运筹学

（绪论）

王焕钢

清华大学自动化系

要点：课程主要内容

运筹学　英国: Operational Research, OR

美国: Operations Research, OR

（中国台湾　作业研究）

运筹学作为一门现代科学，在第二次世界大战期间首先在英美两国发展起来

运筹学用于解决现实生活中各种复杂问题，特别是改善或优化现有系统的效率

1959年国际运筹学协会（International Federation of Operations Research Societies, IFORS）成立

运筹学的主要分支

一般数学规划模型

- 线性规划
- 整数规划
- 非线性规划
- 动态规划
- 网络流优化

……

本课程内容

特定问题的数学模型

- 网络计划
- 排队论
- 存储论
- 决策论
- 对策论

……

运筹学研究的核心问题：建模与求解

建模方面

1）知道常用数学规划模型的结构和特点

2）有助于选择合适的模型解决具体问题

求解方面

1）知道常用模型的求解思路和求解算法

2）有助于对具体问题研发新的求解算法

要点：优化问题建模

例如：人工智能和优化联系非常密切的一系列问题（回归、辨识、估计、训练、学习、拟合、逼近…）

已知某个标量 y 和某个向量 $X=\left(x_1,x_2,\cdots,x_n\right)^T$ 之间若干一一对应的样本数据

$$\left\{y(t),\ X(t)\right\}$$

$$X(t)=\left(x_1(t),x_2(t),\cdots,x_n(t)\right)^T,\ t=1,2,\cdots,N$$

要确定一个函数 $f(X)$，使在包含所有样本数据的某个集合 Ω 里能够用 $f(X)$ 描述 y 和 X 之间的对应关系，即使误差 $\left|y-f(X)\right|$ 对任意的 $\left\{y,X\right\}\in\Omega$ 都尽量小

基本方法：选择含有待定参数的函数 $\hat{f}(X,\theta)$, 通过极小化某种样本误差, 确定待定参数得到所需函数

常用 l_1, l_2, l_∞ 范数的样本误差

$$E_p(\theta)=\begin{cases}\sum_{t=1}^{N}\left|y(t)-\hat{f}\left(X(t),\theta\right)\right| & p=1\\ \sum_{t=1}^{N}\left(y(t)-\hat{f}\left(X(t),\theta\right)\right)^2 & p=2\\ \max_{1\le t\le N}\left|y(t)-\hat{f}\left(X(t),\theta\right)\right| & p=\infty\end{cases}$$

于是，最终要解决的是下述优化问题

$$\min_{\theta} E_p(\theta)$$

这是连续变量无约束优化问题

对于采用 l_∞ 范数形成的优化问题

$$\min_{\theta}\max_{1\le t\le N}\left|y(t)-\hat{f}\left(X(t),\theta\right)\right|$$

为克服目标函数不可导的困难，可以等价转换成下面的连续变量约束优化问题

$$\begin{aligned}&\min\quad \lambda\\ &\text{s.t.}\quad -\lambda\le y(t)-\hat{f}\left(X(t),\theta\right)\le\lambda,\ \ \forall 1\le t\le N\end{aligned}$$

特别是，当 $\hat{f}(X,\theta)$ 是 θ 的线性函数时，即

$$\hat{f}(X,\theta)=\sum_{i=1}^{m}\theta_i\varphi_i(X)$$

上面的优化问题是线性规划问题

通常可以选择足够多的基函数 $\varphi_i(X)$ 使优化问题

$$\min_{\theta} \sum_{t=1}^{N} \left| y(t) - \sum_{i=1}^{m} \theta_i \varphi_i\left(X(t)\right) \right|^p$$

的样本误差任意小, 但这样得到的模型在样本集以外往往会产生很大的预报误差, 这就是所谓过度拟合或过度训练问题

解决该问题的根本途经是同时极小化基函数的个数, 最终要解决连续和离散变量混合的优化问题:

$$\min \sum_{t=1}^{N} \left| y(t) - \sum_{i=1}^{m} \theta_i \sigma_i \varphi_i\left(X(t)\right) \right|^p + w \sum_{i=1}^{m} \sigma_i$$

$$\text{s.t.} \quad \sigma_i \in \{0,1\}, \ \forall 1 \le i \le m$$

其中 w 是设定的正的权值

基于该例子的数学规划模型，可得出如下认识

给定基函数及其数量 $\Rightarrow$ 线性规划

容易，可以找到全局最优解

数量给定优化基函数 $\Rightarrow$ 非线性规划

难，只能获得局部最优解

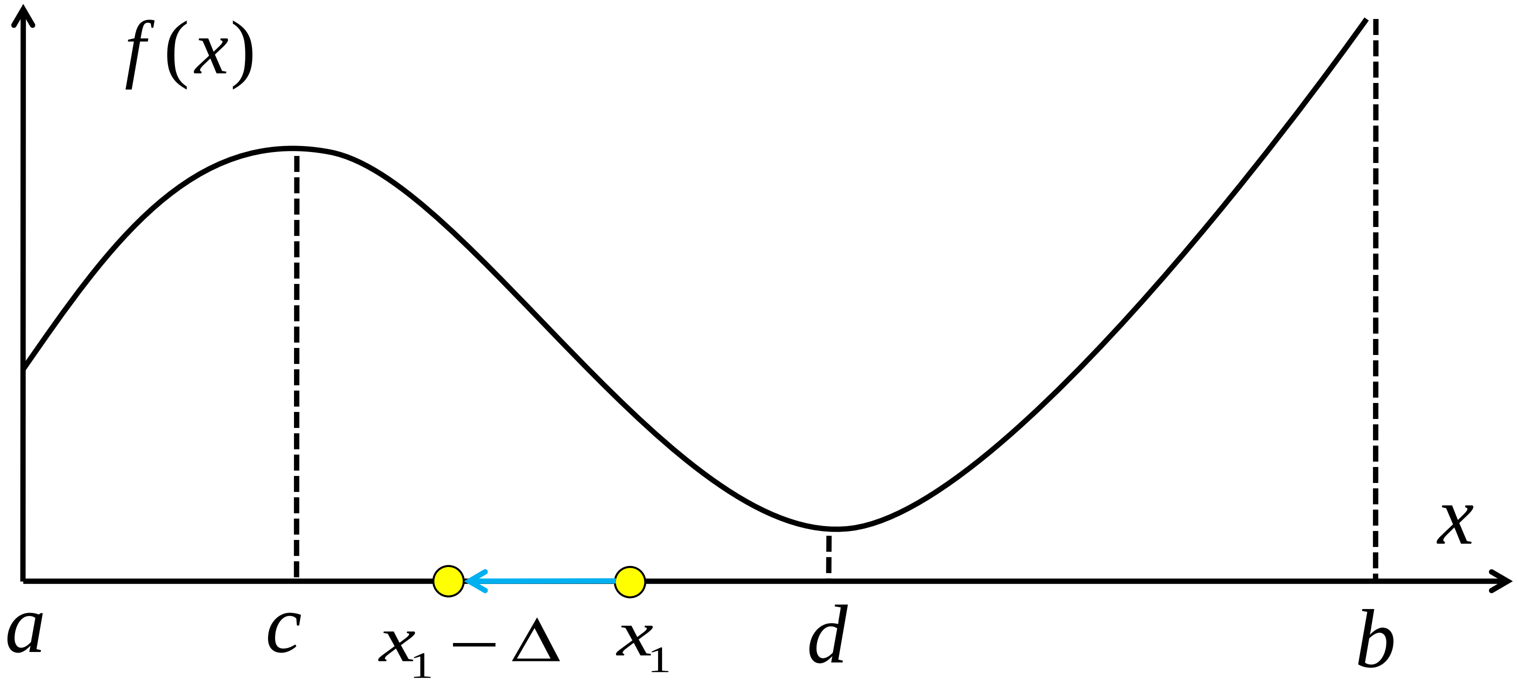
解决过拟合问题 $\Rightarrow$ 非线性规划

很难，只能获得较好的解

在解决实际问题时可参考以上认知选择合适模型

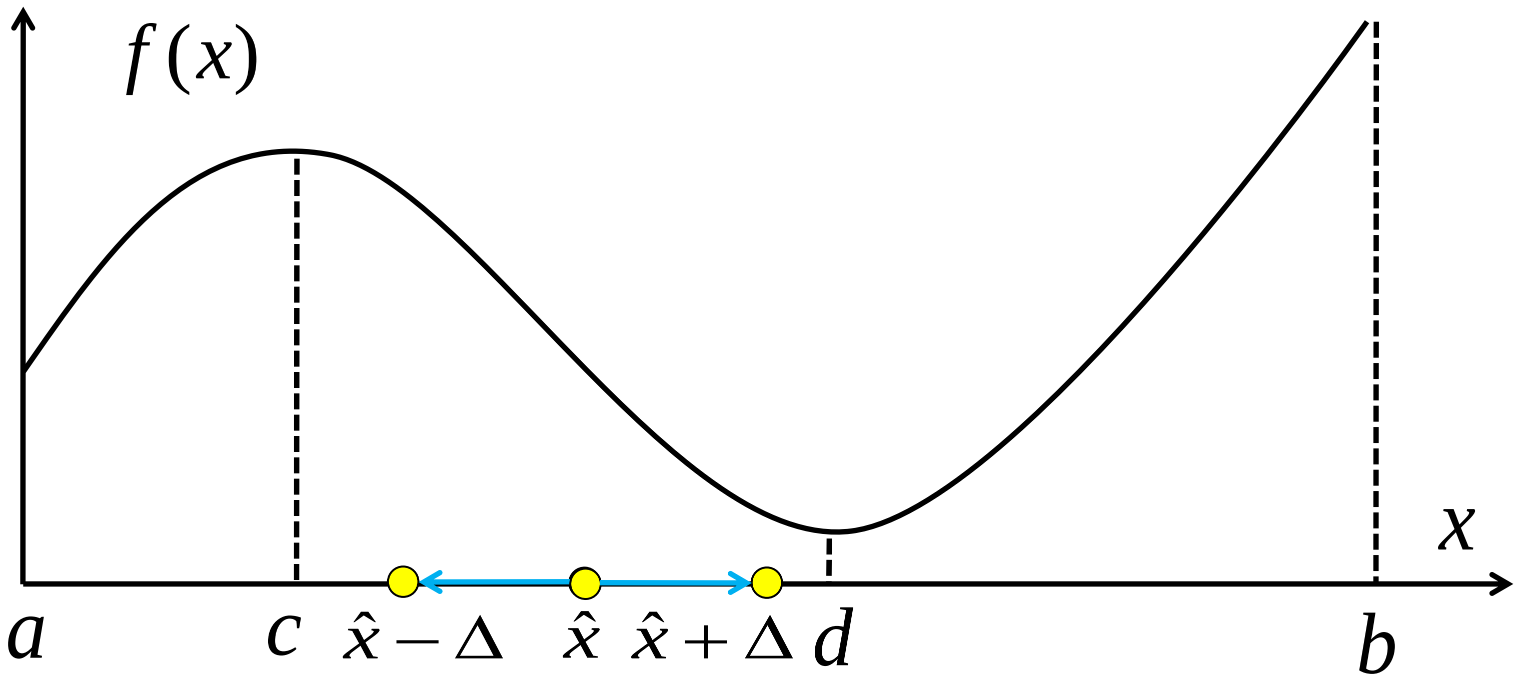
要点：优化方法简述

例 $\max_{a \le x \le b} f(x)$



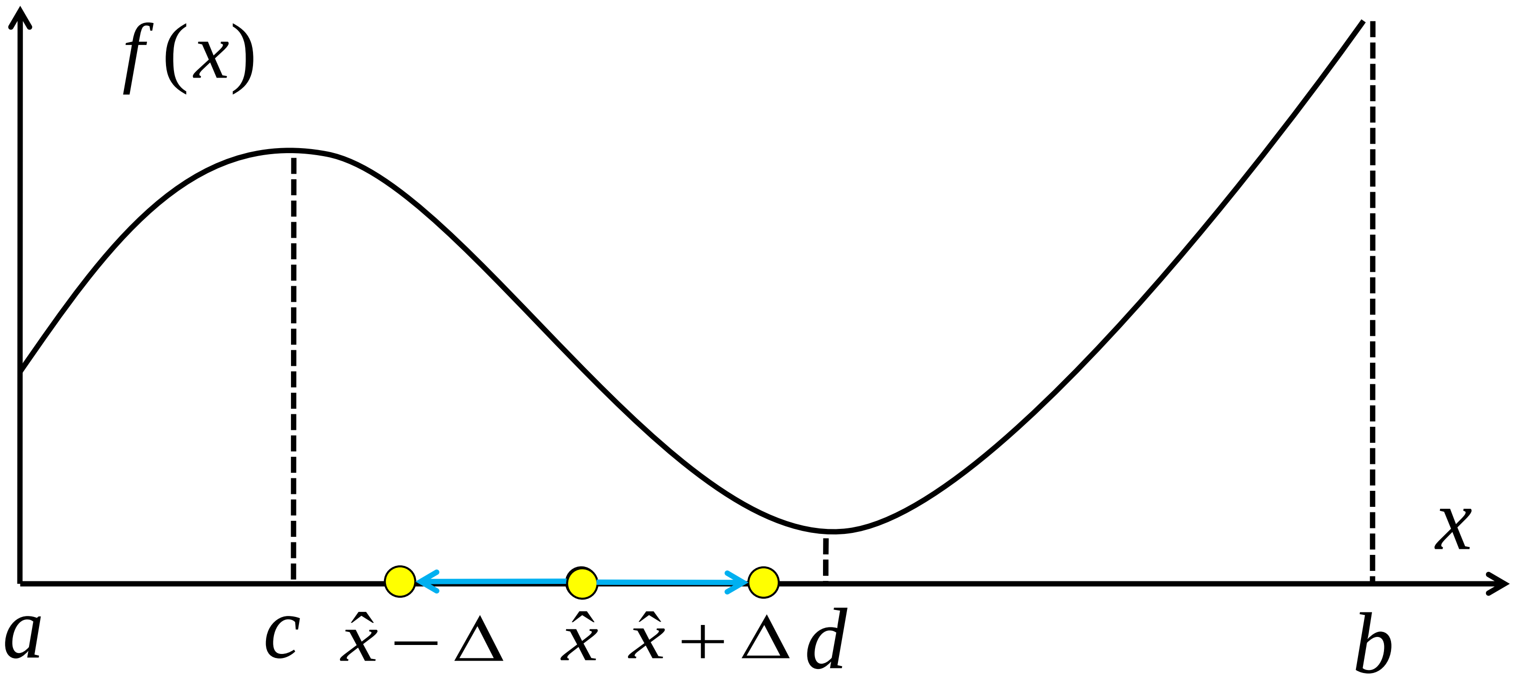
基本方法: 从 a,b 之间的任一点出发, 朝着能改进目标函数的方向搜索前进: $X_{k+1} = X_k + \lambda_k D_k$, 直至不能改进

肯定能够收敛到一个局部最优解, 不能保证全局最优



走出局部最优解的唯一途经是在搜索过程中允许前进到目标函数值变差的点, 例如在c,d之间容许目标函数下降才有可能找到全局最优解

由此产生新问题，无法保证算法收敛



为了使算法收敛，只能引入不确定性，让算法在任何一点以一定的概率前进到邻近的某点，移动概率和相应点的目标函数值正相关，所以图中例子：

$$p\left(\hat{x}+\Delta \mid \hat{x}\right) < p\left(\hat{x}-\Delta \mid \hat{x}\right)$$

由此产生的算法是结果不确定的算法

前面的例子包含了优化方法最基本的类型

确定型搜索　　与　　不确定型搜索

前者是经典的优化教材介绍的主要内容，后者包括模拟退火、禁忌搜索、遗传算法、免疫算法、蚂蚁算法等方法，一般统称为智能算法

本课程主要讨论基于确定性搜索的优化方法