优化理论与算法

第一章
运筹学简介

郭加熠|助理教授

目录

运筹学定义与历史

1. **运筹学**（Operations Research）是一门研究数学模型与工具，从而为复杂的决策问题提供解答与洞察的学科
2. **从古至今**：“优化”（Optimization）一直是生产生活中重要的部分
   –Nothing at all takes place in the Universe in which some rule of maximum or minimum does not appear. –L. Euler, 1707-1783
3. **1940 年代**：运筹学起源于军事：如雷达部署、运输船队的护航、反潜深水炸弹投掷、项目管理问题等等，为盟军最终取得胜利提供保障
4. **1955 年**：Dantzig 提出单纯形法求解线性规划，20 世纪十大算法之一，被称为“运行世界的算法”

运筹学历史

1. **20 世纪 50 到 80 年代**：优化技术应用到生产制造、交通物流、供应链、金融、医疗等行业，并涌现大量数学规划算法
2. **20 世纪 90 年代以来**：计算机技术高速发展，数学规划学的应用范围也取得革命性的突破
3. **最近十年**：大数据时代 - 数据驱动解决超大规模系统的智能数学规划算法。同时，运筹学也是人工智能中不可或缺的关键技术模块。

中国运筹学历史

1. **上世纪 50 年代后期**：现代运筹学被引入中国。中国第一个运筹学小组是在钱学森、许国志先生的推动下，于 1956 年在中科院力学所成立。
2. **上个世纪 80 年代以来**：中国运筹学有了快速发展，在组合优化、生产系统优化、图论和非线性规划领域均有突出贡献
3. **2000 年以后**：运筹学在国民经济规划计划、工业供应链优化、物流网络建设甚至电子商务等方面有着前所未有的深度应用

课程内容

建模技巧：如何将实际问题转换成优化问题，特别是好的优化问题

- 实践案例与系统方法总结

优化算法：如何有效求解优化问题

- 基础算法介绍：单纯形法、梯度下降法、牛顿法等

编程实现：如何用现有工具求解大规模问题

- Python or Matlab, COPT 求解器，CORIDM 平台

目录

基本形式

优化问题通常可以表示为如下形式

$$
\begin{aligned}
\text{maximize/minimize}_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\
\text{subject to} \quad & \boldsymbol{x} \in \Omega
\end{aligned}
$$

其中，

- **决策变量**（能改变什么）：$\boldsymbol{x}$
- **目标函数**（怎么评估）：$f(\boldsymbol{x})$
- **可行域**（什么范围）：Ω

实际中的常用一般形式

$$\begin{aligned} \text{minimize}_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \le 0, \quad \forall i = 1, \ldots, s \\ & h_j(\boldsymbol{x}) = 0, \quad \forall j = 1, \ldots, t \end{aligned}$$

优化问题三要素：决策变量、目标函数、约束

- **决策变量**： $\boldsymbol{x}$
- **目标函数**： $f(\boldsymbol{x})$
- **可行域**: 所有可行解构成的集合
- **约束**：不等式约束函数 $g(\boldsymbol{x})$，等式约束函数 $h(\boldsymbol{x})$
- **可行解**: 某个满足**所有**约束的决策变量的取值
- **最优解**: 取值不比其他可行解对应目标函数值差
- **最优值**: 最优解对应的目标函数值（极限）

最优化问题性质

建模优化问题中，我们通常**不使用严格不等约束**，也就是 $<$ 或 $>$

- 理论上，严格不等约束常常导致优化问题没有最优解
- 实践中，从计算求解角度，将 $x < a$ 转换成 $x \leq a + \epsilon$，其中 ϵ 足够小

优化问题按照**求解结果分类**

- 不可行 (infeasible) 问题：优化问题不存在可行解
- 可行 (feasible) 问题：优化问题至少存在一个可行优解
  - 最优值无界（unbounded）
  - 最优值有界，且能取到
  - 最优值有界，且不能取到

优化问题形式分类

$$
\begin{aligned}
\text{minimize}_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\
\text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, \quad \forall i = 1, \ldots, s \\
& h_j(\boldsymbol{x}) = 0, \quad \forall j = 1, \ldots, t
\end{aligned}
$$

- **无约束优化问题**：当且仅当 $s = t = 0$ 时成立，否则均为约束优化问题。
- **线性规划问题**：目标与约束均为线性函数
- **凸规划问题**：目标与不等约束为凸函数，等式约束函数为线性函数
- **整数规划问题**：部分或全部的决策变量为整数变量
- **连续规划问题**：所有的决策变量均为连续变量

优化问题的求解

一般优化问题通常难以求解：在**次优解**与**长计算时长**中权衡

人们倾向于使用**更容易求解**的模型对问题进行建模与求解

- **线性规划**是研究得最充分、最简单的优化问题
- **凸规划**通常情况求解时间有保障
- **非凸规划**包括整数规划往往会远难于凸规划

目录

目录

线性规划: 生产问题

- 一家公司需要决策生产安排使利润最大化。每种合金的生产要求和利润如下：

	钢	铁	铜	利润
合金一	1	0	1	1
合金二	0	2	1	2
资源	100	200	150	

- **决策变量**：生产合金一的数量 x_1；生产合金二的数量 x_2
- **目标函数**：$x_1 + 2x_2$
- 约束是什么?

线性规划：生产问题

- 该生产问题的优化模型为：

$$
\begin{aligned}
\text{maximize} \quad & x_1 + 2x_2 \\
\text{subject to} \quad & x_1 \le 100 \\
& 2x_2 \le 200 \\
& x_1 + x_2 \le 150 \\
& x_1, x_2 \ge 0
\end{aligned}
$$

- 其中（1,2,100,200,150 等）都叫优化问题的**系数**（注意系数与决策变量的区别）
- 如果张三忘记加入 $x_1, x_2 \ge 0$ 约束，得出最优解 $x_1^* < 0$，这是什么意思呢？（注意：自然语言建模 vs 数学语言建模）

向量符号

在这门课中，**粗体**表示向量，标量不加粗：

$$\boldsymbol{x} = (x_1, x_2, ..., x_n)^T$$

默认所有向量为**列向量**

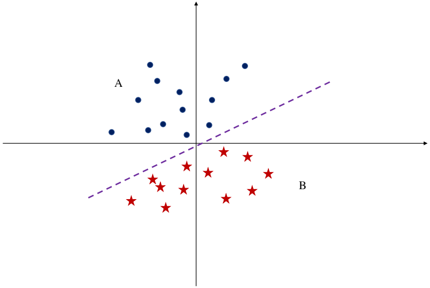
我们用 $\boldsymbol{x}^T$ 表示向量**转置**（transpose）

我们用 $\boldsymbol{a}^T\boldsymbol{x}$ 表示向量 $\boldsymbol{a}$ 与 $\boldsymbol{x}$ 的**内积**，即

$$\boldsymbol{a}^T\boldsymbol{x} = a_1x_1 + ... + a_nx_n = \sum_{i=1}^{n} a_ix_i$$

线性规划：支持向量机（Support Vector Machine）

给定 d 维实数集中的两组数据 $A = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ 及 $B = \{\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_m\}$



可以完全可分的情况

找到一个平面 $\boldsymbol{a}^T\boldsymbol{x} + b = 0$ 将数据点分开：

$$\boldsymbol{a}^T\boldsymbol{x}_i + b > 0, \forall i = 1, ..., n$$

$$\boldsymbol{a}^T\boldsymbol{y}_i + b < 0, \forall j = 1, ..., m$$

线性规划：支持向量机（Support Vector Machine）

- 上述问题等价于寻找 $\boldsymbol{a}$ 和 b 满足下列条件：

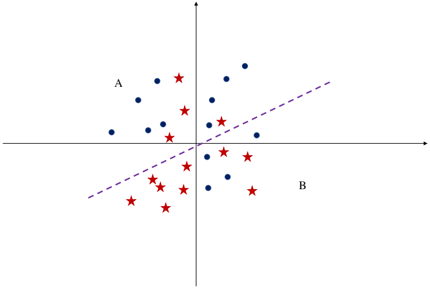
$$\boldsymbol{a}^T\boldsymbol{x}_i + b \geq 1, \forall i = 1, ..., n$$

$$\boldsymbol{a}^T\boldsymbol{y}_i + b \leq -1, \forall j = 1, ..., m$$

- 优化问题形式（此类问题也称为**可行性问题**）：

$$\begin{aligned} \text{minimize}_{\boldsymbol{a},b} \quad & 0 \\ \text{subject to} \quad & \boldsymbol{a}^T\boldsymbol{x}_i + b \geq 1, \forall i = 1, ..., n \\ & \boldsymbol{a}^T\boldsymbol{y}_j + b \leq -1, \forall j = 1, ..., m \end{aligned}$$

线性规划：支持向量机（Support Vector Machine）



不能完全可分的情况

找一个平面使得“错误”最小：

(其中，令 $(w)^+ = \max\{w, 0\}$)

- 对 A 集合中的数据，分类错误可以表示为：$(1 - \boldsymbol{a}^T\boldsymbol{x}_i - b)^+$
- 对 B 集合中的数据，分类错误可以表示为：$(\boldsymbol{a}^T\boldsymbol{y}_i + b + 1)^+$

线性规划：支持向量机（Support Vector Machine）

该问题变成**无约束非线性连续**优化问题，有目标函数：

$$\text{minimize}_{\boldsymbol{a},b} \quad \sum_i (1 - \boldsymbol{a}^T \boldsymbol{x}_i - b)^+ + \sum_j (\boldsymbol{a}^T \boldsymbol{y}_j + b + 1)^+$$

将其等价转换为**线性规划**：

- 令 $\delta_i = (1 - \boldsymbol{a}^T \boldsymbol{x}_i - b)^+, \sigma_j = (\boldsymbol{a}^T \boldsymbol{y}_j + b + 1)^+$，将“$=$”松弛成“$\geq$”则：

$$\begin{aligned}
\text{minimize}_{\boldsymbol{a},b} \quad & \sum_i \delta_i + \sum_j \sigma_j \\
\text{subject to} \quad & \delta_i \geq (1 - \boldsymbol{a}^T \boldsymbol{x}_i - b)^+, \forall i \\
& \sigma_j \geq (\boldsymbol{a}^T \boldsymbol{y}_j + b + 1)^+, \forall j
\end{aligned}$$

线性规划

SVM（Support Vector Machine）支持向量机

- 上述约束条件可以化简成如下形式：

$$\delta_i \geq (1 - \boldsymbol{a}^T \boldsymbol{x}_i - b)^+ \iff \delta_i \geq 1 - \boldsymbol{a}^T \boldsymbol{x}_i - b,\ \delta_i \geq 0$$

$$\sigma_j \geq (\boldsymbol{a}^T \boldsymbol{y}_j + b + 1)^+ \iff \sigma_j \geq \boldsymbol{a}^T \boldsymbol{y}_j + b + 1,\ \sigma_j \geq 0$$

- 优化问题最终化简成**决策变量**为 $\boldsymbol{a}, b, \delta, \sigma$ 的线性规划问题：

$$\begin{aligned}
\text{minimize}_{\boldsymbol{a},b,\delta,\sigma} \quad & \sum_i \delta_i + \sum_j \sigma_j \\
\text{subject to} \quad & \boldsymbol{a}^T \boldsymbol{x}_i + b + \delta_i \geq 1, && \forall i \\
& \boldsymbol{a}^T \boldsymbol{y}_j + b - \sigma_j \leq -1, && \forall j \\
& \delta_i, \sigma_j \geq 0, && \forall i, j
\end{aligned}$$

目录

整数规划: 选址问题

背景: 一家生产企业有 k 个候选建厂地点

目标: 开设部分候选厂，最小化成本

约束:

- 开设第 i 个厂需要固定成本 f_i
- 第 j 个零售商的需求量为 d_j
- 一单位货物从厂 i 到零售商 j 需要的成本为 c_{ij}

决策变量:

- $y_i \in \{0, 1\}$: 是否开设厂 i
- x_{ij}: 从厂 i 到零售商 j 的供应量

目标函数:

$$\text{minimize } \sum_{i=1}^k f_i y_i + \sum_{i,j} c_{ij} x_{ij}$$

约束:

- 零售商需求: $\sum_{i=1}^k x_{ij} \geq d_j$
- 只有开设的厂有产量: $x_{ij} \leq d_j y_i$

现实中的选址问题：永辉超市选址

从 38 家到 300 家的扩张

机器学习做需求预测，考虑圈层效应建模优化问题进行选址

选址模型总体结果：

- 13个行政区；
- 640个推荐区块：包含116个旧店、524个新区块推荐；

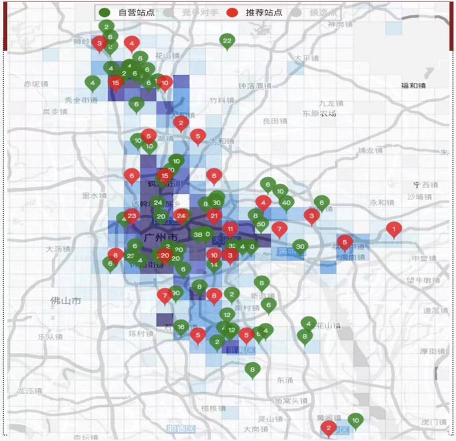
正向验证
- 87%的实际已出点位偏差1格有效命中率

反向验证
- 83%的实际已出点位偏差1格有效命中率

注： 上海除去金山、奉贤、崇明区，共13个行政区；分层级覆盖规则：外环内1km，外环外2km
正向验证：人工点位检验 逆向验证：与屏蔽测试集进行匹配

D Ge, L Hu, B Jiang, G Su, and X Wu. "Intelligent site selection for bricks-and-mortar stores." Modern Supply Chain Research and Applications (2019).

现实中的选址问题：广州充电桩选址与经营



选址推荐模型

$$\max_{j,t} \sum_{t\in T} \omega_t \sum_{j\in J} \alpha_j^t - \gamma \sum_{i\in I} z_i \qquad \text{最大化满足用户的需求}$$

$$
\begin{aligned}
s.t.\quad & \left.\begin{aligned}
& \textstyle\sum_{i\in I} x_i \le M \\
& \alpha_j^t \le r^t * \left(f_j^t + \rho \textstyle\sum_{k\in N_j} g_{kj}^t\right) \quad \forall j\in J \\
& f_i^t + \textstyle\sum_{k\in N_i} g_{ik}^t \le \beta * x_i \quad \forall i\in I
\end{aligned}\right\} \text{供给约束} \\
& \alpha_j^t \le D_j^t \quad \forall j\in J,\ \forall t\in T \\
& \left.\begin{aligned}
& \mu_j^t \le \frac{\alpha_j^t}{D_j^t} \quad \forall j\in J,\ \forall t\in T \\
& C * z_i \ge x_i \quad \forall i\in I
\end{aligned}\right\} \text{需求约束} \\
& \left.\begin{aligned}
& x_i = \hat{x}_i \quad \forall i\in I' \\
& x_i = 0 \quad \forall i\in I^*
\end{aligned}\right\} \text{站点约束} \\
& x_i \in \{0,1,2,\dots,N\} \quad \forall i\in I \quad \} \text{地形约束}
\end{aligned}
$$

图：格栅化处理后建立优化模型

其他典型图优化例子

给定一个图 $G=(V,E)$，其中 V 为点集合，$E\subseteq V\times V$ 为边集合。例如，交通网

- **最短路问题**：寻找图中 s 点到 t 点的最短路
- **最大流问题**：若每条边有流量限制，寻找图中 s 点到 t 点的大流量
- **旅行商问题**：以最短的路径不重复得遍历图中给定点集

目录

二次规划：投资组合问题

Markowitz 投资组合理论：通过分散资产组合来降低风险（鸡蛋不放在一个篮子里）

目标：保证预期回报 r_{min} 情况下，最小化投资组合的风险

决策变量：$\boldsymbol{x} \in \mathbb{R}^n$，其中 x_i 表示投资于证券 i 的资金比例。

- 已知预期收益随机向量 $\boldsymbol{R}$ 及其期望收益向量 $\boldsymbol{u}$；投资预期回报 $\boldsymbol{x}^T\boldsymbol{R}$
- **投资组合回报**：$E(\boldsymbol{x}^T\boldsymbol{R}) = \boldsymbol{x}^T E(\boldsymbol{R}) = \boldsymbol{x}^T\boldsymbol{u}.$
- **投资组合方差**：$Var(\boldsymbol{x}^T\boldsymbol{R}) = \sum_i Cov(R_i, R_j)x_i x_j = \boldsymbol{x}^T\Sigma\boldsymbol{x}$
- 收益的协方差矩阵 Σ，代表与证券相关的风险

二次规划：投资组合问题

- **最小化投资组合风险**，约束表示投资组合收益的下限：

$$\begin{aligned} \text{minimize}_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \Sigma \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{u}^T \boldsymbol{x} \geq r_{min}, \\ & \mathbf{1}^T \boldsymbol{x} = 1. \end{aligned}$$

- **最大化投资者的效用函数**：

$$\begin{aligned} \textit{maximize}_{\boldsymbol{x}} \quad & \boldsymbol{u}^T x - \frac{\delta}{2} \boldsymbol{x}^T \Sigma \boldsymbol{x} \\ \text{subject to} \quad & \mathbf{1}^T \boldsymbol{x} = 1. \end{aligned}$$

目录

无约束优化：多元线性回归

多元线性回归假设因变量 b 与自变量 $\boldsymbol{a} \in \mathbb{R}^n$ 之间是线性关系，其模型可以表示为

$$b = \sum_{i=1}^{n} x_i a_i = \boldsymbol{a}^T \boldsymbol{x}$$

其中，$\boldsymbol{x}$ 为各项特性的**权重向量**。

- **已知**：m 条观测数据 $\{\boldsymbol{a}_i, b_i\}_{i=1}^m$
- **目标**：求解使得误差最小的权重 $\boldsymbol{x}$
- **最小化误差**：$\sum_{i=1}^m (b_i - \boldsymbol{a}_i^T \boldsymbol{x})^2$

以薪资预测为例：b 为薪资水平

- a_1：工作行业
- a_2：工作经验
- a_3：工作职位
- a_4：自身专业

无约束优化：最小二乘问题

矩阵形式表达：$\text{minimize}_{\boldsymbol{x}} \quad \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$

***p*-模问题**：$\text{minimize}_{\boldsymbol{x}} \quad \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \mu\|\boldsymbol{x}\|_p$

- $p = 2$：岭回归
- $p = 1$：Lasso，稀疏性

参考文献：X. Chen, D. Ge, Z. Wang, Y. Ye, “Complexity of unconstrained $L_2 - L_p$ minimization”，Mathematical Programming, 2014

$$A = \begin{bmatrix} a_1^{\mathrm{T}} \\ a_2^{\mathrm{T}} \\ \vdots \\ a_m^{\mathrm{T}} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

向量范数：将向量映射到 $\mathbb{R}_+$

- L_1 范数：$||\boldsymbol{x}||_1 = \sum_i |x_i|$
- L_2 范数：$||\boldsymbol{x}||_2 = \sqrt{\sum_i x_i^2}$
- L_p 范数：$||\boldsymbol{x}||_p = (\sum_i |x_i|^p)^{1/p}$
- L_∞ 范数：$||\boldsymbol{x}||_\infty = \max_i |x_i|$

运筹学与人工智能

- 人工智能问题（包含机器学习、深度学习等），**通常在给定训练数据与约束条件下，针对特定的损失函数进行优化**，所以运筹学是人工智能中不可或缺的关键技术
- **机器学习、深度学习中的常用模型**：线性规划、二次规划、无约束优化、整数规划……
- **强化学习与近似动态规划**：线性规划、整数规划

目录

本课程内容一览

1. 运筹学简介
- 定义，历史
- 一般优化问题
- 各个类型例子

2. 线性规划
- 定义与实践建模
- 单纯形法求解
- 对偶与敏感度分析

3. 网络规划
- 网络规划问题
- 相关建模与算法
- 实践案例

4. 非线性规划
- 凸优化定义与性质
- 各类建模例子
- 拉格朗日对偶分析

5. 无约束优化
- 梯度法与加速方法
- 随即梯度法
- 牛顿法、拟牛顿法
- 实践案例
- 内点法

6. 整数规划
- 定义，算法
- 全单模问题
- 分支定界法
- 割平面法
- 非线性整数规划

7. 其他凸规划
- 二次规划
- 锥规划
- 半正定规划

8. 动态规划与强化学习
- 动态规划概念，算法
- 深入技巧运用
- 强化学习的概念
- 基本模型与算法
- 实践案例

相关学习资料

CORIDM 平台注册：https://coridm.d2d.ai

运筹与优化新编课程：

- 讲义、幻灯片
- 课程视频
- 案例

参考课程：

- COPT教学课程系列：COPT的具体用法与简单案例集
- 供应链管理的理论与实践：一些运筹与管理的实践案例
- 机器学习快速入门：现实中数据如何结合优化的快速讲解

感谢聆听！

Thank You!