小, 有 $Ax^{(l)}=b, x^{(l)}\geq 0, l\geq 0, l=1,2$, 且 $\frac{1}{2}x^{(1)}+\frac{1}{2}x^{(2)}=x$, 这与 x 为极点矛盾. 这就证明了断言.

由断言可知 $k\leq m$. 在 x 的其余分量中选择 $m-k$ 个分量, 使得 A 对应的列向量 $a_{p_{k+1}}, ..., a_{p_m}$ 与 $a_{p_1}, ..., a_{p_k}$ 线性无关 (由 $rank(A)=m$ 保证. 令 $B=(a_{p_1}, ..., a_{p_m})$, 则 $B\begin{pmatrix} x_{p_1} \\ \vdots \\ x_{p_m} \end{pmatrix}=b$. 而且 $x_i=0, \forall i\notin\{p_1, ..., p_m\}$, 故 x 是可行基解.

由凸集表示定理和极点–可行基解等价定理, 我们易得:

当线性规划问题存在最优解时, 目标函数的最优值一定能在可行域的某个极点处达到, 即 (LP) 存在最优解时, 则一定存在一个可行基解是最优解.

这样, 线性规划模型的求解 (最优解) 归结为求最优可行基解. 这一思想正是单纯形方法的基本出发点. 但可行基解的个数往往很多, 不宜一一枚举. 该采取何种策略? 而这正是单纯形算法的实质...

另外, 我们在前面给出了凸集表示定理, 其结论之一是“极点的存在有限性”, 在这里将以另一种方式提出.

可行基解存在性定理:

(LP) 有可行解, 则一定存在可行基解.

2.3 Simplex Method

2.4 Duality Theory

针对任意一个线性规划模型, 可以定义一个称之为**对偶问题** (Dual Problem) 的模型.

对偶理论将揭示原问题与对偶问题之间的内在联系, 为进一步深入研究线性规划的求解算法提供理论依据.

原问题

$$(LP)\begin{cases} min & c^Tx \\ s.t. & Ax \geq b \\ & x \geq 0. \end{cases} \tag{33}$$

对偶问题

$$(DP)\begin{cases} max & b^Tw \\ s.t. & A^Tw \leq c \\ & w \geq 0. \end{cases} \tag{34}$$

对偶关系描述:

(1) 目标函数极小 $(min) \longleftrightarrow$ 极大 (max)

(2) 一方问题的目标函数系数向量成为另一方问题的约束条件右边常数向量

(3) 约束条件矩阵 $A \longleftrightarrow A$, 且约束条件的不等号方向相反

(4) 原–对偶问题的变量都带有非负约束

对偶问题 (等价形式):

$$(DP)\begin{cases} min & (-b)^Tw \\ s.t. & (-A)^Tw \geq -c \\ & w \geq 0. \end{cases} \tag{35}$$

对偶问题的对偶:

$$(DDP)\begin{cases} max & (-c)^Tx \\ s.t. & (-A)^Tx \leq -b \\ & x \geq 0. \end{cases} \tag{36}$$

标准型对偶:

$$\begin{aligned} min \quad & c^Tx \\ s.t. \quad & Ax = b \\ & x \geq 0. \end{aligned} \quad \Rightarrow \quad \begin{aligned} min \quad & c^Tw \\ s.t. \quad & Ax \geq b \\ & -Ax \geq -b \\ & x \geq 0. \end{aligned} \tag{37}$$

$$\begin{aligned} max \quad & b^Tw \\ s.t. \quad & A^Tw \leq c \end{aligned} \quad \Leftarrow \quad \begin{aligned} max \quad & b^Tw' - b^Tw'' \\ s.t. \quad & A^Tw' - A^Tw'' \leq c \\ & w' \geq 0, w'' \geq 0 \end{aligned} \tag{38}$$

一般式对偶:

$$\begin{aligned} min \quad & c^T x \\ s.t. \quad & A_G x \geq b_G \\ & A_E x = b_E \\ & A_L x \leq b_L \\ & x \geq 0. \end{aligned} \tag{39}$$

$$\begin{aligned} max \quad & b_G^T w_G + b_E^T w_E + b_L^T w_L \\ s.t. \quad & A_G^T w_G + A_E^T w_E + A_L^T w_L \leq c \\ & w_G \geq 0, w_E \ free, w_L \leq 0. \end{aligned} \tag{40}$$

原问题与对偶问题之间有着内密切的内在关系, 记 x 和 w 分别为 (LP) 和 (DP) 的任意可行解.

弱对偶定理:

任意线性规划问题 (LP) 及其对偶问题 (DP) 之间成立以下关系

(a) x 是 (LP) 问题的可行解,w 是 (DP) 问题的可行解, $\Rightarrow c^T x \geq b^T w$.

(b) x 是 (LP) 问题的可行解 $\Rightarrow c^T x \geq$ 问题 (DP) 的最大值 $\geq b^T \omega$.

w 是 (DP) 问题的可行解 $\Rightarrow b^T w \leq$ 问题 (LP) 的最小值 $\leq c^T x$.

(c) x 是 (LP) 问题的可行解,w 是 (DP) 问题的可行解, 且 $c^T x = b^T w \Rightarrow$

x 是 (LP) 问题的最优解,w 是 (DP) 问题的最优解.

(d) 问题 (LP) 无界 $\Rightarrow$ 问题 (DP) 无可行解.

问题 (DP) 无界 $\Rightarrow$ 问题 (LP) 无可行解.

最优解存在性定理:

若 (LP) 与 (DP) 都有可行解, 则它们均存在最优解.

对偶定理:

原问题 (LP) 和对偶问题 (DP) 只要一方有最优解, 则另一方也有最优解, 且此时两方的最优值一致.

Pf. 即证明

$$(LP)\begin{cases} min & c^T x \\ s.t. & Ax \geq b \\ & x \geq 0. \end{cases} \qquad (DP)\begin{cases} max & b^T w \\ s.t. & A^T w \leq c \\ & w \geq 0. \end{cases}$$

先将 (LP) 标准化

$$(LP')\begin{cases} min & c^T x \\ s.t. & (A - I)\begin{pmatrix} x \\ y \end{pmatrix} = b \\ & x, y \geq 0. \end{cases}$$

令 $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ 为 (LP') 的最优解, 并考虑 $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ 对应的可行基矩阵分解 B, N.

对任意 $j \in N$, 有 $c_B^T B^{-1} a_j - c_j \le 0$; 对任意 $j \in B$ 有 $c_B^T B^{-1} a_j = c_j$.(对 $j > n$, 记 $c_j = 0, a_j = -e_{j-n}$) 总之有

$$c_B^T B^{-1} a_j - c_j \ 0$$

令 $w = (c_B^T B^{-1})^T$, 则对 $1 \le j \le n$

$$w^T a_j - c_j \le 0 \Leftrightarrow a_j^T w \le c_j$$

$$\Rightarrow A^T w \le c$$

对 $n+1 \le j \le m$

$$w^T a_j \le 0 \Leftrightarrow -w_j \le 0$$

$$\Rightarrow w \ge 0$$

因此 w 是 (DP) 的一个可行解. 又

$$b^T w = w^T b = c_B^T B^{-1} b = \begin{pmatrix} c^T & 0 \end{pmatrix} \begin{pmatrix} x^* \\ y^* \end{pmatrix} = c^T x^*$$

由弱对偶定理知 w 为 (DP) 的最优解.

由 (DP) 最优解推出 (LP) 最优解的过程类似, 不再赘述.

线性规划的 Kuhn-Tucker 条件为:

$$(K-T) \quad \begin{cases} Ax - b \ge 0 \\ x \ge 0 \\ A^T w - c \le 0 \\ w \ge 0 \\ x^T (A^T w - c) = 0 \\ w^T (Ax - b) = 0 \end{cases} \tag{42}$$

互补松弛定理:

设 $\bar{x}$ 和 $\bar{w}$ 分别是 (LP) 和 (DP) 的可行解, 那么 $\bar{x}$ 和 $\bar{w}$ 是对应问题最优解的充要条件是:

$$\begin{cases} \bar{x}^T (A^T \bar{w} - c) = 0 \\ \bar{w}^T (A\bar{x} - b) = 0 \end{cases} \tag{41}$$

Pf.

$$c^T \bar{x} = \bar{\omega}^T A \bar{x} = \bar{x}^T A^T \bar{\omega} = b^T \bar{\omega}$$

因此 $\bar{x}$ 和 $\bar{\omega}$ 为最优解.

利用互补松弛定理, 当知道一个问题的最优解时, 可求出其对偶问题的最优解.

$$\begin{aligned} max \quad & w_1 + 2w_2 \\ s.t. \quad & 3w_1 + w_2 \le 2 \\ & -w_1 + 2w_2 \le 3 \\ & w_1 - 3w_2 \le 1 \\ & w_1 \ge 0, w_2 \ge 0. \end{aligned}$$

的最优解为 $\bar{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1/7 \\ 11/7 \end{pmatrix}$, 则其对偶问题

$$\begin{aligned} min \quad & 2x_1 + 3x_2 + x_3 \\ s.t. \quad & 3x_1 - x_2 + x_3 \ge 1 \\ & x_1 + 2x_2 - 3x_3 \le 2 \\ & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0. \end{aligned}$$

的最优解为 $\bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4/7 \\ 5/7 \\ 0 \end{pmatrix}$.

2.5 Sensitivity Analysis

实际问题中, 很多时候是基于某些采集数据来决定模型的系数. 在这种情况下, 势必会出现系数的扰动及引起的变化.

所谓灵敏度分析就是利用解一个问题得到的结果, 研究当系数有微小变化时最优解的反应.

为简单起见, 考虑问题

$$\begin{aligned} min \quad & c^T x \\ s.t. \quad & Ax = b \\ & x \ge 0. \end{aligned} \tag{43}$$

约束条件右边常数向量作微小变化

$$b + \Delta b = (b_1 + \Delta b_1, ..., b_m + \Delta b_m)^T,$$

得到新的线性规划问题

$$\begin{aligned} min \quad & c^T x \\ s.t. \quad & Ax = b + \Delta b \\ & x \ge 0. \end{aligned} \tag{44}$$

现假设对问题 (43) 利用单纯形法已得到其最优解

$$x^* = \begin{pmatrix} x_B^* \\ x_N^* \end{pmatrix} = \begin{pmatrix} B^{-1} \\ 0 \end{pmatrix}$$

这时对于最优基 B, 根据可行性条件有

$$B^{-1}b \geq 0, \tag{45}$$

根据最优性条件有

$$c_N^T - c_B^T B^{-1} N \geq 0, \tag{46}$$

当 b 变为 $b + \Delta b$ 后, 如果还成立

$$B^{-1}(b + \Delta b) \geq 0, \tag{47}$$

则问题 (43) 的最优基 B 也是问题 (44) 的最优基. 此时问题 (44) 的最优解成为

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}(b + \Delta b) \\ 0 \end{pmatrix}$$

进一步, 记问题 (43) 的目标函数最优值为 $z(b)$, 问题 (44) 的目标函数最优值为 $z(b + \Delta b)$, 则有

$$\begin{aligned} z(b) &= c_B^T B^{-1} b \\ z(b + \Delta b) &= c_B^T B^{-1}(b + \Delta b) \end{aligned} \tag{48}$$

故成立

$$z(b + \Delta b) - z(b) = c_B^T B^{-1} \Delta b \tag{49}$$

事实上,$B^{-T} c_B$ 是对应于基矩阵 B 的单纯形乘子向量 w. 根据对偶理论, 它即为问题 (43) 的对偶问题的最优解.

因此, 记 $w^* = B^{-T} c_B$, 则上式可写成

$$z(b + \Delta b) - z(b) = (w^*)^T \Delta b = \sum_{i=1}^{m} w_i^* \Delta b_i. \tag{50}$$

考虑 $\Delta b_i \to 0$ 时的极限, 就可以得到

$$\frac{\partial z(b)}{\partial b_i} = w_i^*, (i = 1, ..., m) \tag{51}$$

这里, 实际上给出了对偶问题最优解的经济意义 (影子价格)!

针对线性规划问题的多项式时间算法:

• 椭球算法, 最先在理论上明确线性规划问题恒在多项式时间内可解, 因而它是意义深远的.

L.G. Khachiyan. A polynomial algorithm in linear programming. Soviet Mathematics Doklady, Vol. 20, 1979: 191-194.

• 内点算法, 作为可替代单纯形法的有效方法而引入注目, 同时它可被推广并应用到不限于线性规划的更一般问题.

N. Karmarkar. A new polynomial time algorithm for linear programming. Combinatorica, Vol. 4, 1984: 373-395.

3 Network Optimization

运输问题, 最短路问题, 最大流问题, 最小成本流问题等有关网络最优化通常应用于解决实际问题, 并具有重要的经济意义.

就数学模型而言, 它们是线性规划的几个重要特例. 针对线性规划模型已有多项式时间算法, 因为网络模型的特殊数学结构, 利用其结构特性还可以设计出效率更高的求解算法.

3.1 Transport Model

运输模型是一类特殊的线性规划, 即从货源地 (如生产厂家) 装运货物到目的地 (如经销商仓库), 其目标是确定运输表使得总运输成本最小并满足供应和需求的限制条件. 运输模型可扩展应用于其他领域, 包括投资控制, 工作调度, 人员指派等.

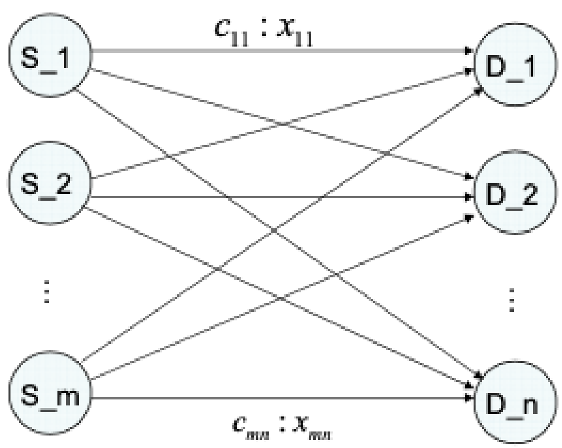
经典的运输问题:

工厂 i 的货物量 $s_i, i = 1, 2, ..., m$.

需求点 j 的需求量 $d_j, j = 1, 2, ..., n$.

从工厂 i 到需求点 j 的单位货运费用 c_{ij} 及其发货量 x_{ij}.

选取一个能使运输总费用达到最小的路径规划.



运输问题的数学描述

$$
\begin{aligned}
min \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\
s.t. \quad & \sum_{j=1}^{n} x_{ij} \le s_i \ , \ i = 1, ..., m \\
& \sum_{i=1}^{m} x_{ij} \ge d_i \ , \ i = 1, ..., n \\
& x_{ij} \ge 0 \ , \ i = 1, ..., m \ \ j = 1, ..., n
\end{aligned} \tag{52}
$$

	Denver	Miami
Los Angles	1000	2690
Detroit	1250	1350
New Orleans	1275	850

表 6: Mileage Chart

	Denver($j=1$)	Miami($j=2$)
Los Angles($i=1$)	\$80	\$215
Detroit($i=2$)	\$100	\$108
New Orleans($i=3$)	\$102	\$68

表 7: Transportation Cost per Car

$$\begin{aligned} min \quad & z = 80x_{11} + 215x_{12} + 100x_{21} + 108x_{22} + 102x_{31} + 68x_{32} \\ s.t. \quad & x_{11} + x_{12} \leq 1000 \\ & x_{21} + x_{22} \leq 1500 \ (Los\ Angeles) \\ & x_{31} + x_{32} \leq 1200 \ (New\ Orleans) \\ & x_{11} + x_{21} + x_{31} \geq 2300 \ (Denver) \\ & x_{12} + x_{22} + x_{32} \geq 1400 \ (Miami) \\ & x_{ij} \geq 0,\ i = 1,2,3,\ j = 1,2 \end{aligned} \tag{53}$$

考虑产销平衡的情形

$$\begin{aligned} min \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\ s.t. \quad & \sum_{j=1}^{n} x_{ij} = s_i \ , \ i = 1, ..., m \\ & \sum_{i=1}^{m} x_{ij} = d_i \ , \ i = 1, ..., n \\ & x_{ij} \geq 0 \ , \ i = 1, ..., m \ \ j = 1, ..., n \end{aligned} \tag{54}$$

运输问题的对偶

$$\begin{aligned} max \quad & \sum_{i=1}^{m} s_i u_i + \sum_{j=1}^{n} d_j v_j \\ s.t. \quad & u_i + v_j \leq c_{ij},\ i = 1, ..., m,\ j = 1, ..., n \\ & \forall u_i, v_j \end{aligned} \tag{55}$$

利用对偶关系:
对基变量 x_{ij} 而言,$u_i + v_j = c_{ij}$, 即 $\sigma_{ij} = u_i + v_j - c_{ij} = 0$.

	Denver	Miami	**Supply**
Los Angles	x_{11} 80	x_{12} 215	**1000**
Detroit	x_{21} 100	x_{22} 108	**1500**
New Orleans	x_{31} 102	x_{32} 68	**1200**
Demand	**2300**	**1400**	

表 8: Mileage Chart

对非基变量 x_{ij} 而言, 若 $\sigma_{ij} = u_i + v_j - c_{ij} \leq 0$, 已对偶可行;
若 $\sigma_{ij} = u_i + v_j - c_{ij} > 0$, 非对偶可行, 则引进基.

运输模型的单纯形法求解步骤:

1. 选取一组 $m+n-1$ 个路径, 作为初始可行基解.
2. 检验当前解是否可改进, 如果可改进, 则引进一个非基变量进行步 3, 否则停止.
3. 当把步 2 中挑选的变量引进时, 确定哪个路径应当由基解中退出.
4. 调整其他基本路径的流量 (满足可行性), 返回到步 2.

Table: **某公司的运输表**

	D(1)	D(2)	D(3)	D(4)	**Supply**
S(1)	10 x_{11}	2 x_{12}	20 x_{13}	11 x_{14}	**15**
S(2)	12 x_{21}	7 x_{22}	9 x_{23}	20 x_{24}	**25**
S(3)	4 x_{31}	14 x_{32}	16 x_{33}	18 x_{34}	**10**
Demand	**5**	**15**	**15**	**15**	

Table: **初始可行基解** (Northwest-Corner Starting Solution)

	D(1)	D(2)	D(3)	D(4)	**Supply**
S(1)	10 5	2 10	20	11	**15**
S(2)	12	7 5	9 15	20 5	**25**
S(3)	4	14	16	18 10	**10**
Demand	**5**	**15**	**15**	**15**	

Table: 迭代-1

	$v_1 = 10$	$v_2 = 2$	$v_3 = 4$	$v_4 = 15$	Supply
$u_1 \equiv 0$	10 5	2 10	20 [-16]	11 [4]	**15**
$u_2 = 5$	12 [3]	7 5	9 15	20 5	**25**
$u_3 = 3$	4 [9]	14 [-9]	16 [-9]	18 10	**10**
Demand	**5**	**15**	**15**	**15**	

Table: 迭代-1 （回路）

	$v_1 = 10$	$v_2 = 2$	$v_3 = 4$	$v_4 = 15$	Supply
$u_1 \equiv 0$	10 $5-\theta$	2 $10+\theta$	20 [-16]	11 [4]	**15**
$u_2 = 5$	12 [3]	7 $5-\theta$	9 15	20 $5+\theta$	**25**
$u_3 = 3$	4 θ [9]	14 [-9]	16 [-9]	18 $10-\theta$	**10**
Demand	**5**	**15**	**15**	**15**	

Table: 迭代-2

	$v_1 = 1$	$v_2 = 2$	$v_3 = 4$	$v_4 = 15$	Supply
$u_1 \equiv 0$	10 [-9]	2 15	20 [-16]	11 [4]	**15**
$u_2 = 5$	12 [-6]	7 0	9 15	20 10	**25**
$u_3 = 3$	4 5	14 [-9]	16 [-9]	18 5	**10**
Demand	**5**	**15**	**15**	**15**	

Table: 迭代-2（回路）

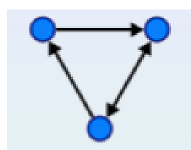
	$v_1 = 1$	$v_2 = 2$	$v_3 = 4$	$v_4 = 15$	Supply
$u_1 \equiv 0$	10 [-9]	2 $15 - \theta$	20 [-16]	11 θ [4]	**15**
$u_2 = 5$	12 [-6]	7 $0 + \theta$	9 15	20 $10 - \theta$	**25**
$u_3 = 3$	4 5	14 [-9]	16 [-9]	18 5	**10**
Demand	**5**	**15**	**15**	**15**	

Table: 迭代-3 （最优解）

	$v_1 = -3$	$v_2 = 2$	$v_3 = 4$	$v_4 = 11$	Supply
$u_1 \equiv 0$	10 [-13]	2 **5**	20 [-16]	11 **10**	**15**
$u_2 = 5$	12 [-10]	7 **10**	9 **15**	20 [-4]	**25**
$u_3 = 7$	4 **5**	14 [-5]	16 [-5]	18 **5**	**10**
Demand	**5**	**15**	**15**	**15**	

3.2 Network Model

设 $G = (V, E)$ 为有向图, 其中 V 是节点的集合,E 是边的集合.



有时把某个节点作为初始点 s, 另一个作为终点 t 而特殊对待.

各边 $e \in E$ 上赋有成本 $c(e)$ 以及容量 $u(e)$, 且都取实数值. 由它们组成的 $|E|$ 维相应向量分别记为

$$c = (c(e)|e \in E), \;\; u = (u(e)|e \in E).$$

统括这些元素的 $N = (G, s, t, c, u)$ 称为网络.

最短路径问题中, 成本 $c(e)$ 解释为边 e 的长度. 最短路径问题的典型形式是:

在网络 $N = (G, s, c)$ 中求出始点 s 到其它各点 $v \in V$ 的最短路径及其长度.

另外, 定义 $w \in V$ 到 $v \in V$ 的路径 (path)

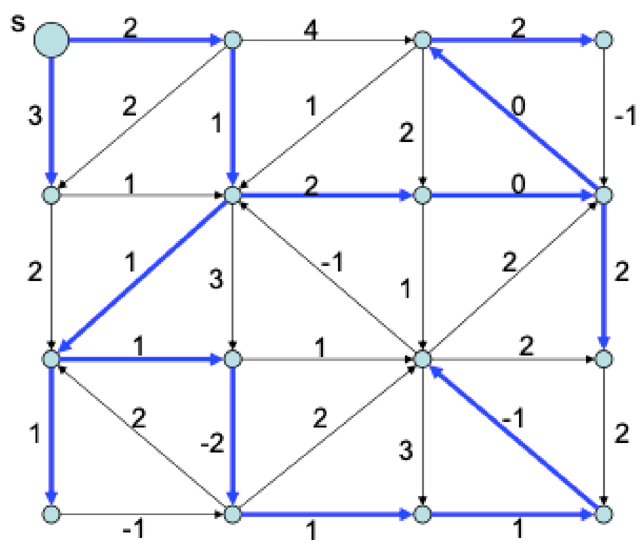
$$\pi = v_{i_0}(= w), v_{i_1}, ..., v_{i_k}(= v)$$

长度为

$$c(\pi) = \sum_{j=0}^{k-1} c(v_{i_j}, v_{i_{j+1}}).$$

当 $\mathcal{N}$ 中没有长度为负的回路时, 存在有从 s 到所有点 v 的最短路径, 它们可用以 s 为根的**最短路树**来表示.

求 s 到 v 的最短路径, 只要沿着 s 到 v 的最短路树的边走下去就行.



最短路径问题的数学描述:

$$
\begin{aligned}
min \quad & \sum_{(i,j)\in E} c_{ij}x_{ij} \\
s.t. \quad & \sum_{(s,j)\in E} x_{sj} = 1 \\
& \sum_{(k,j)\in E} x_{kj} - \sum_{(i,k)\in E} x_{ik} = 0, \forall k \in V - \{s,t\} \\
& -\sum_{(i,t)\in E} x_{it} = -1 \\
& x_{ij} \geq 0 \ , \ \forall (i,j) \in E
\end{aligned}
\tag{56}
$$

最短路径问题的对偶为:

$$
\begin{aligned}
min \quad & (y_s - y_t) \\
s.t. \quad & y_i - y_j \leq c_{ij} \ , \ \forall (i,j) \in E
\end{aligned}
\tag{57}
$$

一般网络的最短路径问题可以看成是一个线性规划模型 (事实上是一个更特殊的运输模型), 可依据对偶性构造其求解算法.

如果基于动态规划的思想, 可给出最短路径问题的强多项式时间算法. 以下仅说明具有代表性的算法之一:

Dijkstra’s algorithm

Input 有向图 $G = (V, E)$, 各边长度 $c : E \to \mathbb{R}_+$, 始点 $s \in V$.
Output 从始点到所有节点 $v \in V$ 的最短路径及其长度 $c^*(v)$.
步一 初始化: 令 $d(s) := 0, d(v) := \infty (v \in V - \{s\})$, 以及 $P := \emptyset$.
步二 迭代: 选取一个满足 $d(v^*) = min\{d(v) | v \in V - P\}$ 的节点 $v^* \in V - P$.

步三 更新:$c^*(v^*) := d(v^*), P := P \cup \{v^*\}$. 进一步对 $w \in V - P$ 的各边 $e = (v^*, w) \in E$ 作如下更新:

$$d(w) := min\{d(w), d(v^*) + c(e)\}.$$

步四 结束判定: 如果 $P = V$ 则结束计算; 否则回到步二.

计算过程中求出的 $d(v)$ 恒为从 s 到 v 的最短路径长度 $c^*(v)$ 的上界, 即 $d(v) \geq c^*(v)$.

但在已执行节点集合 $P \subset V$ 里, 具有性质 $d(v) = c^*(v), v \in P$. 于是, 当 $P = V$ 时即求出了 s 到所有节点的最短路径及其长度, 结束计算.

3.3 Max-Flow Problem

网络 $\mathcal{N} = (G, s, t, u)$ 里, 所谓从源点 s 到汇点 t 的流 (flow) 是指实数向量 $x = (x(e)|e \in E)$, 满足如下流量守恒条件:

$$\sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) = 0, v \in V - \{s, t\} \tag{58}$$

这里 $Out(v)$ 和 $In(v)$ 分别表示 G 中流出和流入 v 节点的边集合.

从 s 的纯流出量

$$x(s) \triangleq \sum_{e \in Out(s)} x(e) - \sum_{e \in In(s)} x(e)$$

与到 t 的纯流入量

$$x(t) \triangleq \sum_{e \in In(t)} x(e) - \sum_{e \in Out(t)} x(e)$$

相等, 即成立

$$x(s) = x(t). \tag{59}$$

这个 $x(s)$ 称为 x 的流值.

当流 x 进一步满足容量约束 (capacity constraint) 条件

$$0 \leq x(e) \leq u(e), e \in E \tag{60}$$

则称它为可行流 (feasible flow).

所谓最大流问题, 就是求出使流值达最大的可行流问题, 数学上可描述如下:

$$\begin{aligned} max \quad & x(s) \\ s.t. \quad & \sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) = 0, v \in V - \{s, t\} \\ & x(s) = \sum_{e \in Out(s)} x(e) - \sum_{e \in In(s)} x(e) \\ & 0 \leq x(e) \leq u(e), e \in E \end{aligned} \tag{61}$$

把一个流网络的顶点集划分成两个集合 S 和 T, 使得源点 $s \in S$ 且汇点 $t \in T$. 分割 (S, T) 的容量为

$$U(S, T) = \sum_{(i,j) \in E, i \in S, j \in T} u(i, j).$$

所谓最小割, 就是指所有分割中容量最小的一个割.

从直观上看, 割集 (S,T) 是从源点 s 到汇点 t 的必经之路, 如果该路堵塞则流从 s 无法到达 t . 于是我们可以得到下面的定理.

最大流最小割定理:

任意一个流网络的最大流量等于该网络的最小割的容量.

Pf.

设 f^* 是一个最大流, 流量为 W, 下定义点集 $\mathcal{S}'$

$$v_j \in \mathcal{S}' \quad if \begin{cases} j = s \\ v_i \in \mathcal{S}' \wedge f^*_{ij} < c_{ij} \\ v_i \in \mathcal{S}' \wedge f^*_{ji} > 0 \end{cases}$$

易得 $v_t \notin \mathcal{S}'$, 否则, 设 $v_t \in \mathcal{S}'$, 得到一条从 v_s 到 v_t 的道路 μ, 规定 v_s 到 v_t 的方向为 μ 的方向, 道路上与 μ 方向一致的边称为前向边, 不一致的称为反向边.

根据 $\mathcal{S}'$ 的定义,μ 中的前向边 (v_i,v_j) 上必有 $f^*_{ij} < c_{ij}$, 后向边 (v_i,v_j) 上必有 $f^*_{ij} > 0$. 令

$$\delta_{ij} = \begin{cases} c_{ij} - f^*_{ij}, & if\ (v_i,v_j)\text{是前向边} \\ f^*_{ij}, & if\ (v_i,v_j)\text{是后向边} \end{cases}$$

取 $\delta = min\{\delta_{ij}\}$, 易得 $\delta_{ij} > 0$

下将 f^* 修改为 f^*_1.

$$f^*_1 = \begin{cases} f^*_{ij} + \delta, & if\ (v_i,v_j)\text{是前向边} \\ f^*_{ij} - \delta, & if\ (v_i,v_j)\text{是后向边} \\ f^*_{ij}, & others. \end{cases}$$

不难验证 f^*_1 仍然满足容量限制条件与平衡条件, 仍为可行流, 但沿着 μ 方向增加了 δ,f^*_1 的总流量比 f^* 增加了 δ, 这与 f^* 为最大流矛盾. 因此 $v_t \notin \mathcal{S}'$.

于是可以得到一个割集 $(\mathcal{S}', \mathcal{S} \setminus \mathcal{S}')$, 有 $v_t \in \mathcal{S} \setminus \mathcal{S}'$.

对于割中的边 (v_i,v_j) 有

$$f^*_{ij} = \begin{cases} c_{ij}, & if\ v_i \in \mathcal{S}', v_j \in \mathcal{S} \setminus \mathcal{S}' \\ 0, & if\ v_j \in \mathcal{S}', v_i \in \mathcal{S} \setminus \mathcal{S}' \end{cases}$$

而 W 满足

$$W = \sum_{v_i \in \mathcal{S}', v_j \in \mathcal{S} \setminus \mathcal{S}'} [f^*_{ij} - f^*_{ij}] = \sum_{v_i \in \mathcal{S}', v_j \in \mathcal{S} \setminus \mathcal{S}'} = c(\mathcal{S}', \mathcal{S} \setminus \mathcal{S}')$$

最大流的流值等于最小割的容量, 证毕.

网络 $\mathcal{N} = (G,s,t,u)$ 中给定一个可行流 x, 通过构造**余网络**和**流扩充路**来迭代更新达到最大流.

从 $\mathcal{N}$ 的各边 $e = (v,w) \in E$ 出发, 按照以下规则生成边 (v,w) 或 (w,v) (根据情况二者仅取其一), 得到的边集合记为 E_x, 同时确定其容量 $\bar{u}$.

Rule 1 如果 $u(e) - x(e) > 0$, 则生成 $(v, w) \in E_x$, 并令其容量 $\bar{u}(v, w) = u(e) - x(e)$.

Rule 2 如果 $x(e) > 0$, 则生成 $(w, v) \in E_x$, 并令其容量 $\bar{u}(w, v) = x(e)$.

所得到的有向图 $G_x = (V, E_x)$ 和容量 $\bar{u}$ 合起来便是余网络 $\mathcal{N}_x = (G_x, s, t, \bar{u})$. 那么,$\mathcal{N}_x$ 中从 s 到 t 的路称为流扩充路.

引理

(流扩充路) 给定网络 $\mathcal{N} = (G, s, t, u)$, 其可行流 x 为最大流的充要条件是, 余网络 $\mathcal{N}_x$ 中不存在有流扩充路.

如果存在有流扩充路, 则 x 可修正为流值更大的流 x' .

假定在 $\mathcal{N}_x$ 中存在从 s 到 t 的路 π, 则利用

$$\Delta := min\{\bar{u}(e)|e \in\}, \tag{62}$$

对各边 $e = (v, w) \in E$ 作如下修正

$$x'(e) = \begin{cases} x(e) + \Delta, & (v, w) \in \pi(\text{由 Rule1 生成的边}) \\ x(e) - \Delta, & (w, v) \in \pi(\text{由 Rule2 生成的边}) \\ x(e), & Others \end{cases} \tag{62}$$

易证 x' 也是 $\mathcal{N}$ 的可行流, 其流值是

$$x'(s) = x(s) + \Delta$$

Exercise.

验证上述 x' 是 $\mathcal{N}$ 的可行流, 并且其流值为 $x'(s) = x(s) + \Delta$.

Pf.

首先验证其为可行流.

因为 x 为一可行流, 所以任给 $e \in E(G)$ 有 $u(e) \geq x(e) \geq 0$.

取 $e' \in E(G)$, 若 e' 不为余网络中 s 到 t 的道路 π, 则有 $x'(e) = x(e)$. 显然 $u(e) \geq x'(e) \geq 0$.

若 e' 为 π 上的正向边, 由 Δ 的定义 $\Delta = min\{\bar{u}(e)|e \in \pi\}$, 则 $\bar{u}(e') \geq \Delta \geq 0$. 而 $\bar{u}(e') = u(e') - x(e')$. 因此有

$$u(e') = x(e') + \bar{u}(e') \geq x(e') + \Delta = x'(e') \geq 0$$

若 e' 为 π 上的反向边,

$$u(e') = x(e') \geq x(e') - \Delta = x'(e') \geq 0$$

综上, 符合容量限制条件.

而对 $\forall v \in V(G) - \{s, t\}$, 设 $\alpha(v)$ 为入度的集合,$\beta(v)$ 为出度的集合.

因为 x 为可行流, 有 $\sum_{e \in \alpha(v)} x(e) - \sum_{e \in \beta(v)} x(e) = 0$

若 v 不为 π 上的顶点, 任取 $e \in \alpha(v)$ 或 $e \in \beta(v)$ 均有 $x'(e) = x(e)$. 故

$$\sum_{e \in \alpha(v)} x'(e) - \sum_{e \in \beta(v)} x'(e) = \sum_{e \in \alpha(v)} x(e) - \sum_{e \in \beta(v)} x(e) = 0$$

若 v 为 π 上的顶点, 不妨设 $\pi = s...e_1ve_2...t$,e_1, e_2 均有可能是正向边或反向边.

i. $e_1 \in \alpha(v)$ 为正向边,$e_2 \in \beta(v)$ 为正向边, 则 $x'(e_1) = x(e_1) + \Delta, x'(e_2) = x(e_2) + \Delta$.

$$\begin{aligned}\sum_{e\in\alpha(v)} x'(e) - \sum_{e\in\beta(v)} x'(e) &= \sum_{e\in\alpha(v)-\{e_1\}} x'(e) + x'(e_1) - \sum_{e\in\beta(v)-\{e_2\}} x'(e) - x'(e_2) \\ &= \sum_{e\in\alpha(v)-\{e_1\}} x(e) + x(e_1) + \Delta - \sum_{e\in\beta(v)-\{e_2\}} x(e) - x(e_2) - \Delta \\ &= \sum_{e\in\alpha(v)} x(e) - \sum_{e\in\beta(v)} x(e) \\ &=0\end{aligned}$$

ii. $e_1 \in \alpha(v)$ 为正向边,$e_2 \in \beta(v)$ 为反向边, 则 $x'(e_1) = x(e_1) + \Delta, x'(e_2) = x(e_2) - \Delta$.

$$\begin{aligned}\sum_{e\in\alpha(v)} x'(e) - \sum_{e\in\beta(v)} x'(e) &= \sum_{e\in\alpha(v)-\{e_1\}} x'(e) + x'(e_1) - \sum_{e\in\beta(v)-\{e_2\}} x'(e) + x'(e_2) \\ &= \sum_{e\in\alpha(v)-\{e_1\}} x(e) + x(e_1) + \Delta - \sum_{e\in\beta(v)-\{e_2\}} x(e) + x(e_2) - \Delta \\ &= \sum_{e\in\alpha(v)} x(e) - \sum_{e\in\beta(v)} x(e) \\ &=0\end{aligned}$$

iii. $e_1 \in \alpha(v)$ 为反向边,$e_2 \in \beta(v)$ 为正向边, 则 $x'(e_1) = x(e_1) - \Delta, x'(e_2) = x(e_2) + \Delta$.

$$\begin{aligned}\sum_{e\in\alpha(v)} x'(e) - \sum_{e\in\beta(v)} x'(e) &= \sum_{e\in\alpha(v)-\{e_1\}} x'(e) - x'(e_1) - \sum_{e\in\beta(v)-\{e_2\}} x'(e) - x'(e_2) \\ &= \sum_{e\in\alpha(v)-\{e_1\}} x(e) - x(e_1) + \Delta - \sum_{e\in\beta(v)-\{e_2\}} x(e) - x(e_2) - \Delta \\ &= \sum_{e\in\alpha(v)} x(e) - \sum_{e\in\beta(v)} x(e) \\ &=0\end{aligned}$$

iv. $e_1 \in \alpha(v)$ 为反向边,$e_2 \in \beta(v)$ 为反向边, 则 $x'(e_1) = x(e_1) - \Delta, x'(e_2) = x(e_2) - \Delta$.

$$\begin{aligned}\sum_{e\in\alpha(v)} x'(e) - \sum_{e\in\beta(v)} x'(e) &= \sum_{e\in\alpha(v)-\{e_1\}} x'(e) - x'(e_1) - \sum_{e\in\beta(v)-\{e_2\}} x'(e) + x'(e_2) \\ &= \sum_{e\in\alpha(v)-\{e_1\}} x(e) - x(e_1) + \Delta - \sum_{e\in\beta(v)-\{e_2\}} x(e) + x(e_2) - \Delta \\ &= \sum_{e\in\alpha(v)} x(e) - \sum_{e\in\beta(v)} x(e) \\ &=0\end{aligned}$$

故满足平衡条件. 综上 x' 为可行流.

设 $\pi = s...v....e_3t$,e_3 为正向边时,$x'(e_3) = x(e_3) + \Delta$

$$\begin{aligned}
x'(s) &= \sum_{e\in\alpha(v)} x'(e) - \sum_{e\in\beta(v)} x'(e) \\
&= \sum_{e\in\alpha(v)-\{e_3\}} x'(e) + x'(e_3) - \sum_{e\in\beta(v)} x'(e) \\
&= \sum_{e\in\alpha(v)-\{e_3\}} x(e) + x(e_3) + \Delta - \sum_{e\in\beta(v)} x(e) \\
&= \sum_{e\in\alpha(v)} x(e) - \sum_{e\in\beta(v)} x(e) + \Delta \\
&= x(s) + \Delta
\end{aligned}$$

e_3 为反向边时,$x'(e_3) = x(e_3) - \Delta$

$$\begin{aligned}
x'(s) &= \sum_{e\in\alpha(v)} x'(e) - \sum_{e\in\beta(v)} x'(e) \\
&= \sum_{e\in\alpha(v)-\{e_3\}} x'(e) - x'(e_3) - \sum_{e\in\beta(v)} x'(e) \\
&= \sum_{e\in\alpha(v)-\{e_3\}} x(e) - x(e_3) + \Delta - \sum_{e\in\beta(v)} x(e) \\
&= \sum_{e\in\alpha(v)} x(e) - \sum_{e\in\beta(v)} x(e) + \Delta \\
&= x(s) + \Delta
\end{aligned}$$

故有 $x'(s) = x(s) + \Delta$, 证毕.

MAX-FLOW 算法

Input 网络 $\mathcal{N} = (G, s, t, u)$, 其中有向图 $G = (V, E)$.

Output 从 s 到 t 的最大流值 f_{max}.

步一 初始化: 令 $x(e) := 0(e \in E)$ 以及 $f_{max} := 0$.

步二 余网络: 构造当前可行流 x 的余网络 $\mathcal{N}_x = (G_x, s, t, \bar{u})$.

步三 流的扩充: 如果 $\mathcal{N}_x$ 中没有流的扩充路则结束计算. 相反, 如果存在有流的扩充路, 则选取其一 π, 根据 (62) 式和 (63) 式修正 x 得到流 x' . 令 $f_{max} := f_{max} + \Delta, x := x'$ 回到步二.

3.4 Min-Cost Problem

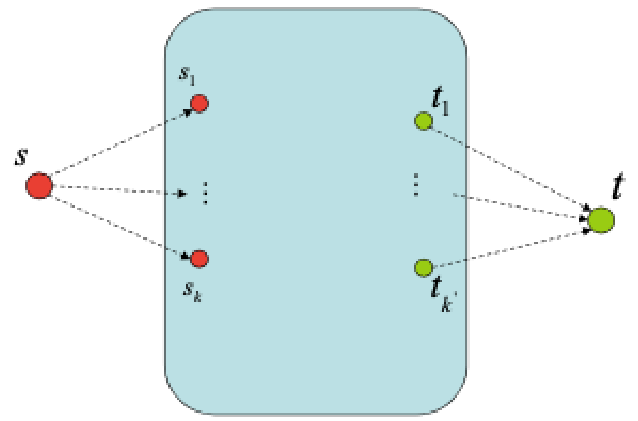
最小成本流问题的数学模型:

$$\begin{aligned}
min \quad & \sum_{e\in E} c(e)x(e) \\
s.t. \quad & \sum_{e\in Out(v)} x(e) - \sum_{e\in In(v)} x(e) = 0, v \in V - \{s,t\} \\
& \sum_{e\in Out(s)} x(e) - \sum_{e\in in(s)} x(e) = f^* \\
& 0 \leq x(e) \leq u(e), e \in E.
\end{aligned} \tag{64}$$

在最小成本流问题中, 可以有多个源点和汇点. 如果最小成本流问题具有 k 个源点 $s_i, i = 1, ..., k$ 以及各自的流出量 $x(s_i) = f^*(s_i)$, 还有 k' 个汇点 $t_j, j = 1, ..., k'$ 以及各自的流出量 $x(t_j) = f^*(t_j)$, 则可以引入哑元源点 s 和哑元汇点 t, 并添加 $(k+k')$ 条边

$$(s, s_i) : c(s, s_i) = 0, u(s, s_i) = f^*(s_i), i = 1, ..., k;$$

$$(t_j, t) : c(t_j, t) = 0, u(t_j, t) = f^*(t_j), j = 1, ..., k'.$$



最小成本流问题具有很好的模型化能力:

通过引入哑元和构造相应的边 (及其量化赋值), 可将最短路径问题和最大流问题转化为求解特殊的最小成本流问题.

(一) 通过引入哑元终点 t , 加入从 $v \in V$ 出发的边 (v, t) 且满足

$$c(v, t') = 0, u(v, t') = 1.$$

那么, 最短路径问题成为求解从 s 到 t' 的具有流值 $f^* = n$ 的最小成本流问题. 这里,$n = |V|$, 并假定原网络各边的容量全为 ∞.

(二) 通过引入哑元始点 s', 并构造如下两条边:

$$(s', s); c(s', s) = 0, u(s', s) = \infty$$

$$(s', t); c(s', t) = 1, u(s', t) = \infty$$

那么, 最大流问题成为求解从 s' 到 t 的最小成本流问题. 这里, 设定原网络里 $c(e) = 0(e \in E)$, 并取最大流值的适当上界 (比如 $\sum_{e \in E} u(e)$) 为从 s' 出发的流值 f^*.

如不考虑特定的节点为源点或者汇点, 并在各边 $e \in E$ 上引入流的下界值 $l = (l(e)|e \in E)$, 从而得到的网络 $\mathcal{N} = (G, c, l, u)$.

求在这个网络上使成本达到最小的流, 即所谓的最小成本循环流问题:

$$\begin{aligned} min \quad & \sum_{e \in E} c(e)x(e) \\ s.t. \quad & \sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) = 0, v \in V \\ & l(e) \le x(e) \le u(e), e \in E. \end{aligned} \tag{65}$$

Exercise.

通过构造说明最小成本流问题作为特殊情况包含最短路问题和最大流问题.

Pf.

最短路径问题

引入哑元终点 t', 对于 $\forall v \in V$, 加入边 (v, t'), 令 $c(v, t') = 0, u(v, t') = 1$

最短路径问题转化为求解从 s 到 t' 的具有流值 $f^* = |V|$ 的最小成本流问题, 源网络容量为 ∞

$u(v, t') = 1, f^* = |V|$ 保证了经过每个点 $v \in V$.

$c(v, t') = 0$ 保证了最短路径问题与最小成本流问题的目标函数相同.

最大流问题

引入哑元始点 s', 构造

$$(s', s); c(s', s) = 0, u(s', s) = \infty$$

$$(s', t); c(s', t) = 1, u(s', t) = \infty$$

最大流问题转化为求解 s' 到 t 的最小成本流问题, 原网络 $c(e) = 0, (e \in E)$ 并取最大流值的适当上界 (如 $\sum_{e \in E} u(e)$) 为 s' 出发的流值 f^*.

f^* 取适当上界保证足以提供原网络中所有流值的总和, 且仅 $c(s', t) = 1$, 最小成本流可求解出沿着 (s', t) 的最小值, 即可得 (s', s) 的最大值, 该值即为最大流.

Exercise.

最小成本循环流问题的模型化能力与最小成本流问题等价.

Pf.

最小成本流等价为最小成本循环流.

引入 $(t, s); c(t, s) = 0, l(t, s) = u(t, s) = f^*, \forall e \in E, l(e) = 0, u(e)$ 保持不变. 最小成本循环流目标函数 $\sum_{e \in E} c(e)x(e)$ 即与最小成本流目标函数相等, 流值均为 f^*.

最小成本循环流等价为最小成本流.

对于每个 $v \in V$,

$$\sum_{e \in Out(v)} l(e) - \sum_{e \in In(v)} l(e) \triangleq \varphi(v)$$

有三种情况.

1.$\varphi(v) > 0$, 记 v 为最小成本流的 s,$f^*(v) = \varphi(v)$

2.$\varphi(v) < 0$, 记 v 为最小成本流的 t,$f^*(v) = -\varphi(v)$

3.$\varphi(v) = 0$, 为其他结点

每条边的 $u(e)$ 更新为 $u(e) - l(e), c(e)$ 不变.

最小成本循环流目标函数变为最小成本流目标函数加上 $\sum_{e \in E} l(e)c(e)$.

最小成本循环流问题 (65) 相比于最小成本流问题, 其约束条件更加简洁, 因而具有更清晰的求解算法描述等优点.

4 Nonlinear Programming

4.1 Definition

考虑问题

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & x \in S \subset \mathbb{R}^n \end{aligned} \tag{66}$$

在此, 目标函数 f 是定义在 $\mathbb{R}^n$ 上的实值函数,S 是决策变量 x 的可取值之集合, 称为问题的**可行域** (feasible region).

最优化问题从属性上可以分为两大类: 一类是具有连续变量的问题, 另一类是离散变量的问题 (即组合优化问题).

非线性规划属于连续型最优化问题的范畴, 通常可行域 S 可由一组方程来描述, 即

$$S = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, ..., m; h_j(x) = 0, j = 1, ..., l\}.$$

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g_i(x) \geq 0, i = 1, ..., m \\ & h_j(x) = 0, j = 1, ..., l \end{aligned} \tag{67}$$

这里,$f(x), g_i(x), h_j(x)$ 都是 n 变量、实值、确定的函数, 且至少有一个是非线性的.

为求解一个非线性规划问题 (即找出其最优解), 与此相关的研究分两个方面: 一是研究最优解的性质, 二是设计有效算法来获得问题的解.

满足约束条件 $x \in S$ 的 x 称为问题的**可行解** (feasible solution), 如果可行解 $x^* \in S$ 进一步满足

$$f(x^*) \leq f(x), \forall x \in S. \tag{68}$$

则称 x^* 为问题 (66) 的**全局最优解** (global optimal solution).

另外, 在包含可行解 $x^* \in S$ 的适当邻域 $U(x^*)$ 里, 成立

$$f(x^*) \leq f(x), \forall x \in S \cap U(x^*). \tag{69}$$

此时称 x^* 为问题 (66) 的**局部最优解** (local optimal solution).

最优性条件: 问题的最优解所满足的必要或者充分条件. 最优性条件将为各种求解算法的设计、分析提供必不可少的理论基础.

一阶必要条件: 设目标函数 $f(x)$ 在点 $\bar{x}$ 处可微, 若 $\bar{x}$ 是局部极小点, 则 $\nabla f(\bar{x}) = 0$.

二阶必要条件: 设目标函数 $f(x)$ 在点 $\bar{x}$ 处二次可微, 若 $\bar{x}$ 是局部极小点, 则 $\nabla f(\bar{x}) = 0$, 并且 Hesse 矩阵 $\nabla^2 f(\bar{x}) \geq 0$.

二阶充分条件: 设目标函数 $f(x)$ 在点 $\bar{x}$ 处二次可微, 若 $\nabla f(\bar{x}) = 0$ 且 $\nabla^2 f(\bar{x}) > 0$, 则 $\bar{x}$ 是局部极小点.

充要条件: 设 $f(x)$ 是定义在 $\mathbb{R}^n$ 上的可微凸函数, 则 $\bar{x}$ 是整体极小点 (全局最优解) 的充要条件是 $\nabla f(\bar{x}) = 0$.

设 $\bar{x} \in S, d \in \mathbb{R}^n$ 是非零向量. 若存在 $\delta > 0$ 使得:

$$\bar{x} + \lambda d \in S, \forall \lambda \in (0, \delta) \tag{70}$$

则称 d 是 S 在 $\bar{x}$ 处的**可行方向**.

记 S 在 $\bar{x}$ 处的所有可行方向的集合为 $F(\bar{x}, S)$.

设 $f(x)$ 是 $\mathbb{R}^n$ 上的实函数,$x \in \mathbb{R}^n$, d 是非零向量. 若存在 $\delta > 0$ 使得:

$$f(\bar{x} + \lambda d) < f(\bar{\lambda}), \forall \lambda \in (0, \delta) \tag{71}$$

则称 d 是 S 在 $\bar{x}$ 处的**下降方向**.

如果 $f(x)$ 是可微函数, 且 $\nabla f(\bar{x})^T d < 0$. 显然, 此处的 d 为 $f(x)$ 在 $\bar{x}$ 处的下降方向. 记这样的方向集合为

$$D(\bar{x}, f) = \{d \mid \nabla f(\bar{x})^T d < 0\}.$$

是**下降方向集的子集.**

易得对于问题 $min\{f(x) \mid x \in S\}$, 设 $\bar{x} \in S$, $f(x)$ 在 $\bar{x}$ 处可微. 如果 $\bar{x}$ 是问题的局部最优解, 则

$$F(\bar{x}, S) \cap D(\bar{x}, f) = \emptyset. \tag{72}$$

记 $\mathcal{I}(\bar{x}) = \{i \in \{1, \cdot\cdot\cdot, m\} | g_i(\bar{x}) = 0\}$,

$$D_f = D(\bar{x}, f) = \{d \mid \nabla f(\bar{x})^T d < 0\}, \tag{73s}$$

$$F_g = F(\bar{x}, g) = \{d \mid \nabla g_i(\bar{x})^T d > 0, i \in \mathcal{I}(\bar{x})\}, \tag{74}$$

$$F_h = F(\bar{x}, h) = \{d \mid \nabla h_j(\bar{x})^T d = 0, j = 1, \cdot\cdot\cdot, l\}. \tag{75}$$

设 $\bar{x}$ 为问题 (67) 的局部最优解,f 和 $g_i, i \in \mathcal{I}(\bar{x})$ 在点 $\bar{x}$ 可微, $g_i, i \notin \mathcal{I}(\bar{x})$ 在点 $\bar{x}$ 连续,h_j 在点 $\bar{x}$ 连续可微, 且 $\{\nabla h_j(\bar{x})\}_{j=1}^l$ 线性无关, 则

$$D_f \cap F_g \cap F_h = \emptyset. \tag{76}$$

4.2 K-T Condition

Fritz-John 条件:

在问题 (67) 中, 设 $\bar{x}$ 为可行点,f 和 $g_i, i \in \mathcal{I}(\bar{x})$ 在点 $\bar{x}$ 可微, $g_i, i \notin \mathcal{I}(\bar{x})$ 在点 $\bar{x}$ 连续,h_j 在点 $\bar{x}$ 连续可微. 如果 $\bar{x}$ 是局部最优解, 则存在不全为零的数 $\lambda_0, \lambda_i, i \in \mathcal{I}(\bar{x})$ 和 $\mu_j, j = 1, ..., l$ 使得

$$\lambda_0 \nabla f(\bar{x}) - \sum_{i \in \mathcal{I}(\bar{x})} \lambda_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} \mu_j \nabla h_j(\bar{x}) = 0 \tag{77}$$

其中 $\lambda_0 \geq 0, \lambda_i \geq 0, i \in \mathcal{I}(\bar{x})$.

Pf.1:

如果 $\{\nabla h_j(\bar{x})\}_{j=1}^l$ 线性相关, 则存在不全为零的数 $\mu_j, j = 1, ..., l$ 使得

$$\mu_j \nabla h_j(\bar{x}) = 0.$$

这时可令 $\lambda_0 = 0, \lambda_i = 0, i \in \mathcal{I}(\bar{x})$, 结论成立.

Pf.2:

如果 $\{\nabla h_j(\bar{x})\}_{j=1}^{l}$ 线性无关, 则必有 $D_f \cap F_g \cap F_h = \emptyset$. 即不等式组

$$\begin{cases} \nabla f(\bar{x})^T d < 0 \\ \nabla g_i(\bar{x})^T d > 0 \quad , i \in \mathcal{I}(\bar{x}) \\ \nabla h_j(\bar{x})^T d = 0 \quad , j = 1, ..., l \end{cases} \tag{78}$$

无解.

令 A 是以 $\{\nabla f(\bar{x}), -\nabla g_i(\bar{x}), i \in \mathcal{I}(\bar{x})\}$ 为列组成的矩阵,B 是以 $\{-\nabla h_j(\bar{x}), j = 1, ..., l\}$ 为列组成的矩阵. 于是得

$$\begin{cases} A^T d < 0 \\ B^T d = 0 \end{cases} \tag{79}$$

无解.

下证

$$\begin{cases} Ap_1 + Bp_2 = 0 \\ p_1 \geq 0 \end{cases} \tag{80}$$

有解.

现定义

$$S_1 = \{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mid y_1 = A^T d, y_2 = B^T d, d \in \mathbb{R}^n\},$$

$$S_2 = \{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mid y_1 < 0, y_2 = 0\}.$$

显然 S_1 和 S_2 为非空凸集, 且 $S_1 \cap S_2 = \emptyset$.

由凸集分离定理知, 对 $\forall d \in \mathbb{R}^n, \forall \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in CL(S_2)$, 存在非零向量 $\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ 使得

$$p_1^T A^T d + p_2^T B^T d \geq p_1^T y_1 + p_2^T y_2.$$

首先令 $y_2 = 0$, 由 d 的任意性 (取 $d = 0$) 及 $y_1 < 0, \Rightarrow p_1 \geq 0$.

再令 $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in CL(S_2), \Rightarrow p_1^T A^T d + p_2^T B^T d \geq 0$.

最后取 $d = -(Ap_1 + Bp_2), \Rightarrow Ap_1 + Bp_2 = 0$.

综上所述, 即得 (80) 有解.

把 p_1 的分量记作 λ_0 和 $\lambda_i, i \in \mathcal{I}(\bar{x})$, p_2 的分量记作 $\mu_j, j = 1, ..., l$. 立即得到

$$\lambda_0 \nabla f(\bar{x}) - \sum_{i \in \mathcal{I}(\bar{x})} \lambda_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} \mu_j \nabla h_j(\bar{x}) = 0 \tag{81}$$

Kuhn-Tucker 条件:

设 $\bar{x}$ 为约束问题 (67) 的可行点,f 和 $g_i, i \in \mathcal{I}(\bar{x})$ 在点 $\bar{x}$ 可微,$g_i, i \notin \mathcal{I}(\bar{x})$ 在点 $\bar{x}$ 连续, h_j 在点

$\bar{x}$ 连续可微, 向量集 $\{\nabla g_i(\bar{x}), i \in \mathcal{I}(\bar{x}); \nabla h_j(\bar{x}), j = 1, ..., l\}$ 线性无关. 如果 $\bar{x}$ 是局部最优解, 则存在数 $\lambda_i \geq 0$ 和 μ_j 使得

$$\lambda_0 \nabla f(\bar{x}) - \sum_{i \in \mathcal{I}(\bar{x})} \lambda_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} \mu_j \nabla h_j(\bar{x}) = 0 \tag{82}$$

定义 Lagrange 函数 $L(x, \lambda, \mu) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x) - \sum_{j=1}^{l} \mu_j h_j(x)$.

若 $\bar{x}$ 为问题局部最优解, 则存在乘子向量 $\bar{\lambda} \geq 0, \bar{\mu}$ 使得

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0.$$

此时, 一阶必要条件可表达为

$$(K-T)\begin{cases} \nabla_x L(x, \lambda, \mu) = 0 \\ g_i(x) \geq 0, i = 1, ..., m \\ \lambda_i g_i(x) = 0, i = 1, ..., m \\ \lambda_i \geq 0, i = 1, ..., m \\ h_j(x) = 0, j = 1, ..., l \end{cases} \tag{83}$$

理论上可以用最优性条件求“非线性规划问题”的最优解, 但在实践中并不切实可行. 在求解最优化问题时最常用的计算方法是“迭代下降算法”.

4.3 Convergence

算法映射:

算法 $\mathcal{A}$ 是定义在空间 X 上的点到集的映射, 即对每一点 $x^{(k)} \in X$, 经算法 $\mathcal{A}$ 作用后产生一个点集 $\mathcal{A}(x^{(k)}) \subset X$, 任意选择一个点 $x^{(k+1)} \in \mathcal{A}(x^{(k)})$ 作为 $x^{(k)}$ 的后续点.

引入所谓算法的闭性, 其实质是点到点映射的连续性的推广.

设 X 和 Y 分别是空间 E_p 和 E_q 中的非空闭集,$\mathcal{A}: X \to Y$ 为点到集的映射. 如果 $x^{(k)} \in X, x^{(k)} \to x, y^{(k)} \in \mathcal{A}(x^{(k)}), y^{(k)} \to y$ 蕴涵着 $y \in \mathcal{A}(x)$, 则称映射 $\mathcal{A}$ 在 $x \in X$ 处是闭的.

为研究算法的收敛性, 首先要明确解集合的概念.

在许多情况下, 要使算法产生的点列收敛于全局最优解是极为困难的. 因此, 一般把满足某些条件的点集定义为解集合. 当迭代点属于这个集合时, 就停止迭代.

例如, 在无约束最优化问题中, 可以定义解集合为

$$\Omega = \{\bar{x} \mid \|\nabla f(\bar{x})\| = 0\},$$

在约束最优化问题中, 解集合取为

$$\Omega = \{\bar{x} \mid \bar{x}\text{是 K-T 点}\}.$$

一般地, **下降算法**总是与某个函数在迭代过程中函数值的减小联系在一起, 因此需要给出下降函数的概念.

设 $\Omega \subset X$ 为解集合,$\mathcal{A}$ 为 X 上的一个算法映射, $\psi(x)$ 是定义在 X 上的连续实函数, 若满足

$$\psi(y) < \psi(x), \qquad if\ x \notin \Omega \wedge y \in \mathcal{A}(x)$$

$$\psi(y) \leq \psi(x), \qquad if\ x \in \Omega \wedge y \in \mathcal{A}(x)$$

则称 ψ 是关于解集合 Ω 和算法 $\mathcal{A}$ 的下降函数.

设 Ω 为解集合,$\mathcal{A}$ 为 X 上的算法映射. 若以任意初始点 $x^{(0)} \in Y \subset X$ 出发, 算法产生的序列的任一收敛子列的极限属于解集合, 则称算法映射 $\mathcal{A}$ 在 Y 上收敛于解集合 Ω.

定理:

设 $\mathcal{A}$ 为 X 上的一个算法,Ω 为解集合, 给定初始点 $x^{(0)} \in X$, 进行如下迭代: 如果 $x^{(k)} \in \Omega$, 则停止迭代; 否则取 $x^{(k+1)} \in \mathcal{A}(x^{(k)})$,$k := k+1$, 重复以上过程. 这样产生迭代序列 $\{x^{(k)}\}$. 又设:

1. 序列 $\{x^{(k)}\}$ 含于 X 的紧子集中;
2. 存在一个连续函数 ψ, 它是关于 Ω 和 $\mathcal{A}$ 的下降函数;
3. 映射 $\mathcal{A}$ 在 Ω 的补集上是闭的.

则序列 $\{x^{(k)}\}$ 的任一收敛子列的极限属于 Ω.

Pf:

先证序列 $\{x^{(k)}\}$ 对应的下降函数值数列 $\{\psi(x^{(k)})\}$ 有极限.

$\{x^{(k)}\}$ 含于 X 的紧子集, 因此有收敛子列 $\{x^{(k)}\}_K$, 设其极限为 $x \in X$. 由 ψ 的连续性得, $\psi(x^{(k)}) \to \psi(x), k \in K$. 即对 $\forall \varepsilon > 0, \exists N$ 使得当 $k \geq N$ 时, 有 $0 < \psi(x^{(k)}) - \psi(x) < \varepsilon, k \in K$. 特别地,$0 < \psi(x^{(N)}) - \psi(x) < \varepsilon$.

又由 ψ 的下降性知,$\psi(x^{(k)}) - \psi(x^{(N)}) < 0, \forall k > N$. 于是有

$$0 < \psi(x^{(k)}) - \psi(x) = \psi(x^{(k)}) - \psi(x^{(N)}) + \psi(x^{(N)}) - \psi(x) < \varepsilon, \forall k > N.$$

即得 $\lim\limits_{k->\infty} \psi(x^{(k)}) = \psi(x)$.

下证 $x \in \Omega$(反证法).

假设 $x \notin \Omega$, 考虑序列 $\{x^{(k+1)}\}_K$. 由于它包含于紧集, 所以也存在收敛子列 $\{x^{(k+1)}\}_{\bar{K}}, \bar{K} \subset K$, 且设其极限为 $\bar{x} \in X$. 显然 (同理前述证明)

$$\lim_{k\to\infty} \psi(x^{(k+1)}) = \psi(\bar{x}).$$

由 $\psi(x^{(k)})$ 极限的唯一性知,

$$\psi(x) = \psi(\bar{x}).$$

另外, 对 $k \in \bar{K} \subset K$ 有

$$x^{(k)} \to x, x^{(k+1)} \in \mathcal{A}(x^{(k)}), x^{(k+1)} \to \bar{x}.$$

算法 A 在 Ω 的补集上是闭的,$x \notin \Omega$, 因此 $\mathcal{A}$ 在 x 处是闭的, 即有 $\bar{x} \in \mathcal{A}(x)$. 由于 ψ 是解集合 Ω 和算法 $\mathcal{A}$ 的下降函数, $x \notin \Omega$, 则有 $\psi(\bar{x}) < \psi(x)$. 这显然矛盾, 所以 $x \in \Omega$.

在迭代下降算法里, 当 $x^{(k)} \in \Omega$ 时才终止迭代. 在实践中许多情况下, 这是一个取极限的过程, 需要无限次迭代. 因此为解决实际问题, 需要规定一些实用的终止迭代过程的准则, 一般称为收敛准则或停机准则.

常用的收敛准则有以下几种:

$$1.\|x^{(k+1)}-x^{(k)}\| < \varepsilon \quad or \quad \frac{\|x^{(k+1)}-x^{(k)}\|}{\|x^{(k)}\|} < \varepsilon$$

$$2.f(x^{(k)})-f(x^{(k+1)}) < \varepsilon \quad or \quad \frac{f(x^{(k)})-f(x^{(k+1)})}{|f(x^{(k)})|} < \varepsilon$$

$$3.\|\nabla f(x^{(k)})\| < \varepsilon$$

在这里,ε 为事先给定的充分小的正数. 除此之外, 还可以根据收敛定理, 制定出其它的收敛准则.

评价算法优劣的标准之一是收敛的快慢, 通常称为**收敛速率**.

一般定义如下: 设序列 $\{x^{(k)}\}$ 收敛于 x^*, 满足

$$0 \leq \lim_{k\to\infty} \frac{\|x^{(k+1)-x^*}\|}{\|x^{(k)}-x^*\|^p} = \beta < \infty \tag{84}$$

的非负数 p 的上确界称为序列 $\{x^{(k)}\}$ 的收敛阶.

若在定义式 (84) 中,$p=1$ 且 $\beta<1$, 则称序列是 (收敛比 β) 线性收敛的.

若在定义式 (84) 中,$p>1$, 或者 $\{p=1,\beta=0\}$, 则称序列是超线性收敛的.

4.4 Iterative Algorithems

最优化可以追溯到十分古老的极值问题, 然而它成为一门独立的学科是在二十世纪四十年代末, 当时 Dantzig 提出了求解一般线性规划问题的单纯形法. 此后各种最优化问题的理论及应用研究得到迅速发展, 特别是线性规划由于其模型的普遍性和实用性, 相关算法的进展引起广泛的重视.

随着实际问题的规模越来越大以及在计算机技术的推动下, 人们开始从复杂性角度研究线性规划和非线性规划的算法.

最优化方法通常采用迭代方法求问题的最优解, 其基本思想是:

给定一个初始点 $x\in\mathbb{R}$, 按照某一迭代规则产生一个点列 $\{x^{(k)}\}$, 使得当 $\{x^{(k)}\}$ 是有穷点列时, 其最后一个点是最优化模型问题的最优解, 当 $\{x^{(k)}\}$ 是无穷点列时, 它有极限点且其极限点是最优化模型问题的最优解.

一个好的迭代算法应具备的典型特征是:

迭代点 $x^{(k)}$ 能稳定地接近局部极小点 x^* 的小邻域, 然后迅速收敛于 x^*. 一般地, 对于某种算法我们需要证明其迭代点列 $\{x^{(k)}\}$ 的聚点 (即子列的极限点) 为一局部极小点. 在实际计算中, 当指定的收敛准则满足时, 迭代即终止.

设 $x^{(k)}$ 为第 k 次迭代点,$d^{(k)}$ 为第 k 次搜索方向,α_k 为第 k 次步长因子, 则第 $k+1$ 次迭代为:

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}. \tag{85}$$

从上述迭代格式可以看出, 不同的搜索方向和不同的步长策略构成不同的方法.

在最优化方法中, 搜索方向 $d^{(k)}$ 一般选取的是某价值函数 (merit function) ψ 在 $x^{(k)}$ 处的下降方向, 即 $d^{(k)}$ 满足

$$\nabla\psi(x^{(k)})^T d^{(k)} < 0. \tag{86}$$

步长因子的确定一般归结为解一维最优化问题

$$\min_{\alpha\geq 0}\psi(x^{(k)}+\alpha d^{(k)}) \tag{87}$$

(a) 给定初始点 $x^{(0)}$

(b) 计算搜索方向 $d^{(k)}$, 即构造某价值函数 ψ 在 $x^{(k)}$ 点处的下降方向作为搜索方向;

(c) 确定步长因子 α_k, 使该价值函数值有某种程度的下降;

(d) 迭代更新, 令 $x^{(k+1)}=x^{(k)}+\alpha_k d^{(k)}$.

(e) 判断停机准则, 若 $x^{(k+1)}$ 满足某种终止条件, 则停止迭代, 得到近似最优解 $\bar{x}=x^{(k+1)}$. 否则, 返回 (b) 重复以上步骤.

5 Unconstrained Optimization

无约束最优化问题

$$\min_{x\in\mathbb{R}^n} \qquad f(x) \tag{88}$$

其目标函数 f 是定义在 $\mathbb{R}^n$ 上的实值函数, 决策变量 x 的可取值之集合是全空间 $\mathbb{R}^n$.

5.1 Gradienet Methods

梯度向量 $\nabla f(x)$ 是函数 f 在点 x 处增加最快的方向, 故它成为最优化时的重要工具. 实际上针对无约束最优化问题, 大家所知的求解算法中大多属于下面的梯度方法类.

GRADIENT **(梯度法类)**

(0) 初始化: 选取适当的初始点 $x^{(0)}\in\mathbb{R}^n$, 令 $k:=0$.

(1) 计算搜索方向: 利用适当的正定对称阵 H_k 计算搜索方向向量 $d^{(k)}:=-H_k\nabla f(x^{(k)})$. (如果 $\nabla f(x^{(k)})=0$, 则结束计算)

(2) 确定步长因子: 解一维最优化问题 $\min\limits_{\alpha\geq 0} f(x^{(k)}+\alpha d^{(k)})$, 求出步长 $\alpha=\alpha_k$ 令 $x^{(k+1)}=x^{(k)}+\alpha_k d^{(k)}, k:=k+1$ 回到第 (1) 步.

对无约束最优化问题: $\min\limits_{x\in\mathbb{R}^n} f(x)$

$$f(x)=f(x^{(k)})+\nabla f(x^{(k)})^T(x-x^{(k)})+O(\|x-x^{(k)}\|^2) \tag{89}$$

取负梯度方向

$$d^{(k)}=-\nabla f(x^{(k)})$$

则当 α_k 足够小时, 总能使

$$f(x^{(k)}+\alpha_k d^{(k)})<f(x^{(k)}).$$

这样的 $d^{(k)}$ 称为**负梯度方向**.

$$f(x)=f(x^{(k)})+\nabla f(x^{(k)})^T(x-x^{(k)})+\frac{1}{2}(x-x^{(k)})^T\nabla^2 f(x^{(k)})(x-x^{(k)})+O(\|x-x^{(k)}\|^3) \tag{90}$$

取搜索方向

$$d^{(k)} = -G_k^{-1}\nabla f(x^{(k)})$$

其中 $G_k = \nabla^2 f(x^{(k)})$ 为函数 f 在 $x^{(k)}$ 点处的 Hesse 矩阵. 这样的 $d^{(k)}$ 称为**牛顿方向**.

5.2 Line Search

在迭代格式中, 通过解一维最优化问题

$$\min_{\alpha\geq 0}\varphi(\alpha) = f(x^{(k)} + \alpha d^{(k)}) \tag{91}$$

确定步长因子的方法称为**一维搜索** (Line Search).

若以问题 (91) 的最优解为步长, 此时称为**精确一维搜索** (Exact Line Search).

经常用到的精确一维搜索有黄金分割法和插值迭代法. 即使说是精确一维搜索, 通过有限次计算求出问题 (91) 的严密解一般也是不可能的, 实际上在得到有足够精度的近似解时, 就采用它作为步长.

黄金分割法

(0) 确定包含极小值的初始单峰区间 $[a_0, b_0]$, 精度 $\varepsilon > 0$. 令 $\lambda = 0.618$. 取 $u_0 = b_0 - \lambda(b_0 - a_0), v_0 = a_0 + \lambda(b_0 - a_0)$, 令 $k := 0$.

(1) 若 $b_k - a_k \leq \varepsilon$, 则得 $\bar{\alpha} = (a_k + b_k)/2$, 停止计算.

(2) 若 $\phi(u_k) = \phi(v_k)$, 则令 $a_{k+1} = u_k, b_{k+1} = v_k, k := 0$, 转步 0.

(3) 若 $\phi(u_k) < \phi(v_k)$, 则令 $a_{k+1} = a_k, b_{k+1} = v_k, v_{k+1} = u_k$. 计算 $u_{k+1} = b_{k+1} - \lambda(b_{k+1} - a_{k+1})$. 令 $k := k + 1$ 转步 1.

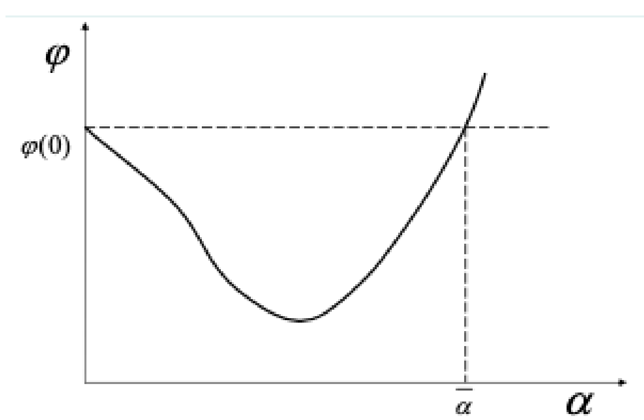
(4) 若 $\phi(u_k) > \phi(v_k)$, 则令 $a_{k+1} = u_k, b_{k+1} = b_k, u_{k+1} = v_k$. 计算 $v_{k+1} = a_{k+1} + \lambda(b_{k+1} - a_{k+1})$. 令 $k := k + 1$ 转步 1.

在实际计算中, 往往不是求解一维最优化问题 (91), 而是找出满足某些适当条件的粗略近似解作为步长, 此时称为**非精确一维搜索** (Inexact Line Search).

与精确一维搜索相比, 在很多情况下采用非精确一维搜索可以提高整体计算效率.

设 $\bar{\alpha}_k$ 是使得 $f(x^{(k)} + \alpha d^{(k)}) = f(x^{(k)})$ 的最小正数 α.

于是, 我们将在区间 $[0, \bar{\alpha_k}]$ 内求得满足适当条件的可接受的步长因子, 即 $\alpha \in [0, \bar{\alpha_k}]$.

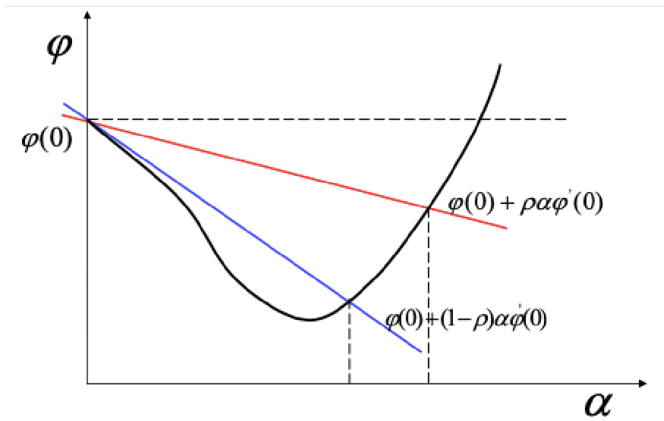


Goldstein(1965) condition:

$$\varphi(\alpha) \leq \varphi(0) + \rho\alpha\varphi'(0) \tag{92}$$

$$\varphi(\alpha) \geq \varphi(0) + (1-\rho)\alpha\varphi'(0) \tag{93}$$

其中 $\rho \in (0, \frac{1}{2})$ 是一个固定参数.

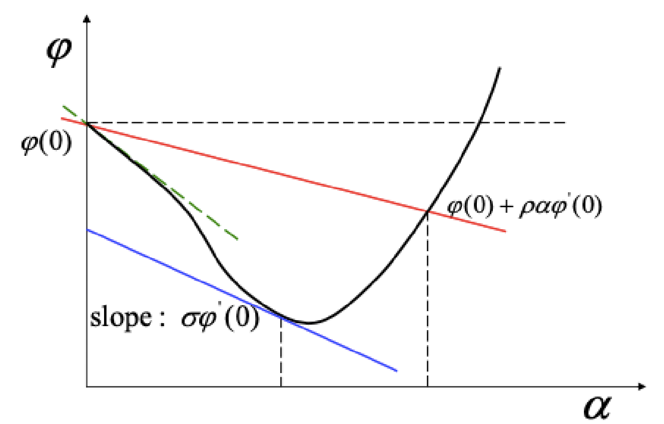


Wolfe(1968)-Powell(1976) conditions:

$$\varphi(\alpha) \leq \varphi(0) + \rho\alpha\varphi'(0) \tag{94}$$

$$\varphi'(\alpha) \geq \sigma\varphi'(0) \tag{95}$$

其中 $\sigma \in (\rho, 1)$ 是另一个固定参数.



在很多实际算法中, 式 (95) 常被强化的双边条件所取代

$$|\varphi'(\alpha)| \leq -\sigma\varphi'(0) \tag{96}$$

基于 Wolfe-Powell 准则的非精确一维搜索算法:

(0) 给定初始一维搜索区间 $[0, \bar{\alpha}]$, 以及 $\rho \in (0, \frac{1}{2}), \sigma \in (\rho, 1)$. 计算 $\varphi_0 = \varphi(0) = f(x^{(k)}), \varphi'_0 =$

$\varphi'(0) = \nabla f(x^{(k)})^T d^{(k)}$. 并令 $a_1 = 0, a_2 = \bar{\alpha}, \varphi_1 = \varphi_0, \varphi_1' = \varphi_0'$. 选取适当的 $\alpha \in (a_1, a_2)$.
(1) 计算 $\varphi = \varphi(\alpha) = f(x^{(k)} + \alpha d^{(k)})$. 若 $\varphi(\alpha) \le \varphi(0) + \rho\alpha\varphi'(0)$, 则转到第 (2) 步. 否则, 由 $\varphi_1, \varphi_1', \varphi$ 构造两点二次插值多项式 $p^{(1)}(t)$, 并得其极小点

$$\hat{\alpha} = a_1 + \frac{1}{2}\frac{(a_1 - \alpha)^2 \varphi_1'}{(\varphi_1 - \varphi) - (a_1 - \alpha)\varphi_1'}$$

于是置 $a_2 = \alpha, \alpha = \hat{\alpha}$, 重复第 (1) 步.
(2) 计算 $\varphi' = \varphi(\alpha) = \nabla f(x^{(k)} + \alpha d^{(k)})^T d^{(k)}$. 若 $\varphi'(\alpha) \ge \sigma\varphi'(0)$, 则输出 $\alpha_k = \alpha$, 并停止搜索. 否则, 由 $\varphi, \varphi', \varphi_1'$ 构造两点二次插值多项式 $p^{(2)}(t)$, 并得其极小点

$$\hat{\alpha} = \alpha - \frac{(a_1 - \alpha)\varphi'}{\varphi_1' - \varphi'}.$$

于是置 $a_1 = \alpha, \alpha = \hat{\alpha}, \varphi_1 = \varphi, \varphi_1' = \varphi'$, 返回第 (1) 步.

从任意初始点出发, 如果某迭代算法产生的点列的极限 (聚点), 在适当假定下可保证恒为问题的最优解 (或者稳定点), 则称该迭代法具有**全局收敛性** (Global Convergence).

与此相对, 如果仅在解的附近选取初始点时, 才可以保证所生成的点列收敛于该解, 则称这样的迭代法有**局部收敛性** (Local Convergence).

5.3 Global Convergence

为了证明迭代法的下降性, 我们应尽量避免搜索方向与负梯度方向几乎正交的情形, 即要求 $d^{(k)}$ 偏离 $g^{(k)} = \nabla f(x^{(k)})$ 的正交方向远一些. 否则,${g^{(k)}}^T d^{(k)}$ 接近于零,$d^{(k)}$ 几乎不是下降方向. 为此, 我们假设 $d^{(k)}$ 与 $-g^{(k)}$ 的夹角 θ^k 满足

$$\theta_k \le \frac{\pi}{2} - \mu, \forall k \tag{97}$$

其中 $\mu > 0$(与 k 无关).

显然 $\theta_k \in [0, \frac{\pi}{2})$, 其定义为

$$cos\ \theta_k = \frac{-{g^{(k)}}^T d^{(k)}}{\|g^{(k)}\|\|d^{(k)}\|} = \frac{-{g^{(k)}}^T s^{(k)}}{\|g^{(k)}\|\|s^{(k)}\|} \tag{98}$$

这里 $s^{(k)} = \alpha_k d^{(k)} = x^{(k+1)} - x^{(k)}$.

下面给出各种步长准则下的下降算法的全局收敛性结论.

全局收敛性定理:

设 $\nabla f(x)$ 在水平集 $L(x^{(0)}) = \{x | f(x) \le f(x^{(0)})\}$ 上存在且一致连续. 下降算法的搜索方向 $d^{(k)}$ 与 $-\nabla f(x^{(k)})$ 之间的夹角 θ_k 满足式 (97), 其中步长 α_k 由三种方法之一确定:

(1) 精确一维搜索

(2) Goldstein 准则 (92),(93)

(3) Wolfe-Powell 准则 (94),(95)

那么, 或者对某个 k 有 $\nabla f(x^{(k)}) = 0$, 或者 $f(x^{(k)}) \to -\infty$, 或者 $\nabla f(x^{(k)}) \to 0$.

全局收敛性证明:

Wolfe-Powell 准则的情形

Pf.

假设对所有的 k, $g^{(k)}=\nabla f(x^{(k)})\neq 0$ 和 $f(x^{(k)})$ 有下界, 故 $f(x^{(k)})-f(x^{(k+1)})\to 0$. 由式 (94) 得,$-g^{(k)^T}s^{(k)}\to 0$.

(反证) 若 $g^{(k)}\to 0$ 不成立, 那么存在 $\varepsilon>0$ 和子列 $\{x^{(k)}\}k\in K$ 使得 $\|g^{(k)}\|\geq\varepsilon$. 从而由

$$-g^{(k)^T}s^{(k)}=\|g^{(k)}\|\|s^{(k)}\|cos\ \theta_k\geq\varepsilon\|s^{(k)}\|sin\ \mu$$

以及式 (97) 有 $\|s^{(k)}\|\to 0$.

又因为 $g(x)=\nabla f(x)$ 在 $L(x^{(0)})$ 上一致连续, 所以

$$\begin{gathered} g^{(k+1)^T}s^{(k)}=g^{(k)^T}s^{(k)}+o(\|s^{(k)}\|) \\ \Downarrow \\ \frac{g^{(k+1)^T}s^{(k)}}{g^{(k)^T}s^{(k)}}\to 1 \end{gathered} \quad . \tag{99}$$

而这与 Wolfe-Powell 准则的式 (95)

$$\frac{g^{(k+1)^T}s^{(k)}}{g^{(k)^T}s^{(k)}}\leq\sigma<1 \tag{100}$$

相矛盾. 因此有 $g^{(k)}\to 0$.

Goldstein 准则的情形

Pf.

假设对所有的 k,$g^{(k)}=\nabla f(x^{(k)})\neq 0$ 和 $f(x^{(k)})$ 有下界. 那么有 $f(x^{(k)})-f(x^{(k+1)})\to 0$ 由式 (92) $\varphi(\alpha)\leq\varphi(0)+\rho\alpha\varphi'(0)$, 得到 $-g^{(k)^T}s^{(k)}\to 0$.

反设 $g^{(k)}\to 0$ 不成立, 则 $\exists\,\varepsilon>0$ 和子列 $\{x^{(k)}\}_{k\in K}$ 使得 $\|g^{(k)}\|>\varepsilon$

$$-g^{(k)^T}s^{(k)}=\|g^{(k)}\|\ \|s^{(k)}\|\ cos\theta_k\geq\varepsilon\|s^{(k)}\|sin\mu.$$

又 $\varphi(0)+\alpha\varphi'(0)-\rho\alpha\varphi'(0)\leq\varphi(\alpha)=\phi(0)+\alpha\varphi'(\theta\alpha),(0<\theta<1)$, 得到 $\varphi'(\theta\alpha)\geq(1-\rho)\varphi'(0)$, 故有

$$\frac{\nabla f(x^{(k)}+\theta\alpha)^T s^{(k)}}{\nabla f(x^{(k)})^T s^{(k)}}\leq(1-\rho)<1$$

又由于 $\|s^{(k)}\|\to 0$

$$\frac{g(x^{(k)}+\theta s^{(k)})^T s^{(k)}}{g(x^{(k)})^T s^{(k)}}\leq(1-\rho)<1$$

但由于 $g(x)=\nabla f(x)$ 在 $L(x^{(0)})$ 上一致连续,

$$g(x^{(k)}+\theta s^{(k)})^T s^{(k)}=g(x^{(k)})^T s^{(k)}+o(\|s^{(k)}\|)$$

$$\frac{g(x^{(k)}+\theta s^{(k)})^T s^{(k)}}{g(x^{(k)})^T s^{(k)}}\to 1$$

这两者矛盾, 故 $d^{(k)}\to 0$.

5.4 Steepest Descent Methods

最速下降法[1]取负梯度作为迭代算法的搜索方向, 其迭代格式为

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}).$$

算法:

(0) 选取初始点 $x^{(0)}$, 设置终止误差 $\varepsilon > 0$, 令 $k := 0$.

(1) 计算 $g^{(k)} = \nabla f(x^{(k)})$. 若 $\|g^{(k)}\| < \varepsilon$,, 则停止迭代并输出 x, 否则进行第 (2) 步.

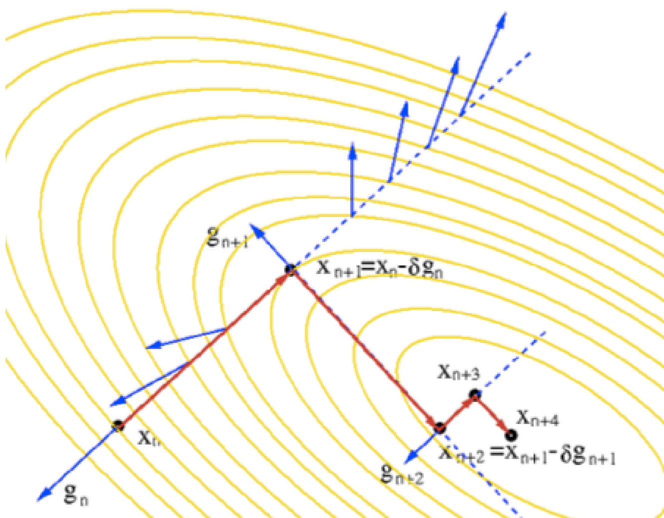
(2) 令 $d^{(k)} = -g^{(k)}$, 并由一维搜索确定步长因子 α_k 使得

$$f(x^{(k)} + \alpha_k d^{(k)}) = \min_{\alpha > 0} f(x^{(k)} + \alpha d^{(k)}).$$

(3) 迭代更新 $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$, 置 $k := k + 1$, 回到第 (1) 步.

最速下降法全局收敛性定理:

设 $f(x) \in C^1$, 在最速下降法中采用 (精确或非精确) 一维搜索, 则产生的迭代点列 $\{x^{(k)}\}$ 的每一个聚点都是驻点.



一般地, 最速下降法只有线性收敛速度. 如下例子是一个非常著名的测试函数 (Rosenbrock function)

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

5.5 Newton Method

设 $f(x)$ 是二次可微实函数, 在 $x^{(k)}$ 附近作二阶 Taylor 展开近似

$$f(x^{(k)} + s) \approx q^{(k)}(s) = f(x^{(k)}) + g^{(k)^T} s + \frac{1}{2} s^T G_k s \tag{101}$$

其中 $g^{(k)} = \nabla f(x^{(k)}), G_k = \nabla^2 f(x^{(k)})$. 将 $q^{(k)}(s)$ 极小化便得

$$s = -G_k^{-1} g^{(k)}. \tag{102}$$

[1]具有线性收敛速率

上式给出的搜索方向 $-G_k^{-1}g^{(k)}$ 称为牛顿方向 (Newton Direction).

在目标函数是正定二次函数

$$f(x)=\frac{1}{2}x^TGx-c^Tx$$

的情况下 (G 为正定阵), 对任意的 x 有 $\nabla^2 f(x)=G$.

在第一次迭代里令 $H_0=G^{-1}$, 则有

$$d^{(0)}=-H_0\nabla f(x^{(0)})=-G^{-1}(Gx^{(0)}-c)=-(x^{(0)}-x^*).$$

这里,$x^*=G^{-1}c$ 是问题的最优解. 若 $x^{(0)}\neq x^*$, 取步长 $\alpha_0=1$, 于是得 $x^{(1)}=x^{(0)}+\alpha_0 d^{(0)}=x^*$. 由此知道, 不管初始点 $x^{(0)}$ 如何取, 在一次迭代后即可到达最优解 x^* .

根据以上事实, 可以认为即使对于一般的非线性函数 $f(x)$, 在迭代中令搜索方向

$$d^{(k)}=-[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)})$$

也是较为合适的.

特别地, 步长 $\alpha_k\equiv 1$ 的迭代公式为

$$x^{(k+1)}=x^{(k)}+d^{(k)}=x^{(k)}-G_k^{-1}g^{(k)}. \tag{103}$$

这就是经典的牛顿迭代法[2].

对于正定二次函数而言, 牛顿法一步即可达到最优解. 对于非二次函数, 牛顿法并不能保证经有限次迭代求得最优解. 但由于目标函数在极小点附近可用二次函数较好地近似, 故当初始点靠近极小点时, 牛顿法的收敛速度一般会很快.

可以证明牛顿法的局部收敛性和二阶收敛速率.

牛顿法收敛定理:

设 $f\in C^2, x^{(k)}$ 充分靠近 x , 其中 $\nabla f(x^*)=0$. 如果 $\nabla^2 f(x^*)$ 正定, 目标函数的 Hesse 矩阵 $G(x)$ 满足 Lipschitz 条件, 即存在 $\beta>0$ 使得对所有 (i,j) 有

$$\|G_{ij}(x)-G_{ij}(y)\|\leq\beta\|x-y\|. \tag{104}$$

则对一切的 k, 牛顿迭代 (103) 有定义, 所得序列 $\{x^{(k)}\}$ 收敛到 x^* , 且具有二阶收敛速率.

Pf 1.

记 $g(x)=\nabla f(x)$, 因为 $f\in C^2$, 我们有

$$g(x-h)=g(x)-G(x)h+O(\|h\|^2).$$

令 $x=x^{(k)}, h=h^{(k)}=x^{(k)}-x^*$ 代入上式得

$$0=g(x^*)=g(x^{(k)}-h^{(k)})=g(x^{(k)})-G(x^{(k)})h^{(k)}+O(\|h(k)\|^2). \tag{105}$$

由于 $G(x)$ 满足 Lipschitz 条件, 易证 $[G(x^{(k)})]^{-1}$ 有界. 方程 (105) 两边同时乘以 $[G(x^{(k)})]^{-1}$ 得

$$\begin{aligned}0=&[G(x^{(k)})]^{-1}g(x^{(k)})-h^{(k)}+O(\|h^{(k)}\|^2)\\=&x^*-(x^{(k)}-[G(x^{(k)})]^{-1}g(x^{(k)}))+O(\|h^{(k)}\|^2)\\=&x^*-x^{(k+1)}+O(\|h^{(k)}\|^2)\\=&-h^{(k+1)}+O(\|h^{(k)}\|^2)\end{aligned}$$

[2]具有二阶收敛速率

所以 $\|h^{(k+1)}\| = O(\|h^{(k)}\|^2)$, 即牛顿迭代法具有二阶收敛速率.

Pf 2.

对于牛顿迭代法, 我们记

$$x^{(k+1)} = x^{(k)} - G_k^{-1} g^{(k)} \triangleq \mathcal{A}(x^{(k)}). \tag{106}$$

注意到 $g(x^*) = 0$, $G(x^*)$ 正定 (非奇异), 有 $\mathcal{A}(x^*) = x^*$. 于是由 $x^{(k+1)} - x^* = \mathcal{A}(x^{(k)}) - \mathcal{A}(x^*)$ 得

$$\begin{aligned}\|x^{(k+1)} - x^*\| =& \|\mathcal{A}(x^{(k)}) - \mathcal{A}(x^*)\| \\ \leq& \|\mathcal{A}'(x^*)(x^{(k)} - x^*)\| + \frac{1}{2}\|\mathcal{A}''(\bar{x})\|\|x^{(k)} - x^*\|^2,\end{aligned}$$

其中 $\bar{x}$ 位于 $x^{(k)}$ 和 x^* 之间的线段上.
显然

$$\mathcal{A}'(x) = [x - G(x)^{-1} g(x)]' = -[G(x)^{-1}]' g(x)$$

所以 $\mathcal{A}'(x^*) = 0$. 从而有

$$\|h^{(k+1)}\| = \|x^{(k+1)} - x^*\| \leq \gamma \|x^{(k)} - x^*\|^2 = \gamma \|h^{(k)}\|^2$$

其中常数 γ 仅依赖于 $f(x)$ 在 x^* 附近的三阶导数.

在式 (103) 的牛顿迭代法里, 如果选取的初始点 $x^{(0)}$ 不在解 x^* 的附近, 那么生成的点列 $\{x^{(k)}\}$ 未必收敛于最优解.

为保证算法的全局收敛性, 有必要对牛顿法作某些改进. 比如, 在牛顿法中也可采用一维搜索来确定步长.

阻尼牛顿法:

(0) 选取初始点 $x^{(0)}$, 设置终止误差 $\varepsilon > 0$, 令 $k := 0$.
(1) 计算 $g^{(k)} = \nabla f(x^{(k)})$. 若 $\|g^{(k)}\| < \varepsilon$ 停止迭代并输出 $x^{(k)}$. 否则进行第 (2) 步.
(2) 解线性方程组 $G_k d = -g^{(k)}$, 求出牛顿方向 $d^{(k)}$
(3) 采用一维搜索确定步长因子 α_k, 令 $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$, 置 $k := k+1$, 回到第 (1) 步.

牛顿法面临的主要困难是 Hesse 矩阵 $G_k = \nabla^2 f(x^{(k)})$ 不正定. 这时二阶近似模型不一定有极小点, 即二次函数 $q^{(k)}(s)$ 是无界的.

为了克服这些困难, 人们提出了很多修正措施.

Goldstein & Price (1967)

$$d(k) = \begin{cases} -G_k^{-1} g^{(k)}, & if\ cos\theta_k > \eta \\ -g^{(k)}, & otherwise \end{cases} \tag{107}$$

Levenberg(1944), Marquardt(1963), Goldfeld et. al(1966)

$$(G_k + \mu_k I) d^{(k)} = -g^{(k)} \tag{108}$$

设 x 是函数 f 的一个不定点, 若方向 d 满足

$$d^T\nabla^2 f(x)d < 0,$$

则称 d 为 f 在 x 处的负曲率方向.

当 Hesse 矩阵 $\nabla^2 f(x^{(k)})$ 不正定时, 负曲率方向法是修正牛顿法的另一种途径.

5.6 Quasi Newton Method

牛顿法的突出优点是局部收敛很快 (具有二阶收敛速率), 但运用牛顿法需要计算二阶导, 而且目标函数的 Hesse 矩阵 $\nabla^2 f(x^{(k)})$ 可能非正定, 甚至奇异.

为了克服这些缺点, 人们提出了拟牛顿法. 其基本思想是: 用不含二阶导数的矩阵 H_k 近似牛顿法中的 Hesse 矩阵的逆 $G(x^{(k)})^{-1}$. 由构造近似矩阵的方法不同, 将出现不同的拟牛顿法.

回顾牛顿法的迭代

$$\begin{cases} G_k d = -g^{(k)} \\ x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)} \end{cases}$$

为了构造 Hesse 矩阵逆 G^{-1} 的近似 H_k, 我们先分析二阶导 $\nabla^2 f(x^{(k)})$ 与一阶导 $\nabla f(x^{(k)})$ 的关系.

设第 k 次迭代后得到 $x^{(k+1)}$, 将目标函数 $f(x)$ 在 x 处二阶 Taylor 展开:

$$\begin{aligned} f(x) \approx & f(x^{(k+1)}) + \nabla f(x^{(k+1)})^T(x - x^{(k+1)}) \\ & + \frac{1}{2}(x - x^{(k+1)})^T\nabla^2 f(x^{(k+1)})(x - x^{(k+1)}), \end{aligned}$$

进一步有

$$\nabla f(x) \approx \nabla f(x^{(k+1)}) + \nabla^2 f(x^{(k+1)})(x - x^{(k+1)}),$$

于是令 $x = x^{(k)}$ 得

$$\nabla f(x^{(k)}) \approx \nabla f(x^{(k+1)}) + \nabla^2 f(x^{(k+1)})(x^{(k)} - x^{(k+1)}).$$

记 $s^{(k)} = x^{(k+1)} - x^{(k)}, y^{(k)} = \nabla f(x^{(k+1)}) - \nabla f(x^{(k)})$, 则有

$$\nabla^2 f(x^{(k+1)})s^{(k)} \approx y^k \quad or \quad \nabla^2 f(x^{(k+1)})^{-1}y^{(k)} \approx s^{(k)}.$$

这样, 计算出 $s^{(k)}$ 和 $y^{(k)}$ 后, 可依上式估计在 $x^{(k+1)}$ 处的 Hesse 矩阵的逆. 我们有理由要求在迭代中构造出 Hesse 矩阵逆的近似 H_{k+1}, 使其满足

$$H_{k+1}y^{(k)} = s^{(k)}. \tag{109}$$

通常把式 (109) 称作正割条件, 也称为拟牛顿条件.

拟牛顿迭代算法的一般格式:

(0) 选取初始点 $x^{(0)}$, 令 $H_0 = I, k := 0$.

(1) 计算搜索方向 $d^{(k)} = -H_k\nabla f(x^{(k)})$.

(2) 采用一维搜索确定步长因子 α_k , 令 $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$.

(3) 基于 $x^{(k)}$ 到 $x^{(k+1)}$ 的梯度变化, 更新 Hesse 矩阵逆的近似, 即确定满足正割条件的 H_{k+1}. 置 $k := k+1$, 返回第 (1) 步.

下面我们就来讨论怎样构造及确定满足拟牛顿条件的 Hesse 矩阵逆的近似 H_{k+1}.

设 H_k 是第 k 次迭代的 Hesse 矩阵逆的近似, 我们希望以 H_k 来产生 H_{k+1}, 即

$$H_{k+1} = H_k + E_k,$$

其中 E_k 是一个低秩的矩阵. 为此, 可采用对称秩一 (SR1) 校正

$$H_{k+1} = H_k + auu^T, (a \in \mathbb{R}, u \in \mathbb{R}^n).$$

由拟牛顿条件 (109) 知

$$H_{k+1}y^{(k)} = H_k y^{(k)} + (au^T y^{(k)})u = s^{(k)}$$

故 u 必与方向 $s^{(k)} - H_k y^{(k)}$ 一致, 且假定 $s^{(k)} - H_k y^{(k)} \neq 0$. 不妨取 $u = s^{(k)} - H_k y^{(k)}$, 此时 $a = \frac{1}{u^T y^{(k)}}$, 从而得到

$$H_{k+1} = H_k + \frac{(s^{(k)} - H_k y^{(k)})(s^{(k)} - H_k y^{(k)})^T}{(s^{(k)} - H_k y^{(k)})^T y^{(k)}}. \tag{110}$$

上式称为对称秩一校正.

对称秩一校正的突出性质:

1. 针对二次函数具有遗传性, 即 $H_k y^{(l)} = s^{(l)}, l = 0, 1, \cdot\cdot\cdot, k-1$.

2. 具有二次终止性, 即对于二次函数不需要进行一维搜索而具有 n 步终止性质, 且 $H_n = [\nabla^2 f(x^*)]^{-1}$.

Pf.

遗传性: $H_k y^{(l)} = s^{(l)}, l = 0, 1, ..., k-1$.

$k = 1$ 时,$H_1 y^{(0)} = s^{(0)}$ 是正割条件, 成立.

假设 $k = 1, 2, ..., n$ 时,$H_k y^{(l)} = s^{(l)}, l = 0, 1, ..., k-1$ 始终成立.

下面讨论 $k = n+1$ 的情况 $l = n$ 时, 由正割条件 $H_{n+1} y^{(n)} = s^{(n)}$;$l < n$ 时

$$\begin{aligned}
(s^{(n)} - H_n y^{(n)})^T y^{(l)} =& {s^{(n)}}^T y^{(l)} - {y^{(n)}}^T H_n y^{(l)} \\
=& {s^{(n)}}^T y^{(l)} - {y^{(n)}}^T s^{(l)} \\
=& {s^{(n)}}^T G s^{(l)} - {s^{(n)}}^T G s^{(l)} \\
=& 0
\end{aligned}$$

因此 $H_{n+1} y^{(l)} = H_n y^{(l)} + \frac{(s^{(n)} - H_n y^{(n)})(s^{(n)} - H_n y^{(n)})^T}{(s^{(n)} - H_n y^{(n)})^T y^{(n)}} y^{(l)} = H_n y^{(l)} = s^{(l)}$ 由数学归纳法证毕.

二次终止性:

假定 $s^{(0)}, s^{(1)}, ..., s^{(n-1)}$ 线性无关.

由遗传性可知

$$H_n y^{(l)} = s^{(l)}, l = 0, ..., n-1$$

$$H_n G s^{(l)} = s^{(l)}, l = 0, ..., n-1$$

$$(H_nG-I)s^{(l)}=0, l=0,...,n-1$$

由于 $s^{(0)}, s^{(1)}, ..., s^{(n-1)}$ 线性无关, 所以 $(s^{(0)}, s^{(1)}, ..., s^{(n-1)})$ 可逆, 所以

$$H_nG-I=0 \Rightarrow H_n=G^{-1}=(\nabla^2 f(x))^{-1}$$

由迭代格式

$$x^{(n+1)}=x^{(n)}-H_n\nabla f(x^{(n)})=x^{(n)}-G^{-1}\nabla f(x^{(n)})$$

又因为

$$\nabla f(x^{(n+1)})-\nabla f(x^{(n)})=G(x^{(n)}-x^{(n+1)})$$

$$x^{(n+1)}=x^{(n)}-G^{-1}\nabla f(x^{(n)})+G^{-1}\nabla f(x^{(n+1)})$$

所以

$$G^{-1}\nabla f(x^{(n+1)})=0 \Rightarrow \nabla f(x^{(n+1)})=0$$

即有限终止且 $H_n=[\nabla^2 f(x^*)]^{-1}$.

对称秩一校正的缺点是, 不能保持迭代矩阵 H_{k+1} 的正定性.

仅当 $(s^{(k)}-H_ky^{(k)})^Ty^{(k)}>0$ 时, 对称秩一校正才能保持正定性. 而这个条件往往很难保证, 即使 $(s^{(k)}-H_ky^{(k)})^Ty^{(k)}>0$ 满足, 它也可能很小从而导致数值上的困难. 这些都使得对称秩一校正的拟牛顿法应用有较大局限性.

采用对称秩二 (SR2) 校正

$$H_{k+1}=H_k+auu^T+bvv^T,$$

并使得拟牛顿条件 (109) 成立, 则有

$$H_{k+1}y^{(k)}=H_ky^{(k)}+(au^Ty^{(k)})u+(bv^Ty^{(k)})v=s^{(k)}.$$

这里 u, v 显然不是唯一确定的, 但有一种明显的选择是:

$$\begin{cases} u=s^{(k)}, \quad au^Ty^{(k)}=1; \\ v=H_ky^{(k)}, bv^Ty^{(k)}=-1. \end{cases}$$

因此有

$$H_{k+1}=H_k+\frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}}-\frac{H_ky^{(k)}y^{(k)^T}H_k}{y^{(k)^T}H_ky^{(k)}}. \tag{111}$$

上式称为 DFP(Davidon-Fletcher-Powell) 校正公式, 由 Davidon(1959) 提出, 后经 Fletcher & Powell(1963) 修改而来.

DFP 校正 (111) 是典型的拟牛顿校正公式, 它有很多重要性质.

(一) 对于二次函数 (采用精确一维搜索)

1 遗传性, 即 $H_ky^{(l)}=s^{(l)}, l=0,1,\cdot\cdot\cdot,k-1$.

2 二次终止性, 即 $H_n=[\nabla f(x^*)]^{-1}$.

3 共轭性, 即当取 $H_0=I$ 时, 迭代产生共轭方向.

(二) 对于一般非线性函数

1 校正保持正定性, 因而 $d^{(k)}$ 总是下降方向.

2 每次迭代需要 $3n^2+O(n)$ 次乘法运算.

3 方法具有超线性收敛速度.

拟牛顿 (正割) 条件:

$$H_{k+1}y^{(k)} = s^{(k)}$$

, 其中 H_{k+1} 是 Hesse 矩阵逆的近似;

$$B_{k+1}s^{(k)} = y^{(k)}$$

, 其中 B_{k+1} 是 Hesse 矩阵的近似.

由对称秩二校正和拟牛顿条件 $H_{k+1}y^{(k)} = s^{(k)}$ 可得到 H_k 的 DFP 校正公式

$$H_{k+1}^{(DFP)} = H_k + \frac{s^{(k)}s^{(k)^T}}{s(k)^T y^{(k)}} - \frac{H_k y^{(k)} y^{(k)^T} H_k}{y^{(k)^T} H_k y^{(k)}}.$$

BFGS (Broyden-Fletcher-Goldfarb-Shanno) 校正

类似地, 我们可从拟牛顿条件 $B_{k+1}s^{(k)} = y^{(k)}$ 得到关于 B_k 的对称秩二校正公式

$$B_{k+1}^{(BFGS)} = B_k + \frac{y^{(k)}y^{(k)^T}}{y^{(k)^T}s^{(k)}} - \frac{B_k s^{(k)} s^{(k)^T} B_k}{s^{(k)^T} B_k s^{(k)}}. \tag{112}$$

把 (112) 式称为关于 B_k 的 BFGS 校正.

如果我们对 B_k 的 BFGS 校正 “求逆”, 就可以得到关于 H_k 的 BFGS 校正公式

$$H_{k+1}^{(BFGS)} = H_k + \left(1 + \frac{y^{(k)^T} H_k y^{(k)}}{s^{(k)^T} y^{(k)}}\right) \frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}} - \frac{H_k y^{(k)} s^{(k)^T} + s^{(k)} y^{(k)^T} H_k}{s^{(k)^T} y^{(k)}}. \tag{113}$$

进一步, 若将 (113) 式中 $\{H \leftrightarrow B, s \leftrightarrow y\}$ 互换, 便得到关于 B_k 的 DFP 校正公式

$$B_{k+1}^{(DFP)} = B_k + \left(1 + \frac{s^{(k)^T} B_k s^{(k)}}{y^{(k)^T} s^{(k)}}\right) \frac{y^{(k)}y^{(k)^T}}{y^{(k)^T}s^{(k)}} - \frac{B_k s^{(k)} y^{(k)^T} + y^{(k)} s^{(k)^T} B_k}{y^{(k)^T} s^{(k)}}. \tag{114}$$

Sherman-Morrison 定理:

设 $A \in \mathbb{R}^{n\times n}$ 是非奇异阵,$u, v \in \mathbb{R}^n$ 是任意向量. 若 $1 + v^T A^{-1} u \neq 0$, 则 A 的秩一校正 $A + uv^T$ 非奇异, 且其逆可以表示为

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}.$$

下给出利用秩二校正的求逆公式, 由 $H_{k+1}^{(DFP)}$ 推导 $B_{k+1}^{(DFP)}$.

$$\begin{aligned}
B_{k+1}^{(DFP)} =&(H_{k+1}^{(DFP)})^{-1} = (H_k + \frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}} - \frac{H_k y^{(k)}y^{(k)^T}H_k}{y^{(k)^T}H_k y^{(k)}})^{-1} \\
=&(H_k + \frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}})^{-1} + (H_k + \frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}})^{-1} \\
&\times (\frac{H_k y^{(k)}y^{(k)^T}H_k}{y^{(k)^T}H_k y^{(k)} - y^{(k)^T}H_k(H_k + \frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}})^{-1}H_k y^{(k)}}) \times (H_k + \frac{s^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}})^{-1} \\
=&H_k^{-1} - \frac{H_k^{-1}s^{(k)}s^{(k)^T}H_k^{-1}}{s^{(k)^T}y^{(k)} + s^{(k)^T}H_k^{-1}s^{(k)}} + (H_k^{-1} - \frac{H_k^{-1}s^{(k)}s^{(k)^T}H_k^{-1}}{s^{(k)^T}y^{(k)} + s^{(k)^T}H_k^{-1}s^{(k)}})^{-1} \\
&\times (\frac{H_k y^{(k)}y^{(k)^T}H_k}{\frac{y^{(k)^T}s^{(k)}s^{(k)^T}y^{(k)}}{s^{(k)^T}y^{(k)}+s^{(k)^T}H_k^{-1}s^{(k)}}}) \times (H_k^{-1} - \frac{H_k^{-1}s^{(k)}s^{(k)^T}H_k^{-1}}{s^{(k)^T}y^{(k)} + s^{(k)^T}H_k^{-1}s^{(k)}})^{-1} \\
=&H_k^{-1} - \frac{H_k^{-1}s^{(k)}s^{(k)^T}H_k^{-1}}{s^{(k)^T}y^{(k)} + s^{(k)^T}H_k^{-1}s^{(k)}} + \frac{y^{(k)}y^{(k)^T}}{y^{(k)^T}s^{(k)}s^{(k)^T}y^{(k)}}
\end{aligned}$$

$$\begin{aligned}
&\times (s^{(k)^T}y^{(k)} + s^{(k)^T}H_k^{-1}s^{(k)}) - \frac{H_k^{-1}s^{(k)}s^{(k)^T}y^{(k)}y^{(k)^T}}{y^{(k)^T}s^{(k)}s^{(k)^T}y^{(k)}} - \frac{y^{(k)}y^{(k)^T}s^{(k)}s^{(k)^T}H_k^{-1}}{y^{(k)^T}s^{(k)}s^{(k)^T}y^{(k)}} \\
&+ \frac{H_k^{-1}s^{(k)}s^{(k)^T}y^{(k)}y^{(k)^T}s^{(k)}s^{(k)^T}H_k^{-1}}{y^{(k)^T}s^{(k)}s^{(k)^T}y^{(k)}(s^{(k)^T}y^{(k)} + s^{(k)^T}H_k^{-1}s^{(k)})} \\
=&B_k - \frac{B_k s^{(k)}s^{(k)^T}B_k}{s^{(k)^T}y^{(k)} + s^{(k)^T}B_k s^{(k)}} + \frac{y^{(k)}y^{(k)^T}}{s^{(k)^T}y^{(k)}} + \frac{y^{(k)}y^{(k)^T}}{(s^{(k)^T}y^{(k)})^2}s^{(k)^T}B_k s^{(k)} \\
&- \frac{B_k s^{(k)}s^{(k)^T}y^{(k)}y^{(k)^T}}{(s^{(k)^T}y^{(k)})^2} - \frac{y^{(k)}y^{(k)^T}s^{(k)}s^{(k)^T}B_k}{(s^{(k)^T}y^{(k)})^2} + \frac{B_k s^{(k)}s^{(k)^T}B_k}{s^{(k)^T}y^{(k)} + s^{(k)^T}B_k s^{(k)}} \\
=&B_k + \frac{y^{(k)}y^{(k)^T}}{s^{(k)^T}y^{(k)}} + \frac{y^{(k)}y^{(k)^T}}{(s^{(k)^T}y^{(k)})^2}s^{(k)^T}B_k s^{(k)} - \frac{B_k s^{(k)}s^{(k)^T}y^{(k)}y^{(k)^T}}{(s^{(k)^T}y^{(k)})^2} - \frac{y^{(k)}y^{(k)^T}s^{(k)}s^{(k)^T}B_k}{(s^{(k)^T}y^{(k)})^2} \\
=&B_k + \frac{y^{(k)}y^{(k)^T}}{s^{(k)^T}y^{(k)}} + \frac{y^{(k)}y^{(k)^T}}{(s^{(k)^T}y^{(k)})^2}s^{(k)^T}B_k s^{(k)} - \frac{B_k s^{(k)}y^{(k)^T}}{s^{(k)^T}y^{(k)}} - \frac{y^{(k)}s^{(k)^T}B_k}{s^{(k)^T}y^{(k)}} \\
=&(I - \frac{y^{(k)}s^{(k)^T}}{s^{(k)^T}y^{(k)}})B_k(I - \frac{s^{(k)}y^{(k)^T}}{s^{(k)^T}y^{(k)}}) + \frac{y^{(k)}y^{(k)^T}}{s^{(k)^T}y^{(k)}}
\end{aligned}$$

5.7 Conjugate Gradient Method

共轭方向

定义: 设 G 是 $n\times n$ 正定阵,$\mathbb{R}^n$ 中的任一组非零向量 $\{d^{(0)}, d^{(1)}, \cdot\cdot\cdot, d^{(k)}\}$, 如果 $d^{(i)^T}Gd^{(j)} = 0 (i \neq j)$, 则称 $d^{(0)}, d^{(1)}, \cdot\cdot\cdot, d^{(k)}$ 是 G− 共轭的.

显然共轭是正交概念的推广, 当取 $G = I$ 时, 共轭即为正交.

共轭方向法 (类):

(0) 给定正定阵 G, 选取初始点 x , 计算 $g = \nabla f(x^{(0)})$ 并构造 $d^{(0)}$ 使得 $g^{(0)^T}d^{(0)}$. 令 k := 0.

(1) 求精确的一维搜索步长 α_k, 即 $\alpha_k = arg \min\limits_{\alpha>0} f(x^{(k)} + \alpha d^{(k)})$.

(2) 更新迭代点 $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$, 并构造 d 得 $d^{(k+1)}$ 使得 $d^{(k+1)^T}Gd^{(j)} = 0, j = 0, 1, ..., k$.

(3) 置 $k := k + 1$, 返回第 (1) 步.

共轭方向法是从研究二次函数的极小化问题中产生的, 但它可以推广到处理非二次函数的极小化问题.

共轭方向法的一个重要性质是, 只要执行精确一维搜索, 迭代算法就具有二次终止性.

共轭方向法基本定理:

严格凸二次函数 $f(x)=\frac{1}{2}x^TGx+c^Tx$, 共轭方向法执行精确一维搜索, 则每步迭代点 $x^{(k+1)}$ 是 $f(x)$ 在线性流形

$$\mathcal{V}=\{x \mid x=x^{(0)}+\sum_{j=0}^{k}\beta_j d^{(j)}, \forall\beta_j\in\mathbb{R}\}$$

中的唯一极小点.

Pf.

设共轭方向法产生的 $G-$ 共轭方向为 $d^{(0)},d^{(1)},...,d^{(k)}$ 由共轭方向的定义知,$\{d^{(0)},d^{(1)},...,d^{(k)}\}$ 线性无关.

下面只要证: 对所有 $k<n$ 成立

$$g^{(k+1)^T}d^{(j)}=0, j=0,1,\cdot\cdot\cdot,k.$$

即在点 $x^{(k+1)}$ 处的函数梯度 $g^{(k+1)}=\nabla f(x^{(k+1)})$ 与子空间 $span\{d^{(0)},d^{(1)},...,d^{(k)}\}$ 正交. 由此易得出定理的结论.

直接由精确一维搜索知, 对 $\forall j$ 成立

$$g^{(j+1)^T}d^{(j)}=0.$$

特别地, 当 $j=k$ 时,$g^{(k+1)^T}d^{(k)}=0$.

事实上, 由于

$$y^{(k)}=g^{(k+1)}-g^{(k)}=G(x^{(k+1)}-x^{(k)})=Gs^{(k)}=\alpha_k Gd^{(k)}.$$

故当 $j<k$ 时有

$$\begin{aligned}g^{(k+1)^T}d^{(j)}&=g^{(j+1)^T}d^{(j)}+\sum_{i=j+1}^{k}y^{(i)^T}d^{(j)}\\&=g^{(j+1)^T}d^{(j)}+\sum_{i=j+1}^{k}\alpha_i d^{(i)^T}Gd^{(j)}\\&=0+0\\&=0\end{aligned}$$

综合上述, 从而证明了

$$g^{(k+1)^T}d^{(j)}=0, j=0,1,...,k.$$

推论:

对于严格凸的二次函数, 若沿着一组共轭方向搜索, 经有限步迭代必达到极小点.

由于共轭方向法具有二次终止性, 人们希望能给出一个具体的算法 (属于共轭方向法类). 通过修改最速下降法, 使其搜索方向具有共轭性质, 这便是共轭梯度法.

下面我们先针对二次函数, 给出共轭梯度法的具体描述.

设二次函数 $f(x)=\frac{1}{2}x^TGx+c^Tx$, 其中 G 是 $n\times n$ 正定阵,c 是 n 维向量. 函数 f 的梯度向量为

$$g(x)=\nabla f(x)=Gx+c.$$

取 $d^{(0)}=-g^{(0)}$, 因为 $x^{(1)}=x^{(0)}+\alpha_0 d^{(0)}$ 中步长 α_0 由精确一维搜索决定, 所以 ${g^{(1)}}^T d^{(0)}=0$.

现设 $d^{(1)}=-g^{(1)}+\beta_0^{(1)}d^{(0)}$, 选择 $\beta_0^{(1)}$ 使 ${d^{(1)}}^T Gd^{(0)}=0$, 即得

$$\beta_0=\frac{{g^{(1)}}^T g^{(1)}}{{g^{(0)}}^T g^{(0)}}.$$

同理, 令 $d^{(2)}=-g^{(2)}+\beta_0^{(2)}d^{(0)}+\beta_1^{(2)}d^{(1)}$, 选择 $\beta_0^{(2)},\beta_1^{(2)}$ 使得 ${d^{(2)}}^T Gd^{(j)}=0, j=0,1.$ 从而有

$$\beta_0^{(2)}=0,$$

$$\beta_1^{(2)}=\frac{{g^{(2)}}^T g^{(2)}}{{g^{(1)}}^T g^{(1)}}$$

一般地, 在第 k 次迭代中, 令

$$d^{(k)}=-g^{(k)}+\sum_{j=1}^{k-1}\beta_j^{(k)}d^{(j)},$$

选择 $\beta_j^{(k)}$ 使得 ${d^{(k)}}^T Gd^{(j)}=0, j=0,1,...,k-1$, 则有

$$\beta_j^{(k)}=\frac{{g^{(k)}}^T Gd^{(j)}}{{d^{(j)}}^T Gd^{(j)}}=\frac{{g^{(k)}}^T (g^{(j+1)}-g^{(j)})}{{d^{(j)}}^T (g^{(j+1)}-g^{(j)})}$$

又由于 ${g^{(k)}}^T g^{(j)}=0, j=0,1,...,k-1$, 故得

$$\beta_j^{(k)}=0, j=0,1,...,k-2$$

$$\beta_{k-1}^{(k)}=\frac{{g^{(k)}}^T g^{(k)}}{{g^{(k-1)}}^T g^{(k-1)}}$$

针对二次函数的共轭梯度算法 (Fletcher & Reeves, 1964)

(0) 给定初始点 $x^{(0)}$, 计算 $g^{(0)}=g(x^{(0)})$, 令 $d^{(0)}=-g^{(0)}, k:=0$.

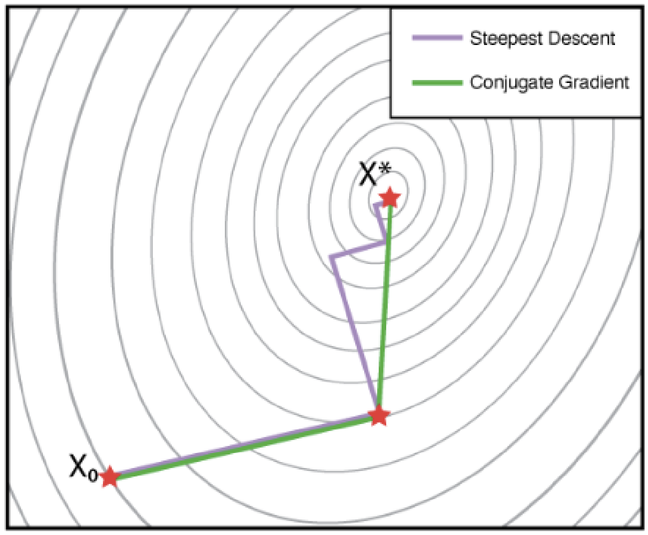
(1) 迭代更新 $x^{(k+1)}=x^{(k)}+\alpha_k d^{(k)},\ \ \alpha_k=\frac{{g^{(k)}}^T g^{(k)}}{{d^{(k)}}^T Gd^{(k)}}$.[3]

(2) 计算 $g^{(k+1)}=g(x^{(k+1)})$, 构造共轭梯度方向 $d^{(k+1)}=-g^{(k+1)}+\beta_k d^{(k)}$, 其中 $\beta_k=\frac{{g^{(k+1)}}^T g^{(k+1)}}{{g^{(k)}}^T g^{(k)}}$

(3) 置 $k:=k+1$, 返回第 (1) 步.

---

[3] 一维精确搜索最小值点



共轭梯度法性质定理:

设目标函数 $f(x)=\frac{1}{2}x^TGx+c^Tx$, 则采用精确一维搜索的共轭梯度法经 $m\le n$ 步迭代后终止, 且对所有的 $1\le k\le m$ 成立下列关系式:

$$d^{(k)^T}Gd^{(j)}=0, g^{(k)^T}g^{(j)}=0, j=0,1,...,k{-}1$$
$$d^{(k)^T}g^{(k)}=-g^{(k)^T}g^{(k)}$$
$$span\{g^{(0)},g^{(1)},...,g^{(k)}\}=span\{g^{(0)},Gg^{(0)},...,G^kg^{(0)}\}$$
$$span\{d^{(0)},d^{(1)},...,d^{(k)}\}=span\{d^{(0)},Gg^{(0)},...,G^kg^{(0)}\}$$

Pf. 记

$$d^{(k)^T}Gd^{(j)}=0, j=0,1,\cdot\cdot\cdot,k{-}1 \tag{CGM.1}$$
$$g^{(k)^T}g^{(j)}=0, j=0,1,\cdot\cdot\cdot,k{-}1 \tag{CGM.2}$$
$$d^{(k)^T}g^{(k)}=-g^{(k)^T}g^{(k)} \tag{CGM.3}$$
$$span\{g^{(0)},g^{(1)},...,g^{(k)}\}=span\{g^{(0)},Gg^{(0)},...,G^kg^{(0)}\} \tag{CGM.4}$$
$$span\{d^{(0)},d^{(1)},...,d^{(k)}\}=span\{d^{(0)},Gg^{(0)},...,G^kg^{(0)}\} \tag{CGM.5}$$

共轭梯度法迭代步骤中得到的条件

$$g^{(k+1)^T}d^{(k)}=0 \qquad (\text{由精确一维搜索}) \tag{CGM.6}$$
$$\alpha_k=\frac{g^{(k)^T}g^{(k)}}{d^{(k)^T}Gd^{(k)}} \qquad (\text{由精确一维搜索}) \tag{CGM.7}$$
$$g^{(k+1)}=g^{(k)}+\alpha_kGd^{(k)} \qquad (\text{由 } \nabla f(x^{(k+1)}) \text{ 直接展开}) \tag{CGM.8}$$
$$\beta_k=\frac{g^{(k+1)^T}g^{(k+1)}}{g^{(k)^T}g^{(k)}} \tag{CGM.9}$$
$$d^{(k+1)}=-g^{(k)}+\beta_kd^{(k)} \tag{CGM.10}$$

对于 $(CGM.3)$, 有

$$\begin{aligned} d^{(k)^T} g^{(k)} &= - g^{(k)^T} g^{(k)} + \beta_k d^{(k-1)^T} g^{(k)} \\ &= - g^{(k)^T} g^{(k)} + 0 \\ &= - g^{(k)^T} g^{(k)} \end{aligned}$$

对于 $(CGM.1),(CGM.2)$, 有

1.$k=1$ 时, 直接验证可得结论成立.

2. 假设 $(CGM.1),(CGM.2)$ 对 k 成立.

3.$(CGM.8)$ 两边转置后同右乘 $g^{(j)}$ 得

$$\begin{aligned} g^{(k+1)^T} g^{(j)} =& g^{(k)^T} g^{(j)} - \alpha_k d^{(k)^T} G g^{(j)} \\ =& g^{(k)^T} g^{(j)} - \alpha_k d^{(k)^T} G (d^{(j)} - \beta_{j-1} d^{(j-1)}) \\ =& g^{(k)^T} g^{(j)} - \alpha_k d^{(k)^T} G d^{(j)} \end{aligned}$$

$j=k$ 时将 $(CGM.7)$ 代入上式为 0;$j<k$ 时, 由归纳假设得上式为 0.

综上 $(CGM.2)$ 成立.

4. 由 $(CGM.10)$,

$$\begin{aligned} d^{(k+1)} G d^{(j)} =& (-g^{(k+1)} + \beta_k d^{(k)}) G d^{(j)} \\ =& - g^{(k+1)} G d^{(j)} + \beta_k d^{(k)} G d^{(j)} \\ =& g^{(k+1)^T} (g^{(j)} - g^{(j+1)}) / \alpha_k + \beta_k d^{(k)} G d^{(j)} \end{aligned}$$

$j=k$ 时将 $(CGM.2),(CGM.7),(CGM.9)$ 代入上式为 0;$j<k$ 时, 由归纳假设得上式为 0.

综上 $(CGM.1)$ 成立.

对于 $(CGM.4)$,

由 $(CGM.10)$ 可知, 存在可逆阵

$$Q = \begin{bmatrix} -1 & \beta_0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & \beta_1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & \beta_{k-1} \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{bmatrix}$$

使得 $(d^{(0)}, d^{(1)}, ..., d^{(k)}) Q = (g^{(0)}, g^{(1)}, ..., g^{(k)})$, 所以

$$span\{g^{(0)}, g^{(1)}, ..., g^{(k)}\} = span\{d^{(0)}, d^{(1)}, ..., d^{(k)}\} \tag{CGM.11}$$

1.$k=1$ 时, 直接由定义可得结论成立.

2. 假设结论对 k 成立, 即 $span\{g^{(0)}, g^{(1)}, ..., g^{(k)}\} = span\{g^{(0)}, Gg^{(0)}, ..., G^k g^{(0)}\}$

3. 对于 $k+1$, 由 $(CGM.8)$ 和归纳假设,

$$g_{k+1} \in span\{g^{(0)}, Gg^{(0)}, ..., G^{k+1} g^{(0)}\}$$

$d^{(0)}, d^{(1)}, ..., d^{(k)}$ 是一组共轭方向, 由共轭方向法基本定理得

$$g^{(k+1)} \perp span\{d^{(0)}, d^{(1)}, ..., d^{(k)}\} \tag{CGM12}$$

所以由 $(CGM.11), (CGM.12)$ 和归纳假设有

$$g^{(k+1)} \notin span\{d^{(0)}, d^{(1)}, ..., d^{(k)}\} = span\{g^{(0)}, g^{(1)}, ..., g^{(k)}\}$$

所以结论对 $k+1$ 成立, 即

$$span\{g^{(0)}, g^{(1)}, ..., g^{(k)}\} = span\{g^{(0)}, Gg^{(0)}, ..., G^k g^{(0)}\}$$

$(CGM.5)$ 证明与 $(CGM.4)$ 同理.

证毕.

将共轭梯度法推广到非二次函数的极小化问题, 其迭代为

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}.$$

步长 α_k 由精确或者非精确一维搜索决定, 而 $d^{(k+1)}$ 的构造如下:

$$d^{(k+1)} = -g^{(k+1)} + \beta_k d^{(k)}.$$

其中,

$$\begin{aligned}
\beta_k &:= \frac{{g^{(k+1)}}^T g^{(k+1)}}{{g^{(k)}}^T g^{(k)}} && (\text{Fletcher} - \text{Reeves}) \\
\beta_k &:= \frac{{g^{(k+1)}}^T (g^{(k+1)} - g^{(k)})}{{d^{(k)}}^T (g^{(k+1)} - g^{(k)})} && (\text{Hestenes} - \text{Stiefel}) \\
\beta_k &:= \frac{{g^{(k+1)}}^T (g^{(k+1)} - g^{(k)})}{{g^{(k)}}^T g^{(k)}} && (\text{Polak} - \text{Ribiere} - \text{Polyak}) \\
\beta_k &:= \frac{{g^{(k+1)}}^T g^{(k+1)}}{-{d^{(k)}}^T g^{(k)}} && (\text{Dixon}) \\
\beta_k &:= \frac{{g^{(k+1)}}^T g^{(k+1)}}{{d^{(k)}}^T (g^{(k+1)} - g^{(k)})} && (\text{Dai} - \text{Yuan})
\end{aligned} \tag{1}$$

对于非二次函数, 共轭梯度法迭代 n 步以后所产生的搜索方向 $d^{(k+1)} = -g^{(k+1)} + \beta_k d^{(k)}$ 可能不再是下降方向 (由非精确一维搜索造成的).

因此,n 步以后我们应该周期性采用最速下降方向作为搜索方向, 即令 $d^{(ln)} = -g^{(ln)}, l = 1, 2, ...$

这种策略称为重启动策略, 这样的共轭梯度法也称作重启动共轭梯度法.

如上所述的共轭梯度法迭代对于一般的非线性函数的最小化也是照样适用的, 但迭代更新的步长因子无法显式表达, 需要执行数值近似的非精确一维搜索.

由于每 n 步迭代执行重启动策略, 若记重新启动时得到的点列为 $\{z^{(j)}\}$, 则可证明这些相隔 n 次的迭代点列超线性收敛. 受实际计算误差的影响, 在很多情形下仅能取得类似线性的收敛速率.

从实际计算效率及稳定性来看, 共轭梯度法未必比拟牛顿法好. 但是, 共轭梯度法中搜索方向的计算仅仅用到目标函数的梯度, 而不必像拟牛顿法那样在每次迭代中更新 Hesse 矩阵 (或

其逆) 的近似阵并记忆之. 所以, 当问题的规模大而且有稀疏结构时, 共轭梯度法有高效执行计算的好处.

在大多数情况下, 为确保共轭梯度法的快速收敛, 预条件处理是必要的.

5.8 Trust-Region Method

为了保证迭代法的全局收敛性, 之前我们采用了一维搜索策略.

一维搜索策略先确定一个搜索方向 $d^{(k)}$, 然后沿着这个方向选择适当的步长因子 α_k, 新的迭代点 $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$.

现在, 我们讨论另一种全局收敛策略—**信赖域方法** (Trust-Region Method).

信赖域方法首先定义当前迭代点 $x^{(k)}$ 的邻域

$$\Omega_k = \{x \in \mathbb{R}^n \mid \|x - x^{(k)}\| \le e_k\},$$

这里 Ω_k 称为信赖域,e_k 是信赖域半径.

假定在这个邻域里, 二次模型 $q^{(k)}(s)$ 是目标函数 $f(x)$ 的一个合适的近似, 则在信赖域中极小化二次模型, 得到近似极小点 $s^{(k)}$, 并取 $x^{(k+1)} = x^{(k)} + s^{(k)}$.

信赖域方法利用二次模型在信赖域内求得方向步 $s^{(k)}$,, 使得目标函数的下降比一维搜索更有效.

信赖域方法不仅具有全局收敛性, 而且不要求目标函数的 Hesse 矩阵 (或其近似) 是正定的.

信赖域子问题

$$\begin{aligned} min \quad & q^{(k)}(s) = f(x^{(k)}) + g^{(k)^T} s + \frac{1}{2} s^T B_k s \\ s.t. \quad & \|s\| \le e_k. \end{aligned} \tag{116}$$

其中 $s = x - x^{(k)}, g^{(k)} = \nabla f(x^{(k)})$, 对称阵 B_k 是 Hesse 矩阵 $\nabla^2 f(x^{(k)})$ 或其近似,$e_k > 0$ 为信赖域半径,$\|\cdot\|$ 为某一范数.

如何选择信赖域半径 e_k ?

我们将根据二次模型 $q^{(k)}(s)$ 对目标函数 $f(x)$ 的拟合程度自适应地调整信赖域半径.

设子问题 (116) 的解 $s^{(k)}$, 令目标函数的下降量

$$Act_k = f(x^{(k)}) - f(x^{(k)} + s^{(k)})$$

为实际下降量, 令二次模型函数的下降量

$$Pre_k = q^{(k)}(0) - q^{(k)}(s^{(k)})$$

为预测下降量. 定义比值

$$r_k = \frac{Act_k}{Pre_k} = \frac{f(x^{(k)}) - f(x^{(k)} + s^{(k)})}{q^{(k)}(0) - q^{(k)}(s^{(k)})}$$

它衡量了二次模型与目标函数之间的一致程度.

当 r_k 越接近 1, 表明二次模型函数 $q^{(k)}(s)$ 与目标函数 f 的一致性程度越好, 此时可以增大半径 e_k 以扩大信赖域.

如果 $r_k > 0$ 但不接近 1, 我们保持信赖域半径 e_k 不变.

如果 r_k 接近零或取负值, 表明 $q^{(k)}(s)$ 与目标函数 f 的一致性程度不理想, 就减小半径 e_k 以缩小信赖域.

信赖域算法

(0) 给定初始点 $x^{(0)}$ ，信赖域半径的上界 $\bar{e}$,$\varepsilon > 0, 0 < \gamma_1 < \gamma_2 < 1, 0 < \eta_1 < 1 < \eta_2$. 取 $e_0 \in (0, \bar{e})$, 令 $k := 0$.

(1) 如果 $\|g^{(k)}\| \le \varepsilon$, 停止迭代. 否则, 求解信赖域子问题 (116) 得到 $s^{(k)}$.

(2) 计算比值 r_k, 更新迭代点

$$x^{(k+1)} = \begin{cases} x^{(k)} + s^{(k)} & if\ r_k > 0, \\ x^{(k)} & otherwise. \end{cases}$$

(3) 调整信赖域半径, 令

$$e_{k+1} = \begin{cases} \eta_1 e_k & if\ r_k < \gamma_1, \\ e_k & if\ \gamma_1 \le r_k < \gamma_2, \\ min(\eta_2 e_k, \bar{e}) & if\ r_k \ge \gamma_2. \end{cases}$$

(4) 置 $k := k + 1$, 返回第 (1) 步.

信赖域方法的全局收敛性定理:

设水平集 $L(x^{(0)}) = \{x \mid f(x) \le f(x^{(0)})\}$ 有界, 且 $f(x)$ 在其上 C^2 连续, 则由信赖域算法产生的迭代序列存在聚点 x^{∞}, 满足一阶和二阶必要条件, 即

$$g^{\infty} = \nabla f(x^{\infty}) = 0,\ G_{\infty} = \nabla^2 f(x^{\infty}) \ge 0.$$

在信赖域算法中关键的一步是, 解信赖域子问题 (116).

这里我们介绍一种求解信赖域子问题的方法, 即由 Powell (1970) 提出的折线法.

所谓折线法, 是连接 Cauchy 点 (由最速下降法产生的极小点) 和牛顿点 (由牛顿法产生的极小点), 其连线与信赖域边界的交点取为 $x^{(k+1)}$.

对于二次模型

$$q^{(k)}(-\alpha g^{(k)}) = f(x^{(k)}) - \alpha \|g^{(k)}\|^2 + \frac{1}{2}\alpha^2 g^{(k)^T} B_k g^{(k)},$$

精确一维搜索的步长因子可表达为

$$\alpha_k = \frac{\|g^{(k)}\|^2}{g^{(k)^T} G_k g^{(k)}}.$$

于是 Cauchy 步为

$$s_C^{(k)} = -\alpha_k g^{(k)} = -\frac{\|g^{(k)}\|^2}{g^{(k)^T} B_k g^{(k)}} g^{(k)}.$$

如果 $\|s^{(k)}\| = \|\alpha_k g^{(k)}\| \ge e_k$, 取

$$s^{(k)} = -\frac{e_k}{\|g^{(k)}\|} g^{(k)}.$$

便得

$$x^{(k+1)} = x^{(k)} - \frac{e_k}{\|g^{(k)}\|} g^{(k)}.$$

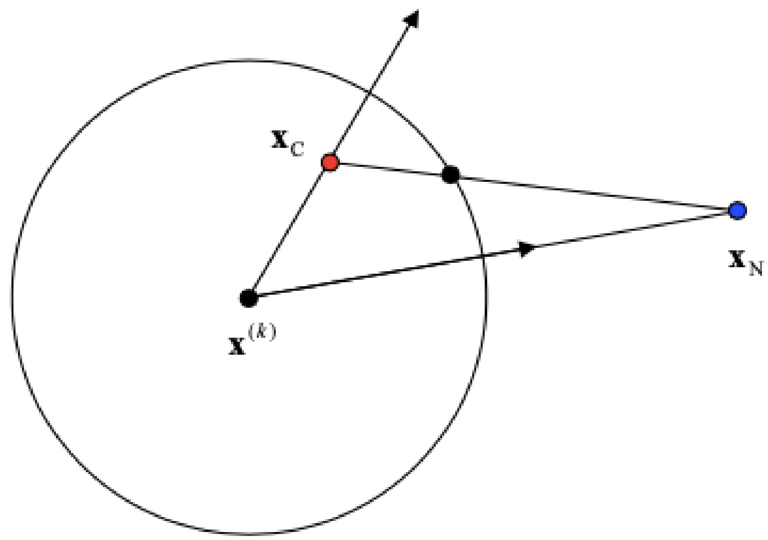


图 1: 折线法图示

如果 $\|s_C^{(k)}\| = \|\alpha_k g^{(k)}\| < e_k$, 再计算牛顿步

$$s_N^{(k)} = -B_k^{-1} g^{(k)}.$$

如果 $s_N^{(k)}\| \leq e_k$, 取

$$s^{(k)} = s_N^{(k)} = -B^{-1} g^{(k)},$$

否则, 取

$$s^{(k)} = s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)}),$$

其中 λ 使得

$$\|s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})\| = e_k.$$

综上所述, 我们得到

$$s^{(k+1)} = \begin{cases} x^{(k)} - \frac{e_k}{\|g^{(k)}\|} g^{(k)} & if\ \|s_C^{(k)}\| \geq e_k \\ x^{(k)} + s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)}) & if\ \|s_C^{(k)}\| < e_k\ \ And\ \ \|s_N^{(k)}\| > e_k \\ x^{(k)} - B_k^{-1} g^{(k)} & if\ \|s_C^{(k)}\| < e_k\ \ And\ \ \|s_N^{(k)}\| \leq e_k \end{cases} \tag{117}$$

折线法满足下列性质:
1) 沿着 Cauchy 点 $x_C^{(k+1)}$ 和牛顿点 $x_N^{(k+1)}$ 的连线, 到 $x^{(k)}$ 的距离单调增加;
2) 沿着 Cauchy 点 $x_C^{(k+1)}$ 和牛顿点 $x_N^{(k+1)}$ 的连线, 子问题模型函数值单调减少.
Pf.

(1).

$$
\begin{aligned}
L(\lambda) \triangleq& \|s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})\|^2 \\
=& (s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)}))^T (s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
=& s_C^{(k)^T} s_C^{(k)} + 2\lambda s_C^{(k)^T} (s_N^{(k)} - s_C^{(k)}) + \lambda^2 (s_N^{(k)} - s_C^{(k)})^T (s_N^{(k)} - s_C^{(k)}) \\
L'(\lambda) =& 2 s_C^{(k)^T} (s_N^{(k)} - s_C^{(k)}) + 2\lambda (s_N^{(k)} - s_C^{(k)})^T (s_N^{(k)} - s_C^{(k)}) \\
\geq& 2 s_C^{(k)^T} (s_N^{(k)} - s_C^{(k)}) \\
=& 2 \frac{\|g^{(k)}\|^2}{g^{(k)T} B_k g^{(k)}} g^{(k)^T} B_k^{-1} g^{(k)} \left(1 - \frac{\|g^{(k)}\|^2 g^{(k)}}{g^{(k)T} B_k g^{(k)} g^{(k)T} B_k^{-1} g^{(k)}}\right)
\end{aligned}
$$

$$
\begin{aligned}
1 - \frac{\|g^{(k)}\|^2 g^{(k)^T} g^{(k)}}{g^{(k^T)} B_k g^{(k)} g^{(k)^T} B_k^{-1} g^{(k)}} =& 1 - \frac{\|g^{(k)}\|^4}{g^{(k)T} B_k g^{(k)} g^{(k)T} B_k^{-1} g^{(k)}} \\
=& 1 - \frac{((\sqrt{B_k} g^{(k)})^T \sqrt{B_k^{-1}} g^{(k)})^2}{g^{(k)T} B_k g^{(k)} g^{(k)T} B_k^{-1} g^{(k)}} \\
\geq& 1 - \frac{\|\sqrt{B_k} g^{(k)})\|^2 \|\sqrt{B_k^{-1}} g^{(k)}\|^2}{g^{(k)T} B_k g^{(k)} g^{(k)T} B_k^{-1} g^{(k)}} \\
=& 1 - \frac{g^{(k)^T} B_k g^{(k)} g^{(k)^T} B_k^{-1} g^{(k)}}{g^{(k)T} B_k g^{(k)} g^{(k)T} B_k^{-1} g^{(k)}} \\
=& 0
\end{aligned}
$$

因此 $L'(\lambda) \geq 0, \lambda \in [0,1]$, 距离单调增加.

(2).

$$
\begin{aligned}
h(\lambda) \triangleq& q^{(k)}(s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
=& f(x^{(k)}) + g^{(k)^T} (s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
&+ \frac{1}{2}(s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)}))^T B_k (s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
h'(\lambda) =& g^{(k)^T} (s_N^{(k)} - s_C^{(k)}) + s_C^{(k)^T} B_k (s_N^{(k)} - s_C^{(k)}) + \lambda (s_N^{(k)} - s_C^{(k)})^T B_k (s_N^{(k)} - s_C^{(k)}) \\
\leq& g^{(k)^T} (s_N^{(k)} - s_C^{(k)}) + s_C^{(k)^T} B_k (s_N^{(k)} - s_C^{(k)}) + (s_N^{(k)} - s_C^{(k)})^T B_k (s_N^{(k)} - s_C^{(k)}) \\
=& (g^{(k)^T} + s_C^{(k)^T} B_k + (s_N^{(k)} - s_C^{(k)}) B_k)(s_N^{(k)} - s_C^{(k)}) \\
=& (g^{(k)^T} + s_N^{(k)^T} B_k)(s_N^{(k)} - s_C^{(k)}) \\
=& 0 \cdot (s_N^{(k)} - s_C^{(k)}) \\
=& 0
\end{aligned}
$$

因此 $h'(\lambda) \leq 0, \lambda \in [0,1]$, 子问题模型函数值单调减少.

6 Quadratic Programming

二次规划 (Quadratic Programming) 是指, 在变量的线性等式和/或不等式限制下求二次函数的极小点问题

$$\begin{aligned} min \quad & Q(x) = \frac{1}{2}x^T G x + c^T x \\ s.t. \quad & a_i^T x = b_i, i \in \mathcal{E} = \{1, \cdot\cdot\cdot, m_e\} \\ & a_i^T x \geq b_i, i \in \mathcal{I} = \{m_e + 1, \cdot\cdot\cdot, m\} \end{aligned} \tag{118}$$

我们假定 G 为对称阵,$a_i(i \in \mathcal{E})$ 是线性无关的.

二次规划的约束可能不相容, 也可能没有有限的最小值, 这时称二次规划问题无解.

如果矩阵 G 半正定, 问题 (118) 是凸二次规划问题, 它的任意局部解也是整体解.

如果矩阵 G 正定, 问题 (118) 是正定二次规划问题, 只要存在解即是唯一的.

如果矩阵 G 不定, 问题 (118) 是一般的二次规划问题, 有可能出现非整体解的局部解.

6.1 Equality Constrained Quadratic Programs Problem

等式约束二次规划问题

$$\begin{aligned} min \quad & Q(x) = \frac{1}{2}x^T G x + c^T x \\ s.t. \quad & Ax = b \end{aligned} \tag{119}$$

这里 $A \in \mathbb{R}^{m \times n}$, 且不失一般性可设 $rank(A) = m$.

设有一种基分解 $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$, 其中 $x_B \in \mathbb{R}^m, x_N \in \mathbb{R}^{n-m}$, 使得其约束矩阵的对应分块 $A = \begin{pmatrix} A_B & A_N \end{pmatrix}$ 中 A_B 可逆. 于是, 等式约束条件可写成

$$x_B = A_B^{-1}(b - A_N x_N),$$

并将上式代入目标函数中得到无约束问题

$$\min_{x_N \in \mathbb{R}^{n-m}} \frac{1}{2} x_N^T \widehat{G}_N x_N + \widehat{c}_N^T x_N. \tag{120}$$

, 其中

$$\widehat{G}_N = G_{NN} - G_{NB}A_B^{-1}A_N - A_N^T G_{BN} + A_N^T A_N^{-T} G_{BB} A_B^{-1} A_N$$
$$\widehat{c}_N = c_N - A_N^T A_B^{-T} c_B + G_{NB} A_B^{-1} b - A_N^T A_B^{-T} G_{BB} A_B^{-1} b$$

以及对应分块形式

$$G = \begin{pmatrix} G_{BB} & G_{BN} \\ G_{NB} & G_{NN} \end{pmatrix}, c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}.$$

(1) 如果 $\widehat{G}_N$ 正定, 则无约束问题的解可唯一地给出

$$x_N^* = -\widehat{G}_N^{-1} \widehat{c}_N,$$

进一步得原问题 (119) 的解为

$$x^* = \begin{pmatrix} x_B^* \\ x_N^* \end{pmatrix} = \begin{pmatrix} A_B^{-1}b \\ 0 \end{pmatrix} + \begin{pmatrix} A_B^{-1}A_N \\ -I \end{pmatrix} \widehat{G}_N^{-1}\widehat{c}_N.$$

设 x^* 对应的 Lagrange 乘子向量为 λ^*, 则有

$$Gx^* + c = A^T\lambda^* \Rightarrow \lambda^* = A_B^{-T}(G_{BB}x_B^* + G_{BN}x_N^* + c_B).$$

(2) 如果 $\widehat{G}_N$ 是半正定的, 则在 $(I-\widehat{G}_N\widehat{G}_N^+)\widehat{c}_N = 0$ 时, 无约束问题有界, 且它的解可表示为

$$c_N^* = -\widehat{G}_N^+\widehat{c}_N + (I - \widehat{G}_N^+\widehat{G}_N)\tilde{y},$$

其中 $\tilde{y} \in \mathbb{R}^{n-m}$ 为任意向量,$\widehat{G}_N^+$ 表示 $\widehat{G}_N$ 的广义逆矩阵. 此时, 原问题的解 x^* 和相应最优乘子 λ^* 可类似确定.
当 $(I-\widehat{G}_N\widehat{G}_N^+)\widehat{c}_N = 0$ 不成立时, 则可推出无约束问题无下界, 从而原问题也无下界.

(3) 如果 $\widehat{G}_N$ 不定 (即存在负的特征根), 显然无约束问题无下界, 故原问题不存在有限最优解.

上述消去法的不足之处是, 当 A_B 接近奇异时, 容易导致数值计算的不稳定.

6.2 Generalized Elimination Method

设 $Z = \{z_{m+1}, ..., z_n\}$ 解空间 $Ker(A)$ 的一组基,$Y = \{y_1, \cdot\cdot\cdot, y_m\}$ 是商空间 $\mathbb{R}^n/Ker(A)$ 的一组基, 则 $\forall x \in \mathbb{R}^n$ 可作如下分解表达

$$x = Yx_Y + Zx_Z$$

从而有

$$Ax = b \Rightarrow AYx_Y + AZx_Z = b \Rightarrow x_Y = (AY)^{-1}b,$$

所以得

$$x = Y(AY)^{-1}b + Zx_Z,$$

其中 $x_Z \in \mathbb{R}^{n-m}$ 是自由变量.

将上式代入目标函数中得无约束问题

$$\min_{x_Z \in \mathbb{R}^{n-m}} \frac{1}{2}x_Z^T(Z^TGZ)x_Z + [Z^TGY(AY)^{-1}b + Z^Tc]^Tx_Z. \tag{121}$$

假定 Z^TGZ 正定, 则有

$$x_Z^* = -(Z^TGZ)^{-1}Z^T[GY(AY)^{-1}b + c].$$

从而得到原问题的最优解

$$x^* = Y(AY)^{-1} - Z(Z^TGZ)^{-1}Z^T[GY(AY)^{-1}b + c],$$

相应的 Lagrange 乘子为

$$\lambda^* = (AY)^{-T}Y^T(Gx^* + c).$$

Lagrange 方法是基于求解可行域内的 (K-T) 点, 即 Lagrange 函数的稳定点. 对于等式约束问题 (119), 其 Lagrange 函数的稳定点就是如下线性方程组的解

$$\begin{cases} Gx + c = A^T\lambda, \\ Ax = b. \end{cases}$$

写成矩阵形式得

$$\begin{pmatrix} G & -A^T \\ -A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = - \begin{pmatrix} c \\ b \end{pmatrix}$$

设矩阵 $\begin{pmatrix} G & -A^T \\ -A & 0 \end{pmatrix}$ 可逆, 则存在矩阵 $U \in \mathbb{R}^{n\times n}, V \in \mathbb{R}^{m\times m}, W \in \mathbb{R}^{m\times n}$ 使得

$$\begin{pmatrix} G & -A^T \\ -A & 0 \end{pmatrix}^{-1} = \begin{pmatrix} U & W^T \\ W & V \end{pmatrix}$$

从而可求得问题的唯一解

$$\begin{cases} x^* = -U_C - W^T b, \\ \lambda^* = -W_C - Vb. \end{cases}$$

上述 Lagrange 方法中的矩阵非奇异性并不一定要求 G^{-1} 存在, 可用不同的方法给出分块矩阵 U, V, W 的表达形式, 从而导致不同的计算公式. 当 G 可逆,A 行满秩, 则 $(AG^{-1}A^T)^{-1}$ 存在, 不难验证

$$\begin{cases} U = G^{-1} - G^{-1}A^T(AG^{-1}AT)^{-1}AG^{-1}, \\ V = -(AG^{-1}A^T)^{-1}, \\ W = -(AG^{-1}A^T)^{-1}AG^{-1}. \end{cases}$$

于是我们得到求解公式

$$\begin{cases} x^* = -G^{-1}c + G^{-1}A^T(AG^{-1}A^T)^{-1}(AG^{-1}c + b), \\ \lambda^* = (AG^{-1}A^T)^{-1}(AG^{-1}c + b). \end{cases}$$

如果取 Y, Z 满足 $\begin{pmatrix} Y & Z \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}^{-1}$, 即 $AY = I_{m\times m}, AZ = 0$. 若另有 Z^TGZ 可逆, 则知 $\begin{pmatrix} G & -A^T \\ -A & 0 \end{pmatrix}$ 可逆. 此时

$$\begin{cases} U = Z(Z^TGZ)^{-1}Z^T, \\ V = -Y^TGP^TY, \\ W = -Y^TP \end{cases}$$

其中 $P = I - GZ(Z^TGZ)^{-1}Z^T$. 基于 A^T 的 QR 分解, 可给出 $\begin{pmatrix} Y & Z \end{pmatrix}$ 的一种特殊取法: 设

$$A^T = Q \begin{pmatrix} R \\ 0 \end{pmatrix} = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R \\ 0 \end{pmatrix}$$

即

$$A = \begin{pmatrix} R^T & 0 \end{pmatrix} \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix}$$

其中 Q 为 $n \times n$ 正交阵,R 为 $m \times m$ 上三角阵. 于是令 $Y = Q_1 R^{-T}, Z = Q_2$, 则有

$$AY = R^T Q_1^T Q_1 R^{-T} = I_{m\times m}, AZ = R^T Q_1^T Q_2 = 0_{m\times(n-m)}.$$

6.3 Active-Set Methods

一般的二次规划

$$\begin{aligned} min \quad & Q(x) = \frac{1}{2}x^T Gx + c^T x \\ s.t. \quad & a_i^T x = b_i, i \in \mathcal{E} = \{1, \cdot\cdot\cdot, me\} \\ & a_i^T x \geq b_i, i \in \mathcal{I} = \{m_{e+1}, \cdot\cdot\cdot, m\} \end{aligned} \tag{122}$$

直观上, 不积极的不等式约束在解的附近不起作用, 可去掉不予考虑; 而积极的不等式约束, 由于它在解处等号成立, 故我们可以用等式约束来代替这些积极的不等式约束.

积极集基本定理:

设 x^* 是一般的二次规划问题 (122) 的局部极小点, 则 x^* 也必是等式约束问题

$$(EQ)\begin{cases} min & Q(x) = \frac{1}{2}x^T Gx + c^T x \\ s.t. & a_i^T x = b_i, i \in \mathcal{E} \cup \mathcal{I}(x^*) \end{cases}$$

的局部极小点. 反之, 如果 x^* 是一般问题 (122) 的可行点, 同时是 (EQ) 的 K-T 点, 且相应的 Lagrange 乘子 λ^* 满足 $\lambda_i^* \geq 0, i \in \mathcal{I}(x^*)$, 则 x^* 必是原问题 (122) 的 K-T 点.

Pf.

当 x^* 是 (122) 的局部极小点时,

$$L(x, \lambda, \mu) \triangleq Q(x) - \sum_{i\in\mathcal{I}} \lambda_i (a_i^T - b_i) - \sum_{j\in\mathcal{E}} \mu_j (a_j^T x - b_j)$$

x^* 满足 K-T 条件

$$\begin{cases} \nabla L(x^*, \lambda, \mu) = Gx^* + c - \sum_{i\in\mathcal{I}} \lambda_i a_i - \sum_{j\in\mathcal{E}} \mu_j a_j = 0, \\ \lambda_i \geq 0, \quad a_i^T - b_i \geq 0, \quad \lambda_i (a_i^T x - b_i) = 0, \quad i \in \mathcal{I}, \\ a_j^T x - b_j = 0, \quad j \in \mathcal{E} \end{cases}$$

当 $i \in \mathcal{I}(x^*)$ 时, 有 $a_i^T - b_i = 0$, 此时 λ_i 不必等于 0.

而当 $i \in \mathcal{I} \setminus \mathcal{I}(x^*)$ 时,$a_i^T - b_i \neq 0, \lambda_i = 0$. 故有

$$\begin{cases} \nabla G x^* + c - \sum_{i\in\mathcal{I}(x^*)} \lambda_i a_i - \sum_{j\in\mathcal{E}} \mu_j a_j = 0, \\ (a_i^T x^* - b_i) = 0, \quad i \in \mathcal{I}(x^*) \cup \mathcal{E} \end{cases}$$

即

$$\begin{cases} \nabla G x^* + c - \sum_{i\in\mathcal{I}(x^*)\cup\mathcal{E}} \lambda_i a_i = 0, \\ (a_i^T x^* - b_i) = 0, \quad i \in \mathcal{I}(x^*) \cup \mathcal{E} \end{cases}$$

x^* 是 (EQ) 问题的局部极小值点 (K-T 点).

反之, 若 x^* 为 (122) 的可行点, 同时是 (EQ) 的 K-T 点, 且 $\lambda_i^* \geq 0$, 则有

$$\begin{cases} \nabla G x^* + c - \sum_{i\in\mathcal{I}(x^*)\cup\mathcal{E}} \lambda_i a_i = 0, \\ \lambda_i^*(a_i^T x^* - b_i) = 0, \qquad i \in \mathcal{I}(x^*)\cup\mathcal{E} \\ a_i^T x^* - b_i = 0, \qquad i \in \mathcal{E} \\ a_i^T x^* - b_i \geq 0, \qquad i \in \mathcal{I} \\ \lambda_i^* \geq 0, i \in \mathcal{I}(x^*) \end{cases}$$

取 $\lambda_i^*(a_i^T x^* - b_i) = 0, i \in \mathcal{I}\setminus\mathcal{I}(x^*)$, , 则有

$$\begin{cases} \nabla G x^* + c - \sum_{i\in\mathcal{I}} \lambda_i a_i - \sum_{i\in\mathcal{E}} \lambda_i a_i = 0, \\ \lambda_i \geq 0, \qquad a_i^T x^* - b_i \geq 0, \qquad \lambda_i^*(a_i^T x^* - b_i) = 0, \qquad i \in \mathcal{I} \\ a_i^T x^* - b_i = 0, \qquad i \in \mathcal{E} \end{cases}$$

这意味着 x^* 是问题 (122) 的 K-T 点, 证毕.

设 $x^{(k)}$ 为当前迭代点, 且是问题 (122) 的可行点. 记 $\mathcal{E}_k = \mathcal{E}\cup\mathcal{I}(x^{(k)})$, 考虑等式约束问题

$$(EQ1)\begin{cases} min \qquad \frac{1}{2}s^T G s + (Gx^{(k)} + c)^T s \\ s.t. \qquad a_i^T s = 0, \qquad i \in \mathcal{E}_k \end{cases}$$

求得 $(EQ1)$ 的解 $s^{(k)}$, 及其相应的 Lagrange 乘子 $\lambda_i^{(k)}, i \in \mathcal{E}_k$.

(a) $s^{(k)} \neq 0$ 时,$x^{(k)}$ 不可能是原问题 (122) 的最优解.

(b) $s^{(k)} = 0$ 时,$x^{(k)}$ 是问题

$$(EQ2)\begin{cases} min \qquad \frac{1}{2}x^T G x + c^T x \\ s.t. \qquad a_i^T x = b_i, i \in \mathcal{E}_k \end{cases}$$

的 K-T 点; 如果 $\lambda^{(k)} \geq 0, i \in \mathcal{I}(x^{(k)})$, 则 $x^{(k)}$ 也是原问题的 K-T 点.

(c) 否则, 由 $\lambda_{i_q}^{(k)} = \min\limits_{i\in\mathcal{I}(x^{(k)})} \lambda_i^{(k)} < 0$ 确定 i_q, 那么如下问题

$$(EQ3)\begin{cases} min \qquad \frac{1}{2}s^T G s + (Gx^{(k)} + c)^T s \\ s.t. \qquad a_i^T s = 0, \qquad i \in \widehat{\mathcal{E}} = \mathcal{E}_k \setminus \{i_q\}. \end{cases}$$

的解 $\hat{s}$ 是原问题在当前点 $x^{(k)}$ 处的可行方向, 即 $a^T \hat{s} \geq 0$.

Pf.

证明 (a).

$s = 0$ 显然是可行解, 若最优解 $s \neq 0$, 则有 $\frac{1}{2}s^T G s + (Gx^{(k)} + c)^T s < 0$.

设 $x^{(k)}$ 是 (122) 的 K-T 点, 则有

$$Gx^{(k)} + c = \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i^* a_i = \sum_{i\in\mathcal{E}_k} \lambda_i^* a_i \tag{ASM1.1}$$

, 其中 λ^* 为 (122) 的 Lagrange 乘子.

$(EQ1)$ 的 K-T 条件为

$$Gs^{(k)} + Gx^{(k)} + c = \sum_{i\in\mathcal{E}_k} \lambda_i^{(k)} a_i \tag{ASM1.2}$$

$(ASM1.1)$ 与 $(ASM1.2)$ 相减, 得到

$$Gs^{(k)} = \sum_{i\in\mathcal{E}_k} (\lambda_i^{(k)} - \lambda_i^*) a_i$$

左乘 $s^{(k)^T}$, 得到

$$s^{(k)^T} Gs^{(k)} = \sum_{i\in\mathcal{E}_k} (\lambda_i^{(k)} - \lambda_i^*) s^{(k)^T} a_i = 0$$

根据 $\frac{1}{2}s^T Gs + (Gx^{(k)} + c)^T s < 0$, 有

$$(Gx^{(k)} + c)^T s < 0$$

而

$$0 > Gx^{(k)} + c)^T s = \sum_{i\in\mathcal{E}_k} \lambda_i^* a_i^T s^{(k)} = 0$$

矛盾.

因此 $s^{(k)} \neq 0$ 时,$x^{(k)}$ 不可能是原问题 (122) 的最优解.

证明 (b).

$x^{(k)}$ 为 $(EQ2)$ 的 K-T 点, 等价于

$$Gx^{(k)} + c = \sum_{i\in\mathcal{E}_k} \lambda_i a_i$$

$x^{(k)}$ 为 $(EQ1)$ 的 K-T 点, 有

$$Gs^{(k)} + Gx^{(k)} + c = \sum_{i\in\mathcal{E}_k} \lambda_i a_i$$

当 $s^{(k)} = 0$ 时,

$$\begin{cases} Gx^{(k)} + c = 0 \\ a_i^T x^{(k)} - b_i = 0, \quad i \in \mathcal{E}_k \end{cases}$$

$x^{(k)}$ 为 $(EQ2)$ 的 K-T 点.

下面设 $\lambda_i \geq 0, i \in \mathcal{I}(x^{(k)})$. 令 $\lambda_j = 0, j \in \mathcal{I} \setminus \mathcal{I}(x^{(k)})$

则有

$$\begin{cases} Gx^{(k)} + c = 0 \\ \lambda_i(a_i^T x^{(k)} - b_i) = 0, & i \in \mathcal{I} \\ \lambda_i \geq 0, & i \in \mathcal{I} \\ a_i^T x^{(k)} - b_i = 0, & i \in \mathcal{E} \end{cases}$$

$x^{(k)}$ 为 (122) 的 K-T 点.

证明 (c).

由 (EQ1),(EQ2),(EQ3), 我们可以得到

$$a_{i_q}^T x^{(k)} = b_{i_q}$$

$$G(\hat{s} + x^{(k)}) + c - \sum_{i \in \mathcal{E}} \lambda_i a_i = 0$$

$$Gx^{(k)} + c - \sum_{i \in \mathcal{E}_k} \lambda_i a_i = 0$$

$$\Rightarrow G\hat{s} + \lambda_{i_q} a_{i_q} = 0, \qquad \hat{s}^T G\hat{s} + \lambda_{i_q} a_{i_q}^T \hat{s} = 0$$

要证明 $a_{i_q}^T \hat{s} \geq 0$, 只需要证明 $\hat{s}^T G\hat{s} \geq 0$.

反设 $\hat{s}^T G\hat{s} < 0$, 则 $\hat{s}$ 在 G 的负特征值方向上, 适当改变 $\hat{s}$ 可以使得 $\hat{s}^T G\hat{s} \to -\infty$, 原问题没有最优解, 矛盾.

综上,$a_{i_q}^T \hat{s} \geq 0$,$\hat{s}$ 为可行方向.

积极集方法 (Active Set Method)

(0) 给出可行点 $x^{(0)}$, 令 $\mathcal{E}_0 = \mathcal{E} \cup \mathcal{I}(x^{(0)}), k := 0$

(1) 求解等式约束问题 $(EQ1)$ 得 $s^{(k)}$, 若 $s^{(k)} \neq 0$, 转第 (3) 步.

(2) 如果 $\lambda_i^{(k)} \geq 0, i \in \mathcal{I}(x^{(k)})$, 则停止; 否则由 $\lambda_{i_q}^{(k)} = \min\limits_{i \in \mathcal{I}(x^{(k)})} \lambda_i^{(k)} < 0$ 确定 i_q 并令 $\mathcal{E}_k := \mathcal{E}_k \setminus \{i_q\}, x^{(k+1)} = x^{(k)}$, 转第 (4) 步.

(3) 由 $\alpha_k = min\{1, \min\limits_{i \notin \mathcal{E}_k, a_i^T s^{(k)} < 0} \frac{b_i - a_i^T x^{(k)}}{a_i^T s^{(k)}}\}$, 计算 $x^{(k+1)} = x^{(k)} + \alpha_k s^{(k)}$. 如果 $\alpha_k = 1$, 转第 (4) 步; 不然定可找到 $p \notin \mathcal{E}_k$ 使得 $a_p^T(x^{(k)} + \alpha_k s^{(k)}) = b_p$, 并令 $\mathcal{E}_k := \mathcal{E}_k \cup \{p\}$.

(4) $\mathcal{E}_{k+1} := \mathcal{E}_k, k := k+1$, 返回第 (1) 步.

7 Nonlinear Constrained Optimization

7.1 Lagrange-Newton Methods

首先考虑等式约束问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & c_i(x) = 0, i \in \mathcal{E} = \{1, ..., m\} \end{aligned} \tag{123}$$

记 $A(x) = [\nabla c(x)]^T = [\nabla c_1(x), ..., \nabla c_m(x)]^T$

由最优性条件知:x 是等式约束问题 (123) 的 K-T 点当且仅当存在乘子 $\lambda \in \mathbb{R}^m$ 使得

$$\nabla f(x) - A(x)^T \lambda = 0,$$

且 x 是一可行点, 即 $c(x) = 0$.

于是得到联立方程组

$$\begin{cases} \nabla f(x) - A(x)^T \lambda = 0, \\ -c(x) = 0. \end{cases}$$

我们可用 Newton-Raphson 迭代法求解上述联立方程组. 记 x 和 λ 的计算增量分别为 δ_x, δ_λ, Newton-Raphson 迭代满足:

$$\begin{pmatrix} W(x,\lambda) & -A(x)^T \\ -A(x) & 0 \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = -\begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ -c(x) \end{pmatrix} \tag{124}$$

, 其中,

$$W(x,\lambda) = \nabla^2 f(x) - \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x).$$

上述方法称为 Lagrange-Newton 法, 最早由 Wilson(1963) 提出的. 其实质上是用 Newton-Raphson 迭代求问题 (123) 的 Lagrange 函数 $L(x,\lambda)$ 的稳定点.

在此, 我们定义价值函数

$$\psi(x,\lambda) = \|\nabla f(x) - A(x)^T\lambda\|^2 + \|c(x)\|^2. \tag{125}$$

显然,$\psi(x,\lambda)$ 是关于 Lagrange-Newton 法的下降函数, 即满足

$$\nabla\psi(x,\lambda)^T \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = -2\psi(x,\lambda) \leq 0$$

Pf.

$$\begin{aligned} \nabla\psi(x,\lambda) =& 2(\nabla^2 f(x) - \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x))(\nabla f(x) - A(x)^T\lambda) \\ &+ 2A(x)^T c(x) - 2A(x)(\nabla f(x) - A(x)^T\lambda) \\ =& 2\begin{pmatrix} W(x,\lambda) & -A(x)^T \\ -A(x) & 0 \end{pmatrix} \begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ -c(x) \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \nabla\psi(x,\lambda)^T \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} =& 2\begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ -c(x) \end{pmatrix}^T \begin{pmatrix} W(x,\lambda) & -A(x)^T \\ -A(x) & 0 \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} \\ \overset{(124)}{=}& -2\begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ -c(x) \end{pmatrix}^T \begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ -c(x) \end{pmatrix} \\ =& -2\|\nabla f(x) - A(x)^T\|^2 - 2\|c(x)\|^2 \\ =& -2\psi(x,\lambda) < 0 \end{aligned}$$

Lagrange-Newton 法迭代步骤:

(0) 给定 $x^{(0)}, \lambda \in \mathbb{R}^m, \beta \in (0,1), \varepsilon \geq 0$, 令 $k := 0$.

(1) 计算价值函数 $\psi(x^{(k)}, \lambda^{(k)})$, 如果 $\psi(x^{(k)}, \lambda^{(k)}) \leq \varepsilon$, 停止计算; 否则在 $(x^{(k)}, \lambda^{(k)})$ 处求解 (124) 得到 $(\delta_{x^{(k)}}, \delta_{\lambda^{(k)}})$, 并令 $\alpha_k = 1$.

(2) 若 $\psi(x^{(k)} + \alpha_k\delta_{x^{(k)}}, \lambda^{(k)} + \alpha_k\delta_{\lambda^{(k)}}) \leq (1 - \beta\alpha_k)\psi(x^{(k)}, \lambda^{(k)})$, 转第 (3) 步; 否则令 $\alpha_k = \frac{1}{4}\alpha_k$, 返回第 (2) 步.

(3) 置 $x^{(k+1)} = x^{(k)} + \alpha_k\delta_{x^{(k)}}, \lambda^{(k+1)} = \lambda^{(k)} + \alpha_k\delta_{\lambda^{(k)}}, k := k+1$, 返回第 (1) 步.

Lagrange-Newton 法的收敛性结果

定理:

设 Lagrange-Newton 法产生的迭代点列 $\{(x^{(k)}, \lambda^{(k)})\}$ 有界, 如果 $f(x)$ 和 $c_i(x)$ 都是二次连续可微, 且逆矩阵

$$\begin{pmatrix} W(x,\lambda) & -A(x)^T \\ -A(x) & 0 \end{pmatrix}^{-1}$$

一致有界, 则 $\{(x^{(k)}, \lambda^{(k)})\}$ 的任何聚点都是方程 $\psi(x,\lambda) = 0$ 的根, 从而 $\{x^{(k)}\}$ 的聚点是问题 (123) 的 K-T 点.

注: 在一定条件下, 还可进一步证明 Lagrange-Newton 法具有二阶收敛速度.

7.2 Sequential Quadratic Programming Methods

Lagrange-Newton 法的一大重要贡献是, 在其基础上发展出了逐步二次规划方法 (Sequential Quadratic Programming Methods). 而后者已成为求解一般非线性约束最优化问题的一类十分重要的方法.

我们可将式 (124) 写成如下形式:

$$\begin{cases} W(x,\lambda)\delta_x + \nabla f(x) = A(x)^T(\lambda + \delta_\lambda). \\ c(x) + A(x)\delta_x = 0. \end{cases}$$

由最优性条件知,$\delta_{x^{(k)}}$ 即为下列二次规划问题

$$\begin{aligned} min \quad & \frac{1}{2} d^T W(x^{(k)}, \lambda^{(k)}) d + \nabla f(x^{(k)})^T d \\ s.t. \quad & c(x^{(k)}) + A(x^{(k)}) d = 0 \end{aligned} \tag{126}$$

的 K-T 点.

Lagrange-Newton 法可以理解为逐步求解上述等式约束二次规划的方法. 设 $d^{(k)}$ 是二次规划问题 (126) 的最优解, 那么可迭代更新

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)},$$

其中 α_k 为第 k 次迭代的步长.

设 $\bar{\lambda}^{(k)}$ 是 (126) 对应的 Lagrange 乘子向量, 那么对 $k \geq 1$ 有

$$\lambda^{(k+1)} = \lambda^{(k)} + \alpha_k(\bar{\lambda}^{(k)} - \lambda^{(k)}).$$

现考虑一般的非线性约束最优化问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & c_i(x) = 0, i \in \mathcal{E} = \{1, ..., m_e\} \\ & c_i(x) \geq 0, i \in \mathcal{I} = \{m_e + 1, ..., m\}. \end{aligned} \tag{127}$$

类似地, 在第 k 次迭代里求解子问题

$$\begin{aligned} min \quad & \frac{1}{2}d^T W_k d + {g^{(k)}}^T d \\ s.t. \quad & c_i(x^{(k)}) + a_i(x^{(k)})^T d = 0, i \in \mathcal{E}, \\ & c_i(x^{(k)}) + a_i(x^{(k)})^T d \geq 0, i \in \mathcal{I}. \end{aligned} \tag{128}$$

在这里,W_k 是原问题 Lagrange 函数的 Hesse 阵或其近似, $g^{(k)} = \nabla f(x^{(k)}), A(x^{(k)}) = (a_1(x^{(k)}), \cdot\cdot\cdot, a_m(x^{(k)})) = [\nabla c(x^{(k)})]^T$.

记子问题 (128) 的解为 $d^{(k)}$, 相应 Lagrange 乘子向量为 $\bar{\lambda}^{(k)}$, 故有

$$\begin{cases} W_k d^{(k)} + g^{(k)} = A(x^{(k)})^T \bar{\lambda}^{(k)}, \\ \bar{\lambda}^{(k)} \geq 0, i \in \mathcal{I}, \\ c(x^{(k)}) + A(x^{(k)}) d^{(k)} = 0. \end{cases}$$

逐步二次规划法的迭代以 $d^{(k)}$ 作为搜索方向. 该搜索方向有很好的性质. 它是许多罚函数的下降方向, 例如 L_1 罚函数

$$P(x,\sigma) = f(x) + \sigma(\sum_{i=1}^{m} |c_i(x)| + \sum_{i=m_e+1}^{m} |c_i(x)_-|)$$

其中 $c(x)_-$ 定义如下:

$$\begin{cases} c_i(x)_- = c_i(x), i \in \mathcal{E}, \\ c_i(x)_- = min\{0, c_i(x)\}, i \in \mathcal{I}. \end{cases}$$

下面的算法是 Han(1977) 提出的逐步二次规划方法:

(0) 给定 $x^{(0)}, W_0 \in \mathbb{R}^{n\times n}, \sigma > 0, \rho \in (0,1), \varepsilon \geq 0$, 令 $k := 0$.

(1) 求解子问题 (128) 给出 $d^{(k)}$, 如果 $\|d^{(k)}\| \leq \varepsilon$, 则停止; 否则求 $\alpha_k \in [0,\rho]$ 使得

$$P(x^{(k)} + \alpha_k d^{(k)}, \sigma) \leq \min_{0\leq\alpha\leq\rho} P(x^{(k)} + \alpha d^{(k)}, \sigma) + \varepsilon_k.$$

(2) 置 $x^{(k+1)} = x^{(k)} + \alpha_k \delta_{x^{(k)}}$, 计算 W_{k+1}, 令 $k := k+1$, 返回第 (1) 步.

可证明前述逐步二次规划法的收敛性结果如下:

定理:

假定 $f(x)$ 和 $c_i(x)$ 连续可微, 且存在常数 $M_1, M_2 > 0$ 使得

$$M_1\|d\|^2 \leq d^T W_k d \leq M_2\|d\|^2, \forall k \in \mathbb{N}, \forall d \in \mathbb{R}^n,$$

如果 $\|\lambda^{(k)}\|_\infty \leq \sigma$ 均成立, 则 Han(1977) 算法产生的点列 $\{x^{(k)}\}$ 的任何聚点都是问题 (127) 的 K-T 点.

7.3 Penalty Function Methods

对于非线性约束最优化问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & c_i(x) = 0, i \in \mathcal{E} = \{1, ..., m_e\} \\ & c_i(x) \geq 0, i \in \mathcal{I} = \{m_e + 1, ..., m\} \end{aligned} \tag{129}$$

的罚函数, 是指利用目标函数 $f(x)$ 和约束方程 $c(x)$ 构造具有惩罚性的函数

$$P(x) = P(f(x), c(x)).$$

对于问题的可行点均有 $P(x) = f(x)$, 而当约束条件被破坏时有 $P(x) > f(x)$. 为了描述约束条件被破坏的程度, 我们定义 $c(x)_-$ 如下:

$$\begin{cases} c_i(x)_- = c_i(x), & i \in \mathcal{E} \\ c_i(x)_- = min\{0, c_i(x)\}, & i \in \mathcal{I} \end{cases}$$

罚函数一般可取为目标函数与“罚项”之和, 即

$$P(x) = f(x) + \phi(c(x)_-)$$

罚项 $\phi(c(x)_-)$ 是定义在 $\mathbb{R}^n$ 上的函数, 它满足

$$\phi(0) = 0, \lim_{\|c\|\to\infty} \phi(c) = +\infty$$

如 Courant 罚函数:

$$P_\sigma(x) = f(x) + \sigma\|c(x)_-\|_2^2$$

, 其中 $\sigma > 0$ 为罚因子. 考虑简单罚函数

$$P_\sigma(x) = f(x) + \sigma\|c(x)_-\|_2^2$$

记 $x(\sigma)$ 是无约束问题 $\min\limits_{x\in\mathbb{R}^n} P_\sigma(x)$ 的最优解, 我们有如下引理.

引理:

若 $x(\sigma)$ 同时是非线性约束最优化问题 (129) 的可行点, 则 $x(\sigma)$ 也是原问题的最优解.

Pf.

由于 $x(\sigma)$ 是无约束问题 $\min\limits_{x\in\mathbb{R}^n} P_\sigma(x)$ 的最优解, 故

$$P_\sigma(x(\sigma)) \le P_\sigma(x), x \in S$$

又 $x(\sigma)$ 是非线性约束最优化问题 (129) 的可行点, 故

$$\sigma\|c(x(\sigma))_-\|_2^2 = 0, P_\sigma(x(\sigma)) = f(x(\sigma))$$

即

$$f(x(\sigma)) < f(x), x \in S$$

$x(\sigma)$ 是原问题的最优解.

上述引理表明, 只要选取充分大的罚因子 σ>0, 则通过求解无约束最优化问题应可找到相应约束最优化问题的最优解. 然而在实际计算中, 确定大小合适的 σ 往往比较困难, 故通常是选取一个单调增的罚因子序列 $\{\sigma_k\}$, 通过求解一系列无约束问题来获得约束最优化问题的解, 这称为序贯无约束极小化技术 (SUMT).

罚函数法的迭代步骤:

(0) 任选初始点 $x^{(0)}$, 给定初始罚因子 $\sigma_0 > 0$ 及 $\beta > 1, \varepsilon > 0$. 令 $k := 0$.

(1) 以 $x^{(k)}$ 作为初始迭代点求解无约束问题的极小点, 即

$$x(\sigma_k) = arg \min_{x \in \mathbb{R}^n} P_{\sigma_k}(x).$$

(2) 若 $\|c(x(\sigma_k))_-\| < \varepsilon$, 则停止迭代并取 $x(\sigma_k)$ 为原约束问题的近似最优解;
否则, 置 $x^{(k+1)} = x(\sigma_k), \sigma_{k+1} = \beta\sigma_k, k := k+1$, 返回步 (1).

易得如下三个引理

引理 1:

设 $\sigma_{k+1} > \sigma_k > 0$, 则有

$$\begin{aligned} P_{\sigma_k}(x(\sigma_k)) &\le P_{\sigma_{k+1}}(x(\sigma_{k+1})) \\ \|c(x(\sigma_k))_-\| &\ge \|c(x(\sigma_{k+1}))_-\| \\ f(x(\sigma_k)) &\le f(x(\sigma_{k+1})) \end{aligned}$$

Pf.

由 $x(\sigma_k)$ 定义, 有

$$f(x(\sigma_k)) + \sigma_k\|c(x(\sigma_k))\|^2 \le f(x(\sigma_{k+1})) + \sigma_k\|c(x(\sigma_{k+1}))\|^2$$

又由于 $\sigma_{k+1} > \sigma_k$, 有

$$f(x(\sigma_{k+1})) + \sigma_k\|c(x(\sigma_{k+1}))\|^2 \le f(x(\sigma_{k+1})) + \sigma_{k+1}\|c(x(\sigma_{k+1}))\|^2$$

则有

$$P_{\sigma_k}(x(\sigma_k)) \le P_{\sigma_{k+1}}(x(\sigma_{k+1}))$$

由 $x(\sigma_{k+1})$ 定义, 有

$$f(x(\sigma_{k+1})) + \sigma_{k+1}\|c(x(\sigma_{k+1}))\|^2 \le f(x(\sigma_k)) + \sigma_{k+1}\|c(x(\sigma_k))\|^2$$

又由于 $\sigma_{k+1} > \sigma_k$, 有

$$f(x(\sigma_{k+1})) + \sigma_k\|c(x(\sigma_{k+1}))\|^2 \le f(x(\sigma_{k+1})) + \sigma_{k+1}\|c(x(\sigma_{k+1}))\|^2$$

即有

$$\begin{aligned} f(x(\sigma_k)) + \sigma_k\|c(x(\sigma_k))\|^2 &\le f(x(\sigma_{k+1})) + \sigma_k\|c(x(\sigma_{k+1}))\|^2 \\ &\le f(x(\sigma_{k+1})) + \sigma_{k+1}\|c(x(\sigma_{k+1}))\|^2 \\ &\le f(x(\sigma_k)) + \sigma_{k+1}\|c(x(\sigma_k))\|^2 \end{aligned}$$

第一、四项相减, 第二、三项相减, 得到

$$(\sigma_{k+1} - \sigma_k)\|c(x(\sigma_k))\|^2 \ge (\sigma_{k+1} - \sigma_k)\|c(x(\sigma_{k+1}))\|^2$$

即

$$\|c(x(\sigma_k))_-\| \geq \|c(x(\sigma_{k+1}))_-\|$$

由上述结论, 我们可以得到

$$\begin{aligned} f(x(\sigma_k)) + \sigma_k \|c(x(\sigma_{k+1}))\|^2 &\leq f(x(\sigma_k)) + \sigma_k \|c(x(\sigma_k))\|^2 \\ &\leq f(x(\sigma_{k+1})) + \sigma_k \|c(x(\sigma_{k+1}))\|^2 \end{aligned}$$

约减后得到

$$f(x(\sigma_k)) \leq f(x(\sigma_{k+1})).$$

引理 2:

设令 $\bar{x}$ 是原问题 (129) 的最优解, 则对任意的 $\sigma_k > 0$ 成立

$$f(\bar{x}) \geq P_{\sigma_k}(x(\sigma_k)) \geq f(x(\sigma_k))$$

Pf.

由于 $\bar{x}$ 是最优解, 则必是可行解, 有

$$f(\bar{x}) = P_{\sigma_k}(\bar{x}) \geq P_{\sigma_k}(x(\sigma_k))$$

又 $\|c(x(\sigma_k))\|^2 \geq 0$

$$P_{\sigma_k}(x(\sigma_k)) \geq f(x(\sigma_k))$$

引理 3:

令 $\delta = \|c(x(\sigma))_-\|$, 则 $x(\sigma)$ 也是约束问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & \|c(x)_-\| \leq \delta \end{aligned}$$

的最优解.

Pf.

设 $\widetilde{x}$ 为约束问题的可行点, 则有

$$P_\sigma(x(\sigma)) \leq P_\sigma(\widetilde{x})$$

$$f(x(\sigma)) + \sigma\|c(x(\sigma))_-\|^2 \leq f(\widetilde{x}) + \sigma\|c(\widetilde{x})_-\|^2$$

$$\|c(x(\sigma))_-\| = \delta > \|c(\widetilde{x})_-\|$$

故有

$$f(x(\sigma)) \leq f(\widetilde{x}).$$

关于罚函数的收敛性, 我们有如下结果

定理 1:

设罚函数法中的 ε 满足

$$\varepsilon > \min_{x \in \mathbb{R}^n} \|c(x)_-\|$$

则算法必有限终止.[4]

Pf.

解的时候, 必存在最优解. 设 x^* 是问题 (129) 的最优解,$x(\sigma_k)$ 是问题 $\min\limits_{x\in\mathbb{R}^n} P_{\sigma_k}(x)$ 的解. 由引理 2 有

反设非有限终止，则存在无穷序列 $\{x^{(k)}\}$ 满足

$$\|c(x(\sigma_k))_-\| \geq \varepsilon,\ \forall\ k \geq 1$$

由于 $\varepsilon > \min\limits_{x\in\mathbb{R}^n} \|c(x)_-\|$, 设 $x^* = arg\min\limits_{x} \|c(x)_-\|$,$\exists\varepsilon_0 > 0,\ \exists K > 0$, 当 $k \geq K$ 时有 $\|c(x(\sigma_k))_-\|^2 \geq \|c(x^*)_-\|^2 + \varepsilon_0$.

$$\begin{aligned} P_{\sigma_k}(x(\sigma_k)) - P_{\sigma_k}(x^*) =& f(x(\sigma_k)) - f(x^*) + \sigma_k(\|c(x(\sigma_k))\|^2 - \|c(x^*)\|^2) \\ \geq& f(x(\sigma_k)) - f(x^*) + \sigma_k\varepsilon_0 \end{aligned}$$

由于 $\sigma_k \to +\infty, k \to +\infty$,$\varepsilon_0 > 0$, 因此 $P_{\sigma_k}(x(\sigma_k)) - P_{\sigma_k}(x^*) > 0$.

这与最优性条件 $P_{\sigma_k}(x(\sigma_k)) - P_{\sigma_k}(x^*) \leq 0$ 矛盾, 故有限终止性得证。

定理 1 表明, 如果原约束问题存在可行点, 则对任意给定的 $\varepsilon > 0$, 算法都将有限终止于问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & \|c(x)_-\| \leq \delta \end{aligned}$$

的解, 且 $\delta \leq \varepsilon$.

定理 2:

如果算法不有限终止, 则必有 $\min\limits_{x\in\mathbb{R}^n} \|c(x)_-\| \geq \varepsilon$, 且 $\lim\limits_{k\to\infty} \|c(x(\sigma_k))_-\| = \min\limits_{x\in\mathbb{R}^n} \|c(x)_-\|$. 此时,$x(\sigma_k)$ 的任何聚点 x^* 都是问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & \|c(x)_-\| = \min_{y\in\mathbb{R}^n} \|c(y)_-\| \end{aligned}$$

的解.

这个定理的证明思路已经在定理 1 的证明过程中阐述过, 不再赘述.

为了叙述简单, 仅考虑等式约束问题

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & c_i(x) = 0, i \in \mathcal{E} \end{aligned} \tag{130}$$

设 x^* 是上述问题的最优解且 λ^* 是相应的 Lagrange 乘子, 由 Kuhn-Tucher 定理知,x^* 必是 Lagrange 函数

$$L(x, \lambda^*) = f(x) - (\lambda^*)^T c(x)$$

[4] 即证明，不存在无穷序列仍然满足条件

的稳定点. 但一般而言,x^* 并不是 Lagrange 函数的极小点.

我们考虑乘子罚函数 (也称增广 Lagrange 函数)

$$P(x,\lambda,\sigma)=L(x,\lambda)+\frac{\sigma}{2}\|c(x)\|_2^2.$$

由于增广 Lagrange 函数的性态, 只要取足够大的罚因子 σ 而不必趋向无穷大, 就可通过极小化 $P(x,\lambda,\sigma)$ 求得原问题的最优解.

记 $A(x)=[\nabla c(x)]^T$, 由于 $c(x^*)=0$, 我们易得

$$\nabla_x P(x^*,\lambda^*,\sigma)=\nabla_x L(x^*,\lambda^*)=0$$

$$\nabla_{xx}^2 P(x^*,\lambda^*,\sigma)=\nabla_{xx}^2 L(x^*,\lambda^*)+\sigma A(x^*)A(x^*)^T.$$

设在 x^* 处满足二阶充分条件, 即对 $\forall d$ 使得 $A(x^*)^T d=0$ 的非零向量, 均有

$$d^T\nabla_{xx}^2 L(x^*,\lambda^*)d>0.$$

那么在二阶充分条件的假定下, 对于充分大的 σ, 可证

$$\nabla_{xx}^2 P(x^*,\lambda^*,\sigma)$$

是正定阵.

定理:

设 x^* 和 λ^* 满足等式约束问题 (130) 局部最优解的二阶充分条件, 则存在 σ_0 使得当 $\lambda>\lambda_0$ 时,x^* 是函数 $P(x,\lambda^*,\sigma)$ 的严格局部极小点.

Pf.

由 x^* 和 λ^* 满足等式约束问题 (130) 局部最优解的二阶充分条件, 有 ①$c(x^*)=0$,②$\nabla_x L(x^*,\lambda^*)=\nabla f(x^*)-\sum_{i=1}^m\lambda_i^*\nabla c_i(x^*)=0$,③对 $\forall d\neq 0\wedge A(x^*)^T d=0$ 有 $d^T\nabla_{xx}^2 L(x^*,\lambda^*)d>0$.

要证明 $\exists\sigma_0$ 使得对 $\forall\sigma>\sigma_0$, 有 $\nabla_x P(X^*,\lambda^*,\sigma)=0\wedge\nabla_{xx}^2 P(x^*,\lambda^*,\sigma)>0$.

Step1.

$$\nabla_x P(x^*,\lambda^*,\sigma)\overset{①}{=}\nabla_x L(x^*,\lambda^*)\overset{②}{=}0$$

Step2.

$$\nabla_{xx}^2 P(x^*,\lambda^*,\sigma)=\nabla_{xx}^2 L(x^*,\lambda^*)+\sigma A(x^*)A(x^*)^T$$

令 $A=A(x^*),A\in\mathbb{R}^{n\times m}$,$rank(A)=r\leq m$,$A=QR,Q\in\mathbb{R}^{n\times r},R\in\mathbb{R}^{r\times m}$,$Q^TQ=I_r,rank(R)=r,R^TR>0$. 对 $\forall u\in\mathbb{R}^n$,$u=d+Qp$, 其中 $Q^Td=0$.

假设 α 为 $\nabla_{xx}^2 L$ 最小特征值,β 为 $\nabla_{xx}^2 LQ$ 最大奇异值,θ 为 $Q^T\nabla_{xx}^2 LQ$ 最小特征值,r 为 RR^T 的最小正特征值.

$$\begin{aligned}u^T\nabla_{xx}^2 P(x^*,\lambda^*,\sigma)u&=(d+Qp)^T(\nabla_{xx}^2+\sigma AA^T)(d+Qp)\\&=d^T\nabla_{xx}^2 Ld+p^TQ^T\nabla_{xx}^2 LQp+2d^T\nabla_{xx}^2 LQp+\sigma p^T(RR^T)p\\&=d^T\nabla_{xx}^2 Ld+2d^T\nabla_{xx}^2 LQp+p^T(Q^T\nabla_{xx}^2 LQ)p+\sigma p^T(RR^T)p\\&\geq\alpha\|d\|^2-2\beta\|d\|\|p\|+(\sigma r+\theta)\|p\|^2\end{aligned}$$

判别式 $\Delta = 4\beta^2 - 4\alpha(\sigma r + \theta) < 0$ 等价于

$$\sigma > \frac{\beta^2}{\alpha r} - \frac{\theta}{r} = \sigma_0$$

此时 $\nabla^2_{xx}P(x^*, \lambda^*, \sigma)$ 正定.x^* 是函数 $P(x^*, \lambda^*, \sigma)$ 的严格局部极小点.

定理:

若 $\bar{x}$ 是等式约束问题 (130) 的可行解, 且对于某个 $\bar{\lambda}, \bar{x}$ 满足 $P(x, \bar{\lambda}, \sigma)$ 的极小点二阶充分条件, 则 $\bar{x}$ 是问题 (130) 的严格局部最优解.

Pf.

$\bar{x}$ 是等式约束问题 (130) 的可行解, 有①$c(\bar{x}) = 0$.

对于 $\bar{\lambda}, \bar{x}$ 满足 $P(x, \bar{\lambda}, \sigma)$ 的极小点二阶充分条件, 有②$\nabla_x P(\bar{x}, \bar{\lambda}, \sigma) = \nabla f(\bar{x}) - \sum_{i=1}^m \bar{\lambda}_i \nabla c_i(\bar{x}) = 0$,③对 $\forall d \neq 0$ 有 $d^T \nabla^2_{xx} P(\bar{x}, \bar{\lambda}, \sigma) d > 0$.

要证明 $\nabla_x L(\bar{x}, \bar{\lambda}) = \nabla f(\bar{x}) - \sum_{i=1}^m \bar{\lambda}_i \nabla c_i(\bar{x}) = 0$, 对 $\forall d \neq 0 \wedge A(\bar{x})^T d = 0$ 有 $d^T \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) d > 0$.

Step1.

$$\nabla_x L(\bar{x}, \bar{\lambda}) = \nabla f(\bar{x}) - \sum_{i=1}^m \bar{\lambda}_i \nabla c_i(\bar{x}) \overset{①}{=} P(\bar{x}, \bar{\lambda}, \sigma) \overset{②}{=} 0$$

Step2.

$$\begin{aligned} d^T \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) d &= d^T (\nabla^2_{xx} P(\bar{x}, \bar{\lambda}, \sigma) - \sigma A(\bar{x}) A(\bar{x})^T) d \\ &= d^T \nabla^2_{xx} P(\bar{x}, \bar{\lambda}, \sigma) d - \sigma d^T A(\bar{x}) A(\bar{x})^T d \\ &= d^T \nabla^2_{xx} P(\bar{x}, \bar{\lambda}, \sigma) d - \sigma (A(\bar{x})^T d)^T A(\bar{x})^T d \\ &= d^T \nabla^2_{xx} P(\bar{x}, \bar{\lambda}, \sigma) d + 0 \\ &\overset{③}{>} 0 \end{aligned}$$

我们事先并不知道最优乘子向量 λ , 因此用乘子 λ 代替, 得到增广 Lagrange 罚函数:

$$P(x, \lambda, \sigma) = f(x) - \lambda^T c(x) + \frac{\sigma}{2} \|c(x)\|_2^2$$

一般的策略是, 先给定充分大的 σ 和乘子向量的初始估计 λ, 然后在迭代过程中修正乘子 λ, 力图使之趋向最优乘子 λ.

如何修正?

基于增广 Lagrange 函数的迭代算法:

(0) 给定初始点 $x^{(0)}$ 和乘子向量初始估计 $\lambda^{(0)}$, 给定罚因子 $\sigma_0 > 0$, 常数 $\alpha > 1, 0 < \beta < 1$ 及容许误差 $\varepsilon > 0$. 令 $k := 0$.

(1) 以 $x^{(k)}$ 为初点求解无约束问题的极小点, 即

$$x^{(k+1)} = arg \min_{x \in \mathbb{R}^n} P(x, \lambda^{(k)}, \sigma)$$

(2) 若 $\|c(x^{(k+1)})\| < \varepsilon$, 则停止迭代并取 $x^{(k+1)}$ 作为原问题的近似最优解; 否则, 更新乘子向量

$$\lambda^{(k+1)} = \lambda^{(k)} - \sigma c(x^{(k+1)}).$$

(3) 如果 $\frac{\|c(x^{(k+1)})\|}{\|c(x^{(k)})\|} \geq \beta$, 则置 $\sigma := \alpha\sigma$. 令 $k := k+1$ 返回第 (1) 步.

上述过程可以认为是从无约束问题开始寻找约束优化问题的最优解, 即从外向内逼近, 直到找到符合约束条件的最优解, 称为**外点罚函数法**.

与外点罚函数法相反, **内点罚函数法**产生的点列从可行域的内部逼近问题的解, 也称**障碍函数法**.

在迭代中总是从内点出发, 并通过引入障碍函数使之保持在可行域内部进行搜索. 因此, 这种方法适用于不等式约束的非线性最优化问题.

$$\begin{array}{ll} min & f(x) \\ s.t. & g_i(x) \geq 0, i \in \mathcal{I} = \{1, ..., m\}. \end{array} \tag{131}$$

现将可行域内部记作 $int\ S$, 其中 $S = \{x | g_i(x) \geq 0, i \in \mathcal{I}\}$. 保持迭代点含于可行域内部的方法是定义如下障碍函数:

$$B(x, \theta) = f(x) + \theta\psi(x)$$

其中障碍因子是 θ 很小的正数 $\psi(x)$ 是连续函数, 当 x 趋于可行域边界时,$\psi(x) \to +\infty$.

两种重要的障碍形式是:

$$\psi_1(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)}, \psi_2(x) = -\sum_{i=1}^{m} log(g_i(x))$$

这样, 当 x 趋向可行域边界时, 函数 $B(x, \theta) \to +\infty$. 否则, 由于 θ 很小, 函数 $B(x, \theta)$ 的取值近似于 $f(x)$.

因此, 我们可通过求解下列问题得到原问题 (131) 的近似解:

$$\begin{array}{ll} min & B(x, \theta) \\ s.t. & x \in int\ S \end{array} \tag{132}$$

由于 $\psi(x)$ 的存在, 在可行域边界形成一道“围墙”, 因此上述障碍问题 (132) 的解 $\bar{x}(\theta)$ 必含于可行域的内部.

需要解释的是, 障碍问题 (132) 表面上看起来仍是带约束的最优化问题, 且它的约束条件比原来的约束还要复杂. 但是, 由于函数 $\psi(x)$ 的障碍阻挡作用是自动实现的, 因此从计算观点看, 求解 (132) 完全可当作无约束问题来处理.

于是, 我们可以给出障碍函数法的计算步骤如下:

(0) 给定初始点 $x^{(0)} \in int\ S$, 初始障碍因子 $\theta_0 > 0, \beta \in (0, 1), \varepsilon > 0$. 令 $k := 0$.

(1) 以 $x^{(k)}$ 作为初始迭代点求解下列问题:

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

记求得的极小点为 $x(\theta_k)$.

(2) 若 $\theta_k\psi(x(\theta_k)) < \varepsilon$, 则停止计算并取 $x(\theta_k)$ 为原问题的近似最优解; 否则置 $x^{(k+1)} = x(\theta_k), \theta_{k+1} = \beta\theta_k$, 令 $k := k+1$, 返回第 (1) 步.

定理

设 $\theta_k > \theta_{k+1} > 0$, 记 $x(\theta) = arg \min_{x \in int\ S} B(x,\theta)$, 则有

$$B(x(\theta_k),\theta_k) \geq B(x(\theta_{k+1}),\theta_{k+1})$$
$$\psi(x(\theta_k)) \leq \psi(x(\theta_{k+1}))$$
$$f(x(\theta_k)) \geq f(x(\theta_{k+1}))$$

Pf.

由 $x(\theta)$ 的定义, 有

$$\begin{aligned} B(x(\theta_{k+1}),\theta_{k+1}) &= f(x(\theta_{k+1})) + \theta_{k+1}\psi(x(\theta_{k+1})) \\ &\leq f(x(\theta_k)) + \theta_{k+1}\psi(x(\theta_k)) \\ &\leq f(x(\theta_k)) + \theta_k\psi(x(\theta_k)) \\ &= B(x(\theta_k),\theta_k) \\ &\leq f(x(\theta_{k+1})) + \theta_k\psi(x(\theta_{k+1})) \end{aligned}$$

由上式, 得到 $B(x(\theta_k),\theta_k) \geq B(x(\theta_{k+1}),\theta_{k+1})$.
将第六项减去第二项和第四项减去第三项进行比较, 我们有

$$(\theta_k - \theta_{k+1})\psi(x(\theta_{k+1})) \geq (\theta_k - \theta_{k+1})\psi(x(\theta_k))$$

即 $\psi(x(\theta_k)) \leq \psi(x(\theta_{k+1}))$.
而

$$f(x(\theta_{k+1})) + \theta_{k+1}\psi(x(\theta_{k+1})) \leq f(x(\theta_k)) + \theta_{k+1}\psi(x(\theta_k))$$

$$f(x(\theta_{k+1})) \leq f(x(\theta_k)) + \theta_{k+1}(\psi(x(\theta_k)) - \psi(x(\theta_{k+1}))) \leq f(x(\theta_k))$$

即 $f(x(\theta_k)) \geq f(x(\theta_{k+1}))$.

Exercise. 给出 (线性) 约束最优化问题的二阶充分最优性条件, 并用于说明增广 Lagrange 函数的极小点与原问题最优解的等价性.

二阶充分条件:

在可行点 x^* 处,∃Lagrange 乘子 λ^* $s.t.$ (x^*,λ^*) 满足 K-T 条件, 若 $d \in ker(A(x^*)), d \neq 0$, 有 $d^T\nabla^2_{xx}L(x,\lambda)d > 0$.

证明等价性等价于证明:

1. 设 x^* 和 λ^* 满足 (130) 的局部最优解的二阶充分条件, 即 $\exists\sigma_0$, 当 $\sigma > \sigma_0$ 时,x^* 为 $P(x,\lambda^*,\sigma)$ 的严格局部极小点.

2.$\bar{x}$ 是 (130) 的可行点, 且对某个 $\bar{\lambda}$,$\bar{x}$ 满足 $P(x,\bar{\lambda},\bar{\sigma})$ 是极小点的二阶充分条件, 则 $\bar{x}$ 是 (130) 的严格局部极小值.

$$L(x,\lambda^*) \triangleq f(x) - (\lambda^*)^T c(x)$$
$$P(x,\lambda,\sigma) \triangleq L(x,\lambda) + \frac{\sigma}{2}\|c(x)\|_2^2$$
$$ker(A(x^*)) \triangleq \{d \mid A(x^*)d = 0\}$$

Pf.

1. 反证, 设不存在这样的 σ_0, 则 $\exists\{\sigma_k\}\ s.t.\ \sigma_k \to +\infty$ 且 x^* 不为 $P(x,\lambda^*,\sigma_k)$ 的严格局部极小点.

对 $\forall d \in ker(A(x^*)), d^T\nabla^2_{xx}Pd = d^T\nabla^2_{xx}Ld + \frac{\sigma_k}{2}d^TA^TAd > 0 + \frac{\sigma_k}{2}\|Ad\|^2 = 0$.

因此 $\exists d \notin ker(A(x^*))$ 有 $d^T\nabla^2_{xx}Pd < 0$.

记 $\Delta_0 = \min\limits_{\|d\|=1, d\notin ker(A(x^*))} \|A(x^*)d\|^2$

$$\begin{aligned} d^T\nabla^2_{xx}P_{\sigma_k}d =& d^T\nabla^2_{xx}Ld + \frac{\sigma_k}{2}\|A(x^*)d\|^2 \\ \geq& d^T\nabla^2_{xx}Ld + \frac{\sigma_k}{2}\Delta_0 \\ =& -\|d\|^2\|\nabla^2_{xx}L\| + \frac{\sigma_k}{2}\Delta_0 \end{aligned}$$

由于 $\|\nabla^2_{xx}L\|, \Delta$ 有限,$\sigma_k \to +\infty$, 那么

$$d^T\nabla^2_{xx}P_{\sigma_k}d \to +\infty$$

与 $d^T\nabla^2_{xx}Pd < 0$ 矛盾, 故存在这样的 σ_0.

2.$\nabla^2_{xx}P(\bar{x},\bar{\lambda},\sigma) = \nabla^2_{xx}L(\bar{x},\bar{\lambda}) + \sigma A^T(\bar{x})A(\bar{x})$ 正定, 则对 $\forall d$ 有

$$d^T\nabla_{xx}Ld + \sigma\|A(\bar{x})d\|^2 > 0$$

那么, 对 $\forall d \neq 0 \wedge d \in ker(A(x^*))$, 由于 $A(\bar{x})d = 0$ 则

$$d^T\nabla^2_{xx}Ld + \sigma\|A(\bar{x})d\|^2 = d^T\nabla^2_{xx}Ld > 0$$

这意味着 $\bar{x}$ 为问题 (130) 的严格局部最优解.

A 约束问题的二阶充分、必要条件汇总

	无约束问题	等式约束问题	一般非线性约束
约束条件	$min \quad f(x)$ $s.t. \quad x \in \mathbb{R}^n$	$min \quad f(x)$ $s.t. \quad c(x) = 0$	$min \quad f(x)$ $s.t. \quad g(x) \geq 0$ $h(x) = 0$
二阶充分条件	$\nabla f(x) = 0$ $\wedge$ $\nabla^2 f(x) > 0$	$\nabla_x L(x,\lambda) = 0$ $\wedge$ $\forall d \neq 0, d \in ker(A(x))$ $d^T\nabla^2_{xx}L(x,\lambda)^Td > 0$	$\exists\bar{\lambda} \geq 0, \bar{\mu}, s.t.\ \nabla_x L(\bar{x},\bar{\lambda},\bar{\mu}) = 0$ $\wedge$ $\forall d \neq 0, d \in \Omega(\bar{\lambda},\bar{\mu})$ $d^T\nabla^2_{xx}L(\bar{x},\bar{\lambda},\bar{\mu})d > 0$
二阶必要条件	$\nabla f(x) = 0$ $\wedge$ $\nabla^2 f(x) \geq 0$	$\nabla_x L(x,\lambda) = 0$ $\wedge$ $\forall d \in ker(A(x))$ $d^T\nabla^2_{xx}L(x,\lambda)^Td \geq 0$	$\exists\bar{\lambda} \geq 0, \bar{\mu}, s.t.\ \nabla_x L(\bar{x},\bar{\lambda},\bar{\mu}) = 0$ $\wedge$ $\forall d \in \Omega(\bar{\lambda},\bar{\mu})$ $d^T\nabla^2_{xx}L(\bar{x},\bar{\lambda},\bar{\mu})d \geq 0$

其中 Ω 定义为

$$\Omega = \left\{ d \mid \begin{array}{c} \nabla g_i(x)^T d = 0,\ i \in \mathcal{I} \wedge \lambda_i > 0 \\ \nabla g_i(x)^T d \geq 0,\ i \in \mathcal{I} \wedge \lambda_i = 0 \\ \nabla h_j(x)^T d = 0,\ j \in \varepsilon \end{array} \right\}$$

B Farkas's 引理

设 $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^n$，那么下面两个论断有且仅有一个是正确的.

1.$\exists x \in \mathbb{R}^n,\ s.t.\ Ax = b \wedge x \geq 0$

2.$\exists y \in \mathbb{R}^m,\ s.t.\ A^T y \geq 0 \wedge b^T y < 0$

C 关于 KKT 条件的补充说明

KKT 条件的含义是这个优化问题的极值点一定满足这组方程组。(不是极值点也可能会满足，但是不会存在某个极值点不满足的情况）它也是原来的优化问题取得极值的必要条件，解出来了极值点之后还是要代入验证的。但是因为约束比较多, 情况比较复杂, KKT 条件并不是对于任何情况都是满足的。要满足 KKT 条件需要有一些规范性条件（Regularity conditions)，就是要求约束条件的质量不能太差，常见的比如：

1.LCQ：如果 $h(x)$ 和 $g(x)$ 都是形如 $Ax + b$ 的仿射函数，那么极值一定满足 KKT 条件。

2.$LICQ$：起作用的 $g(x)$ 函数（即 $g(x)$ 相当于等式约束的情况）和 $h(x)$ 函数在极值点处的梯度线性无关，那么极值一定满足 KKT 条件。

3.$Slater$ 条件：如果优化问题是个凸优化问题，且至少存在一个点满足 $h(x) = 0$ 和 $g(x) = 0$，极值一定满足 KKT 条件, 并且满足强对偶性质.