最优化理论与方法主要知识点的梳理综述和拓展

2023 至 2024 学年第 1 学期

学生姓名：曹陆铖
学院专业班级：数学与统计学院数学与应用数学强基 2 班
学生学号：20210684

重庆大学数学与统计学院

1 线性规划主要知识点

1.0.1 标准型

线性规划问题的标准形为

$$\max z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$
$$\text{s.t.} \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{nn}x_n = b_m \\ x_1, x_2, \cdots, x_n \geqslant 0 \end{cases}$$

称 $c_1, c_2, \cdots, c_n$ 为价值系数, $b_1, b_2, \cdots, b_m$ 为资源限制量, $a_{ij}(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$ 为第 i 个产品对第 j 种资源的单位消耗量.

1.0.2 化标准型

(1) 将 $\min \alpha = a$ 变成 $\max \beta = -a$

(2) 添加松弛变量或剩余变量，将大括号中不等式约束都变成等式约束

(3) 使所有式子最右边的数都 ≥ 0

(4) 将大括号中单变量的式子都变成啥 $x_n \geq 0$ 的形式：若某变量 $x_m \leq 0$, 则设一个 ≥ 0 的新变量 $x_m{}'$ 使 $x_m{}' = -x_m$, 这样可以将 $x_m \leq 0$ 变成 $x_m{}' \geq 0$

(5) 若某变量 x_p 是自由变量, 则设两个 ≥ 0 新变量 $x_p{}'$、$x_p{}''$ 使 $x_p{}' - x_p{}'' = x_p$

(6) 把大括号里所有单变量的式子, 写到最下面一行

1.1 单纯形法+大 M 法、两阶段法

1.1.1 单纯形法

(1) 将题目给出的线性规划问题化为标准型

(2) 根据标准型填写初始单纯形表

(3) 找出可行解

(4) 求出检验数 $\sigma_{\mathrm{j}} = \mathrm{c_j} - (\mathrm{c_{B1}} \cdot \mathrm{x_{j1}} + \mathrm{c_{B2}} \cdot \mathrm{x_{j2}} + \cdots)$

(5) 观察一下 σ_{j} 这一行的数字看一下是否都 ≤ 0 若这些数字都 ≤ 0, 则该可行解就是最优解若这些数字有 > 0 的, 则该可行解不是最优解继续进行 (6)

(6) 找到 σ_{j} 行最大的数字那一列对应的变量 $\mathrm{x_a}$ (进基变量) 求出 $\theta_{\mathrm{i}} = \mathrm{b_i} \div \mathrm{x_{a_i}}$

(7) 找到表中 θ_{i} 最小值对应 $\mathrm{X_B}$ 列的变量 $\mathrm{X_b}$ (出基变量) 找到 $\mathrm{x_a}$ 的列和 $\mathrm{x_b}$ 的行交叉的数字 m

(8) 用 $\mathrm{x_a}$ 上面的数字替代 $\mathrm{x_b}$ 前面的数字, 用 $\mathrm{x_a}$ 替代 $\mathrm{x_b}$, 清空 σ_{j} 行与 θ_{i} 列

(9) 对 $\mathrm{x_1}\ \mathrm{x_2} \cdots \mathrm{x_n}$ 与 b 列组成的矩阵进行运算, 将 $\mathrm{m}_{变}$ 成 1 , 峒列其他元素变成 0 , 形成一个新的矩阵, 将该矩阵中的数字填入表格中对应的位置形成新的单纯形表并进行步骤 (3)

1.1.2 大 M 法

大 M 法与接下来介绍的两阶段法均针对于无法在初始矩阵中找到一个单位矩阵的情况。

例：求解

$$\max Z = 3x_1 + 2x_2 - x_3$$
$$\text{s.t.} \begin{cases} -4x_1 + 3x_2 + x_3 \geq 4 \\ x_1 - x_2 + 2x_3 \leq 10 \\ -2x_1 + 2x_2 - x_3 = -1 \\ x_i \geq 0 \quad i = 1, 2, 3 \end{cases}$$

由于标准化后不存在一个单位矩阵, 无法建立初始单纯形矩阵, 我们考虑加入一个无穷大的实数 M(−M 称为罚因子) 来人为引进新的变量达到单位矩阵的效果:

$$\max Z = 3x_1 + 2x_2 - x_3 - Mx_6 - Mx_7$$
$$\text{s.t.} \begin{cases} -4x_1 + 3x_2 + x_3 - x_4 + x_6 = 4 \\ x_1 - x_2 + 2x_3 + x_5 = 10 \\ 2x_1 - 2x_2 + x_3 + x_7 = 1 \\ x_i \geq 0 \quad i = 1, 2, \ldots, 7 \end{cases}$$

接下来按照先前普通单纯形法的步骤去做即可。值得注意的是, 若目标函数求最小值, 则罚因子应为 M 而不是 −M; 若最终结果函数值仍含有 M 则表明原问题无可行解。

1.1.3 两阶段法

例：求解

$$\max Z = 3x_1 + 2x_2 - x_3$$

$$\text{s.t.}\quad \begin{cases} -4x_1 + 3x_2 + x_3 - x_4 = 4 \\ x_1 - x_2 + 2x_3 + x_5 = 10 \\ 2x_1 - 2x_2 + x_3 = 1 \\ x_i \geq 0 \quad i = 1,2,3,4,5 \end{cases}$$

第一阶段为:

$$\min W = x_6 + x_7$$

$$\text{s.t.}\quad \begin{cases} -4x_1 + 3x_2 + x_3 - x_4 + x_6 = 4 \\ x_1 - x_2 + 2x_3 + x_5 = 10 \\ 2x_1 - 2x_2 + x_3 + x_7 = 1 \\ x_i \quad \geq 0 \quad i = 1,2,\ldots,7 \end{cases}$$

按照单纯形法, 我们可以得到一组最优解, 而第一阶段的目标函数值为 0 (如果不为 0 则说明无可行解)。第二阶段, 我们直接将第一阶段最后矩阵修改检验数即可按照一般步骤算出结果。

1.2 对偶问题＋互补松驰定理

1.2.1 对偶问题

1. 对称形式的对偶

对原线性规划问题

$$\max z = c^{\mathrm{T}} x$$
$$\text{s. t.} \begin{cases} Ax \leqslant b \\ x \geqslant 0 \end{cases}$$

其对偶规划问题为

$$\min w = y^{\mathrm{T}} b$$
$$\text{s. t.} \quad \begin{cases} y^{\mathrm{T}} A \geqslant c^{\mathrm{T}} \\ y \geqslant 0 \end{cases}$$

它们之间的关系是对称关系.

2. 一般线性规划问题的对偶问题

<table>
<tr><td colspan="2">原问题 (或对偶问题)</td><td colspan="2">对偶问题 (或原问题)</td></tr>
<tr><td colspan="2">约束条件矩阵 A $\left(A^T\right)$</td><td colspan="2">约束条件矩阵 A^T(A)</td></tr>
<tr><td colspan="2">目标函数 max</td><td colspan="2">目标函数 min</td></tr>
<tr><td rowspan="4">约束条件</td><td>m 个</td><td>m 个</td><td rowspan="4">变量</td></tr>
<tr><td>$\leq$</td><td>≥ 0</td></tr>
<tr><td>$\geq$</td><td>≤ 0</td></tr>
<tr><td>$=$</td><td>无约束</td></tr>
<tr><td rowspan="4">变量</td><td>n 个</td><td>n 个</td><td rowspan="4">约束条件</td></tr>
<tr><td>$\geq$</td><td>≥ 0</td></tr>
<tr><td>$\leq$</td><td>≤ 0</td></tr>
<tr><td>无约束</td><td>$=$</td></tr>
</table>

1.2.2 互补松弛定理

原线性规划

$$\max z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$$
$$\text{s. t. } \begin{cases} \boldsymbol{Ax} = \boldsymbol{b} \\ \boldsymbol{x} \geqslant 0 \end{cases}$$

对偶规划

$$\min w = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{b}$$
$$\text{s. t.} \quad \boldsymbol{y}^{\mathrm{T}}\boldsymbol{A} \geqslant \boldsymbol{c}^{\mathrm{T}}$$

设 x, y 分别是原线性规划和对偶规划的可行解, 则 x, y 分别是原线性规划和对偶规划的最优解的充分必要条件是

$$\left(y^{\mathrm{T}}A - c^{\mathrm{T}}\right)x = 0$$

1.3 对偶单纯形法

已知线性规划问题的一个基解 x, 对应的基矩阵 $\boldsymbol{B}, \boldsymbol{x_B} = \boldsymbol{B}^{-1}\boldsymbol{b} = \boldsymbol{\alpha}$, 该基解的所有检验数非正. (1) 若 $\alpha \geqslant 0$, 则得线性规划问题的最优解 $\boldsymbol{x}^* = \begin{pmatrix} \boldsymbol{\alpha} \\ 0 \end{pmatrix}$, 停止计算; 若存在 $\boldsymbol{\alpha}$ 的某个分量 $\alpha_i < 0$, 而 $\beta_{ij} \geqslant 0 (j = 1, 2, \cdots, n)$, 则线性规划问题无可行解, 停止计算; (2) 取 $\alpha_l = \min_{i\in S}\{\alpha_i \mid \alpha_i < 0\}$, 确定基变量 x_l 为换出变量; (3) 令 $\theta_k = \min_{j\in T}\left\{\frac{\sigma_j}{\beta_{lj}} \mid \beta_{lj} < 0\right\} = \frac{\sigma_k}{\beta_{lk}}$; (4) 以 β_{kk} 为主元素, 施展旋转变换得到新的基解, 转 (1).

1.4 整数线性规划中割平面法、分支定界法

2 非线性规划主要知识点

2.1 下降方向、可行下降方向

2.1.1 下降方向

若搜索方向 $\boldsymbol{p}_k$ 满足 $f\left(\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k\right)<f\left(\boldsymbol{x}^{(k)}\right)$, 则称 $\boldsymbol{p}_k$ 为 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(k)}$ 处的下降方向 (descent direction).

2.2 凸集、凸函数、凸规划

2.2.1 凸集

设集合 $D\subset\mathbf{R}^n$, 若 $\forall\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)}\in D$ 均有 $\alpha x^{(1)}+(1-\alpha)x^{(2)}\in D(\forall\alpha\in[0,1])$, 则称 D 为凸集.

2.2.2 凸函数

设 $D\subset\mathbf{R}^n$ 是非空凸集, $\alpha\in(0,1)$, 若

$$f\left(\alpha x^{(1)}+(1-\alpha)x^{(2)}\right)\leqslant\alpha f\left(x^{(1)}\right)+(1-\alpha)f\left(x^{(2)}\right),\quad\forall x^{(1)},x^{(2)}\in D$$

则称 $f(\boldsymbol{x})$ 为 D 上的凸函数 (convex function); 若

$$f\left(\alpha x^{(1)}+(1-\alpha)x^{(2)}\right)<\alpha f\left(x^{(1)}\right)+(1-\alpha)f\left(x^{(2)}\right),\quad\forall x^{(1)},x^{(2)}\in D \text{ 且} x^{(1)}\neq x^{(2)}$$

则称 $f(x)$ 为 D 上的严格凸函数 (strictly convex function) ; 若 $-f(x)$ 是 D 上的凸 (严格凸) 函数, 则称 $f(x)$ 为 D 上的凹 (严格凹) 函数 (concave function, strictly concave function).

2.3 无约束优化的一阶和二阶最优性条件

定理 5.6 (一阶必要条件) 设 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 的某个邻域内具有连续的偏导数,$\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的局部最优解的必要条件是 $\nabla f\left(\boldsymbol{x}^*\right)=0$.

证明用反证法. 若 $\nabla f\left(\boldsymbol{x}^*\right) \neq 0$, 取搜索方向为 $\boldsymbol{p}=-\nabla f\left(\boldsymbol{x}^*\right)$, 则 $\nabla f\left(\boldsymbol{x}^*\right)^{\mathrm{T}} p=-\left\|\nabla f\left(\boldsymbol{x}^*\right)\right\|^2<0$. 由定理 5.5, p 是 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 处的下降方向, 当 $t>0$ 充分小时就有 $f\left(\boldsymbol{x}^*+t \boldsymbol{p}\right)<f\left(\boldsymbol{x}^*\right)$, 这与 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的局部最优解矛盾.

若 $\nabla f\left(\boldsymbol{x}^*\right)=0$, 则称 $\boldsymbol{x}^*$ 为 $f(\boldsymbol{x})$ 的驻点 (stagnation point).

定理 5.7 (二阶必要条件) 设 $f(x)$ 在 x^* 的某个邻域内具有连续的二阶偏导数, 则 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的局部最优解的必要条件是 $\nabla f\left(\boldsymbol{x}^*\right)=0$ 且 $\nabla^2 f\left(\boldsymbol{x}^*\right)$ 半正定.

证明由定理 5.6 有 $\nabla f\left(\boldsymbol{x}^*\right)=0$, 将 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 处进行二阶 Taylor 展开有

$$
\begin{aligned}
f(\boldsymbol{x})= & f\left(\boldsymbol{x}^*\right)+\nabla f\left(\boldsymbol{x}^*\right)^{\mathrm{T}}\left(\boldsymbol{x}-\boldsymbol{x}^*\right) \\
& +\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}^*\right)^{\mathrm{T}} \nabla^2 f\left(\boldsymbol{x}^*\right)\left(\boldsymbol{x}-\boldsymbol{x}^*\right)+o\left(\left\|\boldsymbol{x}-\boldsymbol{x}^*\right\|^2\right)
\end{aligned}
$$

令 $x=x^*+t p(p \neq 0)$, 则有

$$
f\left(\boldsymbol{x}^*+t \boldsymbol{p}\right)=f\left(\boldsymbol{x}^*\right)+\frac{1}{2} t^2 \boldsymbol{p}^{\mathrm{T}} \nabla^2 f\left(\boldsymbol{x}^*\right) \boldsymbol{p}+o\left(\|t \boldsymbol{p}\|^2\right)
$$

当 $|t|$ 充分小时有 $\boldsymbol{p}^{\mathrm{T}} \nabla^2 f\left(\boldsymbol{x}^*\right) \boldsymbol{p} \geqslant 0$, 即 $\nabla^2 f\left(\boldsymbol{x}^*\right)$ 半正定.

定理 5.8 (二阶充分条件) 设 $f(x)$ 在 x^* 的某个邻域内具有连续的二阶偏导数, 若 $\nabla f\left(\boldsymbol{x}^*\right)=0$ 且 $\nabla^2 f\left(\boldsymbol{x}^*\right)$ 正定, 则 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的严格局部是最优解.

证明由二阶 Taylor 公式有

$$
\begin{aligned}
f(\boldsymbol{x})= & f\left(\boldsymbol{x}^*\right)+\nabla f\left(\boldsymbol{x}^*\right)^{\mathrm{T}}\left(\boldsymbol{x}-\boldsymbol{x}^*\right)+\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}^*\right)^{\mathrm{T}} \nabla^2 f\left(\boldsymbol{x}^*\right)\left(\boldsymbol{x}-\boldsymbol{x}^*\right) \\
& +o\left(\left\|\boldsymbol{x}-\boldsymbol{x}^*\right\|^2\right) \\
= & f\left(\boldsymbol{x}^*\right)+\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}^*\right)^{\mathrm{T}} \nabla^2 f\left(\boldsymbol{x}^*\right)\left(\boldsymbol{x}-\boldsymbol{x}^*\right) \\
& +o\left(\left\|\boldsymbol{x}-\boldsymbol{x}^*\right\|^2\right)
\end{aligned}
$$

当 $\boldsymbol{x}$ 充分接近 $\boldsymbol{x}^*$ 时, 由 $\nabla^2 f\left(\boldsymbol{x}^*\right)$ 正定有 $f(\boldsymbol{x})>f\left(\boldsymbol{x}^*\right)$, 即 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的严格局部最优解.

2.4 正定二次函数的一维搜索、Newton 切线法、黄金分割法（0.618 法）

对无约束非线性规划问题, 若已知迭代点 $\boldsymbol{x}^{(k)}$ 和下降方向 $\boldsymbol{p}_k$, 则需要确定适当的搜索步长 t_k, 使得 $f\left(\boldsymbol{x}^{(k+1)}\right)=f\left(\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k\right)<f\left(\boldsymbol{x}^{(k)}\right)$, 即对一元函数 $\varphi(t)=f\left(\boldsymbol{x}^{(k)}+t\boldsymbol{p}_k\right)$, 在 $[0,+\infty)$ 上, 选取 $t=t_k$, 使得

$$\varphi\left(t_k\right)=f\left(x^{(k)}+t_kp_k\right)<f\left(x^{(k)}\right)=\varphi(0)$$

从 $\boldsymbol{x}^{(k)}$ 出发, 求 $f(\boldsymbol{x})$ 的下一个近似最优解 $\boldsymbol{x}^{(k+1)}$ 的问题, 就转化为沿直线 $\boldsymbol{x}^{(k)}+t\boldsymbol{p}_k$ 求一元函数 $\varphi(t)$ 的极小值点问题了, 称之为一维搜索 (linear search). 无约束非线性规划问题 (5.1) 就等价于

$$\min\varphi(t),t\in[0,+\infty)$$

2.4.1 Newton 切线法

输入初值 t_0, 误差精度 $\varepsilon>0$, 最大迭代次数 N.

(1) 对 $k=1,2,\cdots,N$ 做到 (6);

(2) 计算 $\varphi'\left(t_0\right),\varphi''\left(t_0\right)$; (3) 若 $\varphi''\left(t_0\right)=0$, 则停止计算;

(4) 计算 $t\leftarrow t_0-\frac{\varphi'(t_0)}{\varphi''(t_0)}$; (5) 若 $|t-t_0|<\varepsilon$ 或 $|\varphi'(t)|<\varepsilon$, 则输出 t, 停机; 否则, 转 (6);

(6) $t_0\leftarrow t$; (7) 若迭代 N 次后仍没达到精度要求, 则停止.

2.4.2 黄金分割法（0.618 法）

给定初值 a,b, 黄金分割数 $\omega=0.618033988$, 误差精度 $\varepsilon>0$.

(1) 计算 $t_1\leftarrow a+(1-\omega)(b-a),t_2\leftarrow a+\omega(b-a),\varphi_1\leftarrow\varphi\left(t_1\right),\varphi_2\leftarrow\varphi\left(t_2\right)$; (2) 若 $b-a<\varepsilon$, 则得最优解 $t^*=\frac{a+b}{2}$; 否则, 转 (3);

(3) 若 $\varphi_1<\varphi_2$, 则 $b\leftarrow t_2,t_2\leftarrow t_1,\varphi_2\leftarrow\varphi_1$, 转 (4); 若 $\varphi_1>\varphi_2$, 则 $a\leftarrow t_1,t_1\leftarrow t_2$, $\varphi_1\leftarrow\varphi_2$, 转 (5); 若 $\varphi_1=\varphi_2$, 则 $a\leftarrow t_1,b\leftarrow t_2$, 转 (1);

(4) 计算 $t_1\leftarrow a+(1-\omega)(b-a),\varphi_1\leftarrow\varphi\left(t_1\right)$, 转 (2);

(5) 计算 $t_2\leftarrow a+\omega(b-a),\varphi_2\leftarrow\varphi\left(t_2\right)$, 转 (2).

2.5 最速下降法、Newton 法、拟 Newton 法(DFP、BFGS)、共轭梯度法（FR、PRP）

2.5.1 最速下降法

给定初始点 $\boldsymbol{x}^{(0)}$, 误差精度 $\varepsilon > 0$.

(1) $k \leftarrow 0$;

(2) 求 $\nabla f\left(\boldsymbol{x}^{(k)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(k)}\right)\right\| < \varepsilon$, 则停止计算, 输出 $\boldsymbol{x}^* = \boldsymbol{x}^{(k)}$;

(3) 一维搜索求最优步长 t_k 满足 $f\left(\boldsymbol{x}^{(k+1)}\right) = f\left(\boldsymbol{x}^{(k)} - t_k \nabla f\left(\boldsymbol{x}^{(k)}\right)\right) = \min_{t \geqslant 0} f\left(\boldsymbol{x}^{(k)} - t \nabla f\left(\boldsymbol{x}^{(k)}\right)\right)$, 取

$$\boldsymbol{x}^{(k+1)} \leftarrow \boldsymbol{x}^{(k)} - t_k \nabla\left(f\left(\boldsymbol{x}^{(k)}\right)\right.$$

(4) 若 $\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\right\| < \varepsilon$, 则得最优解 $\boldsymbol{x}^* = \boldsymbol{x}^{(k+1)}$, 停止计算; 否则, 令 $k \leftarrow k+1$, 转 (2).

2.5.2 Newton 法

给定初始点 $x^{(0)}$, 误差精度 $\varepsilon > 0$.

(1) $k \leftarrow 0$;

(2) 求 $\nabla f\left(\boldsymbol{x}^{(k)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(k)}\right)\right\| <_{\varepsilon}$, 则停止计算并输出 $\boldsymbol{x}^* = \boldsymbol{x}^{(k)}$;

(3) 计算 $\left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1}$, 迭代计算 $x^{(k+1)} \leftarrow x^{(k)} - \left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right)$

(4) 若 $\left\|x^{(k+1)} - x^{(k)}\right\| < \varepsilon$, 则停止计算并输出 $x^* = x^{(k+1)}$; 否则, 令 $k \leftarrow k+1$, 转 (2).

2.5.3 DFP 方法

(1) 选定初始点 $\boldsymbol{x}^{(0)}$, 初始矩阵 $\boldsymbol{H}_0=\boldsymbol{I}$, 精度要求 $\varepsilon>0$;

(2) 计算 $\nabla f\left(\boldsymbol{x}^{(0)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(0)}\right)\right\|<\varepsilon$, 则停止计算并输出 $\boldsymbol{x}^*=\boldsymbol{x}^{(0)}$; 否则, 转 (3);

(3) 取 $\boldsymbol{p}_1=-\boldsymbol{H}_0\nabla f\left(\boldsymbol{x}^{(0)}\right), k\leftarrow 0$, 转 (4);

(4) 一维搜索求 t_k 满足 $f\left(\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k\right)=\min_{t\geqslant 0} f\left(\boldsymbol{x}^{(k)}+t\boldsymbol{p}_k\right)$, 令 $\boldsymbol{x}^{(k+1)}\leftarrow\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k$, 转 (5);

(5) 计算 $\nabla f\left(\boldsymbol{x}^{(k+1)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(k+1)}\right)\right\|<\varepsilon$, 停止计算并输出 $\boldsymbol{x}^*\leftarrow\boldsymbol{x}^{(k+1)}$; 否则, 转 (6);

(6) 若 $k+1=n$, 令 $\boldsymbol{x}^{(0)}\leftarrow\boldsymbol{x}^{(n)}$, 转 (3); 否则, 转 (7);

(7) 计算 $\boldsymbol{H}_{k+1}=\boldsymbol{H}_k+\frac{\Delta\boldsymbol{x}_k\Delta\boldsymbol{x}_k^{\mathrm{T}}}{\Delta\boldsymbol{x}_k^{\mathrm{T}}\Delta\boldsymbol{g}_k}-\frac{\boldsymbol{H}_k\Delta\boldsymbol{g}_k\Delta\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_k}{\Delta\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_k\Delta\boldsymbol{g}_k},\quad \boldsymbol{p}_{k+1}=-\boldsymbol{H}_{k+1}\nabla f\left(\boldsymbol{x}^{(k+1)}\right)$, $k\leftarrow k+1$, 转 (4).

2.5.4 BFGS 方法

(1) 选定初始点 $\boldsymbol{x}^{(0)}$, 初始矩阵 $\boldsymbol{H}_0$, 误差精度 $\varepsilon>0$;

(2) 计算 $\nabla f\left(\boldsymbol{x}^{(0)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(0)}\right)\right\|<\varepsilon$, 则停止计算并输出 $\boldsymbol{x}^*=\boldsymbol{x}^{(0)}$; 否则, 转 (3);

(3) 取 $\boldsymbol{p}_0=-\boldsymbol{H}_0\nabla f\left(\boldsymbol{x}^{(0)}\right), k\leftarrow 0$, 转 (4);

(4) 一维搜索求 t_k 满足 $f\left(\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k\right)=\min_i f\left(\boldsymbol{x}^{(k)}+t\boldsymbol{p}_k\right)$, 令 $\boldsymbol{x}^{(k+1)}\leftarrow\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k$, 转 (5);

(5) 计算 $\nabla f\left(\boldsymbol{x}^{(k+1)}\right)$, 若 $\left\|\nabla f\left(x^{(k+1)}\right)\right\|<\varepsilon$, 则停止计算并输出 $\boldsymbol{x}^*\leftarrow\boldsymbol{x}^{(k+1)}$; 否则, 转 (6);

(6) 若 $k+1=n$, 令 $\boldsymbol{x}^{(0)}\leftarrow\boldsymbol{x}^{(n)}$, 转 (3); 否则, 转 (7)

(7) 计算

$$\boldsymbol{H}_{k+1}=\boldsymbol{H}_k+\frac{\Delta\boldsymbol{x}_k\Delta\boldsymbol{x}_k^{\mathrm{T}}}{\Delta\boldsymbol{x}_k^{\mathrm{T}}\Delta\boldsymbol{g}_k}\left(1+\frac{\Delta\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_k\Delta\boldsymbol{g}_k}{\Delta\boldsymbol{x}_k^{\mathrm{T}}\Delta\boldsymbol{g}_k}\right)-\frac{1}{\Delta\boldsymbol{x}_k^{\mathrm{T}}\Delta\boldsymbol{g}_k}\left(\Delta\boldsymbol{x}_k\Delta\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_k+\boldsymbol{H}_k\Delta\boldsymbol{g}_k\Delta\boldsymbol{x}_k^{\mathrm{T}}\right)$$

$$\boldsymbol{p}_{k+1}=-\boldsymbol{H}_{k+1}\nabla f\left(\boldsymbol{x}^{(k+1)}\right)$$

$$k\leftarrow k+1,\ \text{转 (4).}$$

2.5.5 FR 方法

给定初始点 $x^{(0)}$, 误差精度 $\varepsilon>0$.

(1) 计算 $\nabla f\left(\boldsymbol{x}^{(0)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(0)}\right)\right\|<_{\varepsilon}$, 则停止计算并输出 $\boldsymbol{x}^{*}=\boldsymbol{x}^{(0)}$; 否则, 转 (2);

(2) 取 $\boldsymbol{p}_{0}=-\nabla f\left(\boldsymbol{x}^{(0)}\right), k \leftarrow 0$, 转 (3);

(3) 用一维搜索法求 t_{k} 满足 $f\left(\boldsymbol{x}^{(k)}+t_{k} \boldsymbol{p}_{k}\right)=\min_{i \geqslant 0} f\left(\boldsymbol{x}^{(k)}+t \boldsymbol{p}_{k}\right)$, 令 $\boldsymbol{x}^{(k+1)} \leftarrow \boldsymbol{x}^{(k)}+t_{k} \boldsymbol{p}_{k}$, 转 (4);

(4) 计算 $\nabla f\left(\boldsymbol{x}^{(k+1)}\right)$, 若 $\left\|\nabla f\left(\boldsymbol{x}^{(k+1)}\right)\right\|<\varepsilon$, 则停止计算并输出 $\boldsymbol{x}^{*}=\boldsymbol{x}^{(k+1)}$; 否则, 转 (5);

(5) 若 $k+1=n$, 令 $\boldsymbol{x}^{(0)} \leftarrow \boldsymbol{x}^{(n)}$, 转 (2); 否则, 转 (6);

(6) 计算 $\boldsymbol{p}_{k+1}=-\nabla f\left(\boldsymbol{x}^{(k+1)}\right)+\beta_{k} \boldsymbol{p}_{k}$, 其中 $\beta_{k}=\frac{\left\|\nabla f\left(x^{(k+1)}\right)\right\|^{2}}{\left\|\nabla f\left(x^{(k)}\right)\right\|^{2}}$, 令 $k \leftarrow k+1$, 转 (3).

2.6 Kuhn-Tuckcr 最优性条件

设 $\boldsymbol{x}^{*} \in D, f, g_{i}(i \in I)$ 在 $\boldsymbol{x}^{*}$ 处可微, $g_{i}(i \notin I)$ 在 $\boldsymbol{x}^{*}$ 连续, $\left\{\nabla g_{i}\left(x^{*}\right) \mid i \in I\right\}$ 线性无关, 若 $\boldsymbol{x}^{*}$ 是不等式约束问题 (6.2) 的最优解, 则存在 $\lambda_{i}^{*} \geqslant 0(i \in I)$, 使得

$$\nabla f\left(\boldsymbol{x}^{*}\right)+\sum_{i \in I} \lambda_{i}^{*} \nabla g_{i}\left(\boldsymbol{x}^{*}\right)=0$$

证明由定理 6.1 , 存在不全为零的非负实数 $\lambda_{0}, \lambda_{i}(i \in I)$, 使得

$$\lambda_{0} \nabla f\left(\boldsymbol{x}^{*}\right)+\sum_{i \in I} \lambda_{i} \nabla g_{i}\left(\boldsymbol{x}^{*}\right)=0$$

若 $\lambda_{0}=0$, 则 $\lambda_{i}(i \in I)$ 不全为零, 于是得出 $\left\{\nabla g_{i}\left(x^{*}\right) \mid i \in I\right\}$ 线性相关, 与已知条件矛盾, 故 $\lambda_{0}>0$. 令 $\lambda_{i}^{*}=\frac{\lambda_{i}}{\lambda_{0}}$, 就有

$$\nabla f\left(x^{*}\right)+\sum_{i \in I} \lambda_{i}^{*} \nabla g_{i}\left(x^{*}\right)=0, \quad \lambda_{i}^{*} \geqslant 0, i \in I$$

称满足 Kuhn-Tucker 条件的点为 Kuhn-Tucker 点. 在定理 6.2 中, 若增加条件: $g_i(i \notin I)$ 在 x^* 可微, 则有如下等价形式的 Kuhn-Tucker 条件:

$$\begin{cases} \nabla f\left(\boldsymbol{x}^*\right)+\sum_{i=1}^m \lambda_i^* \nabla g_i\left(\boldsymbol{x}^*\right)=0 & \\ \lambda_i^* g_i\left(\boldsymbol{x}^*\right)=0, & i=1,2,\cdots,m \\ \lambda_i^* \geqslant 0, & i=1,2,\cdots,m \end{cases}$$

2.7 外罚函数法、内罚函数法

2.7.1 外罚函数法

(1) 选择初始点 $\boldsymbol{x}^{(0)}$, 罚参数列 $\{\mu_k\}\,(k=1,2,\cdots)$, 误差精度 $\varepsilon<0, k \leftarrow 1$;

(2) 构造辅助函数 $F\left(\boldsymbol{x}, \mu_k\right)$;

(3) 用某种无约束优化算法, 以 $\boldsymbol{x}^{(k-1)}$ 为初值求解 $\min F\left(\boldsymbol{x}, \mu_k\right)$ 得最优解 $\boldsymbol{x}^{(k)}$;

(4) 若 $\boldsymbol{x}^{(\boldsymbol{k})}$ 满足终止条件, 则停止计算并输出 $\boldsymbol{x}^*=\boldsymbol{x}^{(k)}$; 否则, 令 $k \leftarrow k+1$, 转 (2).

2.7.2 内罚函数法

(1) 选初始点 $\boldsymbol{x}^{(0)} \in D_0$, 罚参数列 $\{d_k\}\,(k=1,2,\cdots)$, 误差精度 $\varepsilon<0$, 令 $k \leftarrow 1$;

(2) 构造辅助函数

$$F\left(\boldsymbol{x}, d_k\right)=f(\boldsymbol{x})-d_k \sum_{i=1}^m \frac{1}{g_i(\boldsymbol{x})} \text{ 或} F\left(\boldsymbol{x}, d_k\right)=f(\boldsymbol{x})-d_k \sum_{i=1}^m \ln \left[-g_i(\boldsymbol{x})\right]$$

(3) 选用某种无约束优化方法, 以 $x^{(k-1)}$ 为初值, 求解 $\min F\left(x, d_k\right)$ 得最优解 $\boldsymbol{x}^{(k)}$;

(4) 若 $\boldsymbol{x}^{(k)}$ 满足终止条件, 则停止计算并输出最优解 $\boldsymbol{x}^*=\boldsymbol{x}^{(k)}$; 否则, 令 $k \leftarrow k+1$, 转 (2).

2.8 乘子法

2.8.1 等式约束问题的乘子法 (P-H 乘子法)

(1) 给定初值 $\boldsymbol{x}^{(0)}$, 初始估计 $\lambda^{(1)}$, 罚参数 μ, 误差精度 $\varepsilon > 0$, 常数 $\alpha > 1, \beta \in (0,1), k \leftarrow 0$;

(2) 以 $\boldsymbol{x}^{(k-1)}$ 为初值, 利用无约束优化算法求解 $\min F\left(\boldsymbol{x}, \boldsymbol{\lambda}^{(k)}, \mu\right)$ 得最优解 $\boldsymbol{x}^{(k)}$;

(3) 若 $\left\|\boldsymbol{h}\left(\boldsymbol{x}^{(k)}\right)\right\| < \varepsilon$ 或 $\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\| < \boldsymbol{\varepsilon}$, 则停止计算并输出近似解 $\boldsymbol{x}^{*} = \boldsymbol{x}^{(k)}$; 否则, 转 (4);

(4) 若 $\frac{\left\|\boldsymbol{h}\left(\boldsymbol{x}^{(k)}\right)\right\|}{\left\|\boldsymbol{h}\left(\boldsymbol{x}^{(k-1)}\right)\right\|} \geqslant \beta$, 则令 $\mu \leftarrow \alpha\mu$, 转 (5);

(5) 计算 $\lambda_j^{(k+1)} = \lambda_j^{(k)} + \mu h_j\left(x^{(k)}\right) (j = 1, 2, \cdots, l), k \leftarrow k+1$, 转 (2).

2.8.2 不等式约束问题的乘子法 (Rockafellar 乘子法)

(1) 选取初值 $\boldsymbol{x}^{(0)}$, 初始估计 $\boldsymbol{\lambda}^{(1)}$, 罚参数 μ, 误差精度 $\varepsilon > 0, k \leftarrow 1$;

(2) 以 $\boldsymbol{x}^{(k-1)}$ 为初值求解无约束优化问题 $\min F\left(\boldsymbol{x}, \lambda^{(k)}, \mu\right)$, 得到最优解 $\boldsymbol{x}^{(k)}$;

(3) 若 $\boldsymbol{x}^{(k)}$ 满足终止条件, 则停止计算并输出近似最优解 $\boldsymbol{x}^{*} = \boldsymbol{x}^{(k)}$; 否则, 转 (4);

(4) 计算 $\lambda_i^{(k+1)} = \max\left\{0, \mu g_i\left(x^{(k)}\right) + \lambda_i^{(k)}\right\}, k \leftarrow k+1$, 转 (2).

2.9 Frank-Wolfe 方法、Zoutendijk 方法（可行方向法）

2.9.1 Frank-Wolfe 方法

(1) 给定初始可行点 $\boldsymbol{x}^{(0)}$, 误差精度 $\epsilon > 0, k \leftarrow 1$;

(2) 构造线性规划 (6.10), 并求得最优解 $\boldsymbol{y}^{(k)}$;

(3) 若 $\left|\nabla f\left(\boldsymbol{x}^{(k)}\right)^{\mathrm{T}}\left(\boldsymbol{y}^{(k)} - \boldsymbol{x}^{(k)}\right)\right| < \varepsilon$, 则停止计算并输出解 $\boldsymbol{x}^{*} = \boldsymbol{x}^{(k)}$; 否则, 转 (4);

(4) 从 $\boldsymbol{x}^{(k)}$ 出发, 沿方向 $\boldsymbol{p}_k=\boldsymbol{y}^{(k)}-\boldsymbol{x}^{(k)}$ 进行一维搜索, 求 t_k, 满足

$$f\left(\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k\right)=\min_{0\leqslant 1\leqslant 1} f\left(\boldsymbol{x}^{(k)}+t\boldsymbol{p}_k\right)$$

(5) 令 $\boldsymbol{x}^{(k+1)}\leftarrow\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}_k, k\leftarrow k+1$, 转 (2).

2.9.2 Zoutendijk 方法

(1) 给定初始点 $\boldsymbol{x}^{(0)}$, 误差精度 $\varepsilon>0, k\leftarrow 0$;

(2) 分解 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{1k}\\ \boldsymbol{A}_{2k}\end{pmatrix}, \boldsymbol{b}=\begin{pmatrix}\boldsymbol{b}_{1k}\\ \boldsymbol{b}_{2k}\end{pmatrix}$, 使得 $\boldsymbol{A}_{1k}\boldsymbol{x}^{(k)}=\boldsymbol{b}_{1k}, \boldsymbol{A}_{2k}\boldsymbol{x}^{(k)}<\boldsymbol{b}_{2k}$;

(3) 构造线性规划

$$\begin{aligned}&\min \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p}\\ &\text{s. } t.\begin{cases}\boldsymbol{A}_{1k}\boldsymbol{p}\leqslant 0\\ \boldsymbol{C}\boldsymbol{p}=0\\ -1\leqslant p_i\leqslant 1, \quad i=1,2,\cdots,n\end{cases}\end{aligned}$$

求得最优解 $\boldsymbol{p}^{(k)}$;

(4) 若 $\left|\nabla f\left(\boldsymbol{x}^{(k)}\right)^{\mathrm{T}}\boldsymbol{p}^{(k)}\right|<\boldsymbol{\varepsilon}$, 则停止计算并输出 $\boldsymbol{x}^*=\boldsymbol{x}^{(k)}$; 否则, 得到可行下降方向 $\boldsymbol{p}^{(k)}$, 转 (5);

(5) 计算 $\boldsymbol{\xi}^{(k)}=\left(\xi_1^{(k)},\xi_2^{(k)},\cdots,\xi_{r_k}^{(k)}\right)^{\mathrm{T}}=\boldsymbol{b}_{2k}-\boldsymbol{A}_{2k}\boldsymbol{x}^{(k)}, \boldsymbol{\eta}^{(k)}=\left(\eta_1^{(k)},\eta_2^{(k)},\cdots,\eta_k^{(k)}\right)^{\mathrm{T}}=\boldsymbol{A}_{2k}\boldsymbol{p}^{(k)}$;

(6) 求解 $\min_{0\leqslant\leqslant_{\leq}^{l(k)}} f\left(\boldsymbol{x}^{(k)}+t\boldsymbol{p}^{(k)}\right)$ 得步长 t_k, 其中

$$t_{\max}^{(k)}=\begin{cases}+\infty, & \boldsymbol{\eta}^{(k)}\leqslant 0\\ \min_i\left\{\frac{\xi_i^{(k)}}{\eta_i^{(k)}}\mid \boldsymbol{\eta}_i^{(k)}>0\right\}, & \boldsymbol{\eta}^{(k)}\leqslant 0\end{cases}$$

(7) 令 $\boldsymbol{x}^{(k+1)}\leftarrow\boldsymbol{x}^{(k)}+t_k\boldsymbol{p}^{(k)}, k\leftarrow k+1$, 转 (2).

3 动态规划、多目标规划主要知识点

建立动态规划模型的步骤一般如下:

(1) 根据时间或空间的自然特征, 把实际问题恰当地划分为若干个阶段, 使问题能够转化为一个多阶段决策问题;

(2) 正确地选择状态变量 x_k, 确定状态集合, 使状态变量既能描述过程的演变特征, 又能满足无后效性及可知性, 可知性是指各阶段的状态变量值可直接或间接地获得;

(3) 定决策变量 d_k 及每个阶段的允许决策集合 $D_k\left(x_k\right)$;

(4) 写出状态转移方程 $x_{k+1}=T_k\left(x_k, d_k\right)$;

(5) 正确地写出指标函数 $v_{k,n}$;

(6) 写出动态规划函数的基本方程, 即最优值函数满足的递推方程及端点条件;

(7) 逆序递推计算求解.

3.1 最短路径问题的动态规划解法

3.2 利用动态规划方法求解机器生产分配问题、背包问题、非线性规划问题（教材-例 8.2、例 8.3、例 8.4、例 8.5）

3.3 多目标规划的绝对最优解、有效解、弱有效解概念及性质

3.3.1 绝对最优解

设 $\boldsymbol{x}^* \in D$, 若对 $\forall \boldsymbol{x} \in D$ 均有 $f_i\left(\boldsymbol{x}^*\right) \leqslant f_i(\boldsymbol{x})(i=1, 2, \cdots, p)$, 则称 $\boldsymbol{x}^*$ 为 (7.1) 的绝对最优解, 称 $\boldsymbol{f}\left(\boldsymbol{x}^*\right)=\left(f_1\left(\boldsymbol{x}^*\right), f_2\left(\boldsymbol{x}^*\right), \cdots, f_p\ \left(\boldsymbol{x}^*\right)\right)^{\mathrm{T}}$ 为 (7.1) 的绝对最优值.

3.3.2 有效解

对多目标规划问题 (7.1), 设 $x^* \in D$, 若不存在 $x \in D$, 使 $f(x)\ f\left(x^*\right)$, 则称 x^* 为 (7.1) 的有效解. 所有的有效解组成的集合称为有效解集, 记为

$P(f, D)$.

3.3.3 弱有效解

设 $x^* \in D$, 若不存在 $x \in D$, 使 $f(x) < f\left(x^*\right)$, 则称 x^* 为 (7.1) 的弱有效解, 又称 x^* 为弱 Pareto 最优解. 所有弱有效解组成的集合称为弱有效解集, 记为 $P_w(f, D)$.

3.4 评价函数法（线性加权和法、极大极小法、理想点法）

3.4.1 线性加权和法

(1) 给出权系数 $\omega_1, \omega_2, \cdots, \omega_p$, 使得 $\sum_{i=1}^p \omega_i = 1$ 且 $\omega_i \geqslant 0$;

(2) 构造单目标规划问题 $\min_{x \in D} \sum_{i=1}^p \omega_i f_i(\boldsymbol{x})$;

(3) 用约束非线性规划的方法求解单目标规划问题, 得出多目标规划问题 (7. 1) 的有效解或弱有效解.

3.4.2 极大极小法

(1) 给定权系数 $\omega_1, \omega_2, \cdots, \omega_p > 0$ 且 $\sum_{i=1}^p \omega_i = 1$;

(2) 构造辅助问题 (7.3);

(3) 求解辅助规划问题 (7.3) 得最优点 $\boldsymbol{x}^*, \boldsymbol{v}^*$;

(4) 输出多目标规划问题 (7.1) 的弱有效解 $\boldsymbol{x}^*$.

3.4.3 最短距离法

(1) 求理想点 $f_i^* = f_i\left(\boldsymbol{x}_i^*\right) = \min_{\boldsymbol{x} \in D_i}(\boldsymbol{x})(i = 1, 2, \cdots, p), \boldsymbol{f}^* = \left(f_1^*, f_2^*, \cdots, f_p^*\right)^{\mathrm{T}}$;

(2) 当 $\boldsymbol{x}_1^* = \boldsymbol{x}_2^* = \cdots = \boldsymbol{x}_p^*$ 时, 绝对最优解为 $\boldsymbol{x}^* = \boldsymbol{x}_1^*$, 停止计算; 否则, 转 (3);

(3) 求解单目标规划问题 $\min_{\boldsymbol{x} \in D}(\boldsymbol{f}) = \sum_{i=1}^p \left[f_i(\boldsymbol{x}) - f_i^*\right]^2$, 得到最优解 $\boldsymbol{y}^*$;

(4) 输出解 $\boldsymbol{y}^*$, 停止计算.