西安电子科技大学
XIDIAN UNIVERSITY

IPIU
智能感知与图像理解

最优化理论
第二节：最优化数学建模

人工智能学院
智能感知与图像理解实验室

第二节：最优化数学建模

1 最优化问题数学模型

2 最优化问题的算法

3 最优化问题分类

4 二维最优化问题图解法

第二节：最优化问题数学建模

1. 最优化问题数学模型
2. 最优化问题的算法
3. 最优化问题分类
4. 二维最优化问题图解法

1. 最优化问题抽象数学模型

最优化理论与方法是近二十多年来发展十分迅速的一个数学分支。在数学上，最优化是一种求极值的方法。最优化已经广泛的渗透到工程、经济、电子技术等领域。

相关概念

最优化是从所有可能方案中选择最合理的一种以发到最优目标的学科。

最优方案是达到最优目标的方案。

最优化方法是搜寻最优方案的方法。

最优化理论是最优化方法的理论。

最优化问题：凡是追求最优目标的数学问题都属于最优化问题。


1. 最优化问题抽象数学模型

最优化方法

为了使系统达到最优的目标所提出的各种求解方法称为最优化方法。最优化方法是在第二次世界大战前后，在军事领域中对导弹、雷达控制的研究中逐渐发展起来的。

最优化方法解决问题一般步骤：

（1）提出需要进行最优化的问题，开始收集有关资料和数据；

（2）建立求解最优化问题的有关数学模型，确定变量，列出目标函数和有关约束条件；

（3）分析模型，选择合适的最优化方法；

（4）求解方程。一般通过编制程序在电子计算机上求得最优解；

（5）最优解的验证和实施。

随着系统科学的发展和各个领域的需求，最优化方法不断地应用于经济、自然、军事和社会研究的各个领域。


1. 最优化问题抽象数学模型

数学模型

数学模型就是对现实事物或问题的数学抽象或描述。

建立数学模型时要尽可能简单，而且要能完整地描述所研究的系统，具体建立怎样的数学模型需要丰富的经验和熟练的技巧。即使在建立了问题的数学模型之后，通常也必须对模型进行必要的数学简化以便于分析、计算。

食谱问题

我每天要求一定量的两种维生素，Vc和Vb。假设这些维生素可以分别从牛奶和鸡蛋中得到。

维生素	奶中含量	蛋中含量	每日需求
V*c*(mg)	2	4	40
V*b*(mg)	3	2	50
单价(US$)	3	2.5	

需要确定每天喝奶和吃蛋的量，

目标以便以最低可能的花费购买这些食物，

而满足最低限度的维生素需求量。



食谱问题数学模型

令x表示要买的奶的量，y为要买的蛋的量。食谱问题可以写成如下的数学形式：

Min $3x + 2.5y$	极小化目标函数
s.t. $2x + 4y \geq 40$	
$3x + 2y \geq 50$	可行区域（单纯形）
$x, y \geq 0.$	可行解



1. 最优化问题抽象数学模型

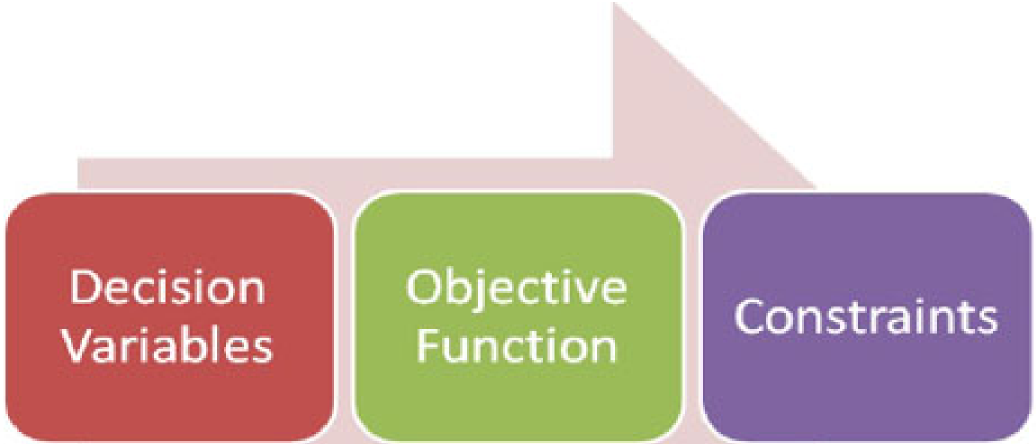
最优化数学模型

建立最优化问题数学模型的三要素：

（1）决策变量和参数x；

（2）目标函数$f(x)$；

（3）约束或限制条件Ω；



我要做的决策是什么？

我要达到的目标是什么？

我的决策有什么约束条件？

1. 最优化问题抽象数学模型

建立最优化问题数学模型的三要素：

（1）决策变量和参数。

决策变量是由数学模型的解确定的未知数。参数表示系统的控制变量，有确定性的也有随机性的。

（2）约束或限制条件。

由于现实系统的客观物质条件限制，模型必须包括把决策变量限制在它们可行值之内的约束条件，而这通常是用约束的数学函数形式来表示的。

（3）目标函数。

这是作为系统决策变量的一个数学函数来衡量系统的效率，即系统追求的目标。

1. 最优化问题抽象数学模型

最优化问题求解的两个关键步骤：

建模：将问题准确有效的通过数学形式表达 <——三个基本要素

求解：获得最优化目标函数的决策 <——优化算法

模型 → 实际问题 → 优化模型 → 解 → 应用

建模
- 决策（变量）
- 目标（函数）
- 约束（条件）

求解
- 优化算法


1. 最优化问题抽象数学模型

例1. 设每天需要混合饲料的批量为100磅，饲料中必须包含至少0.8%而不超过1.2%的钙，至少22%的蛋白质，至多5%的粗纤维。假设主要配料包括石灰石、谷物和大豆粉。以最低成本确定满足动物所需营养的最优混合饲料 。这些配料的主要营养成分如表所示：

配料	每磅配料中的营养含量			每磅成本（元）
	钙	蛋白质	纤维	
石灰石	0.380	0.00	0.00	0.0164
谷物	0.001	0.09	0.02	0.0463
大豆粉	0.002	0.50	0.08	0.1250



1. 最优化问题抽象数学模型

解：设x_1, x_2, x_3为生产100磅混合饲料所需的石灰石、谷物和大豆粉的量（磅）。该问题的数学模型如下

$$\min Z = 0.0164x_1 + 0.0463x_2 + 0.1250x_3$$

$$s.t. \quad x_1 + x_2 + x_3 = 100$$

$$0.380x_1 + 0.001x_2 + 0.002x_3 \le 0.012 \times 100$$

$$0.380x_1 + 0.001x_2 + 0.002x_3 \ge 0.008 \times 100$$

$$0.09x_2 + 0.50x_3 \ge 0.22 \times 100$$

$$0.02x_2 + 0.08x_3 \le 0.05 \times 100$$

$$x_1 \ge 0 \qquad x_2 \ge 0 \qquad x_3 \ge 0$$

极小化目标函数

等式约束

不等式约束

x_1, x_2, x_3 为决策变量



1. 最优化问题抽象数学模型

最优化问题数学模型的标准形式

$$
\begin{cases}
\min f(x) & & \longrightarrow \text{目标函数} \\
s.t. \quad h_i(x) = 0 & i = 1,2,\cdots,m & \longrightarrow \text{等式约束} \\
\qquad g_j(x) \geq 0 & j = 1,2,\cdots p & \longrightarrow \text{不等式约束}
\end{cases}
$$

其中 $x = [x_1, x_2, \ldots, x_n]^T \in R^n$ 为决策变量，满足所有约束的变量成为**可行解**或**可行点**，可行点的集合成为可行域*D*。

$$
D = \{x | h_i(x) = 0,\ i = 1,2,\cdots m, \quad g_j(x) \geq 0, \quad j = 1,2,\cdots p,\ x \in R^n\}
$$

若 $h_i(\mathrm{x}), \mathrm{g}_i(\mathrm{x})$ 是连续函数，则*D* 是闭集。



1. 最优化问题抽象数学模型

在可行域中找一点x^*，使目标函数$f(x)$在该点取最小值，即满足：$f(x^*)=\min f(x)$. $s.t.$ $g_j(x^*)\geq 0.$ $h_i(x)=0$ 的过程即为最优化的求解过程。

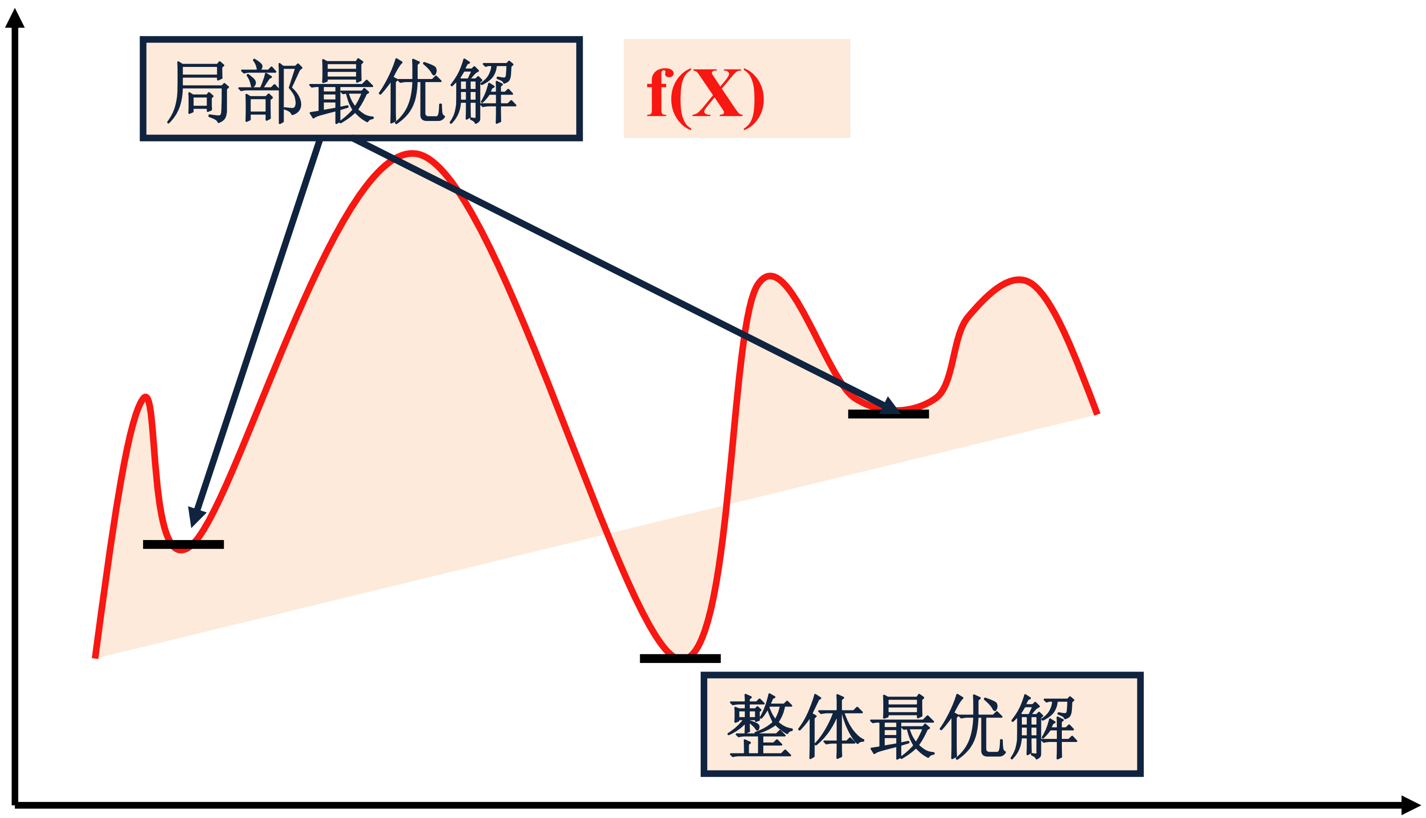
x^*称为问题的最优点或最优解，$f(x^*)$称为最优值。

定义1：整体（全局）最优解：若$x^*\in D$，对于一切$x\in D$，恒有$f(x^*)\leq f(x)$，则称x^*是最优化问题的整体最优解。

定义2：局部最优解：若$x^*\in D$，存在某邻域$N_\varepsilon(x^*)$，使得对于一切$x\in N_\varepsilon(x^*)\cap D$，恒有$f(x^*)\leq f(x)$ 则称x^*是最优化问题的局部最优解。其中$N_\varepsilon(x^*)=\{x \mid \|x-x^*\|<\varepsilon, \varepsilon>0\}$

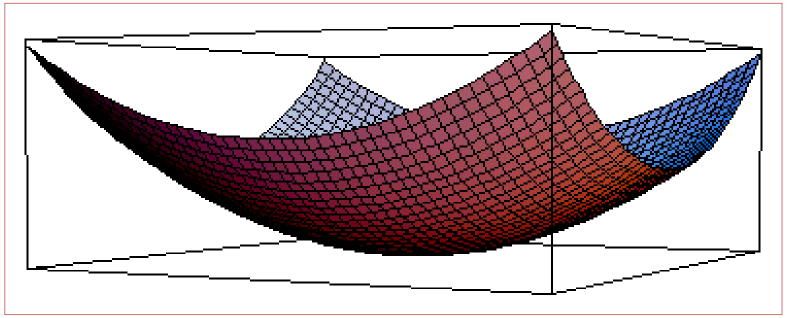
严格最优解：当$x\neq x^*$，有$f(x^*)<f(x)$，则称x^*为问题的严格最优解。

1. 最优化问题抽象数学模型



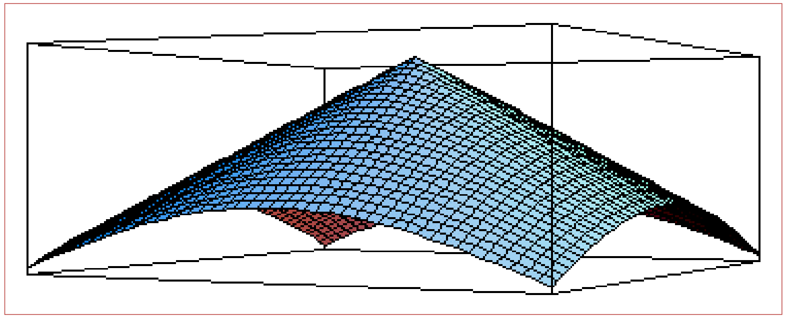
1. 最优化问题抽象数学模型

函数 $z=3x^2+4y^2$
在 (0,0) 处有极小值.



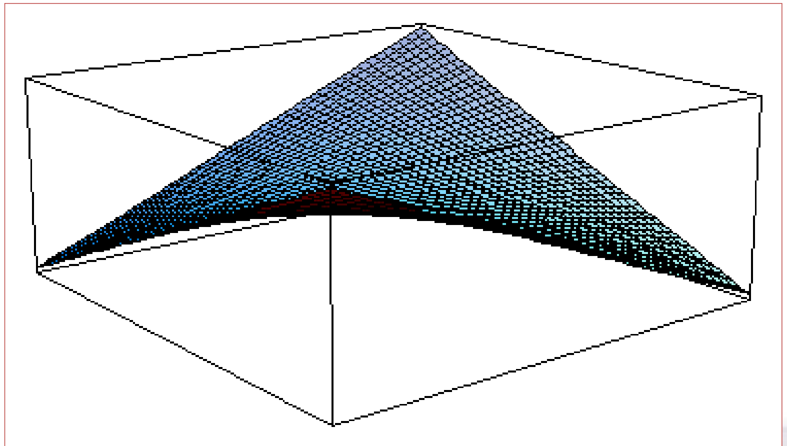
(1)

函数 $z=-\sqrt{x^2+y^2}$
在 (0,0) 处有极大值.



(2)

函数 $z=xy$ 在
(0,0) 处无极值.



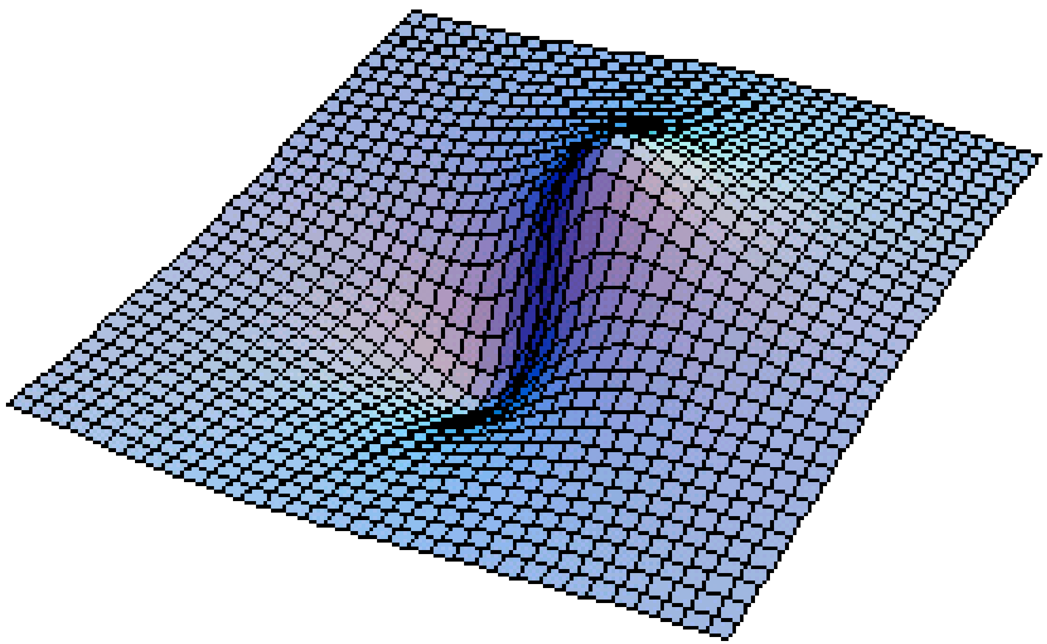
(3)

1. 最优化问题抽象数学模型

$$f(x,y)=\frac{x+y}{x^2+y^2+1}$$

极大值：$f(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})=\frac{1}{\sqrt{2}}$

极小值$f(-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})=-\frac{1}{\sqrt{2}}$



极大值问题与极小值问题的关系

对于求解极大值问题，我们也可将其转化为极小值问题求解：

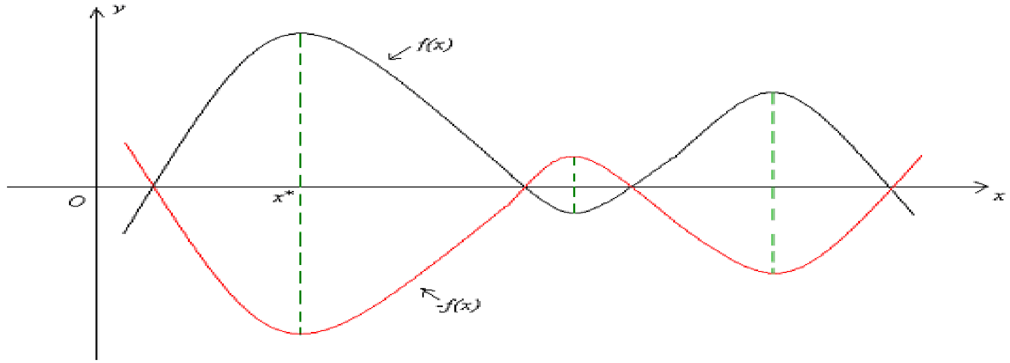
$$\max f(X) \quad \Leftrightarrow \quad -\min[-f(X)]$$

$$\begin{aligned} &\max f(\vec{x}) \\ &s.t.\ \ \vec{h}(\vec{x}) = \vec{0} \\ &\qquad \vec{s}(\vec{x}) \geq \vec{0} \end{aligned} \quad \Rightarrow \quad \begin{aligned} &\min -f(\vec{x}) \\ &s.t.\ \ \vec{h}(\vec{x}) = \vec{0} \\ &\qquad \vec{s}(\vec{x}) \geq \vec{0} \end{aligned}$$

$$\Downarrow$$

$$\vec{x}^* = \vec{x}^*$$

$$f = -\left(-f\left(\vec{x}^*\right)\right)$$



第二节：最优化问题数学建模

1 最优化问题数学模型

2 最优化问题的算法

3 最优化问题分类

4 二维最优化问题图解法

最优化问题的求解算法

（1） 公式解算法（精确解）

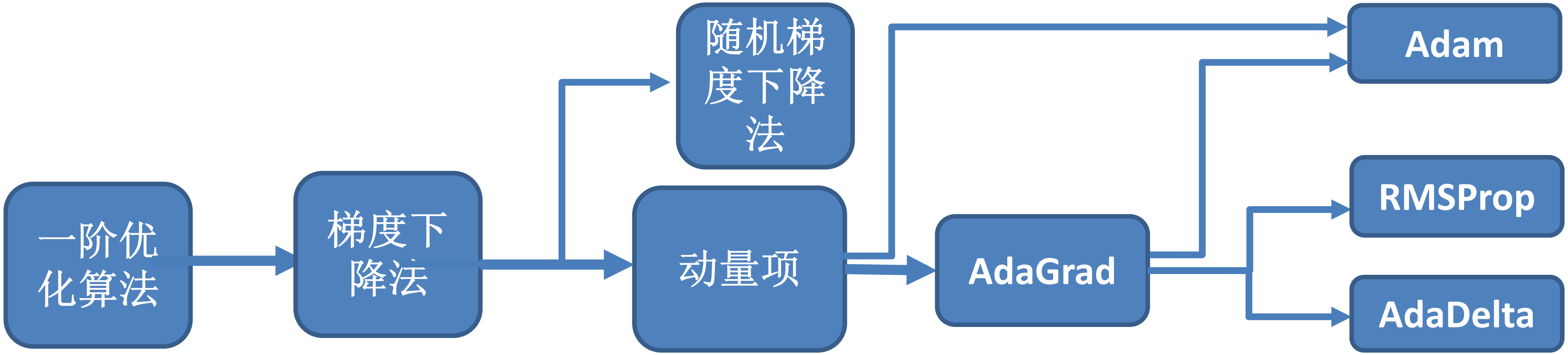
费马定理：对于一个可导函数，寻找其极值的统一做法是寻找导数为0的点；

拉格朗日乘数法：对于带等式约束的极值问题，该方法是经典的解决方案；

KKT条件： 拉格朗日乘数法的推广，用于求解既带有等式约束，又带有不等式约束的函数极值。

最优化问题的求解算法

（2） 数值优化算法——一阶优化算法（利用一阶导数信息即梯度求解）



最优化问题的求解算法

（2） 数值优化算法——一阶优化算法（利用一阶导数信息即梯度求解）

梯度下降法：优化思想是用当前位置负梯度方向作为搜索方向；

动量项：加速梯度下降法的收敛速度，减少震荡；

AdaGrad：相比于标准梯度下降法可以自动变更学习率；

RMSProp: 缓解了AdaGrad的学习率下降较快的问题；

AdaDelta: 避免了长期累积梯度值所导致的学习率趋于0的问题；

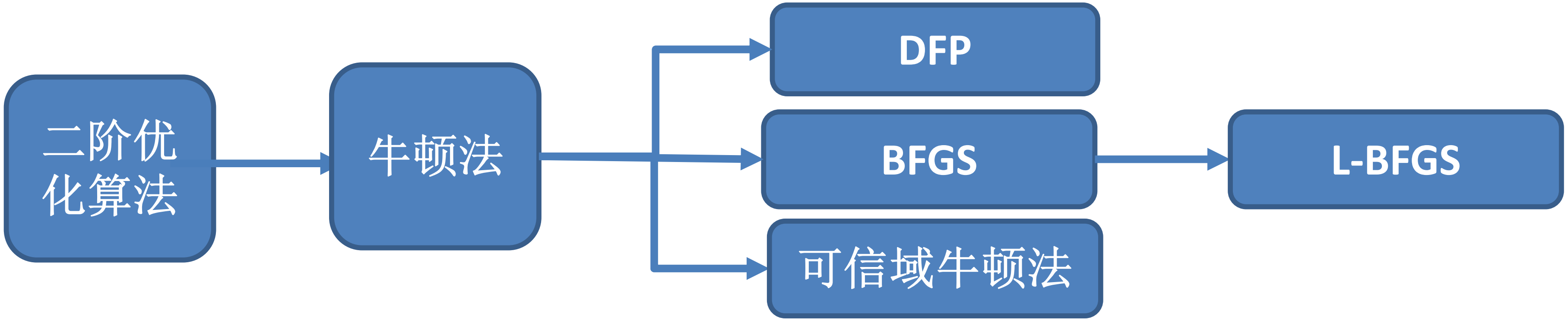
Adam: 整合了自适应学习率v（累积了梯度的平方和）与动量项m;

随机梯度下降法： 每个数据都计算算一下损失函数，然后求梯度更新参数，计算速度快。



最优化问题的求解算法

（2） 数值优化算法—二阶优化算法（利用二阶导数信息求解）



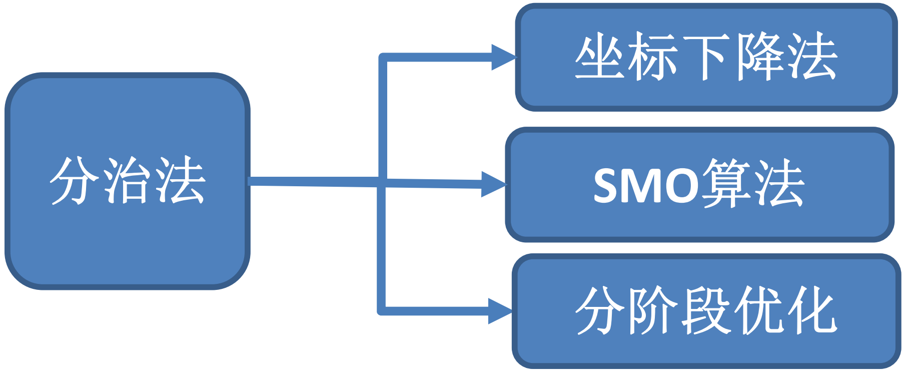
牛顿法：利用一阶和二阶导数的无约束目标最优化方法；

可信域牛顿法：调整牛顿方向的步长来实现收敛到最优解和序列递减；

拟牛顿法（DFP，DFGS，LBFGS）：改善牛顿法每次需要求解复杂的Hessian矩阵的逆矩阵的缺陷，它使用正定矩阵来近似Hessian矩阵的逆，从而简化了运算的复杂度。

最优化问题的求解算法

（3）分治法（一种算法设计思想，它将一个大的问题分解成子问题进行求解）



坐标下降法： 每次对一个变量进行优化；

SMO算法： 用于求解支持向量机的对偶问题；

分阶段优化： 在每次迭代时，先固定住优化变量x一部分分量a不动，对另外一部分变量b进行优化；然后再固定住b不动，对b进行优化。如此反复，直至收敛到最优解处。代表算法AdaBoost。

最优化问题的求解算法

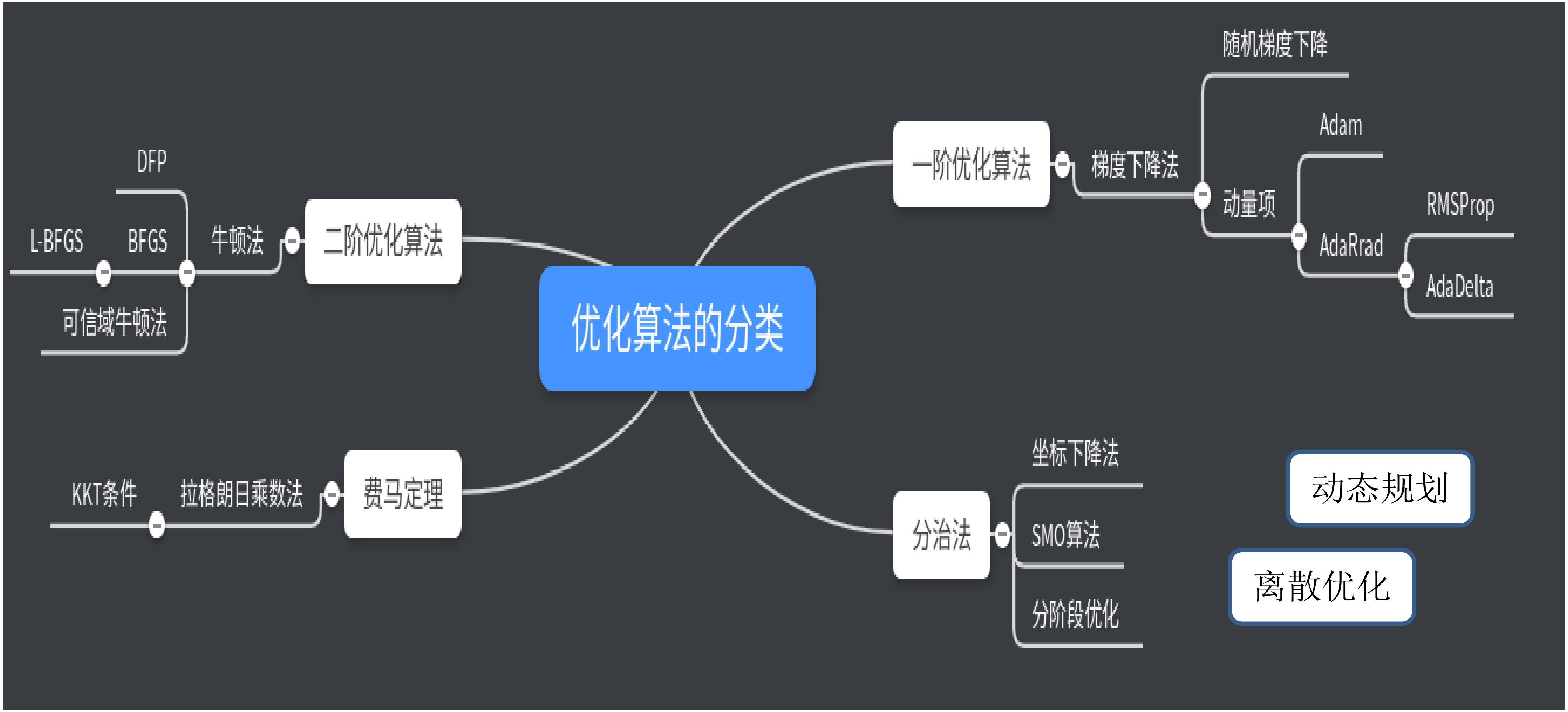
（4）动态规划

它将一个问题分解成子问题求解，如果整个问题的某个解是最优的，则这个解的任意一部分也是子问题的最优解。这样通过求解子问题，得到最优解，逐步扩展，最后得到整个问题的最优解。

（5）离散优化

变量只能取一些离散的值，不是连续函数。

最优化模型的求解算法分类



第二节：最优化问题概论

1. 最优化问题数学模型
2. 最优化问题的算法
3. 最优化问题分类
4. 二维最优化问题图解法

最优化问题分类

- 最优化问题
  - 静态问题
    - 无约束问题
      - 线性问题
      - 非线性问题
        - 一维问题
        - n维问题
    - 约束问题
      - 线性规划
      - 非线性规划
  - 动态问题
    - 无约束
    - 约束

3. 最优化问题分类

最优化模型分类

最优化模型分类方法有很多，可按变量、约束条件、目标函数个数、目标函数和约束条件的是否线性是否依赖时间等分类。

根据目标函数，约束条件的特点将最优化模型包含的主要内容大致如下划分：

线性规划　　整数规划

非线性规划　　多目标规划

动态规划　　对策论



第二节：最优化问题概论

1 最优化问题数学模型

2 最优化问题的算法

3 最优化问题分类

4 二维最优化问题图解法

最优化问题图解法

➢确定可行域：

画约束直线，确定满足约束条件的半平面，所有半平面的交集，即为线性规划的可行域。

➢确定目标函数的等值线及优化方向：

画一条目标函数等值线，并确定目标函数优化的方向。

➢平行移动目标函数等值线，通过观察得到线性规划的最优解。

图解法的步骤

最优化问题图解法例题

例1. 用图解法求解线性规划问题

$$\max\ z = 2x_1 + x_2$$

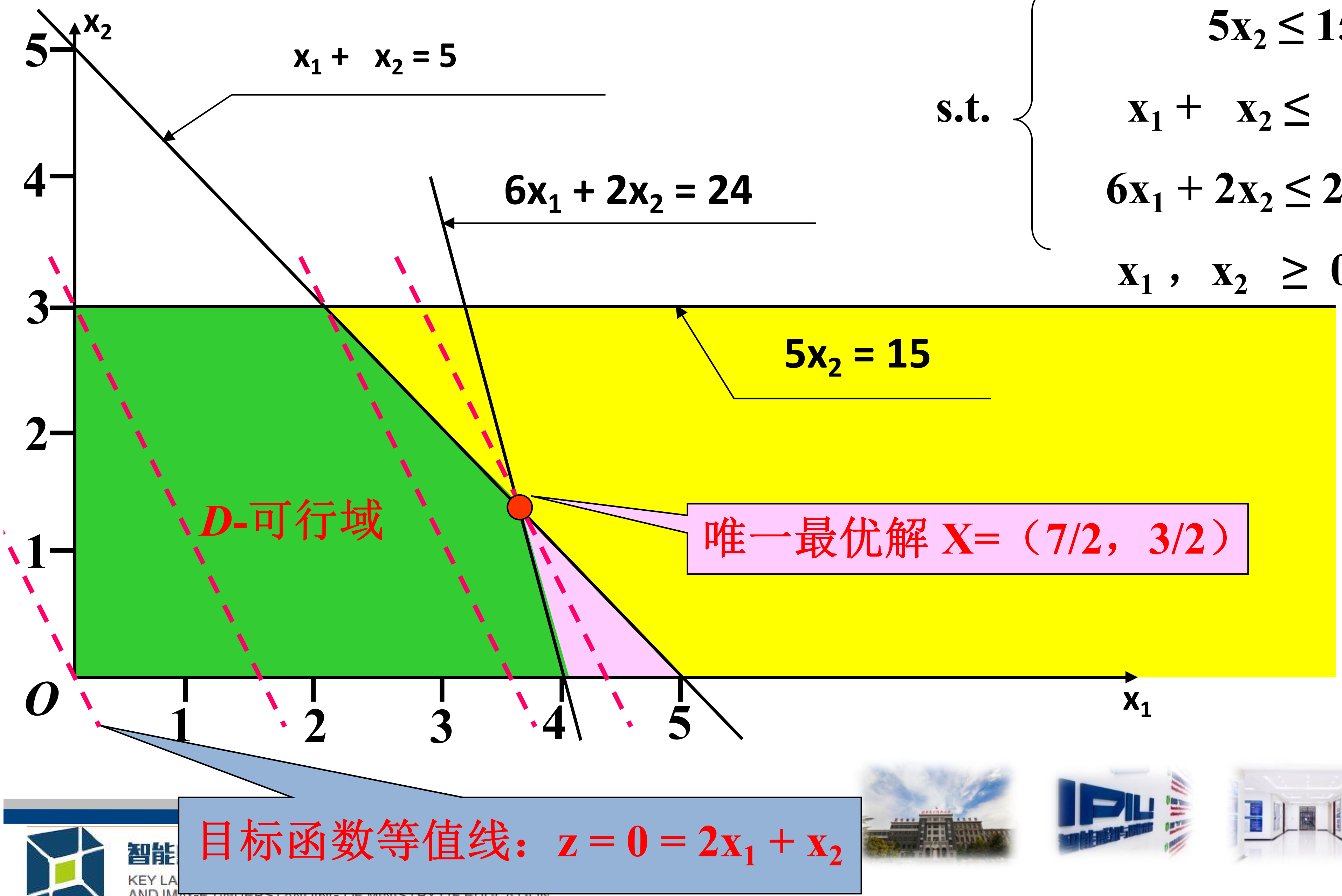
$$\text{s.t.}\ \begin{cases} 5x_2 \le 15 \\ x_1 + x_2 \le 5 \\ 6x_1 + 2x_2 \le 24 \\ x_1,\ x_2 \ge 0 \end{cases}$$

最优化问题图解法例题

$$\max \quad z = 2x_1 + x_2$$

$$\text{s.t.} \begin{cases} 5x_2 \le 15 \\ x_1 + x_2 \le 5 \\ 6x_1 + 2x_2 \le 24 \end{cases}$$

$$x_1,\ x_2 \ge 0$$



例2.用图解法求解线性规划问题

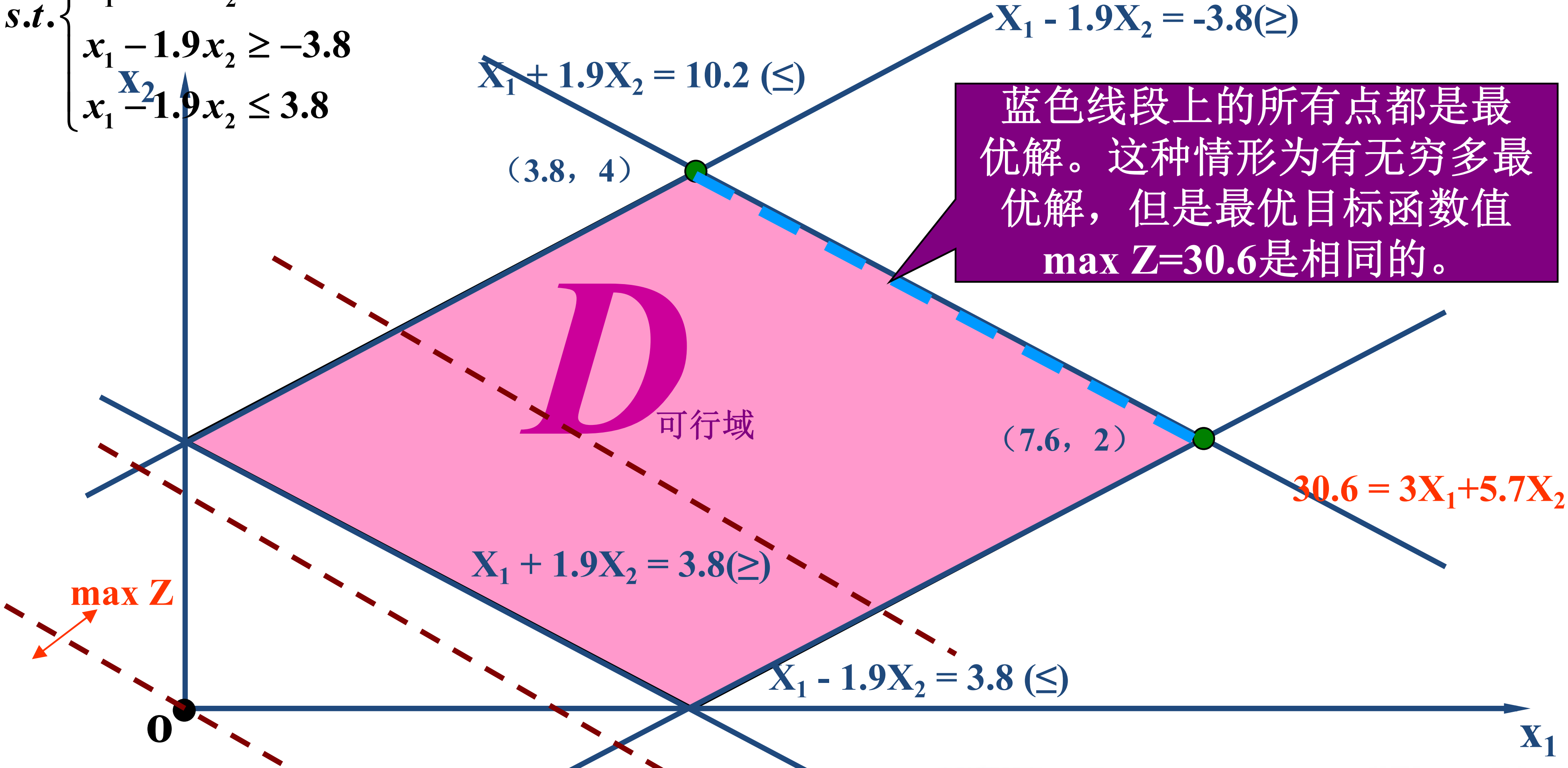
$$\max \quad z = 3x_1 + 5.7x_2$$

$$s.t.\begin{cases} x_1 + 1.9x_2 \le 10.2 \\ x_1 + 1.9x_2 \ge 3.8 \\ x_1 - 1.9x_2 \ge -3.8 \\ x_1 - 1.9x_2 \le 3.8 \end{cases}$$

最优化问题图解法例题

$$\max\ z = 3x_1 + 5.7x_2$$

$$s.t.\begin{cases} x_1 + 1.9x_2 \le 10.2 \\ x_1 + 1.9x_2 \ge 3.8 \\ x_1 - 1.9x_2 \ge -3.8 \\ x_1 - 1.9x_2 \le 3.8 \end{cases}$$

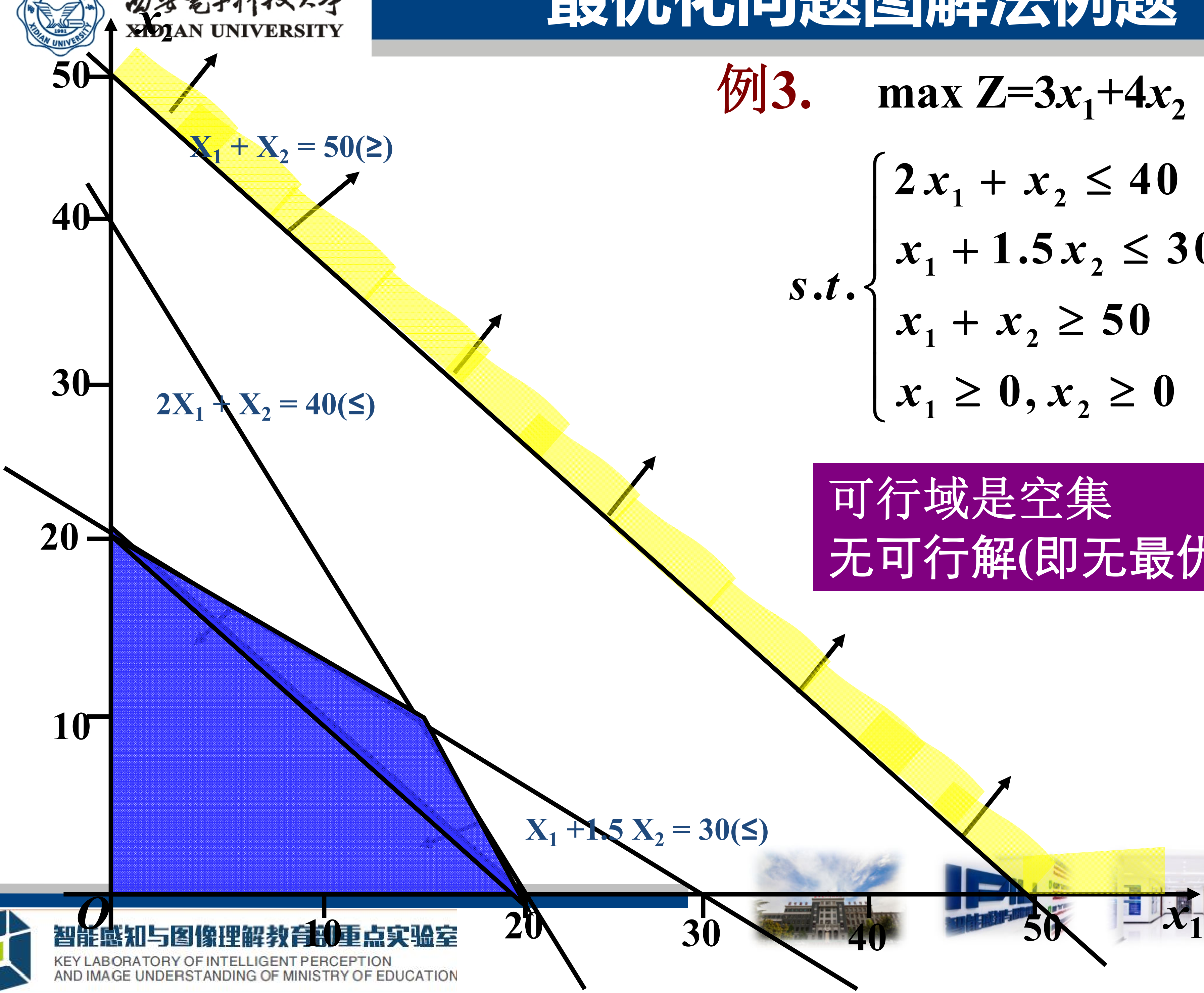




最优化问题图解法例题

例3. $\max Z=3x_1+4x_2$

$$
s.t.\begin{cases} 2x_1 + x_2 \le 40 \\ x_1 + 1.5x_2 \le 30 \\ x_1 + x_2 \ge 50 \\ x_1 \ge 0, x_2 \ge 0 \end{cases}
$$



可行域是空集
无可行解(即无最优解)

4. 二维最优化问题图解法

4.1 二维问题图解法

对于只有两个变量的线性规划问题，可以采用在平面上作图的方法求解，称为图解法。

图解法中解的概念

- **无穷多解**: 图解法中，此情况出现在目标函数等值直线向优化方向平移时，最后与可行域边界的一条边重合，此时，除该直线段的两个端点（即可行域的两个顶点）外，直线段上所有点的目标函数值都达到最优。
- **无界解**：图解法中，此情况出现在可行域为无界区域，且目标函数等值直线向优化方向平移时，始终无法脱离可行域。发生这种情况往往是建模时遗漏了某些约束条件所至。
- **无解**：当可行域为空集时，问题不存在可行解。发生此情况是因为模型中出现了相互矛盾的约束条件。


最优化问题图解法中出现情况

图解法

图解法的启示

1) **线性规划问题解的可能情况**

a.唯一最优解

b.无穷多最优解

c.无解（没有有界最优解，无可行解）

2) **若线性规划问题的可行域非空，则可行域是一个凸集；**

3) **若线性规划问题的最优解存在，则一定可以在可行域的凸集的某个顶点上达到。**

图解法的优缺点

优缺点

- 简单、直观、便于初学者理解和记忆；
- 仅适用于低维情况，通常适用于含两个或三个变量的情况。

对于高维情况，怎么求解呢？ ----- 单纯形法

4.2 极小点的定义

下面针对多元函数的情形给出各类极小点的定义。

定义1.1 对于任意给定的实数$\delta > 0$，满足不等式 $\|X - X_0\| < \delta$的X的集合称为点X_0的邻域，记为

$$\mathrm{N}(X_0, \delta) = \{\mathrm{X} |\ \|X - X_0\| < \delta, \delta > 0\}$$

定义1.2 设$\mathrm{f}: \mathrm{D} \subseteq \boldsymbol{R}^n \rightarrow \boldsymbol{R}^1$,若存在点$X^* \epsilon D$和数$\delta > 0$，$\forall \mathrm{X} \in \mathrm{N}(X^*, \delta) \cap \mathrm{D}$都有$\mathrm{f}(X^*) \leq \mathrm{f}(\mathrm{X})$,则称$X^*$为$\mathrm{f}(\mathrm{X})$的局部极小点（非严格的）。

定义1.3设$\mathrm{f}: \mathrm{D} \subseteq \boldsymbol{R}^n \rightarrow \boldsymbol{R}^1$,若存在点$X^* \epsilon D$和数$\delta > 0$，$\forall \mathrm{X} \in \mathrm{N}(X^*, \delta) \cap \mathrm{D}$但$\mathrm{X} \neq X^*$，都有$\mathrm{f}(X^*) < f(X)$，则称$X^*$为$\mathrm{f}(\mathrm{X})$的严格局部极小点。



4.2 极小点的定义

定义1.4设$f: D \subseteq \boldsymbol{R}^n \to \boldsymbol{R}^1$,若存在点$X^* \epsilon D$，$\forall X \in D$都有$f(X^*) \leq f(X)$，则称$X^*$为f(X)的全局极小点（非严格的）。

定义1.5设$f: D \subseteq \boldsymbol{R}^n \to \boldsymbol{R}^1$,若存在点$X^* \epsilon D$，$\forall X \in D$但$X \neq X^*$，都有$f(X^*) < f(X)$，则称$X^*$为f(X)的严格全局极小点。

由以上定义看到，X^*是局部极小点，是指在以X^*为中心的一个邻域中$f(X)$在点X^*处取得最小的值；而X^*是全局极小点，是指在定义域D中$f(X)$在点X^*处取得最小的值。全局极小点可能在某个局部极小点处取得，也可能在D的边界上取得。实际问题通常是求全局极小点，但是直到目前为止，最优化中绝大多数方法都是求局部极小点的，解决这一矛盾的一种方法是先求出所有的局部极小点，再求全局极小点。

课后练习 Tutorial #2

图解法练习：

(1) $\min f(X) = -x_1 + x_2,$

$$s.t.\ \begin{cases} 3x_1 - 7x_2 \geq 8 \\ x_1 - x_2 \leq 5 \\ \quad x_1, x_2 \geq 0. \end{cases}$$

(2) $\min f(X) = 13x_1 + 5x_2,$

$$s.t.\ \begin{cases} 7x_1 + 3x_2 \geq 19 \\ 10x_1 + 2x_2 \leq 11 \\ \quad x_1, x_2 \geq 0. \end{cases}$$

西安電子科技大學
XIDIAN UNIVERSITY

IPIU
智能感知与图像理解

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education

THE END

Thanks for your participation!