第一章 最优化方法概述

§1.1 最优化问题及举例

最优化是一门应用十分广泛的学科，它所研究的问题是讨论在众多的方案中什么样的方案是最优的，以及如何找出最优的方案。事实上，最优化问题对我们而言并不陌生，其在管理、金融、经济、工程等多个领域都有应用。

生产计划

- 资源材料有限
- 人员安排合理
- 产品利润最大

投资决策

- 有限的资金
- 最大化收益
- 最小化风险

机器学习

- 分类问题
- 回归问题
- 聚类问题

§ 1.1 最优化问题及举例

例1 某公司生产甲、乙两种商品，每件甲商品要消耗A种原料3kg，B种原料3kg，产值为180元；每件乙商品要消耗A种原料4.5kg，B种原料1.5kg，产值为150元。现该公司共有A种原料900kg，B种原料600kg，试确定甲、乙两种商品各生产多少件，可使该公司总产值最大。

原料 产品	A	B	
甲	3kg	3kg	180元
乙	4.5kg	1.5kg	150元
	900kg	600kg	

§ 1.1 最优化问题及举例

设该公司生产甲商品x_1件，生产乙商品x_2件，则该公司的总产值 $z = 180x_1+150x_2$

受原材料总量的限制，需要满足如下约束条件：

$$\begin{cases} 3x_1 + 4.5x_2 \le 900 \\ 3x_1 + 1.5x_2 \le 600 \end{cases} \qquad (1)$$

此外，x_1和x_2应是非负数，即 $x_1 \ge 0, x_2 \ge 0$ (2)

因此，在满足约束条件(1)和(2)下，使该公司总产值达到最大的数学模型为：

$$\max\ z = 180x_1+150x_2$$

$$\begin{aligned} \text{s.t.}\quad & 3x_1 + 4.5x_2 \le 900 \\ & 3x_1 + 1.5x_2 \le 600 \\ & x_1 \ge 0, x_2 \ge 0 \end{aligned}$$

§ 1.1 最优化问题及举例

例2 设有n个市场，第j个市场的位置为(a_j, b_j)，对某种货物的需量为$p_j(j=1, ..., n)$. 现计划建立m个货栈，第i个货栈的容量为$c_i(i=1, ..., m)$. 试确定货栈的位置，使各货栈到各市场的运输量与距离的乘积之和最小。

货栈 \ 市场	1	…	n
1	W_{11}	…	W_{1n}
⋮	⋮	⋱	⋮
m	W_{m1}	…	W_{mn}

$$\sum_{j=1}^{n} W_{1j} \le c_1$$

$$\sum_{i=1}^{m} W_{i1} = p_1$$

§ 1.1 最优化问题及举例

设第i个货栈的位置为(x_i, y_i)，则第i个货栈到第j个市场的距离为 $d_{ij}=\sqrt{(x_i-a_j)^2+(y_i-b_j)^2}$

又设第i个货栈供给第j个市场的货物量为W_{ij}

受货栈容量和市场需求量的限制，需要满足如下约束条件：

$$\begin{cases} \sum_{j=1}^{n} W_{ij} \le c_i,\ i=1,\ldots,m \\ \sum_{i=1}^{m} W_{ij} = p_j,\ j=1,\ldots,n \\ W_{ij} \ge 0 \end{cases}$$

§ 1.1 最优化问题及举例

在满足约束条件下，使运输量与距离乘积之和最小的数学模型为：

$$\boldsymbol{min} \quad \sum_{i=1}^{m}\sum_{j=1}^{n} W_{ij}\sqrt{(\boldsymbol{x}_i - \boldsymbol{a}_j)^2 + (\boldsymbol{y}_i - \boldsymbol{b}_j)^2}$$

$$s.t. \quad \sum_{j=1}^{n} W_{ij} \le c_i, \; i = 1, \dots, m$$

$$\sum_{i=1}^{m} W_{ij} = p_j, \; j = 1, \dots, n$$

$$W_{ij} \ge 0, \; i = 1, \dots, m, \; j = 1, \dots, n.$$

§ 1.1 最优化问题及举例

例3 机器学习中的优化问题-数据拟合

假设给定一个训练集 $T=\{(x_1,y_1),(x_1,y_1),\cdots,(x_N,y_N)\}$ **,其中** $\boldsymbol{x_i}$ **是输入，** $\boldsymbol{y_i}$**是输出。**

利用多项式函数对输入和输出之间的关系进行拟合：

$$f(x,w)=w_0+w_1x+w_2x^2+\cdots+w_Mx^M$$

为了最小化拟合误差，构建如下目标函数：

$$\min L(w)=\frac{1}{2}\sum_{i=1}^{N}(f(x_i,w)-y_i)^2$$

§ 1.2 模型的分类

最优化问题数学模型的一般形式：

$$\begin{aligned} & \min\ f(x) \\ \text{s.t.}\quad & g_i(x) \geq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = m+1, \ldots, p \end{aligned}$$

其中，x为决策变量，$f(x)$为目标函数，$g_i(x)$和$h_i(x)$为约束条件。

根据实际问题的需求，最优化模型可能有不同的形式，但经过适当的变换都可转换成上述的一般形式。

§ 1.2 模型的分类

根据决策变量、约束条件、目标函数的不同形式，可对最优化模型分类：

(1) 按约束条件

- 约束最优化模型(有约束条件)
  - 等式约束 $h_i(x) = 0$
  - 不等式约束 $g_i(x) \geq 0$
  - 混合约束 $h_i(x) = 0$ 和 $g_i(x) \geq 0$
  - 盒式约束 $l_i \leq x_i \leq u_i$
- 无约束最优化模型(无约束条件)

§ 1.2 模型的分类

(2) 按目标函数个数

- 单目标最优化模型　（只有一个目标函数）
- 多目标最优化模型　（有多个目标函数）

(3) 按约束条件和目标函数

- 线性规划模型　（目标函数和约束条件均是线性的）
- 非线性规划模型　（模型中存在非线性函数）

§ 1.2 模型的分类

(4) 按决策变量

- 离散最优化模型　（变量取离散值）
- 连续最优化模型　（变量连续取值）

§ 1.3 预备知识

(1)基本概念

- 可行点：如果点x满足最优化模型中的所有约束条件，则x为可行点。
- 可行域：所有可行点的集合称为可行域

$$D = \{x \mid g_i(x) \geq 0,\ h_i(x) = 0\}$$

如果可行域是整个空间，则最优化模型为无约束最优化模型。

- 全局最优解：一个可行点$x^* \in D$，$\forall\, x \in D$，如果有$f(x^*) \leq f(x)$，则x^*为全局最优解。
- 局部最优解：对于可行点x^*，如果存在一个邻域

$$N(x^*) = \{x \mid \|x - x^*\| \leq \delta\}$$

使得$f(x^*) \leq f(x)$，$\forall\, x \in N(x^*) \cap D$，则$x^*$为局部最优解。

§ 1.3 预备知识

(2)凸集

- **定义：设集合$D\subseteq R^n$，如果对于任意x，$y\in D$有**

$$\lambda x+(1-\lambda)y\in D,\quad \lambda\in[0,1]$$

换言之，如果任意两点x，$y\in D$，连接x与y的直线仍在D内，则称D为凸集，且$\lambda x+(1-\lambda)y$为x和y的凸组合。

- **性质：设D_1，$D_2\subseteq R^n$为两个凸集，则**

  (i) 对于任意的非零实数β，$\beta D_1=\{\beta x|x\in D_1\}$为凸集。

  (ii) $D_1\cap D_2=\{x|x\in D_1$ 且$x\in D_2\}$为凸集。

  (iii) $D_1+D_2=\{x+y|x\in D_1, y\in D_2\}$为凸集。

  (iv) $D_1-D_2=\{x-y|x\in D_1, y\in D_2\}$为凸集。

§ 1.3 预备知识

➢ **集合$D\subseteq R^n$是凸集的充分必要条件为D中任意m个点的凸组合仍属于D.**

证明：充分性。令$m=2$，对于任意2个点x_1和x_2有

$$\alpha_1 x_1+\alpha_2 x_2\in D\quad(\alpha_1+\alpha_2=1,\alpha_1,\alpha_2\geq 0)$$

则由定义可知，D是凸集。

§ 1.3 预备知识

必要性(归纳法)。根据定义，在m=2时显然成立。

不妨假设当m=k时，结论也成立。

那么，当m=k+1时，设x_i∈ D，i=1, ..., k+1

对于任意一组实数$\alpha_i \geq 0$，$\sum \alpha_i = 1$.

有

$$\sum_{i=1}^{k+1} \alpha_i x_i = \sum_{i=1}^{k} \alpha_i x_i + \alpha_{k+1} x_{k+1}$$

$$= (1-\alpha_{k+1}) \boxed{\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} x_i} + \alpha_{k+1} x_{k+1}$$

由于$1-\alpha_{k+1} = \sum_{i=1}^{k} \alpha_i$，则 $\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} = 1$，且 $\frac{\alpha_i}{1-\alpha_{k+1}} \geq 0$

故 $\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} x_i \in D$

综上，根据归纳假设和凸集的定义，有

$$\sum_{i=1}^{k+1} \alpha_i x_i \in D$$

得证。

(3)凸函数

- **定义：设$f(x)$在凸集D上有定义，如果对于任意x，$y\in D$和 $\forall\lambda\in[0,1]$，有**

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)$$

则$f(x)$是D上的凸函数。

如果对任意x和y，$x\neq y$ 和$\forall\lambda\in[0,1]$，有

$$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y)$$

则$f(x)$是D上的严格凸函数。

§ 1.3 预备知识

- 性质：

  (i) $f(x)$是定义在凸集D上的凸函数，实数$\beta \geqslant 0$，则$\beta f(x)$是凸函数。

  (ii) $f_1(x)$和$f_2(x)$是D上的凸函数，则$f_1(x)+f_2(x)$也是凸函数。

凸集刻画了可行域的特点，凸函数强调了目标函数的性质。对于最优化问题，如果可行域是凸集，目标函数是凸函数，则该优化问题称为凸规划。

凸规划的局部最优解必是全局最优解。

§ 1.3 预备知识

定理：设x^*是凸规划问题的一个局部最优解，则x^*也是全局最优解。

证明：设 x^*是局部最优解，但不是全局最优解，则存在可行点y，满足$f(y)<f(x^*)$.

由可行域是凸的，对于$\forall\lambda\in[0,1]$，$\lambda^* x+(1-\lambda)y$ 也是可行点。

根据目标函数的凸性，有

$$
\begin{aligned}
f(\lambda x^*+(1-\lambda)y) &\le \lambda f(x^*)+(1-\lambda)f(y) \\
&< \lambda f(x^*)+(1-\lambda)f(x^*) \\
&= f(x^*)
\end{aligned}
$$

说明在x^*的任意小邻域内都存在函数值小于$f(x^*)$的可行点，与x^*是局部最优解矛盾，故x^*必是全局最优解。

课堂练习

1. 某石油钻井钻探的负责人必须从10个可行位置中挑选5个费用最低的点进行钻探。设10个点的标号为s_1，$s_2 \dots s_{10}$，各点钻探费用分别为c_1，$c_2 \dots c_{10}$. 按区域发展规划，有如下规定：(1) 同时钻了s_1和s_7后就不能钻s_8

(2) 钻了s_3或s_4后，就不能钻s_6

(3) s_5, s_6, s_7, s_8中只能钻2个

试构造一个在满足上述要求的前提下寻求钻探费用最小方案的数学规划模型。

2. 验证集合$H = \{x \mid p^T x = \alpha\}$为凸集，其中$p$为$n$维列向量，$\alpha$为实数。