



最优化方法

东南大学 计算机&人工智能学院 宋沫飞 songmf@seu.edu.cn



引言



- 数学优化
- 最小二乘和线性规划
- 凸优化
- 例题
- 课程目标和内容
- 非线性优化
- 发展历史









□ (数学) 优化问题/数学规划





- □(数学)优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素





- □ (数学) 优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素
- □数学形式





- □ (数学) 优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素
- □数学形式

```
minimize f_0(x)
subject to f_i(x) \leq b_i, \quad i = 1, \dots, m
```





- □ (数学) 优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素
- ■数学形式

```
minimize f_0(x)
subject to f_i(x) \leq b_i, \quad i = 1, \dots, m
```

• $x = (x_1, ..., x_n)$: 优化变量





- □ (数学) 优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素
- □数学形式

```
minimize f_0(x)
subject to f_i(x) \leq b_i, \quad i = 1, \dots, m
```

- $x = (x_1, ..., x_n)$: 优化变量
- $f_0: \mathbf{R}^n \to \mathbf{R}:$ 目标函数





- □ (数学) 优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素
- □数学形式

minimize $f_0(x)$ subject to $f_i(x) \leq b_i, \quad i = 1, \dots, m$

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





- □ (数学) 优化问题/数学规划
 - ❖ 从一个可行解集合中,寻找出最优的元素
- □数学形式

minimize $f_0(x)$ subject to $f_i(x) \leq b_i, \quad i = 1, \dots, m$

- $x = (x_1, ..., x_n)$: 优化变量
- $f_0: \mathbf{R}^n \to \mathbf{R}:$ 目标函数
- $f_i: \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$: 约束函数
- □ 最优解: 在所有满足约束的向量中,向量*x**对应的目标函数值最小。









- □投资组合优化
 - 优化变量: 各资产分配的资本数
 - 约束: 总预算、每一份资产的资本范围、最小收益
 - 目标: 总风险或回报





- □投资组合优化
 - 优化变量: 各资产分配的资本数
 - 约束: 总预算、每一份资产的资本范围、最小收益
 - 目标: 总风险或回报
- □电子设计中的器件尺寸
 - 优化变量: 器件的长和宽
 - 约束: 工程约束、时间要求、最大面积
 - 目标: 总功耗





- □投资组合优化
 - 优化变量: 各资产分配的资本数
 - 约束: 总预算、每一份资产的资本范围、最小收益
 - 目标: 总风险或回报
- □电子设计中的器件尺寸
 - 优化变量: 器件的长和宽
 - 约束: 工程约束、时间要求、最大面积
 - 目标: 总功耗
- ■数据拟合
 - 优化变量: 模型参数
 - 约束: 先验信息、参数限制
 - 目标: 预测误差









□一般形式的优化问题





- □一般形式的优化问题
 - *难以求解





- □一般形式的优化问题
 - *难以求解
 - ❖折中:长时间的计算代价,找不到解





- □一般形式的优化问题
 - *难以求解
 - ❖折中:长时间的计算代价,找不到解

□ 例外: 特定问题存在高效且可靠的解法





- □一般形式的优化问题
 - *难以求解
 - ❖折中:长时间的计算代价,找不到解

- □ 例外: 特定问题存在高效且可靠的解法
 - ❖最小二乘





- □一般形式的优化问题
 - *难以求解
 - ❖折中:长时间的计算代价,找不到解

- □ 例外: 特定问题存在高效且可靠的解法
 - ❖最小二乘
 - *线性规划





- □一般形式的优化问题
 - *难以求解
 - ❖折中:长时间的计算代价,找不到解

- □ 例外: 特定问题存在高效且可靠的解法
 - ❖最小二乘
 - *线性规划
 - * 凸优化问题









minimize
$$||Ax - b||_2^2$$





minimize $||Ax - b||_2^2$

□最小二乘





minimize $||Ax - b||_2^2$

□最小二乘

A 解析解: $x^* = (A^T A)^{-1} A^T b$





- □最小二乘
 - *****解析解: $x^* = (A^T A)^{-1} A^T b$
 - ❖具有可靠且有效的求解算法和软件





- □最小二乘
 - *解析解: $x^* = (A^T A)^{-1} A^T b$
 - ❖具有可靠且有效的求解算法和软件
 - **◇** 计算时间: 正比于 n^2k ($A \in \mathbf{R}^{k \times n}$); 若A具有特殊结构,求解更快。





- □最小二乘
 - *****解析解: $x^* = (A^T A)^{-1} A^T b$
 - ❖具有可靠且有效的求解算法和软件
 - * 计算时间: 正比于 n^2k $(A \in \mathbf{R}^{k \times n})$; 若A具有特殊结构,求解更快。
 - *成熟技术





- □最小二乘
 - *****解析解: $x^* = (A^T A)^{-1} A^T b$
 - ❖具有可靠且有效的求解算法和软件
 - **◇** 计算时间: 正比于 n^2k ($A \in \mathbf{R}^{k \times n}$); 若A具有特殊结构,求解更快。
 - *成熟技术
- □最小二乘的使用





- □最小二乘
 - *****解析解: $x^* = (A^T A)^{-1} A^T b$
 - ❖具有可靠且有效的求解算法和软件
 - **◇** 计算时间: 正比于 n^2k ($A \in \mathbf{R}^{k \times n}$); 若A具有特殊结构,求解更快。
 - ❖ 成熟技术
- □最小二乘的使用
 - *判别十分简单





- □最小二乘
 - *****解析解: $x^* = (A^T A)^{-1} A^T b$
 - ❖具有可靠且有效的求解算法和软件
 - * 计算时间: 正比于 n^2k ($A \in \mathbf{R}^{k \times n}$); 若A具有特殊结构,求解更快。
 - ❖ 成熟技术
- □最小二乘的使用
 - *判别十分简单
 - ❖ 使用标准方法增强灵活性(加权、正则化)



线性规划





线性规划



minimize
$$c^T x$$
 subject to $a_i^T x \leq b_i, \quad i = 1, \dots, m$



线性规划



□线性规划





minimize
$$c^T x$$
 subject to $a_i^T x \leq b_i, \quad i = 1, \dots, m$

- □线性规划
 - * 没有解析解





minimize
$$c^T x$$
 subject to $a_i^T x \leq b_i, \quad i = 1, \dots, m$

- □线性规划

 - ❖ 没有解析解❖ 具有可靠且有效的求解算法和软件





- □线性规划





- □线性规划

 - ❖ 成熟技术





- □线性规划

 - * 成熟技术
- ■线性规划的使用





- □线性规划

 - *没有解析解 *具有可靠且有效的求解算法和软件 *计算时间:正比于 n^2m if $m \ge n$;若具有特殊结构,求解更
 - ❖ 成熟技术
- ■线性规划的使用
 - ❖ 判别难于最小二乘





- □线性规划

 - *没有解析解 *具有可靠且有效的求解算法和软件 *计算时间:正比于 n^2m if $m \ge n$;若具有特殊结构,求解更
 - ❖ 成熟技术
- ┙线性规划的使用

 - ❖判别难于最小二乘❖一些标准的技巧可用于将某些问题转化为线性规划(分段 线性方程、包含范数的问题)









minimize
$$f_0(x)$$

subject to $f_i(x) \leq b_i, \quad i = 1, \dots, m$





minimize
$$f_0(x)$$

subject to $f_i(x) \leq b_i, \quad i = 1, \dots, m$

□目标函数和约束函数均为凸函数:





minimize
$$f_0(x)$$

subject to $f_i(x) \leq b_i, \quad i = 1, \dots, m$

□目标函数和约束函数均为凸函数:

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y)$$

*若
$$\alpha + \beta = 1$$
, $\alpha \ge 0$, $\beta \ge 0$





minimize
$$f_0(x)$$

subject to $f_i(x) \leq b_i, \quad i = 1, \dots, m$

□目标函数和约束函数均为凸函数:

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y)$$

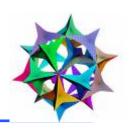
*若
$$\alpha + \beta = 1$$
, $\alpha \ge 0$, $\beta \ge 0$

□最小二乘和线性规划均为特殊的凸优化问题









□凸优化问题求解





- □凸优化问题求解
 - *没有解析解





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法
 - ❖ 计算时间: 约正比于 $\max\{n^3, n^2m, F\}$, F是对各函数的值、一阶和二阶导数的代价





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法
 - ❖ 计算时间: 约正比于 $\max\{n^3, n^2m, F\}$, F是对各函数的值、一阶和二阶导数的代价
 - ❖ 几乎成熟





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法
 - ❖ 计算时间: 约正比于 $\max\{n^3, n^2m, F\}$, F是对各函数的值、一阶和二阶导数的代价
 - ❖ 几乎成熟
- □凸优化的使用





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法
 - ❖ 计算时间: 约正比于 $\max\{n^3, n^2m, F\}$, F是对各函数的值、一阶和二阶导数的代价
 - ❖ 几乎成熟
- □凸优化的使用
 - ◆通常很难判別





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法
 - ❖ 计算时间:约正比于 $\max\{n^3, n^2m, F\}$,F是对各函数的值、一阶和二阶导数的代价
 - ❖ 几乎成熟
- □凸优化的使用
 - ❖通常很难判别
 - ❖具有很多技巧可以将其他问题转化为凸优化形式





- □凸优化问题求解
 - *没有解析解
 - ❖具有可靠且有效的解法
 - ❖ 计算时间:约正比于 $\max\{n^3, n^2m, F\}$,F是对各函数的值、一阶和二阶导数的代价
 - ❖ 几乎成熟
- □凸优化的使用
 - ❖通常很难判别
 - ❖具有很多技巧可以将其他问题转化为凸优化形式
 - ❖很多问题可以通过凸优化进行求解







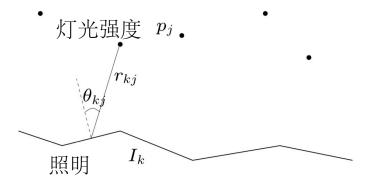


 $\square M$ 盏灯用于n块(面积小、平面)区域的照明:第k 块区块的照明强度 I_k 和灯光强度 p_i 线性相关





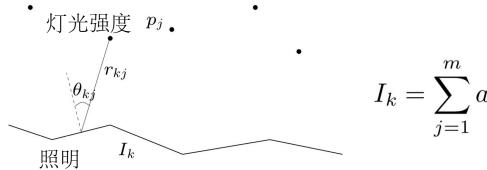
 $\square M$ 盡灯用于n块(面积小、平面)区域的照明:第k 块区块的照明强度 I_k 和灯光强度 p_i 线性相关







 $\square M$ 盡灯用于n块(面积小、平面)区域的照明:第k 块区块的照明强度 I_k 和灯光强度 p_i 线性相关

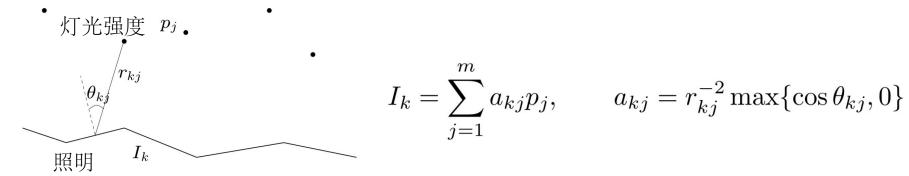


$$I_k = \sum_{j=1}^m a_{kj} p_j, \qquad a_{kj} = r_{kj}^{-2} \max\{\cos \theta_{kj}, 0\}$$





 $\square M$ 盏灯用于n块(面积小、平面)区域的照明:第k 块区块的照明强度 I_k 和灯光强度 p_i 线性相关

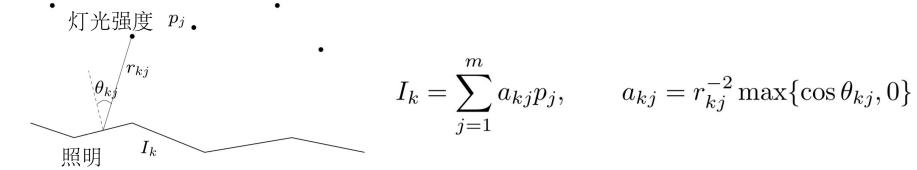


□问题: 在有约束的灯光强度下获取理想的照明强度





 $\square M$ 盡灯用于n块(面积小、平面)区域的照明:第k 块区块的照明强度 I_k 和灯光强度 p_i 线性相关



回题: 在有约束的灯光强度下获取理想的照明强度 minimize $\max_{k=1,...,n} |\log I_k - \log I_{\text{des}}|$ subject to $0 \le p_j \le p_{\text{max}}, \quad j=1,\ldots,m$









■使用最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2$$

*若得到的 p_j 超出范围,强制修正





■使用最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{des})^2$$

- *若得到的 p_i 超出范围,强制修正
- ■使用带权的最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2 + \sum_{j=1}^{m} w_j (p_j - p_{\text{max}}/2)^2$$





■使用最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2$$

- *若得到的 p_i 超出范围,强制修正
- ■使用带权的最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2 + \sum_{j=1}^{m} w_j (p_j - p_{\text{max}}/2)^2$$

*迭代调整权重,直到 $0 \le p_j \le p_{\text{max}}$





■使用最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2$$

- *若得到的 p_i 超出范围,强制修正
- ■使用带权的最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2 + \sum_{j=1}^{m} w_j (p_j - p_{\text{max}}/2)^2$$

- *迭代调整权重,直到 $0 \le p_j \le p_{\text{max}}$
- ■使用线性规划

minimize
$$\max_{k=1,...,n} |I_k - I_{\text{des}}|$$

subject to $0 \le p_j \le p_{\text{max}}, \quad j = 1,...,m$





■使用最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2$$

- *若得到的 p_i 超出范围,强制修正
- ■使用带权的最小二乘

minimize
$$\sum_{k=1}^{n} (I_k - I_{\text{des}})^2 + \sum_{j=1}^{m} w_j (p_j - p_{\text{max}}/2)^2$$

- *迭代调整权重,直到 $0 \le p_j \le p_{\text{max}}$
- ■使用线性规划

minimize
$$\max_{k=1,...,n} |I_k - I_{\text{des}}|$$
 subject to $0 \le p_j \le p_{\text{max}}, \quad j=1,\ldots,m$

□ 显然, 上述解法均为近似解法



使用凸优化





使用凸优化



minimize
$$f_0(p) = \max_{k=1,...,n} h(I_k/I_{\text{des}})$$
 subject to $0 \le p_j \le p_{\text{max}}, \quad j=1,...,m$





minimize
$$f_0(p) = \max_{k=1,...,n} h(I_k/I_{\text{des}})$$
 subject to $0 \le p_j \le p_{\text{max}}, \quad j=1,...,m$

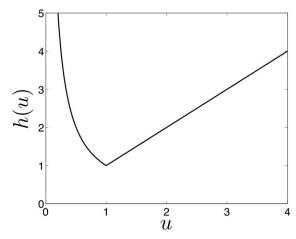
山比处: $h(u) = \max\{u, 1/u\}$





minimize
$$f_0(p) = \max_{k=1,...,n} h(I_k/I_{\text{des}})$$
 subject to $0 \le p_j \le p_{\text{max}}, \quad j=1,...,m$

山此处: $h(u) = \max\{u, 1/u\}$



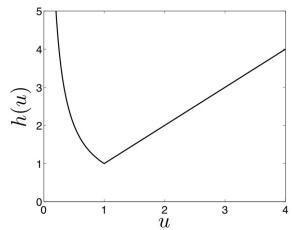




minimize
$$f_0(p) = \max_{k=1,...,n} h(I_k/I_{\text{des}})$$

subject to $0 \le p_j \le p_{\text{max}}, \quad j=1,...,m$

山此处: $h(u) = \max\{u, 1/u\}$



□ 该函数为凸函数:多个凸函数的最大化函数仍为凸函数

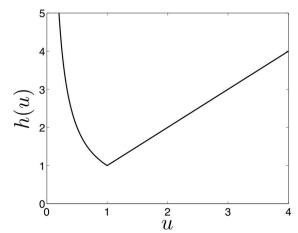




minimize
$$f_0(p) = \max_{k=1,...,n} h(I_k/I_{\text{des}})$$

subject to $0 \le p_j \le p_{\text{max}}, \quad j = 1,...,m$

山此处: $h(u) = \max\{u, 1/u\}$



- □ 该函数为凸函数: 多个凸函数的最大化函数仍为凸函数
- □精确求解的代价比最小二乘法略大









- ■额外约束: 约束增加会否使问题复杂?
 - ❖任意10盏灯的照明强度不会超过总能量的一半
 - ❖打开的灯不超过总数的一半





- ■额外约束: 约束增加会否使问题复杂?
 - ❖任意10盏灯的照明强度不会超过总能量的一半
 - ❖打开的灯不超过总数的一半
- □答案:增加条件1,问题仍然很容易求解;增加 条件2,大大增加问题的难度





- ■额外约束:约束增加会否使问题复杂?
 - ❖任意10盏灯的照明强度不会超过总能量的一半
 - ❖打开的灯不超过总数的一半
- □答案:增加条件1,问题仍然很容易求解;增加条件2,大大增加问题的难度
- □ 寓意: 直觉不总是正确; 没有背景知识, 非常容易的问题和非常困难的问题看起来会十分相似。



课程目标和内容





课程目标和内容



- □目标
 - *能够将一些问题形式化为凸优化问题
 - *能够求解具有一定规模的优化问题
 - *对求解过程的性能进行分析



课程目标和内容



- ┛目标
 - ❖能够将一些问题形式化为凸优化问题
 - ❖能够求解具有一定规模的优化问题
 - *对求解过程的性能进行分析
- □内容
 - ◆理论: 凸集、凸函数、凸优化问题
 - ❖应用
 - **❖**算法









□非凸的优化问题求解思路





- □非凸的优化问题求解思路
- □局部优化





- □非凸的优化问题求解思路
- □局部优化
 - *在可行解中寻找近似最小的点





- □非凸的优化问题求解思路
- □局部优化
 - ❖在可行解中寻找近似最小的点
 - *速度快





- □非凸的优化问题求解思路
- □局部优化
 - *在可行解中寻找近似最小的点
 - *速度快
 - *需要初始解





- □非凸的优化问题求解思路
- □局部优化
 - ❖在可行解中寻找近似最小的点
 - ❖速度快
 - **◇**需要初始解
 - ❖不能提供和全局最优解的差距信息





- □非凸的优化问题求解思路
- □局部优化
 - ❖在可行解中寻找近似最小的点
 - ❖速度快
 - **◇**需要初始解
 - ❖不能提供和全局最优解的差距信息
- □全局优化





- □非凸的优化问题求解思路
- □局部优化
 - ❖在可行解中寻找近似最小的点
 - ❖速度快
 - *需要初始解
 - ❖不能提供和全局最优解的差距信息
- □全局优化
 - * 计算时间和问题规模呈指数增长关系





- □非凸的优化问题求解思路
- □局部优化
 - ❖在可行解中寻找近似最小的点
 - ❖速度快
 - *需要初始解
 - ❖不能提供和全局最优解的差距信息
- □全局优化
 - ❖计算时间和问题规模呈指数增长关系
- □上述方法大多基于求解凸的子问题









□ 凸分析理论: 1900-1970





□ 凸分析理论: 1900-1970

算法





- □ 凸分析理论: 1900-1970
- 算法
 - *1939:线性规划





- □ 凸分析理论: 1900-1970
- □算法
 - *1939:线性规划
 - ❖ 1947:单纯形法





- □ 凸分析理论: 1900-1970
- □算法
 - *1939:线性规划
 - ***1947:**单纯形法
 - ❖ 1960年代:早期的内点法





- □ 凸分析理论: 1900-1970
- 算法
 - *1939:线性规划
 - ❖ 1947:单纯形法
 - ❖ 1960年代:早期的内点法
 - ❖ 1970年代: 椭球法和其他次梯度法





- □ 凸分析理论: 1900-1970
- 算法
 - ❖ 1939:线性规划
 - **❖ 1947:**单纯形法
 - ❖ 1960年代:早期的内点法
 - ❖ 1970年代: 椭球法和其他次梯度法
 - ❖ 1980年代:面向线性规划的多项式时间的内点法





- □ 凸分析理论: 1900-1970
- □算法
 - ❖ 1939:线性规划
 - ❖ 1947:单纯形法
 - ❖ 1960年代:早期的内点法
 - ❖ 1970年代: 椭球法和其他次梯度法
 - ❖ 1980年代:面向线性规划的多项式时间的内点法
 - **❖ 1980**年代后期至今:面向非线性凸优化的多项式时间的内点法





- □ 凸分析理论: 1900-1970
- □算法
 - ***1939:**线性规划
 - ❖ 1947:单纯形法
 - ❖ 1960年代:早期的内点法
 - ❖ 1970年代: 椭球法和其他次梯度法
 - ❖ 1980年代:面向线性规划的多项式时间的内点法
 - **❖ 1980**年代后期至今:面向非线性凸优化的多项式时间的内点法
- □应用





- □ 凸分析理论: 1900-1970
- □算法
 - ❖ 1939:线性规划
 - ❖ 1947:单纯形法
 - ❖ 1960年代:早期的内点法
 - ❖ 1970年代: 椭球法和其他次梯度法
 - ❖ 1980年代:面向线性规划的多项式时间的内点法
 - **❖ 1980**年代后期至今:面向非线性凸优化的多项式时间的内点法
- □应用
 - **❖ 1990**前:工程应用很少





- □ 凸分析理论: 1900-1970
- □算法
 - ❖ 1939:线性规划
 - ❖ 1947:单纯形法
 - ❖ 1960年代:早期的内点法
 - ❖ 1970年代: 椭球法和其他次梯度法
 - ❖ 1980年代:面向线性规划的多项式时间的内点法
 - **❖ 1980**年代后期至今:面向非线性凸优化的多项式时间的内点法
- □应用
 - **❖ 1990**前: 工程应用很少
 - ❖ 自1990年:工程应用开始增多(控制、信号处理、通讯、电路设计);新的问题类型(半正定规划等)