最优化方法导论

主讲人：杨林

https://optimization-2025.github.io/

第0讲：课程导论

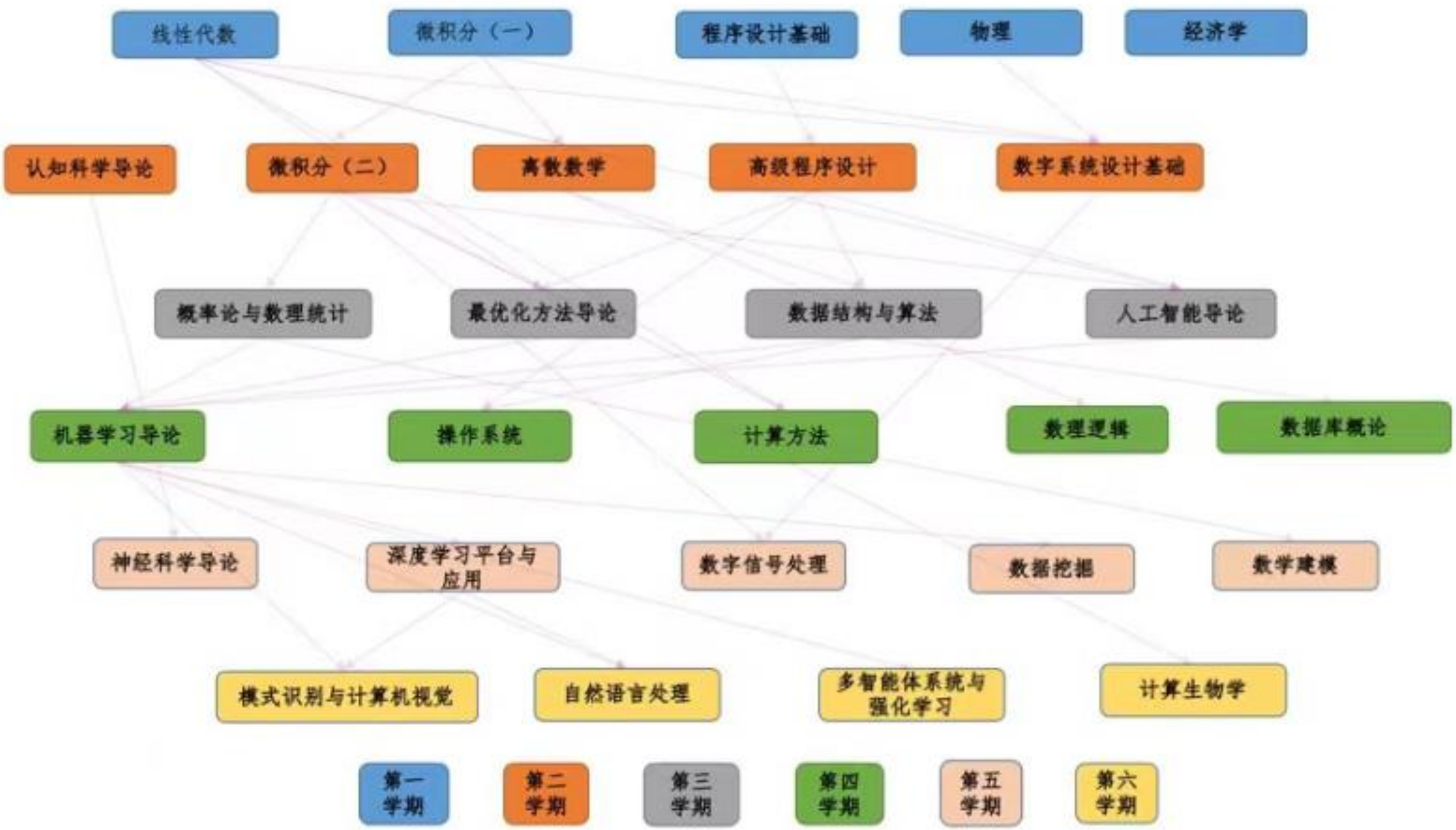
杨 林

任课老师

杨林，助理教授，特聘研究员，博导主要研究机器学习、大模型（推理、部署）、计算机系统建模与优化

邮箱：linyang@nju.edu.cn

关于本课程?



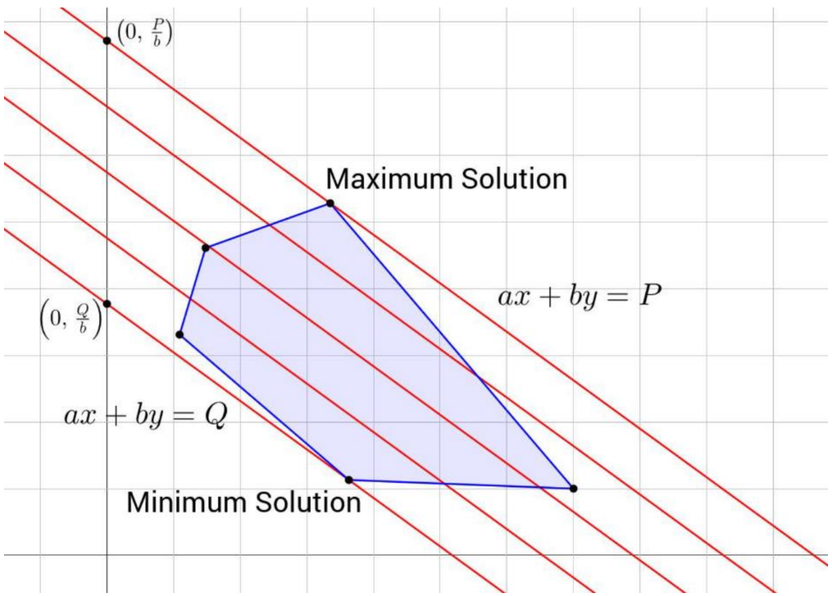
线性代数
微积分（一）
程序设计基础
物理
经济学
认知科学导论
微积分（二）
离散数学
高级程序设计
数字系统设计基础
概率论与数理统计
最优化方法导论
数据结构与算法
人工智能导论
机器学习导论
操作系统
计算方法
数理逻辑
数据库概论
神经科学导论
深度学习平台与应用
数字信号处理
数据挖掘
数学建模
模式识别与计算机视觉
自然语言处理
多智能体系统与强化学习
计算生物学
第一学期
第二学期
第三学期
第四学期
第五学期
第六学期

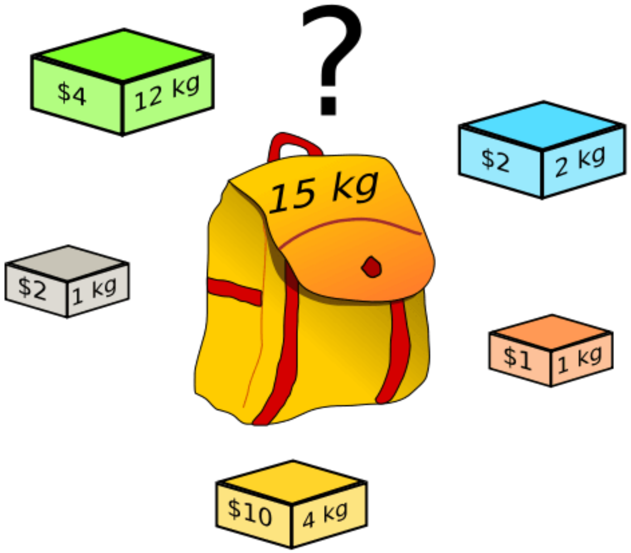
生活中的优化问题

在体积$V=\pi r^2h=330cm^3$的约束条件下，求表面积$S=2\pi r^2+2\pi rh$的最小值？

进一步考虑：人手握的舒适度（约束条件增加）、罐子叠放运输的稳定性、成本

优化问题的可解性





$$f(x, y) = ax + by$$

$$x_i \in \{0,1\} \quad f(x) = \sum x_i v_i$$

多项式时间内可以找到精确的最优解

要找到最优解，需要执行大量的“操作”

凸优化问题

“In fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity.” ------Rockafellar（洛克菲拉）,1993

关于本课程?

1 助教：

李怡璠、李紫微、陈浩旭、唐恒俊

2 课程时间：

周三、周五（双）5-6 节

3 课程考核（适量微调）：

- 期末考试（闭卷）50%
- 期中考试（闭卷）25%
- 课堂测验（4次）10%；作业（6次，optimize2023@163.com）15%

4 参考书：

“Convex Optimization” Stephen Boyd and Lieven Vandenberghe
关键字“优化”相关教材

第一讲：数学优化的要素

杨 林

大 纲

1.数学优化的要素

2.优化方法建模实例

2.1.生活中的优化问题

2.2.运筹学中经典的优化的例子

2.3.网络优化

2.4.大模型加速推理

大 纲

1.数学优化的要素

2. 优化方法建模实例

2.1.生活中的优化问题

2.2.运筹学中经典的优化的例子

2.3.网络优化

2.4.大模型加速推理

1 数学优化的要素

“优化”作为工程上的一种工具：
“在给定的决策集中（通过某种方式）寻找最优解”

最小化	$f_0(x)$
约束条件	$f_i(x) \leq b_i,\ i = 1, 2, \ldots, m$

- $x = (x_1, x_2, \ldots, x_n)$，优化变量
- $f_0: \mathbb{R}^n \to \mathbb{R}$，目标函数（目标）
- $f_i: \mathbb{R}^n \to \mathbb{R}$，$i = 1, 2, \ldots, m$，约束函数（约束）

最优解 x^* 即为满足所有约束条件的向量中使 $f(x)$ 取得最小值的向量.

大 纲

1.数学优化的要素

2.优化方法建模实例

2.1.生活中的优化问题

2.2.运筹学中经典的优化的例子

2.3.网络优化

2.4.大模型加速推理

2 应用实例

■ 2.1 生活中的优化问题

□ **例 1(零食采购问题)：** 为周末社团团建采购多种零食，在满足来自五湖四海同学的口味需求以及有限的预算和储物空间的前提下，最大化全体同学的总体满意度。假设：

1. 选择 n 种食物的数量分别为 $x_1, \cdots, x_n$
2. s_j 表示采购一单位第 j 种零食所带来的满意度评分；c_j 表示第 j 种零食一单位的价格；v_j 表示第 j 种零食一单位的体积
3. 团建零食采购的预算为 W
4. 团建场地零食柜储物空间有限，最大储存空间为 V
5. 需要满足社员的最低糖果类零食x_c需求
6. 避免某种零食过多(薯片x_p不能超过总量的 30%)

2 应用实例

可以得到

$$
\begin{aligned}
&\text{最大化} \quad && \textstyle\sum_{j=1}^{n} s_j x_j \\
&\text{约束条件} && \textstyle\sum_{j=1}^{n} c_j x_j \le W,\ \ j = 1, 2, \dots, n \\
& && \textstyle\sum_{j=1}^{n} v_j x_j \le V,\ \ j = 1, 2, \dots, n \\
& && x_c \ge \alpha \\
& && x_p - 0.3 \textstyle\sum_{j=1}^{n} x_j \le 0 \\
& && x_j \ge 0, j = 1,2, \cdots, n
\end{aligned}
$$

- **决策变量**：$x_1, x_2, \dots, x_{\mathrm{n}}$（每种零物的购买量）
- **目标函数**：最大化 $\sum_{j=1}^{n} s_j x_j$（最大化满意度）

2 应用实例

最大化	$\sum_{j=1}^{n} s_j x_j$
约束条件	$\sum_{j=1}^{n} c_j x_j \leq W,\ j = 1, 2, \ldots, n$
	$\sum_{j=1}^{n} v_j x_j \leq V,\ j = 1, 2, \ldots, n$
	$x_c \geq \alpha$
	$x_p - 0.3 \sum_{j=1}^{n} x_j \leq 0$
	$x_j \geq 0, j = 1,2, \cdots, n$

- **约束条件:**
  - **预算约束:** $\sum_{j=1}^{n} c_j x_j \leq W$（不超过零食采购预算）
  - **空间约束:** $\sum_{j=1}^{n} v_j x_j \leq V$（不超过团建场地零食放置空间）
  - **最低糖果约束:** $x_c \geq \alpha$
  - **最大薯片约束:** $x_p - 0.3 \sum_{j=1}^{n} x_j \leq 0$
  - **非负约束:** $x_j \geq 0, j = 1,2, \cdots, n$（零食量非负）

2 应用实例

■ 2.2 运筹学中的经典优化例子

□ **例 2(饮食问题):** 设计一份满足所有基本营养需求的每日饮食计划，要求对于给定食物在满足所有营养需求的前提下，使得每日的饮食总成本最低。假设:

1. 选择 n 种食物的数量分别为 $x_1, \cdots, x_n$
2. 食物 j 的一单位成本为 c_j ,包含营养素 i 的数量为 a_{ij}
3. 健康饮食要求营养素 i 至少为 b_i

2 应用实例

可以得到

最大化 $z=\sum_{j=1}^{n} c_j x_j$

约束条件 $\sum_{j=1}^{n} a_{ij} x_j \geq b_i,\ \ i=1,2,\ldots,m$

$x_j \geq 0, j=1,2,\cdots,n$

- **决策变量**：$x_1, x_2, \ldots, x_\mathrm{n}$（每种零物的购买量）
- **目标函数**：最小化 $z=\sum_{j=1}^{n} c_j x_j$（最小化总成本）
- **约束条件**：
- **营养约束**：$\sum_{j=1}^{n} a_{ij} x_j \geq b_i,\ \ i=1,2,\ldots,m$（确保每种营养达标）
- **非负约束**：$x_j \geq 0, j=1,2,\cdots,n$（食物量非负）

2 应用实例

■ 2. 3 网络优化

□ **例 3(互联网链路带宽分配):** 一个简单网络有 3 个用户 ($User$ 1,2,3)共享两条通信链路($Link\ A, B$),在链路容量限制下，最大化所有用户的总效用(全局满意度)。假设:

1. 链路容量

- $Link\ A$:带宽容量 $C_A = 10\ Mbps$
- $Link\ B$:带宽容量 $C_B = 15\ Mbps$

2. 用户路径

- $User$ 1:仅使用 $Link\ A \rightarrow$ 速率 x_1
- $User$ 2:仅使用 $Link$ B $\rightarrow$ 速率 x_2
- $User$ 3:同时使用 $Link\ A$ 和 $Link$ B $\rightarrow$ 速率 x_3

3. 用户 i 的满意度由其对速率的效用函数 $U_i(x_i)$ 衡量，假设采用公平性模型：$U_i(x_i) = \log(x_i)$(对数效用函数)

2 应用实例

可以得到

最大化	$\sum_{i=1}^{3} U_i(x_i) = \log(x_1) + \log(x_2) + \log(x_3)$
约束条件	$x_1 + x_3 \leq 10$
	$x_2 + x_3 \leq 15$
	$x_i \geq 0, j = 1,2,3$

- **决策变量**：x_1, x_2, x_3（分别表示 $User$ 1, 2, 3 的分配速率（Mbps））
- **目标函数**：$\log(x_1) + \log(x_2) + \log(x_3)$
- **约束条件**：

□ **链路容量约束**：

- $Link\ A$：$User$ 1 和 $User$ 3 共享→ $x_1 + x_3 \leq 10$
- $Link\ B$：$User$ 2 和 $User$ 3 共享→ $x_2 + x_3 \leq 15$

□ **非负约束**：$x_j \geq 0, j = 1,2,3$

2 应用实例

■ 2.4 大模型加速推理

□ 例 4: 在一个异构的多节点推理系统中:

- 每个节点 N_i 拥有:
  - 单位调用推理时延 c_{ij}
  - 单位调用计算时延 t_c^i
  - 单位调用传输时延 t_t^i
  - 最大缓存参数量 m_i
- 每个参数 P_j 具有:
  - 参数量 ω_j
  - 参数调用频率 λ_j

□ 每个参数只能分配给一个节点处理

□ 每个节点分配的总参数量不得超过最大缓存参数量 m_i

□ 目标是在分配这些参数给不同节点的同时，**最小化**所有节点中完成任务所需的**最大耗时**(缓解瓶颈)

2 应用实例

可以得到

$$
\begin{array}{ll}
\text{最小化} & T_{\max} \\
\text{约束条件} & \sum_{i=1}^{n} x_{ij} = 1, \forall j = 1,\ \dots, m \\
& \sum_{j=1}^{m} x_{ij}\omega_j \le m_i, \forall i = 1,\ \dots, n \\
& \sum_{j=1}^{m} x_{ij}c_{ij} \le T_{\max}1, \forall i = 1,\ \dots, n \\
& c_{ij} = \lambda_j t_c^i + \omega_j t_t^i\ \ \forall i,j \\
& x_{ij} \in \{0,1\}, \forall i,j
\end{array}
$$

- **决策变量：** $x_{ij} \in \{0,1\}$:表示 P_j 是否分给配节点 N_i
- **中间变量：**
  - c_{ij} :节点 N_i 的运行P_j耗时
  - $T_{\max}$:所有节点中的最大耗时
- **目标函数：** 最小化 $T_{\max}$

2 应用实例

可以得到

最小化	$T_{\max}$
约束条件	$\sum_{i=1}^{n} x_{ij} = 1, \forall j = 1, \ \ldots, m$
	$\sum_{j=1}^{m} x_{ij}\omega_j \leq m_i, \forall i = 1, \ \ldots, n$
	$\sum_{j=1}^{m} x_{ij}c_{ij} \leq T_{\max}1, \forall i = 1, \ \ldots, n$
	$c_{ij} = \lambda_j t_c^i + \omega_j t_t^i \ \ \forall i, j$
	$x_{ij} \in \{0,1\}, \forall i, j$

■ **约束条件：**

□ **参数唯一分配约束：** $\sum_{i=1}^{n} x_{ij} = 1, \forall j = 1, \ \ldots, m$（每个参数只能分配给一个节点）

□ **缓存容量约束：** $\sum_{j=1}^{m} x_{ij}\omega_j \leq m_i, \forall i = 1, \ \ldots, n$（节点所处理参数的缓存总量不得超过其缓存上限）

□ **最大耗时定义约束：** $\sum_{j=1}^{m} x_{ij}c_{ij} \leq T_{\max}1, \forall i = 1, \ \ldots, n$

2 应用实例

可以得到

最小化	$T_{\max}$
约束条件	$\sum_{i=1}^{n} x_{ij} = 1, \forall j = 1, \ \ldots, m$
	$\sum_{j=1}^{m} x_{ij}\omega_j \le m_i, \forall i = 1, \ \ldots, n$
	$\sum_{j=1}^{m} x_{ij}c_{ij} \le T_{\max}1, \forall i = 1, \ \ldots, n$
	$c_{ij} = \lambda_j t_c^i + \omega_j t_t^i \ \ \forall i, j$
	$x_{ij} \in \{0,1\}, \forall i, j$

■ **约束条件:**

□ **（物理）等式约束:** $c_{ij} = \lambda_j t_c^i + \omega_j t_t^i$

谢 谢 ！