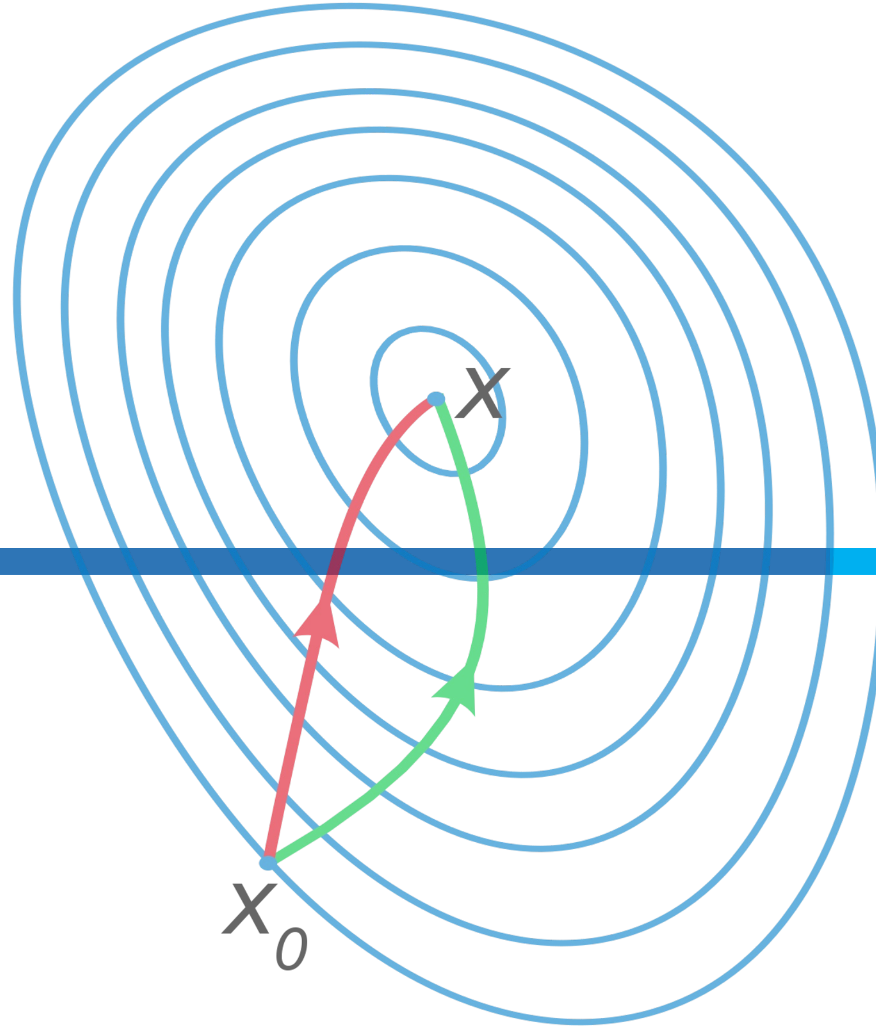
最优化方法

第二周



计算机学院

余皓然

2024/3/4

本周内容

Part1 如何从复杂度的角度衡量算法好坏？
算法复杂度、P问题与NP问题与NP难问题
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； **线性规划问题、单纯形法、内点法；** **有约束凸优化问题、KKT条件、内点法**
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
局部搜索、模拟退火、遗传算法、差分进化算法、蚁群算法、粒子群优化算法、人工蜂群算法
Part4 针对较复杂的多目标最优化问题，如何找到较好解？
多目标优化问题、NSGA-II算法

梯度下降法

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； **线性规划问题、单纯形法、内点法；** **有约束凸优化问题、KKT条件、内点法**
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

下降法

$\min \quad f(\boldsymbol{x})$

$\text{var.} \quad \boldsymbol{x} \in \mathcal{R}^n.$

其中，$f(\boldsymbol{x})$是二阶连续可微的凸函数

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t = 0$

1 确定下降方向$\boldsymbol{d}^t$ 向量

2 确定步长$\alpha^t > 0$ 标量

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若满足停止条件则结束；否则，$t \leftarrow t + 1$且返回1

通常采用?

下降法

$\min \quad f(\boldsymbol{x})$
var. $\quad \boldsymbol{x} \in \mathcal{R}^n.$
其中，$f(\boldsymbol{x})$是二阶连续可微的凸函数

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t = 0$
1 确定下降方向$\boldsymbol{d}^t$
2 确定步长$\alpha^t > 0$
3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$
4 若满足停止条件则结束；否则，$t \leftarrow t + 1$且返回1

通常采用：$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于某预设阈值

回顾：全局最优解的充要条件是$\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$

下降法

下降法（一般形式）

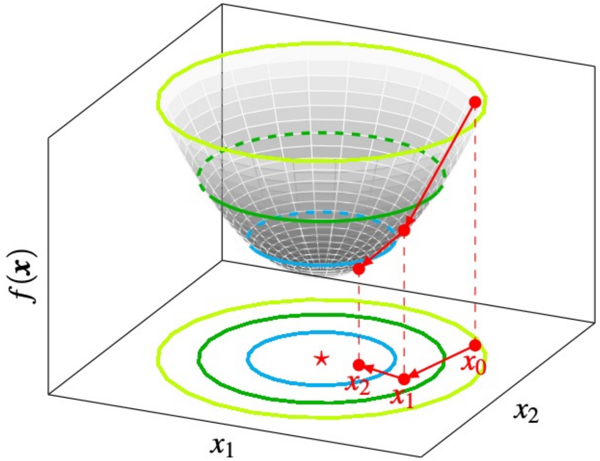
0 初始化当前解$\boldsymbol{x}^0$及$t = 0$

1 确定下降方向$\boldsymbol{d}^t$

2 确定步长$\alpha^t > 0$

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t + 1$且返回1



二维变量时解搜索路径示例　　示图取自Bo Jiang <CS257 Linear and Convex Optimization>

下降法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boldsymbol{d}^t$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

为确保$\boldsymbol{d}^t$是使函数值下降的方向，$\boldsymbol{d}^t$应满足：

$$\left(\nabla f(\boldsymbol{x}^t)\right)^T\boldsymbol{d}^t<0$$

一维时可化简为：$f'(x^t)\,d^t<0$，即d^t与$f'(x^t)$异号

下降法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boldsymbol{d}^t$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

为确保$\boldsymbol{d}^t$是使函数值下降的方向，$\boldsymbol{d}^t$应满足：

$$\left(\nabla f(\boldsymbol{x}^t)\right)^T\boldsymbol{d}^t<0$$

利用凸函数的一阶条件可证明若$f(\boldsymbol{x}^{t+1})<f(\boldsymbol{x}^t)$，则上述条件必须成立

回顾

若$f(\boldsymbol{x})$是定义在凸集C上的一阶可微函数，那么$f(\boldsymbol{x})$是凸函数的充要条件是：对任意$\boldsymbol{x},\boldsymbol{y}\in C$，有$f(\boldsymbol{y})\geq f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})$。

下降法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boldsymbol{d}^t$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

为确保$\boldsymbol{d}^t$是使函数值下降的方向，$\boldsymbol{d}^t$应满足：

$$\left(\nabla f(\boldsymbol{x}^t)\right)^T\boldsymbol{d}^t<0$$

梯度下降法：选择 $\boldsymbol{d}^t=-\nabla f(\boldsymbol{x}^t)$（在$\nabla f(\boldsymbol{x}^t)\neq\boldsymbol{0}$时，满足不等式）

梯度下降法

梯度下降法

0 初始化当前解$\boldsymbol{x}^0$及$t = 0$

1 确定下降方向$\boxed{\boldsymbol{d}^t = -\nabla f(\boldsymbol{x}^t)}$

2 确定步长$\alpha^t > 0$

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t+1$且返回1

函数的泰勒展开式（一维）

Taylor's theorem[4][5][6] — Let $k \geq 1$ be an integer and let the function $f : \mathbf{R} \to \mathbf{R}$ be k times differentiable at the point $a \in \mathbf{R}$. Then there exists a function $h_k : \mathbf{R} \to \mathbf{R}$ such that

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + h_k(x)(x-a)^k,$$

and

$$\lim_{x \to a} h_k(x) = 0.$$

This is called the **Peano form of the remainder**.

$f(x)$在 $x = a$附近的取值可利用一阶导信息近似：

$$f(x) \approx f(a) + (x-a)f'(a)$$

梯度下降法

梯度下降法

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boxed{\boldsymbol{d}^t = -\nabla f(\boldsymbol{x}^t)}$

2 确定步长$\alpha^t > 0$

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t+1$且返回1

函数的泰勒展开式（一维）

$f(x)$在 $x = a$附近的取值可利用一阶导信息近似：

$$f(x) \approx f(a) + (x-a)f'(a)$$

函数的泰勒展开式（多维）

$f(\boldsymbol{x})$在 $\boldsymbol{x} = \boldsymbol{a}$附近的取值可利用梯度信息近似：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T (\boldsymbol{x} - \boldsymbol{a})$$

梯度下降法

$f(\boldsymbol{x})$在 $\boldsymbol{x}=\boldsymbol{a}$附近的取值可利用梯度信息近似：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a})+\left(\nabla f(\boldsymbol{a})\right)^T(\boldsymbol{x}-\boldsymbol{a})$$

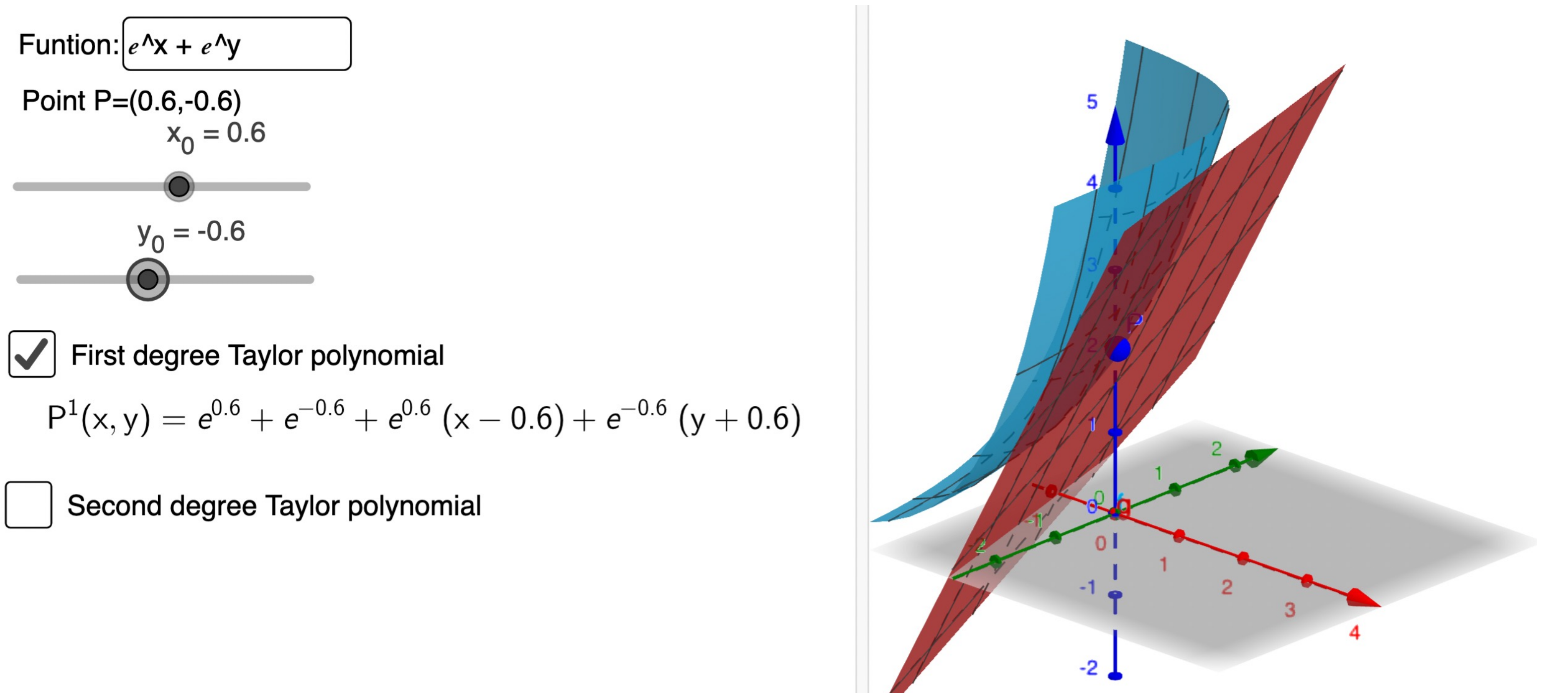
例：对$f(\boldsymbol{x})=e^{x_1}+e^{x_2}$，$\nabla f(\boldsymbol{x})=(e^{x_1},e^{x_2})^T$，考查在$\boldsymbol{x}=(0.6,-0.6)$的展开

梯度下降法

$f(\boldsymbol{x})$在 $\boldsymbol{x}=\boldsymbol{a}$附近的取值可利用梯度信息近似：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T (\boldsymbol{x}-\boldsymbol{a})$$

例：对$f(\boldsymbol{x}) = e^{x_1} + e^{x_2}$，$\nabla f(\boldsymbol{x}) = (e^{x_1}, e^{x_2})^T$，考查在$\boldsymbol{x} = (0.6, -0.6)$的展开



示图取自https://www.geogebra.org/m/Ehnz3hGb

梯度下降法

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T (\boldsymbol{x} - \boldsymbol{a})$$

当α^t很小时，$\boldsymbol{x}^{t+1}$在$\boldsymbol{x}^t$附近，利用$\nabla f(\boldsymbol{x}^t)$对$f(\boldsymbol{x}^{t+1})$作泰勒近似：

$$f(\boldsymbol{x}^{t+1}) \approx f(\boldsymbol{x}^t) + \alpha^t \left(\nabla f(\boldsymbol{x}^t)\right)^T \boldsymbol{d}^t$$

假设$\|\boldsymbol{d}^t\|_2 = 1$，如何选择$\boldsymbol{d}^t$可以最小化$f(\boldsymbol{x}^{t+1})$？

梯度下降法

0 初始化当前解$\boldsymbol{x}^0$及$t = 0$

1 确定下降方向$\boxed{\boldsymbol{d}^t = -\nabla f(\boldsymbol{x}^t)}$

2 确定步长$\alpha^t > 0$

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t + 1$且返回1

梯度下降法

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T (\boldsymbol{x} - \boldsymbol{a})$$

梯度下降法

0 初始化当前解$\boldsymbol{x}^0$及$t = 0$

1 确定下降方向$\boxed{\boldsymbol{d}^t = -\nabla f(\boldsymbol{x}^t)}$

2 确定步长$\alpha^t > 0$

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t + 1$且返回1

当α^t很小时，$\boldsymbol{x}^{t+1}$在$\boldsymbol{x}^t$附近，利用$\nabla f(\boldsymbol{x}^t)$对$f(\boldsymbol{x}^{t+1})$作泰勒近似：

$$f(\boldsymbol{x}^{t+1}) \approx f(\boldsymbol{x}^t) + \alpha^t \left(\nabla f(\boldsymbol{x}^t)\right)^T \boldsymbol{d}^t$$

假设$\|\boldsymbol{d}^t\|_2 = 1$，如何选择$\boldsymbol{d}^t$可以最小化$f(\boldsymbol{x}^{t+1})$？

$$\left(\nabla f(\boldsymbol{x}^t)\right)^T \boldsymbol{d}^t = \|\nabla f(\boldsymbol{x}^t)\|_2 \|\boldsymbol{d}^t\|_2 \cos(\text{夹角})$$

向量内积公式

若$\|\boldsymbol{d}^t\|_2 = 1$，$\boldsymbol{d}^t$需要取$\dfrac{-\nabla f(\boldsymbol{x}^t)}{\|\nabla f(\boldsymbol{x}^t)\|_2}$

梯度下降法

梯度下降法的收敛性

设$f(\boldsymbol{x})$是一阶连续可微凸函数且梯度满足：

对所有$\boldsymbol{x}, \boldsymbol{y} \in \mathcal{R}^n$有：$\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|_2 \leq L\|\boldsymbol{x} - \boldsymbol{y}\|_2$。

利普希茨连续
（Lipschitz continuity）

若$\boldsymbol{x}^*$是最小化$f(\boldsymbol{x})$的全局最优解，当$\alpha^t = \alpha \in \left(0, \frac{1}{L}\right]$，在梯度下降法下有：

$$f(\boldsymbol{x}^t) - f(\boldsymbol{x}^*) \leq \frac{1}{2t\alpha}\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2^2 \text{ 。}$$

梯度下降法

梯度下降法的收敛性

设$f(\boldsymbol{x})$是一阶连续可微凸函数且梯度满足：

对所有有$\boldsymbol{x}, \boldsymbol{y} \in \mathcal{R}^n$有：$\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|_2 \leq L\|\boldsymbol{x} - \boldsymbol{y}\|_2$。

利普希茨连续
（Lipschitz continuity）

若$\boldsymbol{x}^*$是最小化$f(\boldsymbol{x})$的全局最优解，当$\alpha^t = \alpha \in \left(0, \frac{1}{L}\right]$，在梯度下降法下有：

$$f(\boldsymbol{x}^t) - f(\boldsymbol{x}^*) \leq \frac{1}{2t\alpha}\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2^2 \text{。}$$

初始解$\boldsymbol{x}^0$的选择很重要

收敛速度慢：速度是$O\left(\frac{1}{t}\right)$

若要求$f(\boldsymbol{x}^t) - f(\boldsymbol{x}^*) \leq 10^{-8}$，$t$应大于等于$\frac{10^8}{2\alpha}\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2^2$

梯度下降法

线性回归（L2正则化）	
min	$\frac{1}{M}\|\boldsymbol{Aw}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{w}\|_2^2$
var.	$\boldsymbol{w}\in\mathcal{R}^N.$

考虑$\boldsymbol{w}=(w_1,w_2)^T$的简化情况

（不考虑特征归一化和常数项w_0）

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
...	...	...
4.1149	-2.5394	25.1117

人工根据$b=2a_1+5a_2+\mathcal{N}(0,20)$ 生成30条数据

梯度下降法

人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|\boldsymbol{Aw}-\boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w} \in \mathcal{R}^N.$$

```
M = 30;
vectorw = [2; 5];
MatrixA = normrnd(0,10,[M,2]);
vectorb = MatrixA * vectorw + normrnd(0,20,[M,1]);

lambda = 10;
resolution = 50;
x = linspace(-5,5,resolution);
y = linspace(-10,10,resolution);
for i = 1:resolution
for j = 1:resolution
z(i,j) = 1/M * (MatrixA * [x(j), y(i)]' - vectorb)' * (MatrixA * [x(j), y(i)]' - vectorb) + (x(j)^2 + y(i)^2) * lambda;
end
end
surf(x,y,z)
figure(2)
contour(x,y,z,50)
```

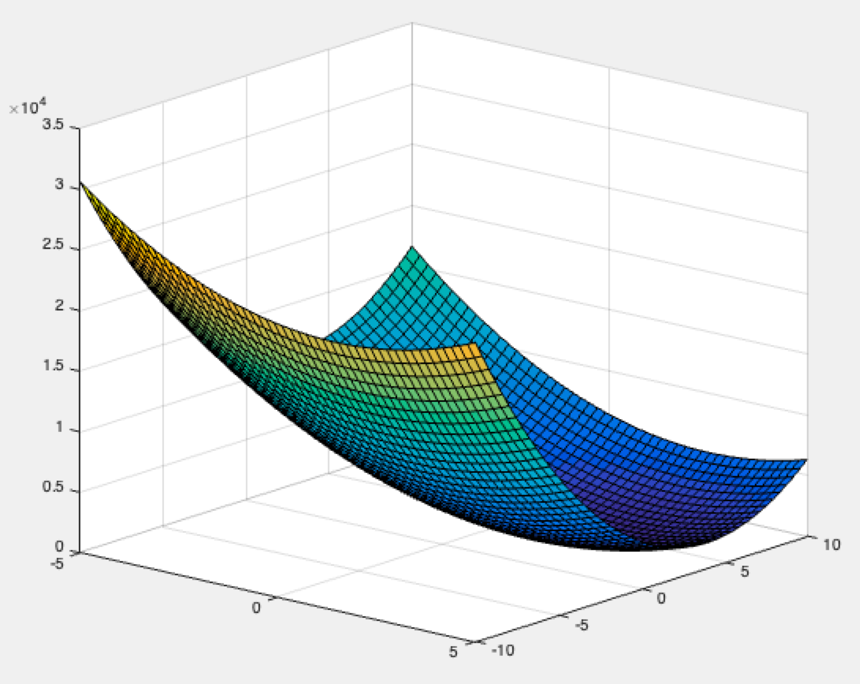
梯度下降法

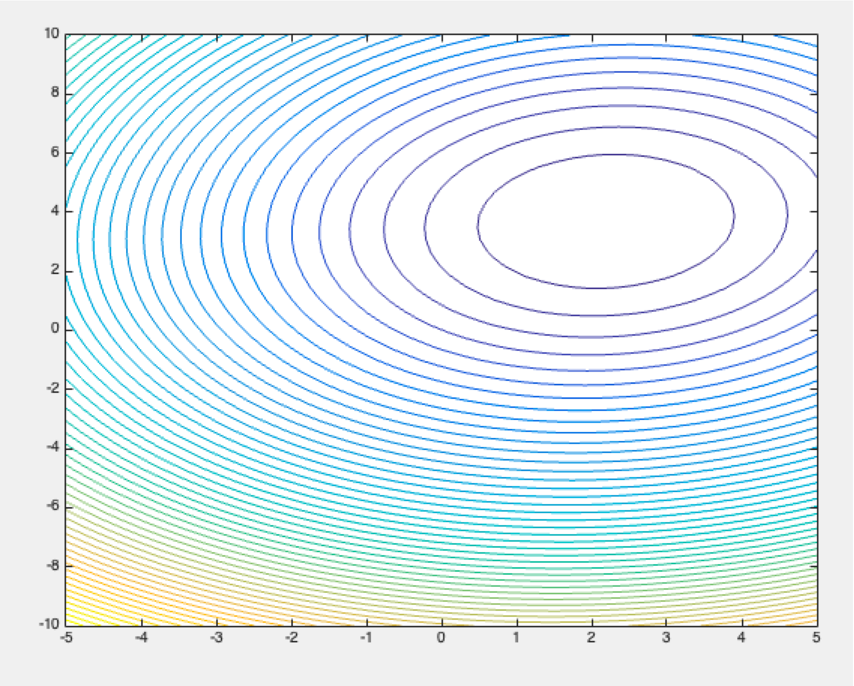
人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|\boldsymbol{Aw}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w}\in\mathcal{R}^N.$$

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117





梯度下降法

人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

线性回归（L2正则化）	
min	$\frac{1}{M}\|\mathbf{A}\boldsymbol{w} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{w}\|_2^2$
var.	$\boldsymbol{w} \in \mathcal{R}^N.$

$$\nabla f(\boldsymbol{w}) = \frac{2}{M}\mathbf{A}^T(\mathbf{A}\boldsymbol{w} - \boldsymbol{b}) + 2\lambda\boldsymbol{w}$$

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
...	...	...
4.1149	-2.5394	25.1117

梯度下降法

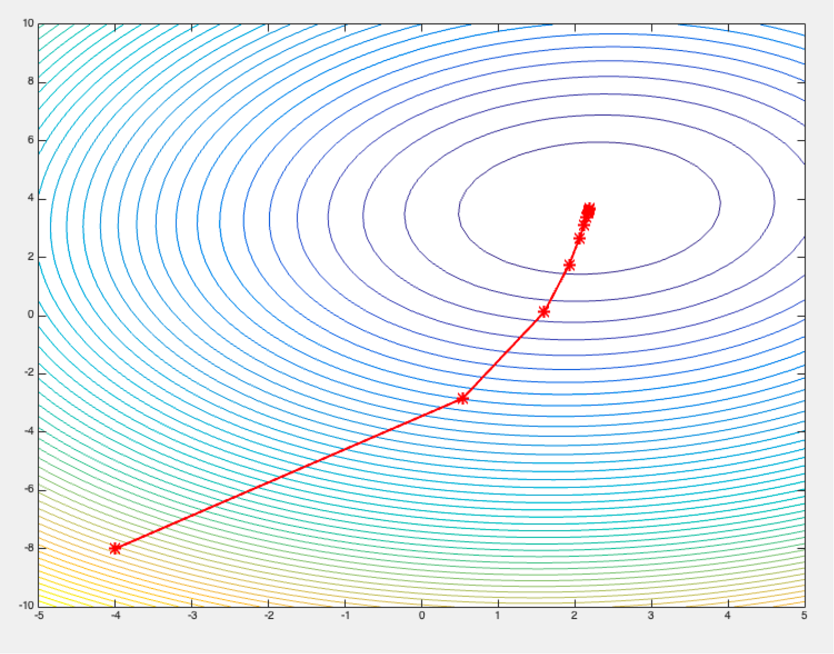
人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|\boldsymbol{A}\boldsymbol{w} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w} \in \mathcal{R}^N.$$

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117

```
iterations = 10;
alpha = 0.002;
solutionw = [-4; -8];
w1history = zeros(1,iterations+1);
w2history = zeros(1,iterations+1);
w1history(1) = solutionw(1);
w2history(1) = solutionw(2);
for i = 1:iterations
derivative = 2 / M * MatrixA' * (MatrixA * solutionw -
vectorb) + 2 * lambda * solutionw;
solutionw = solutionw - alpha * derivative;
w1history(i+1) = solutionw(1);
w2history(i+1) = solutionw(2);
end
hold on;
plot(w1history, w2history,'r*-')
```



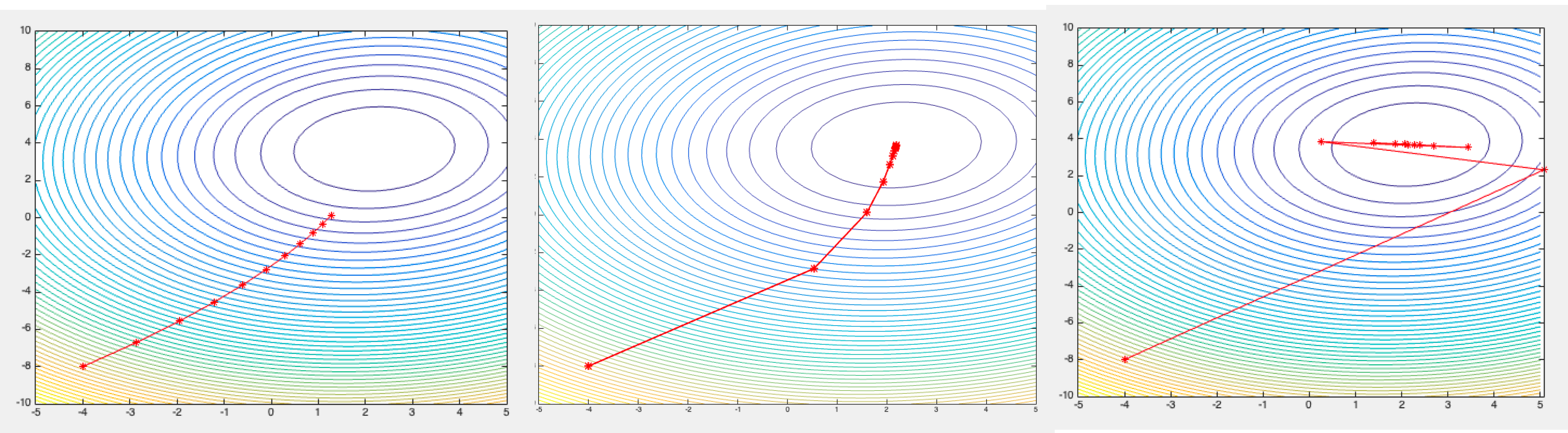
梯度下降法

人工根据$b=2a_1+5a_2+\mathcal{N}(0,20)$ 生成30条数据

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|A\boldsymbol{w}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w}\in\mathcal{R}^N.$$

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117



$\alpha=0.0005$ $\alpha=0.002$ $\alpha=0.004$

梯度下降法

梯度下降法

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boxed{\boldsymbol{d}^t=-\nabla f(\boldsymbol{x}^t)}$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

如何改进算法从而改善收敛速度慢的问题?

$$f(\boldsymbol{x}^{t+1})\approx f(\boldsymbol{x}^t)+\alpha^t\left(\nabla f(\boldsymbol{x}^t)\right)^T\boldsymbol{d}^t$$

$\boldsymbol{d}^t=-\nabla f(\boldsymbol{x}^t)$已经是最小化$f(\boldsymbol{x}^{t+1})$的选择

牛顿法

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、KKT条件、内点法
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

牛顿法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boldsymbol{d}^t$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

利用函数的二阶泰勒展开式做近似，再确定$\boldsymbol{d}^t$

牛顿法

函数的泰勒展开式（一维）

Taylor's theorem[4][5][6] — Let $k \geq 1$ be an integer and let the function $f : \mathbf{R} \to \mathbf{R}$ be k times differentiable at the point $a \in \mathbf{R}$. Then there exists a function $h_k : \mathbf{R} \to \mathbf{R}$ such that

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + h_k(x)(x-a)^k,$$

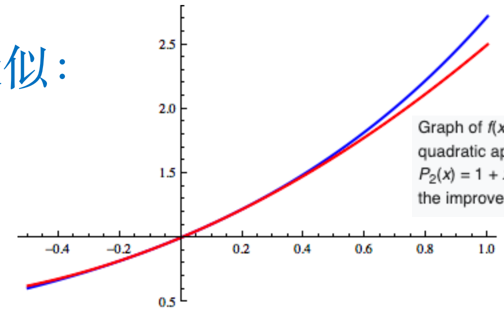
and

$$\lim_{x \to a} h_k(x) = 0.$$

This is called the **Peano form of the remainder**.

$f(x)$在 $x = a$附近的取值可利用一阶和二阶导信息近似：

$$f(x) \approx f(a) + (x-a)f'(a) + \frac{1}{2}(x-a)^2 f''(a)$$



Graph of $f(x) = e^x$ (blue) with its quadratic approximation $P_2(x) = 1 + x + x^2/2$ (red) at $a = 0$. Note the improvement in the approximation.

示图取自Wikipedia Taylor‘s theorem词条

牛顿法

函数的泰勒展开式（一维）

$f(x)$在 $x=a$附近的取值可利用一阶和二阶导信息近似：

$$f(x) \approx f(a) + (x-a)f'(a) + \frac{1}{2}(x-a)^2 f''(a)$$

函数的泰勒展开式（多维）

$f(\boldsymbol{x})$在 $\boldsymbol{x}=\boldsymbol{a}$附近的取值可利用梯度和海森矩阵信息近似：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T(\boldsymbol{x}-\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T \nabla^2 f(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})$$

牛顿法

$f(\boldsymbol{x})$在 $\boldsymbol{x}=\boldsymbol{a}$附近的取值可利用梯度和海森矩阵信息近似：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T(\boldsymbol{x}-\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T\nabla^2 f(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})$$

例：对$f(\boldsymbol{x}) = e^{x_1} + e^{x_2}$，考查在$\boldsymbol{x} = (0.6, -0.6)$的展开

Funtion: e^x + e^y

Point P=(0.6,-0.6)

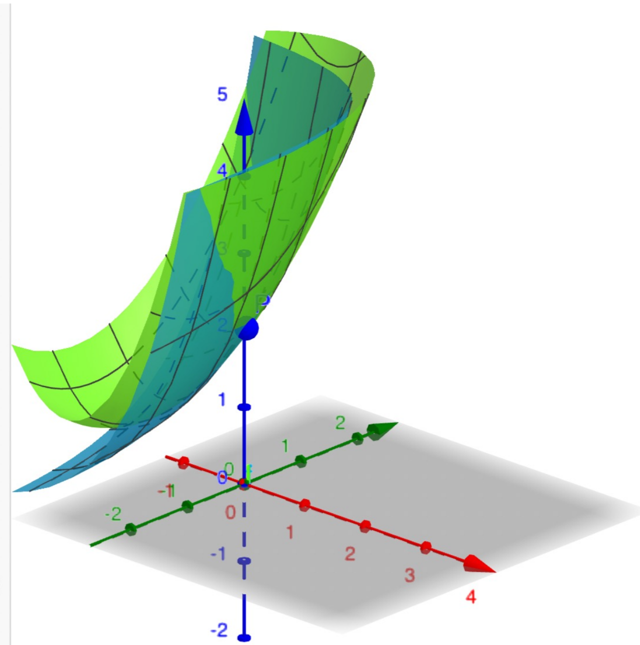
$x_0 = 0.6$

$y_0 = -0.6$

☐ First degree Taylor polynomial

☑ Second degree Taylor polynomial

$$P^2(x,y) = e^{0.6} + e^{-0.6} + e^{0.6}(x-0.6) + e^{-0.6}(y+0.6) +$$
$$+\frac{1}{2}\left(e^{0.6}(x-0.6)^2 + 2\cdot 0\,(x-0.6)(y+0.6) + e^{-0.6}(y+0.6)^2\right.$$



示图取自https://www.geogebra.org/m/Ehnz3hGb

牛顿法

$f(\boldsymbol{x})$在 $\boldsymbol{x} = \boldsymbol{a}$附近的取值可利用梯度和海森矩阵信息近似：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T (\boldsymbol{x}-\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T \nabla^2 f(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})$$

例：对$f(\boldsymbol{x}) = e^{x_1} + e^{x_2}$，考查在$\boldsymbol{x} = (0.6, -0.6)$的展开

Funtion: e^x + e^y

Point P=(0.6,-0.6)

$x_0 = 0.6$

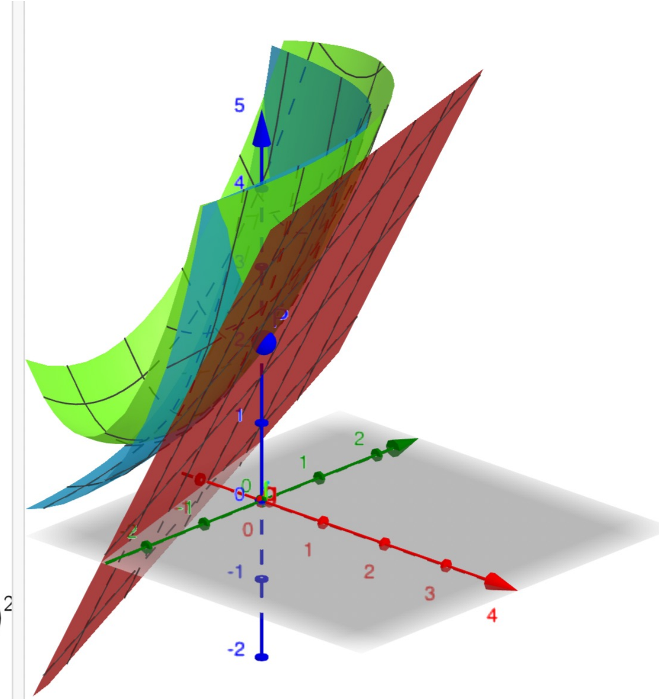
$y_0 = -0.6$

☑ First degree Taylor polynomial

$$P^1(x,y) = e^{0.6} + e^{-0.6} + e^{0.6}(x-0.6) + e^{-0.6}(y+0.6)$$

☑ Second degree Taylor polynomial

$$P^2(x,y) = e^{0.6} + e^{-0.6} + e^{0.6}(x-0.6) + e^{-0.6}(y+0.6) + \\ + \frac{1}{2}\left(e^{0.6}(x-0.6)^2 + 2\cdot 0\,(x-0.6)(y+0.6) + e^{-0.6}(y+0.6)^2\right.$$



示图取自https://www.geogebra.org/m/Ehnz3hGb

牛顿法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boxed{\boldsymbol{d}^t}$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\nabla f(\boldsymbol{a})\right)^T(\boldsymbol{x}-\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T\nabla^2 f(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})$$

$$f(\boldsymbol{x}^{t+1}) \approx f(\boldsymbol{x}^t) + \left(\nabla f(\boldsymbol{x}^t)\right)^T(\alpha^t\boldsymbol{d}^t) + \frac{1}{2}(\alpha^t\boldsymbol{d}^t)^T\nabla^2 f(\boldsymbol{x}^t)(\alpha^t\boldsymbol{d}^t)$$

$$f(\boldsymbol{x}^{t+1}) \approx f(\boldsymbol{x}^t) + \left(\nabla f(\boldsymbol{x}^t)\right)^T\boldsymbol{s} + \frac{1}{2}\boldsymbol{s}^T\nabla^2 f(\boldsymbol{x}^t)\boldsymbol{s}$$

如何选择$\boldsymbol{s}$可以最小化$f(\boldsymbol{x}^{t+1})$？

牛顿法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boxed{\boldsymbol{d}^t}$

2 确定步长$\alpha^t > 0$

3 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \alpha^t \boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t+1$且返回1

$$f(\boldsymbol{x}^{t+1}) \approx f(\boldsymbol{x}^t) + \left(\nabla f(\boldsymbol{x}^t)\right)^T \boldsymbol{s} + \frac{1}{2} \boldsymbol{s}^T \nabla^2 f(\boldsymbol{x}^t) \boldsymbol{s}$$

右侧是关于$\boldsymbol{s}$的凸函数

Example 3.2 *Quadratic functions.* Consider the quadratic function $f : \mathbf{R}^n \to \mathbf{R}$, with $\mathbf{dom}\, f = \mathbf{R}^n$, given by

$$f(x) = (1/2)x^T P x + q^T x + r,$$

with $P \in \mathbf{S}^n$, $q \in \mathbf{R}^n$, and $r \in \mathbf{R}$. Since $\nabla^2 f(x) = P$ for all x, f is convex if and only if $P \succeq 0$ (and concave if and only if $P \preceq 0$).

二次函数描述取自Stephen Boyd <Convex Optimization> (优化理论的经典著作，引用量>67000)

牛顿法

下降法（一般形式）

0 初始化当前解$\boldsymbol{x}^0$及$t=0$

1 确定下降方向$\boxed{\boldsymbol{d}^t}$

2 确定步长$\alpha^t>0$

3 更新当前解：$\boldsymbol{x}^{t+1}=\boldsymbol{x}^t+\alpha^t\boldsymbol{d}^t$

4 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t\leftarrow t+1$且返回1

$$f(\boldsymbol{x}^{t+1})\approx f(\boldsymbol{x}^t)+\left(\nabla f(\boldsymbol{x}^t)\right)^T\boldsymbol{s}+\frac{1}{2}\boldsymbol{s}^T\nabla^2 f(\boldsymbol{x}^t)\boldsymbol{s}$$

右侧是关于$\boldsymbol{s}$的凸函数　　右侧对$\boldsymbol{s}$求梯度，找令梯度取零的$\boldsymbol{s}$值

梯度为：$\nabla f(\boldsymbol{x}^t)+\nabla^2 f(\boldsymbol{x}^t)\boldsymbol{s}$

若海森矩阵$\nabla^2 f(\boldsymbol{x}^t)$可逆，有$\boldsymbol{s}^*=-\left(\nabla^2 f(\boldsymbol{x}^t)\right)^{-1}\nabla f(\boldsymbol{x}^t)$

即$\alpha^t\boldsymbol{d}^t=-\left(\nabla^2 f(\boldsymbol{x}^t)\right)^{-1}\nabla f(\boldsymbol{x}^t)$

牛顿法

牛顿法

0 初始化当前解$\boldsymbol{x}^0$及$t = 0$

1 确定下降方向$\boldsymbol{d}^t = -\left(\nabla^2 f(\boldsymbol{x}^t)\right)^{-1} \nabla f(\boldsymbol{x}^t)$

2 更新当前解：$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \boldsymbol{d}^t$

3 若$\|\nabla f(\boldsymbol{x}^{t+1})\|_2$小于阈值则结束；否则，$t \leftarrow t + 1$且返回1

收敛性能分析略

优点：收敛速度很快

缺点：初始解$\boldsymbol{x}^0$需要离$\boldsymbol{x}^*$足够近（即仅有局部收敛性）；
需要$f(\cdot)$二阶连续可微且$\left(\nabla^2 f(\boldsymbol{x}^t)\right)^{-1}$存在（$\nabla^2 f(\boldsymbol{x}^t)$是正定矩阵）

可先用梯度下降得到低精度解，再用牛顿法

牛顿法

人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
...	...	...
4.1149	-2.5394	25.1117

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|\mathbf{A}\boldsymbol{w}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w}\in\mathcal{R}^N.$$

$$\nabla f(\boldsymbol{w})=\frac{2}{M}\mathbf{A}^T(\mathbf{A}\boldsymbol{w}-\boldsymbol{b})+2\lambda\boldsymbol{w}$$

$$\boldsymbol{\nabla}^2 f(\boldsymbol{w})=\frac{2}{M}\mathbf{A}^T\mathbf{A}+2\lambda\mathbf{I}$$

牛顿法

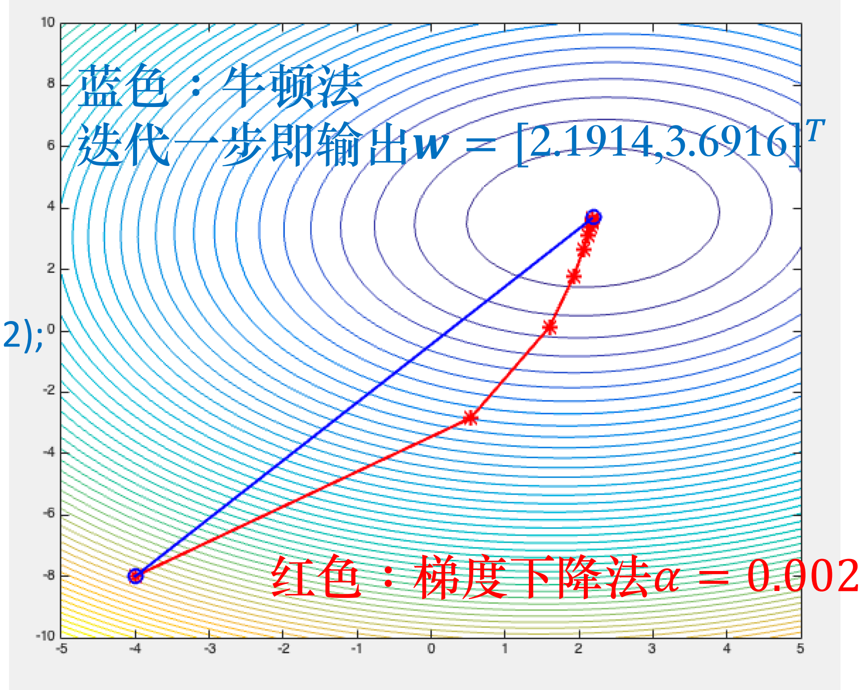
人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|A\boldsymbol{w} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w} \in \mathcal{R}^N.$$

```
iterations = 10;
solutionw = [-4; -8];
w1history = zeros(1,iterations+1);
w2history = zeros(1,iterations+1);
w1history(1) = solutionw(1);
w2history(1) = solutionw(2);
for i = 1:iterations
derivative = 2 / M * MatrixA' * (MatrixA * solutionw -
vectorb) + 2 * lambda * solutionw;
Hessian = 2 / M * MatrixA' * MatrixA + 2 * lambda * eye(2);
inverse = inv(Hessian);
solutionw = solutionw - inverse * derivative;
w1history(i+1) = solutionw(1);
w2history(i+1) = solutionw(2);
end
hold on;
plot(w1history, w2history,'bo-')
```



牛顿法

人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117

线性回归（L2正则化）

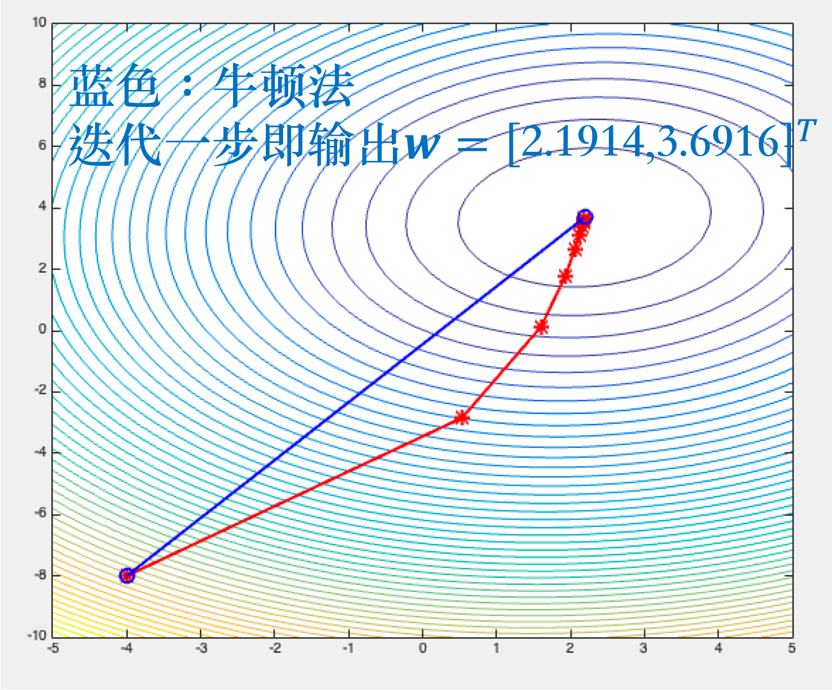
$$\min \quad \frac{1}{M}\|\mathbf{A}\boldsymbol{w} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w} \in \mathcal{R}^N.$$

$$\nabla f(\boldsymbol{w}) = \frac{2}{M}\mathbf{A}^T(\mathbf{A}\boldsymbol{w} - \boldsymbol{b}) + 2\lambda\boldsymbol{w}$$
$$\boldsymbol{\nabla}^2 f(\boldsymbol{w}) = \frac{2}{M}\mathbf{A}^T\mathbf{A} + 2\lambda\mathbf{I}$$

回顾二阶泰勒展开：

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + \left(\boldsymbol{\nabla} f(\boldsymbol{a})\right)^T(\boldsymbol{x} - \boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T\boldsymbol{\nabla}^2 f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})$$

可以证明针对当前目标函数，≈可取=



牛顿法

人工根据$b = 2a_1 + 5a_2 + \mathcal{N}(0,20)$ 生成30条数据

特征 a_1	特征 a_2	标签b
0.7993	-6.2909	-39.2102
-9.4848	-12.0385	-81.6599
…	…	…
4.1149	-2.5394	25.1117

线性回归（L2正则化）

$$\min \quad \frac{1}{M}\|\mathbf{A}\boldsymbol{w} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{w}\|_2^2$$
$$\text{var.} \quad \boldsymbol{w} \in \mathcal{R}^N.$$

$$\nabla f(\boldsymbol{w}) = \frac{2}{M}\mathbf{A}^T(\mathbf{A}\boldsymbol{w} - \boldsymbol{b}) + 2\lambda\boldsymbol{w}$$

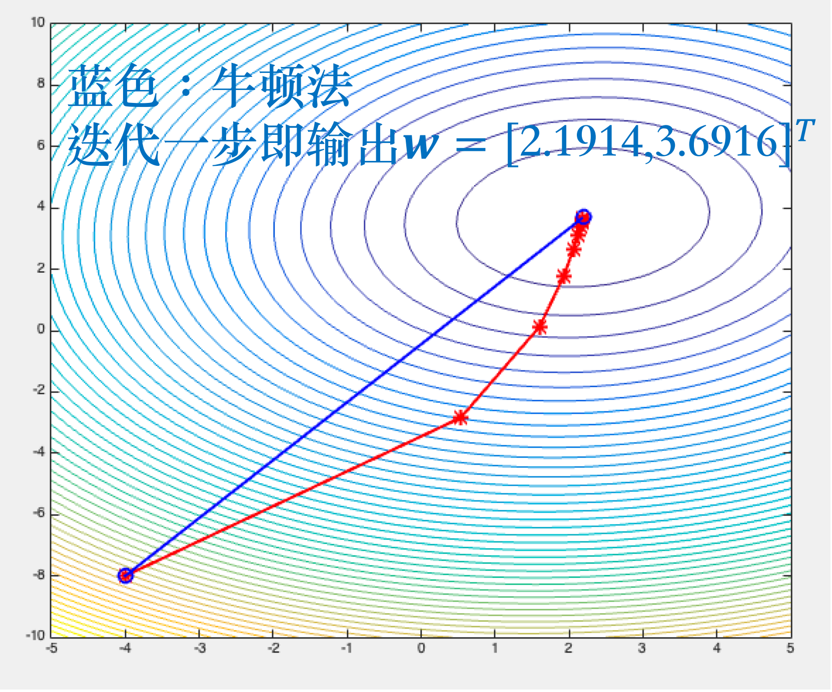
$$\nabla^2 f(\boldsymbol{w}) = \frac{2}{M}\mathbf{A}^T\mathbf{A} + 2\lambda\mathbf{I}$$

证明对任何初始解$\boldsymbol{w}^0$，牛顿法可一步到达上述问题的$\boldsymbol{w}^*$

（1） 由$\nabla f(\boldsymbol{w}^*) = \mathbf{0}$得到$\boldsymbol{w}^*$的闭式解

$$\boldsymbol{w}^* = \left(\frac{1}{M}\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I}\right)^{-1}\frac{1}{M}\mathbf{A}^T\boldsymbol{b}$$

（2） 证明$\boldsymbol{d}^t = -\boldsymbol{w}^0 + \boldsymbol{w}^*$



无约束非凸优化问题

$$\min \quad f(\boldsymbol{x})$$
$$\text{var.} \quad \boldsymbol{x} \in \mathcal{R}^n.$$
其中，$f(\boldsymbol{x})$是二阶连续可微的函数

当函数非凸时，也可用梯度下降法/牛顿法搜索局部最小值（无理论保障）

当函数非凸时 $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$ $\overset{\Leftarrow}{\nRightarrow}$ $\boldsymbol{x}$是局部最小值

无约束非凸优化问题

min $\quad f(\boldsymbol{x})$
var. $\quad \boldsymbol{x} \in \mathcal{R}^n$.
其中，$f(\boldsymbol{x})$是二阶连续可微的函数

当函数非凸时，也可用梯度下降法/牛顿法搜索局部最小值（无理论保障）

当函数非凸时 $\nabla f(\boldsymbol{x}) = \mathbf{0}$ $\overset{\Leftarrow}{\not\Rightarrow}$ $\boldsymbol{x}$是局部最小值

例，$f(x) = x^3$在$x = 0$处梯度为0

基于梯度下降法的改进算法（如随机梯度下降法、批量梯度下降法）在深度学习中常用于优化网络权重、最小化损失函数

牛顿法需要计算海森矩阵、对其求逆，复杂度高

线性规划问题

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、KKT条件、内点法
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

线性规划

最优化问题的一般形式

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & g_i(\boldsymbol{x}) \le 0, i = 1, \ldots, m, \\ & h_i(\boldsymbol{x}) = 0, j = 1, \ldots, p, \\ \text{var.} \quad & \boldsymbol{x} \in \mathcal{D}. \end{aligned}$$

当目标函数和所有约束（包括$\mathcal{D}$对$\boldsymbol{x}$的约束）都是线性时，称为线性规划问题

线性函数的一般形式：$f(\boldsymbol{x}) = \boldsymbol{a}_0^T\boldsymbol{x} + b_0$ （注意$\boldsymbol{a}_0$是向量）

线性规划

最优化问题的一般形式

$$
\begin{aligned}
\min \quad & f(\boldsymbol{x}) \\
\text{s.t.} \quad & g_i(\boldsymbol{x}) \le 0, i = 1, \dots, m, \\
& h_i(\boldsymbol{x}) = 0, j = 1, \dots, p, \\
\text{var.} \quad & \boldsymbol{x} \in \mathcal{D}.
\end{aligned}
$$

当目标函数和所有约束（包括$\mathcal{D}$对$\boldsymbol{x}$的约束）都是线性时，称为线性规划问题

线性函数的一般形式：$f(\boldsymbol{x}) = \boldsymbol{a}_0^T \boldsymbol{x} + b_0$ （注意$\boldsymbol{a}_0$是向量）

$f(\boldsymbol{x}) = x_1 x_2 - x_3$

$f(\boldsymbol{x}) = \frac{x_1}{x_2} + x_3 - 2$

$f(\boldsymbol{x}) = 3x_1 - 5x_2 - x_3 + 9$

线性规划

最优化问题的一般形式

$$
\begin{aligned}
\min \quad & f(\boldsymbol{x}) \\
\text{s.t.} \quad & g_i(\boldsymbol{x}) \le 0, i = 1, \dots, m, \\
& h_i(\boldsymbol{x}) = 0, j = 1, \dots, p, \\
\text{var.} \quad & \boldsymbol{x} \in \mathcal{D}.
\end{aligned}
$$

当目标函数和所有约束（包括$\mathcal{D}$对$\boldsymbol{x}$的约束）都是线性时，称为线性规划问题

线性函数的一般形式：$f(\boldsymbol{x}) = \boldsymbol{a}_0^T \boldsymbol{x} + b_0$ （注意$\boldsymbol{a}_0$是向量）

$f(\boldsymbol{x}) = x_1 x_2 - x_3$	$f(\boldsymbol{x}) = \dfrac{x_1}{x_2} + x_3 - 2$	$f(\boldsymbol{x}) = 3x_1 - 5x_2 - x_3 + 9$
✗	✗	✓

线性规划

最优化问题的一般形式

$$
\begin{array}{ll}
\min & f(\boldsymbol{x}) \\
\text{s.t.} & g_i(\boldsymbol{x}) \le 0, i = 1, \dots, m, \\
 & h_i(\boldsymbol{x}) = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x} \in \mathcal{D}.
\end{array}
$$

当目标函数和所有约束（包括$\mathcal{D}$对$\boldsymbol{x}$的约束）都是线性时，称为线性规划问题

线性规划问题 (linear programming)

$$
\begin{array}{ll}
\min & \boldsymbol{a}_0^T\boldsymbol{x} + b_0 \\
\text{s.t.} & \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
 & \boldsymbol{d}_j^T\boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x}.
\end{array}
$$

（ $\boldsymbol{a}_0, \boldsymbol{a}_i, \boldsymbol{d}_j$是向量）

线性规划

线性规划问题

$$\begin{array}{ll} \min & \boldsymbol{a}_0^T\boldsymbol{x} + b_0 \\ \text{s.t.} & \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\ & \boldsymbol{d}_j^T\boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\ \text{var.} & \boldsymbol{x}. \end{array}$$

判断是否是线性规划问题

$$\begin{array}{ll} \min & x_1x_2 + x_3 \\ \text{s.t.} & x_1 + 3x_2 \ge x_3, \\ \text{var.} & x_1, x_2, x_3. \end{array}$$

$$\begin{array}{ll} \min & x_1 - x_2 + x_3 \\ \text{s.t.} & x_1 + 3\sqrt{x_2} \ge x_3, \\ \text{var.} & x_1, x_2, x_3. \end{array}$$

$$\begin{array}{ll} \max & x_1 - x_2 + x_3 + 5 \\ \text{s.t.} & x_1 + 3x_2 \ge x_3, \\ & x_2 - x_1 \le 6, \\ \text{var.} & x_1, x_2, x_3. \end{array}$$

线性规划

线性规划问题

$$\begin{array}{ll} \min & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\ \text{s.t.} & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\ & \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\ \text{var.} & \boldsymbol{x}. \end{array}$$

判断是否是线性规划问题

$$\begin{array}{ll} \min & x_1 x_2 + x_3 \\ \text{s.t.} & x_1 + 3x_2 \ge x_3, \\ \text{var.} & x_1, x_2, x_3. \end{array}$$

×

$$\begin{array}{ll} \min & x_1 - x_2 + x_3 \\ \text{s.t.} & x_1 + 3\sqrt{x_2} \ge x_3, \\ \text{var.} & x_1, x_2, x_3. \end{array}$$

×

$$\begin{array}{ll} \max & x_1 - x_2 + x_3 + 5 \\ \text{s.t.} & x_1 + 3x_2 \ge x_3, \\ & x_2 - x_1 \le 6, \\ \text{var.} & x_1, x_2, x_3. \end{array}$$

✓

线性规划

线性规划问题

$$
\begin{array}{ll}
\min & \boldsymbol{a}_0^T\boldsymbol{x} + b_0 \\
\text{s.t.} & \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
 & \boldsymbol{d}_j^T\boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x}.
\end{array}
$$

判断是否是线性规划问题

$$
\begin{array}{ll}
\min & x_1 \\
\text{s.t.} & 5 - x_1 < 0, \\
\text{var.} & x_1.
\end{array}
$$

严格不等式无法整理成≤约束

线性规划

线性规划问题

$$
\begin{array}{ll}
\min & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \leq 0, i = 1, \dots, m, \\
 & \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x}.
\end{array}
$$

判断是否是线性规划问题

$$
\begin{array}{ll}
\min & x_1 \\
\text{s.t.} & 5 - x_1 < 0, \\
\text{var.} & x_1.
\end{array}
$$

严格不等式无法整理成≤约束

一般教材只讨论对带有≤约束的线性规划的求解，不讨论带有<约束的问题

各类求解线性规划的算法也只考虑≤约束，不考虑<约束

为什么？

线性规划

线性规划问题

$$\begin{array}{ll} \min & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\ \text{s.t.} & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\ & \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\ \text{var.} & \boldsymbol{x}. \end{array}$$

判断是否是线性规划问题

$$\begin{array}{ll} \min & x_1 \\ \text{s.t.} & 5 - x_1 < 0, \\ \text{var.} & x_1. \end{array}$$

严格不等式无法整理成≤约束

实际应用中很少有真正需要严格不等式约束的情况（即严格要求 < 而不允许取 = ）

比如 购买材料问题中的总预算约束、交通运输问题中的运载量约束 等 都允许取=

线性规划

线性规划问题	
min	$\boldsymbol{a}_0^T \boldsymbol{x} + b_0$
s.t.	$\boldsymbol{a}_i^T \boldsymbol{x} + b_i \leq 0, i = 1, \dots, m,$
	$\boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p,$
var.	$\boldsymbol{x}.$

判断是否是线性规划问题

$$\begin{array}{ll} \text{min} & x_1 \\ \text{s.t.} & 5 - x_1 < 0, \\ \text{var.} & x_1. \end{array}$$

严格不等式无法整理成≤约束

如果真遇到带有严格不等式约束的问题怎么办？

引入取值极小的参数，把<约束改写成≤约束

线性规划标准形式

线性规划问题

$$
\begin{array}{ll}
\min & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
 & \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x}.
\end{array}
$$

通过重新定义决策变量和参数，线性规划都可以化简成如下形式：

线性规划问题的标准形式

$$
\begin{array}{ll}
\min & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} & \boldsymbol{x} \ge \mathbf{0}.
\end{array}
$$

注意：在化简的过程中，决策变量和参数等都已经重新定义，比如决策变量$\boldsymbol{x}$可以不同于原式中的$\boldsymbol{x}$

线性规划标准形式

线性规划问题

$$\begin{aligned} \min \quad & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\ \text{s.t.} \quad & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\ & \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

通过重新定义决策变量和参数，线性规划都可以化简成如下形式：

线性规划问题的标准形式

$$\begin{aligned} \min \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x} \ge \mathbf{0}. \end{aligned}$$

非矩阵向量表示

目标

$$\min c_1 x_1 + \ldots + c_n x_n$$

约束

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

变量

$$x_1 \ge 0, \ \ldots, \ x_n \ge 0 .$$

线性规划标准形式

线性规划问题

$$\begin{aligned}
\min\quad & \boldsymbol{a}_0^T\boldsymbol{x} + b_0 \\
\text{s.t.}\quad & \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
& \boldsymbol{d}_j^T\boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.}\quad & \boldsymbol{x}.
\end{aligned}$$

线性规划问题的标准形式

$$\begin{aligned}
\min\quad & \boldsymbol{c}^T\boldsymbol{x} \\
\text{s.t.}\quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.}\quad & \boldsymbol{x} \ge \mathbf{0}.
\end{aligned}$$

为什么要变换成标准形式？

标准形式便于设计算法求解（类比于将一元二次方程化为标准形式再用公式求解）

如何进行变换？

1. 怎么把不等式约束变成等式约束？

2. 怎么通过变换使所有决策变量的定义域都为正实数集合？

线性规划标准形式

线性规划问题

$$
\begin{aligned}
\text{min} \quad & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} \quad & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
& \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} \quad & \boldsymbol{x}.
\end{aligned}
$$

线性规划问题的标准形式

$$
\begin{aligned}
\text{min} \quad & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} \quad & \boldsymbol{x} \ge \boldsymbol{0}.
\end{aligned}
$$

1. 怎么把不等式约束变成等式约束？

$$
\begin{aligned}
\text{min} \quad & x_1 - 3x_2 \\
\text{s.t.} \quad & 5 - x_1 \le 0, \\
& x_2 \ge 3, \\
\text{var.} \quad & x_1 \ge 0, x_2 \ge 0.
\end{aligned}
$$

线性规划标准形式

线性规划问题

$$
\begin{array}{ll}
\min & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
 & \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x}.
\end{array}
$$

$\Rightarrow$

线性规划问题的标准形式

$$
\begin{array}{ll}
\min & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} & \boldsymbol{x} \ge \mathbf{0}.
\end{array}
$$

1. 怎么把不等式约束变成等式约束？

构造新的变量（称为松弛变量 即slack variable）

$$
\begin{array}{ll}
\min & x_1 - 3x_2 \\
\text{s.t.} & 5 - x_1 \le 0, \\
 & x_2 \ge 3, \\
\text{var.} & x_1 \ge 0, x_2 \ge 0.
\end{array}
$$

线性规划标准形式

线性规划问题

$$\begin{aligned} \min \quad & \boldsymbol{a}_0^T\boldsymbol{x}+b_0 \\ \text{s.t.} \quad & \boldsymbol{a}_i^T\boldsymbol{x}+b_i \le 0, i=1,\dots,m, \\ & \boldsymbol{d}_j^T\boldsymbol{x}+c_j=0, j=1,\dots,p, \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

$\Rightarrow$

线性规划问题的标准形式

$$\begin{aligned} \min \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{x}=\boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x} \ge \mathbf{0}. \end{aligned}$$

1. 怎么把不等式约束变成等式约束？

构造新的变量（称为松弛变量 即**slack variable**）

$$\begin{aligned} \min \quad & x_1-3x_2 \\ \text{s.t.} \quad & 5-x_1 \le 0, \\ & x_2 \ge 3, \\ \text{var.} \quad & x_1 \ge 0, x_2 \ge 0. \end{aligned}$$

$$\begin{aligned} \min \quad & x_1-3x_2 \\ \text{s.t.} \quad & 5-x_1+x_3=0, \\ & 3-x_2+x_4=0, \\ \text{var.} \quad & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0. \end{aligned}$$

线性规划标准形式

线性规划问题

$$
\begin{aligned}
\min \quad & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} \quad & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
& \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} \quad & \boldsymbol{x}.
\end{aligned}
$$

线性规划问题的标准形式

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} \quad & \boldsymbol{x} \ge \mathbf{0}.
\end{aligned}
$$

2. 怎么使决策变量的定义域变为正实数集合？

$$
\begin{aligned}
\min \quad & x_1 - 3x_2 \\
\text{s.t.} \quad & 5 - x_1 + x_2 = 0, \\
\text{var.} \quad & x_1 \ge 0, x_2.
\end{aligned}
$$

线性规划标准形式

线性规划问题

$$
\begin{aligned}
\min \quad & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} \quad & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
& \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} \quad & \boldsymbol{x}.
\end{aligned}
$$

线性规划问题的标准形式

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} \quad & \boldsymbol{x} \ge \mathbf{0}.
\end{aligned}
$$

2. 怎么使决策变量的定义域变为正实数集合？

$$
\begin{aligned}
\min \quad & x_1 - 3x_2 \\
\text{s.t.} \quad & 5 - x_1 + x_2 = 0, \\
\text{var.} \quad & x_1 \ge 0, x_2.
\end{aligned}
$$

将相应变量改写成两个新的非负变量（称为人工变量 即**artificial variable**）的差值

线性规划标准形式

线性规划问题

$$\begin{aligned} \min \quad & \boldsymbol{a}_0^T\boldsymbol{x}+b_0 \\ \text{s.t.} \quad & \boldsymbol{a}_i^T\boldsymbol{x}+b_i \le 0, i=1,\dots,m, \\ & \boldsymbol{d}_j^T\boldsymbol{x}+c_j = 0, j=1,\dots,p, \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

$\Rightarrow$

线性规划问题的标准形式

$$\begin{aligned} \min \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{x}=\boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x} \ge \mathbf{0}. \end{aligned}$$

2. 怎么使决策变量的定义域变为正实数集合？

$$\begin{aligned} \min \quad & x_1 - 3x_2 \\ \text{s.t.} \quad & 5 - x_1 + x_2 = 0, \\ \text{var.} \quad & x_1 \ge 0, x_2. \end{aligned}$$

将相应变量改写成两个新的非负变量（称为人工变量 即**artificial variable**）的差值

利用$x_2 = x_3 - x_4, x_3 \ge 0, x_4 \ge 0$进行改写

$$\begin{aligned} \min \quad & x_1 - 3x_3 + 3x_4 \\ \text{s.t.} \quad & 5 - x_1 + x_3 - x_4 = 0, \\ \text{var.} \quad & x_1 \ge 0, x_3 \ge 0, x_4 \ge 0. \end{aligned}$$

线性规划标准形式

线性规划问题的标准形式

$$
\begin{aligned}
\text{min} \quad & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} \quad & \boldsymbol{x} \geq \mathbf{0}.
\end{aligned}
$$

练习：将以下问题改写成线性规划的标准形式

$$
\begin{aligned}
\text{max} \quad & 5x_1 - 3x_2 \\
\text{s.t.} \quad & 2 - 3x_1 + 5x_2 \geq 0, \\
\text{var.} \quad & x_1 \leq 0, x_2.
\end{aligned}
$$

线性规划标准形式

线性规划问题的标准形式

$$\begin{aligned} \min \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x} \geq \mathbf{0}. \end{aligned}$$

练习：将以下问题改写成线性规划的标准形式

$$\begin{aligned} \max \quad & 5x_1 - 3x_2 \\ \text{s.t.} \quad & 2 - 3x_1 + 5x_2 \geq 0, \\ \text{var.} \quad & x_1 \leq 0, x_2. \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \min \quad & -5x_1 + 3x_2 \\ \text{s.t.} \quad & 2 - 3x_1 + 5x_2 \geq 0, \\ \text{var.} \quad & x_1 \leq 0, x_2. \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \min \quad & -5x_1 + 3x_2 \\ \text{s.t.} \quad & -2 + 3x_1 - 5x_2 + x_3 = 0, \\ \text{var.} \quad & x_1 \leq 0, x_2, x_3 \geq 0. \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \min \quad & 5x_4 + 3x_2 \\ \text{s.t.} \quad & -2 - 3x_4 - 5x_2 + x_3 = 0, \\ \text{var.} \quad & x_4 \geq 0, x_2, x_3 \geq 0. \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \min \quad & 5x_4 + 3x_5 - 3x_6 \\ \text{s.t.} \quad & -2 - 3x_4 - 5x_5 + 5x_6 + x_3 = 0, \\ \text{var.} \quad & x_4 \geq 0, x_3 \geq 0, x_5 \geq 0, x_6 \geq 0. \end{aligned}$$

线性规划的例子

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、**KKT**条件、内点法
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

生产销售

某豆腐店用黄豆制作两种不同口感的豆腐出售。制作口感较鲜嫩的豆腐每份需要0.3千克一级黄豆及0.5千克二级黄豆，售价10元；制作口感较厚实的豆腐每份需要0.4千克一级黄豆及0.2千克二级黄豆，售价5元。现小店购入9千克一级黄豆和8千克二级黄豆。问：如何安排制作计划才能获得最大收益。

例子修改自《数学建模》全国赛培训课件

生产销售

某豆腐店用黄豆制作两种不同口感的豆腐出售。制作口感较鲜嫩的豆腐每份需要0.3千克一级黄豆及0.5千克二级黄豆，售价10元；制作口感较厚实的豆腐每份需要0.4千克一级黄豆及0.2千克二级黄豆，售价5元。现小店购入9千克一级黄豆和8千克二级黄豆。问：如何安排制作计划才能获得最大收益。

问题分析：

决策变量是制作口感鲜嫩和厚实的豆腐的份数；

约束是不超出制作两种不同口感豆腐所需黄豆总量；

目标是使得售出各种豆腐能获得最大收益

模型假设：假设制作的豆腐能全部售出；假设豆腐售价无波动

生产销售

某豆腐店用黄豆制作两种不同口感的豆腐出售。制作口感较鲜嫩的豆腐每份需要0.3千克一级黄豆及0.5千克二级黄豆，售价10元；制作口感较厚实的豆腐每份需要0.4千克一级黄豆及0.2千克二级黄豆，售价5元。现小店购入9千克一级黄豆和8千克二级黄豆。问：如何安排制作计划才能获得最大收益。

设口感鲜嫩和厚实的豆腐的份数为x_1和x_2

$$
\begin{aligned}
&\max\ 10x_1 + 5x_2 \\
&\text{s.t. } 0.3x_1 + 0.4x_2 \le 9, \\
&\qquad\ 0.5x_1 + 0.2x_2 \le 8, \\
&\text{var. } x_1 \ge 0,\ x_2 \ge 0.
\end{aligned}
$$

（假设决策变量不必取整数）

生产销售

某豆腐店用黄豆制作两种不同口感的豆腐出售。制作口感较鲜嫩的豆腐每份需要0.3千克一级黄豆及0.5千克二级黄豆，售价10元；制作口感较厚实的豆腐每份需要0.4千克一级黄豆及0.2千克二级黄豆，售价5元。现小店购入9千克一级黄豆和8千克二级黄豆。现可按照3元每千克的价格最多加购4千克一级黄豆。问：如何安排制作计划才能获得最大收益。

设口感鲜嫩和厚实的豆腐的份数为x_1和x_2

$$
\begin{aligned}
&\max\ 10x_1 + 5x_2 \\
&\text{s.t. } 0.3x_1 + 0.4x_2 \le 9, \\
&\qquad\ 0.5x_1 + 0.2x_2 \le 8, \\
&\text{var. } x_1 \ge 0, x_2 \ge 0.
\end{aligned}
$$

生产销售

某豆腐店用黄豆制作两种不同口感的豆腐出售。制作口感较鲜嫩的豆腐每份需要0.3千克一级黄豆及0.5千克二级黄豆，售价10元；制作口感较厚实的豆腐每份需要0.4千克一级黄豆及0.2千克二级黄豆，售价5元。现小店购入9千克一级黄豆和8千克二级黄豆。现可按照3元每千克的价格最多加购4千克一级黄豆。问：如何安排制作计划才能获得最大收益。

设口感鲜嫩和厚实的豆腐的份数为x_1和x_2，设加购的黄豆为x_3千克

$$
\begin{aligned}
\max\ & 10x_1 + 5x_2 \\
\text{s.t.}\ & 0.3x_1 + 0.4x_2 \le 9, \\
& 0.5x_1 + 0.2x_2 \le 8, \\
\text{var.}\ & x_1 \ge 0, x_2 \ge 0.
\end{aligned}
\Rightarrow
\begin{aligned}
\max\ & 10x_1 + 5x_2 - 3x_3 \\
\text{s.t.}\ & 0.3x_1 + 0.4x_2 \le 9 + x_3, \\
& 0.5x_1 + 0.2x_2 \le 8, \\
& x_3 \le 4, \\
\text{var.}\ & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0.
\end{aligned}
$$

物流运输

要从甲城调出蔬菜2000吨，从乙城调出蔬菜2500吨，从丙城调出3000吨，分别供应A地2000吨，B地2300吨、C地1800吨、D地1400吨，已知每吨运费如下表。问如何调拨才能最节省运费？

目的单位 调出单位	A	B	C	D
甲	21	27	13	40
乙	45	51	37	20
丙	32	35	20	30

物流运输

要从甲城调出蔬菜2000吨，从乙城调出蔬菜2500吨，从丙城调出3000吨，分别供应A地2000吨，B地2300吨、C地1800吨、D地1400吨，已知每吨运费如下表。问如何调拨才能最节省运费?

目的单位 / 调出单位	A	B	C	D
甲	21	27	13	40
乙	45	51	37	20
丙	32	35	20	30

决策变量：x_{ij}为从第i城运往第j地的蔬菜数量

参数：a_{ij}为从第i城运往第j地的单位运费

b_i为从第i城调出的蔬菜总量

c_j为第j地所需蔬菜总量

物流运输

a_{ij}为从第i城运往第j地的单位运费

b_i为从第i城调出的蔬菜总量

c_j为第j地所需蔬菜总量

要从甲城调出蔬菜2000吨，从乙城调出蔬菜2500吨，从丙城调出3000吨，分别供应A地2000吨，B地2300吨、C地1800吨、D地1400吨，已知每吨运费如下表。问如何调拨才能最节省运费?

目的单位 调出单位	A	B	C	D
甲	21	27	13	40
乙	45	51	37	20
丙	32	35	20	30

$$\min \sum_{i=1}^{3} \sum_{j=1}^{4} (a_{ij} x_{ij})$$

$$\text{s.t.} \sum_{j=1}^{4} x_{ij} = b_i, i = 1,2,3,$$

$$\sum_{i=1}^{3} x_{ij} = c_i, j = 1,2,3,4,$$

$$\text{var.}\ x_{ij} \geq 0, i = 1,2,3, j = 1,2,3,4.$$

排班调度

某新商店招聘员工，每天在岗的员工人数需要分别满足如下下限：

周一 27；周二 24；周三 23；周四 20；周五 25；周六 27；周日 28。

每名员工都是连续工作五天、休息两天（如可以周二到周六工作，周日周一休息）。求最少需要招聘多少员工？

例子修改自Andras London <Applications of Linear Programming>

排班调度

某新商店招聘员工，每天在岗的员工人数需要分别满足如下下限：

周一 27；周二 24；周三 23；周四 20；周五 25；周六 27；周日 28。

每名员工都是连续工作五天、休息两天（如可以周二到周六工作，周日周一休息）。求最少需要招聘多少员工？

$$
\begin{aligned}
\min \quad & z = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \\
& x_1 + x_4 + x_5 + x_6 + x_7 \geq 27 \\
& x_1 + x_2 + x_5 + x_6 + x_7 \geq 24 \\
& x_1 + x_2 + x_3 + x_6 + x_7 \geq 23 \\
& x_1 + x_2 + x_3 + x_4 + x_7 \geq 20 \\
& x_1 + x_2 + x_3 + x_4 + x_5 \geq 25 \\
& x_2 + x_3 + x_4 + x_5 + x_6 \geq 27 \\
& x_3 + x_4 + x_5 + x_6 + x_7 \geq 28 \\
& x_i \geq 0 (i = 1, \ldots, 7)
\end{aligned}
$$

例子修改自Andras London <Applications of Linear Programming>

博弈论

玩石头-剪刀-布的游戏，赢了得1元、输了付1元，已知对手一定会从“石头”、“剪刀”、“布”中选一个。你可以采取随机的策略，即以p_R, p_P, p_S的概率分别出“石头”、“布”、“剪刀”。问什么策略可以最大化在最差情况下的收益（即maximize worst-case payoff）？

你的收益矩阵

你 \ 对方	石头	布	剪刀
石头	0	-1	1
布	1	0	-1
剪刀	-1	1	0

例子取自Ron Parr CPS 570 <Linear Programming and Game Theory>

博弈论

你 \ 对方	石头	布	剪刀
石头	0	-1	1
布	1	0	-1
剪刀	-1	1	0

已知对手一定会从“石头”、“剪刀”、“布”中选一个。你可以以p_R, p_P, p_S的概率分别出“石头”、“布”、“剪刀”。什么策略可以最大化在最差情况下的收益?

对方出“石头”，收益为 $p_P - p_S$

对方出“布”，收益为 $-p_R + p_S$

对方出“剪刀”，收益为 $p_R - p_P$

博弈论

对方 / 你	石头	布	剪刀
石头	0	-1	1
布	1	0	-1
剪刀	-1	1	0

已知对手一定会从“石头”、“剪刀”、“布”中选一个。你可以以p_R, p_P, p_S的概率分别出“石头”、“布”、“剪刀”。什么策略可以最大化在最差情况下的收益？

对方出“石头”，收益为 $p_P - p_S$
对方出“布”，收益为 $-p_R + p_S$
对方出“剪刀”，收益为 $p_R - p_P$

$$
\begin{aligned}
\max\quad & u \\
\text{s.t.}\quad & u \le p_P - p_S, \\
& u \le -p_R + p_S, \\
& u \le p_R - p_P, \\
& p_P + p_S + p_R = 1, \\
\text{var.}\quad & 1 \ge p_P \ge 0, 1 \ge p_S \ge 0, 1 \ge p_R \ge 0, u.
\end{aligned}
$$

博弈论

你 \ 对方	石头	布	剪刀
石头	0	-1	1
布	1	0	-1
剪刀	-1	1	0

已知对手一定会从“石头”、“剪刀”、“布”中选一个。你可以以p_R, p_P, p_S的概率分别出“石头”、“布”、“剪刀”。什么策略可以最大化在最差情况下的收益?

$$
\begin{aligned}
\max \quad & u \\
\text{s.t.} \quad & u \le p_P - p_S, \\
& u \le -p_R + p_S, \\
& u \le p_R - p_P, \\
& p_P + p_S + p_R = 1, \\
\text{var.} \quad & 1 \ge p_P \ge 0, 1 \ge p_S \ge 0, 1 \ge p_R \ge 0, u.
\end{aligned}
$$

求解可得$p_P^* = p_S^* = p_R^* = \frac{1}{3}$以及$u^* = 0$

例子

选址问题：已知在一个平面坐标系内已经放置了四个机器，坐标为(3, 0), (0, -3), (-2, 1), (1, 4)。需要为一个新的机器选址，最小化它到其余四个机器距离之和。要求考虑曼哈顿距离，比如(x, y)到(3, 0)距离为$|x-3|+|y|$

思考是否可能整理成线性规划问题？

线性规划

在所有类型的优化问题中，线性规划或许是最兼具易解性与适用性的问题

- 易解性：局部最优解一定是全局最优解；最优解若存在，可行域顶点中一定有最优解 ...

- 适用性：广泛存在于交通运输、生产销售、排班调度等实际应用中

得益于工业界和学术界的广泛关注，已有多种多项式时间算法可以高效求得大规模（如包含上千万个决策变量）线性规划问题的全局最优解

二元线性规划的求解（图解法）

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、KKT条件、内点法
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

栏目索引

考向二　线性目标函数的最值问题

例2　(2018课标Ⅱ,14,5分)若x,y满足约束条件$\begin{cases}x+2y-5\geq 0,\\ x-2y+3\geq 0,\\ x-5\leq 0,\end{cases}$则$z=x+y$的最大值为________.

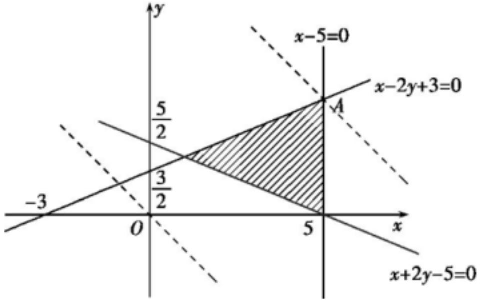
二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

考向二　线性目标函数的最值问题

例2　(2018课标Ⅱ,14,5分)若x,y满足约束条件$\begin{cases}x+2y-5\geq 0,\\x-2y+3\geq 0,\\x-5\leq 0,\end{cases}$则$z=x+y$的最大值为________.

解析　由线性约束条件画出可行域(如图中阴影部分所示).



当直线$x+y-z=0$经过点$A(5,4)$时,$z=x+y$取得最大值,最大值为9.

答案　9

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

画可行域 → 作直线 → 平移直线找最优解 → 得到目标函数最优值

讨论最优解的存在性和唯一性？

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

画可行域 → 作直线 → 平移直线找最优解 → 得到目标函数最优值

讨论最优解的存在性和唯一性？(1) 存在且唯一

$$
\begin{aligned}
\max \quad & x_1+x_2 \\
\text{s.t.} \quad & x_1+2x_2-5 \geq 0, \\
& x_1-2x_2+3 \geq 0, \\
& x_1-5 \leq 0. \\
\text{var.} \quad & x_1, x_2
\end{aligned}
$$

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

画可行域 → 作直线 → 平移直线找最优解 → 得到目标函数最优值

讨论最优解的存在性和唯一性？(1) 存在且唯一；(2) 存在但不唯一 (无穷个)

$$
\begin{aligned}
\text{max}\quad & x_1 \\
\text{s.t.}\quad & x_1+2x_2-5\geq 0, \\
& x_1-2x_2+3\geq 0, \\
& x_1-5\leq 0. \\
\text{var.}\quad & x_1, x_2
\end{aligned}
$$

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

画可行域 → 作直线 → 平移直线找最优解 → 得到目标函数最优值

讨论最优解的存在性和唯一性？(1) 存在且唯一；(2) 存在但不唯一 (无穷个)；(3) 不存在-可行域非空但无界

$$
\begin{aligned}
\text{max}\quad & x_1+x_2\\
\text{s.t.}\quad & x_1+2x_2-5\ge 0,\\
& x_1-2x_2+3\ge 0.\\
\text{var.}\quad & x_1, x_2
\end{aligned}
$$

注意：不是所有可行域非空无界时解都不存在

二元线性规划

当线性规划问题只有两个变量时，可以用图解法求解

画可行域	→	作直线	→	平移直线找最优解	→	得到目标函数最优值

讨论最优解的存在性和唯一性？(1) 存在且唯一；(2) 存在但不唯一 (无穷个)；

(3) 不存在-可行域非空但无界；(4) 不存在-可行域为空

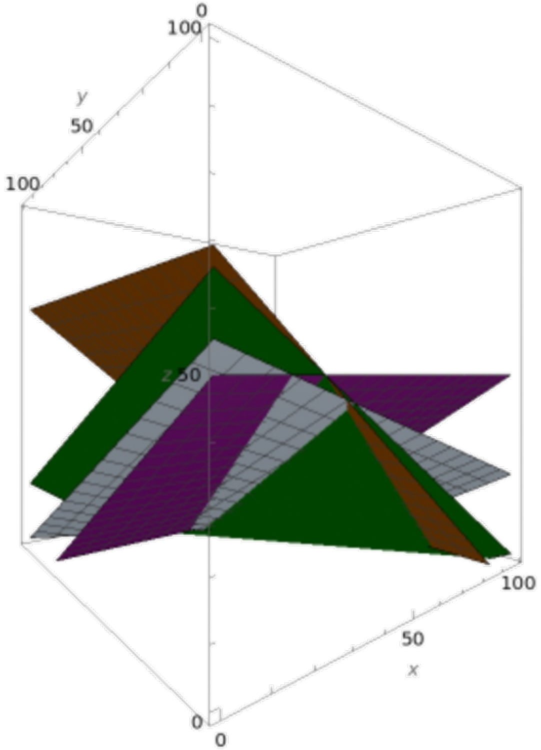
$$
\begin{aligned}
\text{max} \quad & x_1+x_2 \\
\text{s.t.} \quad & x_1+2x_2-5 \geq 0, \\
& x_1-2x_2+3 \geq 0, \\
& x_1 \leq 0. \\
\text{var.} \quad & x_1, x_2
\end{aligned}
$$

二元线性规划的求解（单纯形法）

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、KKT条件、内点法
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

图解法的缺点

图解法只能用于求解二元线性规划



例如，三个变量时，每个约束对应一个平面

远远无法解决大量实际中的线性规划问题

图解法的缺点

实际中的许多线性规划问题规模巨大

Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods

Robert E. Bixby[1] John W. Gregory[2] Irvin J. Lustig[3]
Roy E. Marsten[4] David F. Shanno[5]

May, 1991[6]

Abstract

Experience with solving a 12,753,313 variable linear program is described. This problem is the linear programming relaxation of a set partitioning problem arising from an airline crew scheduling application. A scheme is described that requires successive solutions of small subproblems, yielding a procedure that has little growth in solution time in terms of the number of variables. Experience using the simplex method as implemented in CPLEX, an interior point method as implemented in OB1, and a hybrid interior point/simplex approach is reported. The resulting procedure illustrates the power of an interior point/simplex combination for solving very large-scale linear programs.

航空公司机组人员调度问题

一个航空公司平均每月要对超过2000个驾驶员、5000个空乘进行排班

- 需要完整覆盖往来100多个城市的100多个航班
- 每个员工根据工作合同有不同的工作/休息天数要求，如工作四天休两天
- 航线飞行时长影响休息天数

雇佣机组人员的开销是航空公司的第二大开销

例子取自Bruno F. Santos “AE4423 Lect 6.1 Crew Scheduling Context”视频

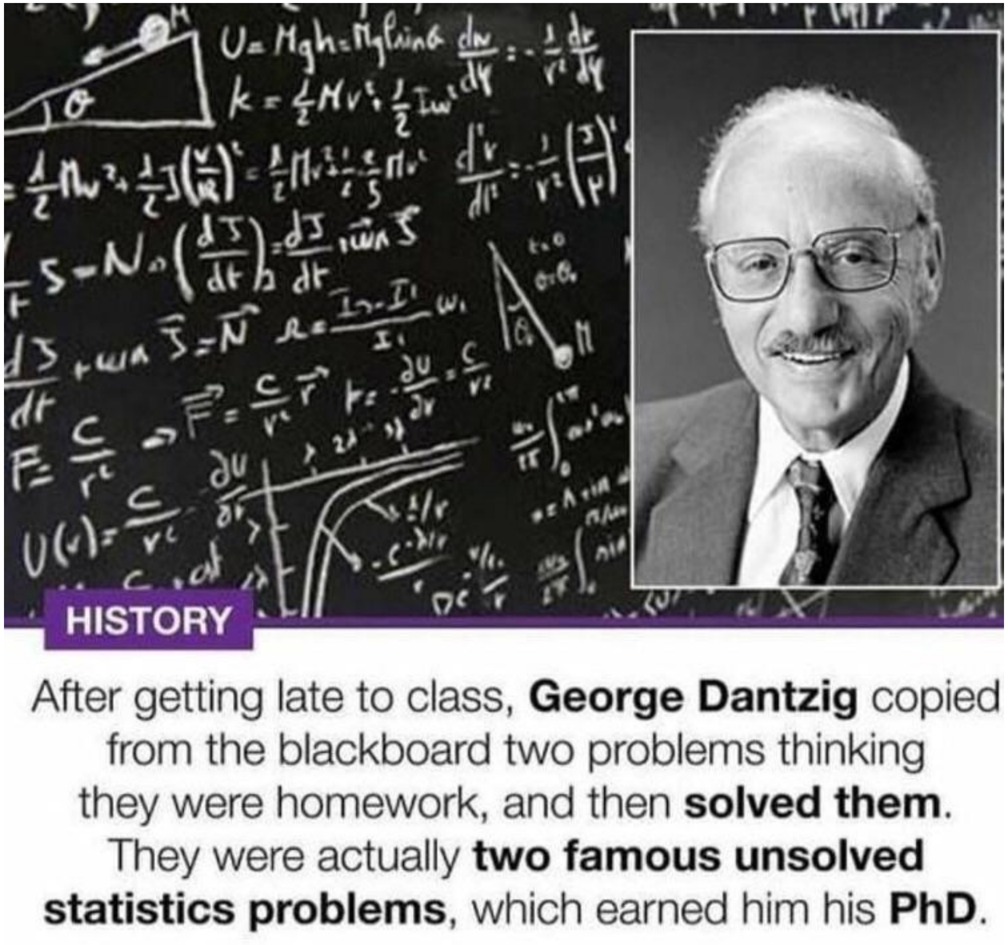
航空公司机组人员调度问题

Example of a crew roster for multiple crews (rows).

Note: NA - a previous pairing was assigned in the past

	MON	TUE	WED	THU	FRI	SAT	SUN	MON	TUE	WED	THU	FRI	SAT	SUN	MON	TUE
L_001	NA	NA	NA	NA	NA	PA_0059	CA_004	CA_004	RP	RP	RP	RP				
L_002	NA	PA_0011	ME_002	ME_002	RP	RP	RP									
L_003	NA	NA	NA	PA_0034	CA_002	CA_002	CA_002	RP	RP	RP						
L_004	PA_0009	NA_002	NA_002	NA_002	NA_002	RP	RP									
L_005	NA	NA	NA	NA	PA_0048	CA_001	CA_001	CA_001	RP	RP	RP					
L_006	NA	NA	NA	NA	NA	PA_0054	NA_005	NA_005	NA_005	RP	RP	RP				
L_007	NA	NA			PA_0045	AF_002	AF_002	AF_002	AF_002	AF_002	RP	RP	RP			
L_008	NA	NA	NA	NA	NA	NA	PA_0068	AF_004	AF_004	AF_004	AF_004	RP	RP			
L_009			PA_0025	ME_002	ME_002	RP	RP	RP								
L_010	PA_0010	AF_001	AF_001	AF_001	RP	RP	RP									
L_011	NA	NA	NA				PA_0065	NA_003	NA_003	NA_003	NA_003	NA_003	RP	RP	RP	RP
L_012	NA	NA	NA	NA	PA_0047	ME_002	ME_002	RP	RP	RP						
L_013	NA	NA	NA	NA	NA	PA_0051	ME_002	ME_002	ME_002	RP	RP	RP	RP			
L_014	NA	NA	NA	PA_0037	NA_002	NA_002	NA_002	RP	RP	RP						
L_015	NA	NA	NA	NA	PA_0046	AF_001	AF_001	AF_001	RP	RP	RP					
L_016	NA	NA	NA	NA												
L_017	NA	NA	PA_0026	CA_004	CA_004	RP	RP	RP	RP							
L_018	NA	NA	NA	PA_0038	ME_003	ME_003	ME_003	ME_003	RP	RP						
L_019			PA_0023	AF_003	AF_003	RP	RP									
L_020	NA	NA	NA	NA	NA	NA	PA_0070	NA_001	NA_001	NA_001	RP	RP	RP			
L_021	NA	NA	NA	NA	NA	NA										
L_022	NA	NA	NA	NA	PA_0050	NA_004	NA_004	NA_004	RP	RP	RP					
L_023	NA	NA	NA	NA	NA	NA	PA_0063	AF_003	AF_003	RP	RP					
L_024			PA_0020	NA_005	NA_005	NA_005	NA_005	RP	RP	RP						
L_025		PA_0013	ME_003	ME_003	ME_003	ME_003	ME_003	RP	RP	RP						
L_026	NA	NA	NA	NA	NA	PA_0056	NA_002	NA_002	NA_002	RP	RP	RP				
L_027	PA_0004	CA_004	CA_004	RP	RP	RP	RP									
L_028	NA	NA	PA_0029	CA_003	CA_003	RP	RP	RP	RP							
L_029	NA	NA	NA	NA	NA	PA_0052	AF_001	AF_001	AF_001	RP	RP	RP				
L_030	NA	NA	NA	NA	NA	NA										
L_031	NA	PA_0015	ME_001	ME_001	RP	RP	RP									
L_032	NA	NA	NA	NA	NA	NA	NA									
L_033	NA	NA	NA	PA_0032	ME_001	ME_001	RP	RP	RP							
L_034	NA	NA	NA	NA	NA	PA_0058	CA_002	CA_002	CA_002	RP	RP	RP				
L_035		PA_0017	AF_002	AF_002	AF_002	AF_002	AF_002	AF_002	RP	RP	RP	RP				
L_036					PA_0041	CA_003	CA_003	CA_003	RP	RP	RP					

单纯形法



1947年，George Dantzig提出解决线性规划问题的单纯形法

单纯形法

Stanford | News

Home Find Stories For Journalists Contact

Stanford Report, May 25, 2005

George B. Dantzig, operations research professor, dies at 90

BY DAWN LEVY

During his first year as a doctoral student at the University of California-Berkeley, Dantzig arrived late to the class of Jerzy Neyman, one of the great founders of modern statistics. On the blackboard were two problems that Dantzig assumed to be homework.

"A few days later I apologized to Neyman for taking so long to do the homework—the problems seemed harder to do than usual," Dantzig once recalled. It turned out the conundrums, which Dantzig solved, were two famous unsolved problems in statistics.

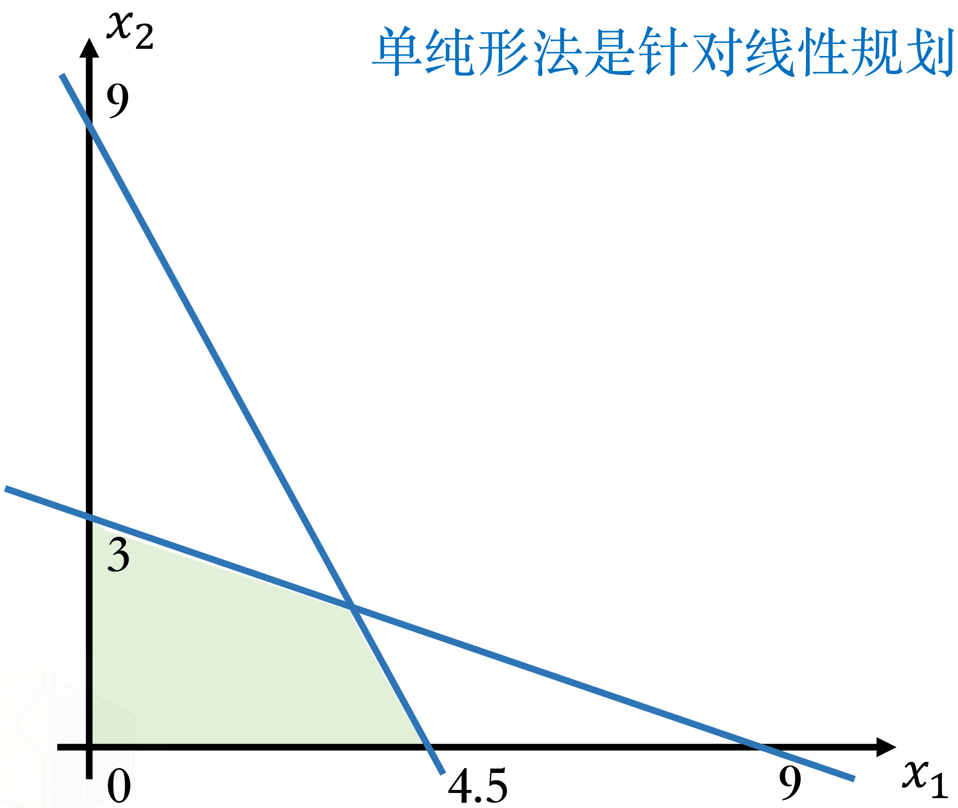
单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

单纯形法解二元线性规划

$$
\begin{aligned}
\min\quad & -3x_1-2x_2\\
\text{s.t.}\quad & 2x_1+x_2\le 9,\\
& x_1+3x_2\le 9,\\
\text{var.}\quad & x_1\ge 0, x_2\ge 0.
\end{aligned}
\quad\Rightarrow\quad
\begin{aligned}
\min\quad & -3x_1-2x_2\\
\text{s.t.}\quad & 2x_1+x_2+x_3=9,\\
& x_1+3x_2+x_4=9,\\
\text{var.}\quad & x_1\ge 0, x_2\ge 0, x_3\ge 0, x_4\ge 0.
\end{aligned}
$$

单纯形法是针对线性规划的标准形式求解

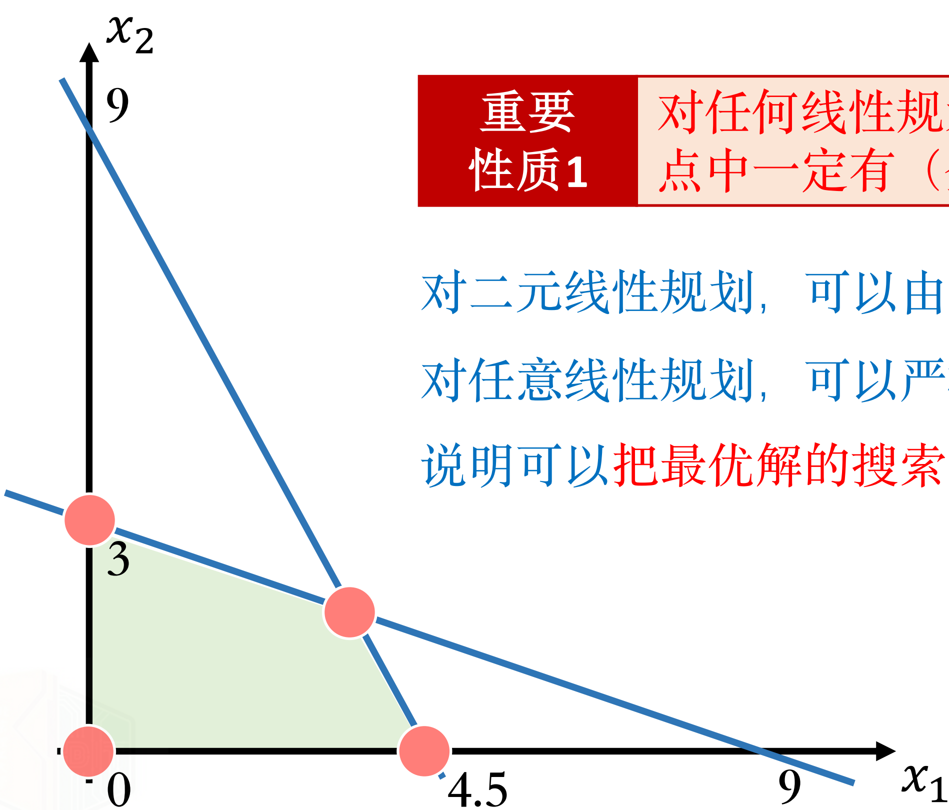


例子修改自Yong Wang

“Operations Research-The Simplex Method”视频

单纯形法解二元线性规划

$$
\begin{aligned}
\min\quad & -3x_1-2x_2\\
\text{s.t.}\quad & 2x_1+x_2 \le 9,\\
& x_1+3x_2 \le 9,\\
\text{var.}\quad & x_1 \ge 0, x_2 \ge 0.
\end{aligned}
\quad\Rightarrow\quad
\begin{aligned}
\min\quad & -3x_1-2x_2\\
\text{s.t.}\quad & 2x_1+x_2+x_3 = 9,\\
& x_1+3x_2+x_4 = 9,\\
\text{var.}\quad & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0.
\end{aligned}
$$



重要性质1 对任何线性规划问题，最优解若存在，可行域的顶点中一定有（全局）最优解。

对二元线性规划，可以由图解法理解该结论

对任意线性规划，可以严格证明该性质

说明可以把最优解的搜索范围缩小到所有顶点构成的集合

单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

重要性质1	对任何线性规划问题，最优解若存在，可行域的顶点中一定有最优解。

单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

重要性质1	对任何线性规划问题，最优解若存在，可行域的顶点中一定有最优解。

如何根据标准形式找到所有顶点？

单纯形法解二元线性规划

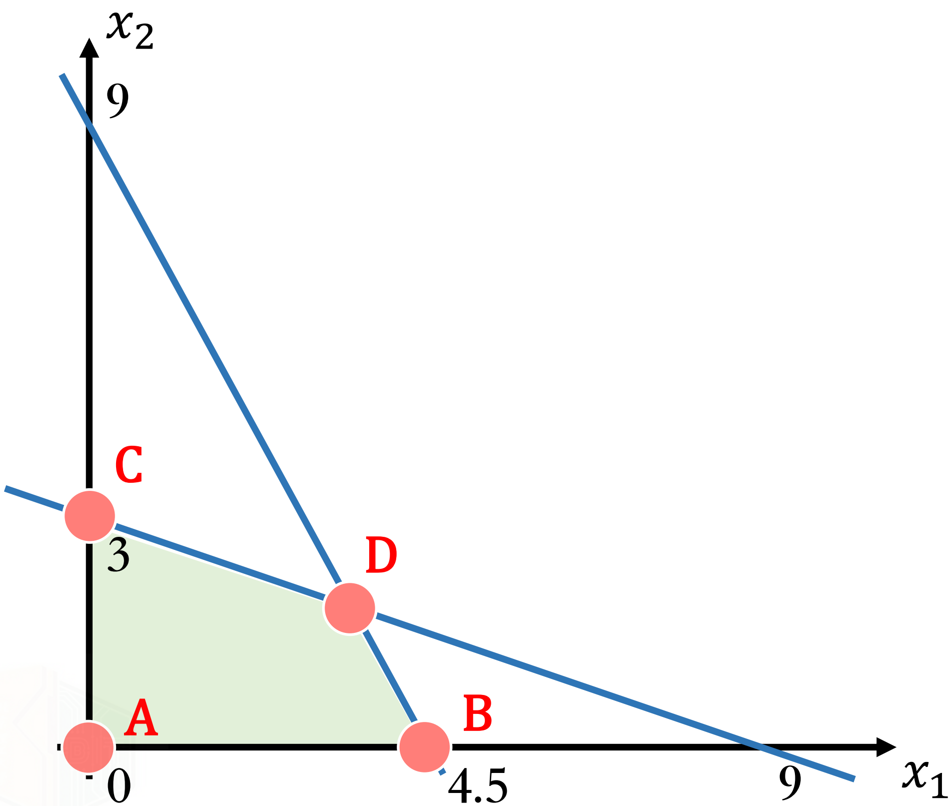
min $\quad -3x_1 - 2x_2$

s.t. $\quad 2x_1 + x_2 + x_3 = 9,$

$\quad x_1 + 3x_2 + x_4 = 9,$

var. $\quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0.$

如何根据标准形式找到所有顶点？



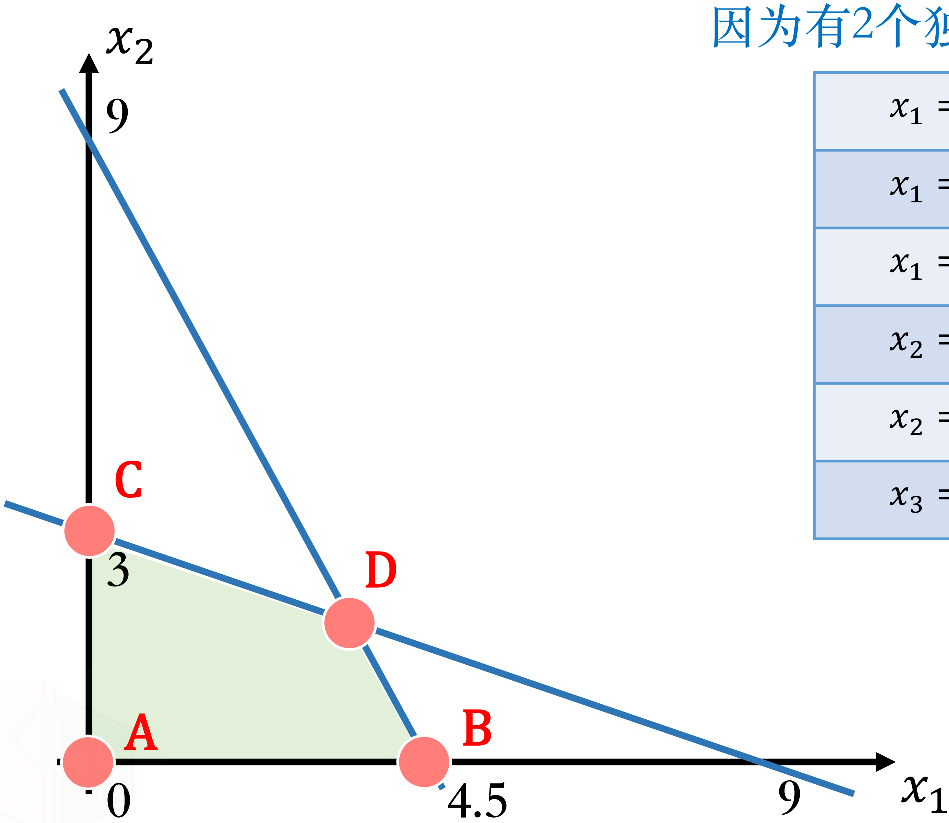
单纯形法解二元线性规划

$$\begin{aligned}
\text{min}\quad & -3x_1 - 2x_2 \\
\text{s.t.}\quad & 2x_1 + x_2 + x_3 = 9, \\
& x_1 + 3x_2 + x_4 = 9, \\
\text{var.}\quad & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0.
\end{aligned}$$

如何根据标准形式找到所有顶点？

因为有2个独立等式4个变量，将任意4-2=2个变量设为0

$x_1 = x_2 = 0$	$x_3 = 9, x_4 = 9$	对应顶点A
$x_1 = x_3 = 0$	$x_2 = 9, \boxed{x_4 = -18}$	非可行解
$x_1 = x_4 = 0$	$x_2 = 3, x_3 = 6$	对应顶点C
$x_2 = x_3 = 0$	$x_1 = 4.5, x_4 = 4.5$	对应顶点B
$x_2 = x_4 = 0$	$x_1 = 9, \boxed{x_3 = -9}$	非可行解
$x_3 = x_4 = 0$	$x_1 = 18/5, x_2 = 9/5$	对应顶点D



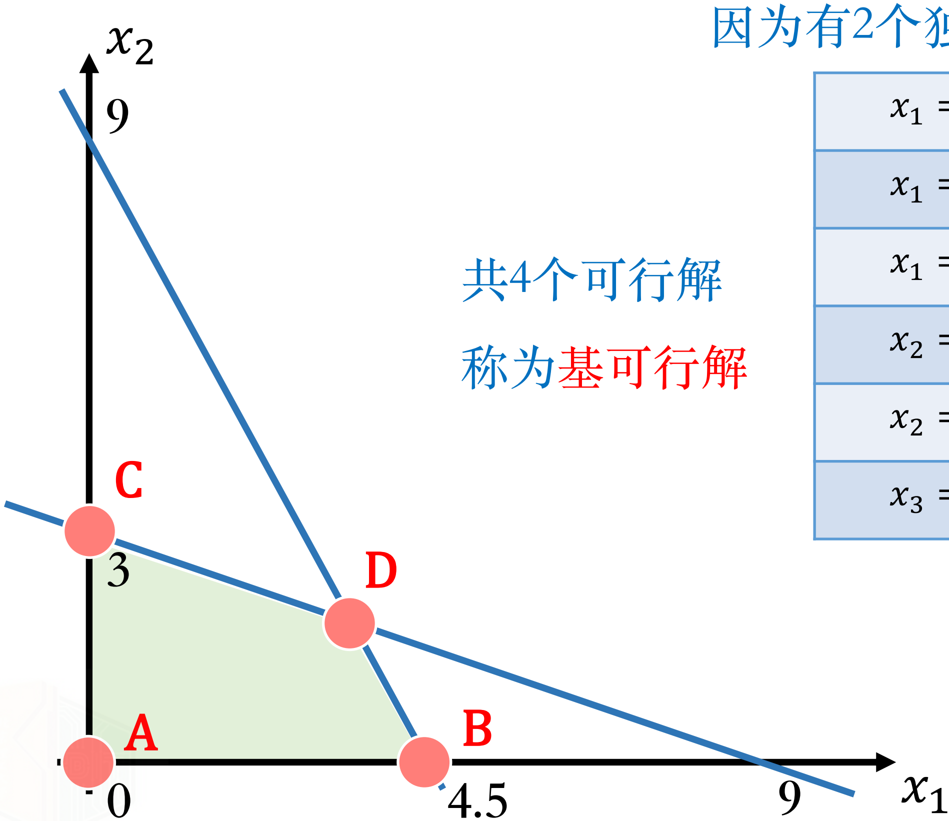
单纯形法解二元线性规划

$$\begin{aligned}
\min\quad & -3x_1 - 2x_2 \\
\text{s.t.}\quad & 2x_1 + x_2 + x_3 = 9, \\
& x_1 + 3x_2 + x_4 = 9, \\
\text{var.}\quad & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0.
\end{aligned}$$

如何根据标准形式找到所有顶点？

因为有2个独立等式4个变量，将任意4-2=2个变量设为0

$x_1 = x_2 = 0$	$x_3 = 9, x_4 = 9$	对应顶点A
$x_1 = x_3 = 0$	$x_2 = 9, \boxed{x_4 = -18}$	非可行解
$x_1 = x_4 = 0$	$x_2 = 3, x_3 = 6$	对应顶点C
$x_2 = x_3 = 0$	$x_1 = 4.5, x_4 = 4.5$	对应顶点B
$x_2 = x_4 = 0$	$x_1 = 9, \boxed{x_3 = -9}$	非可行解
$x_3 = x_4 = 0$	$x_1 = 18/5, x_2 = 9/5$	对应顶点D



共4个可行解

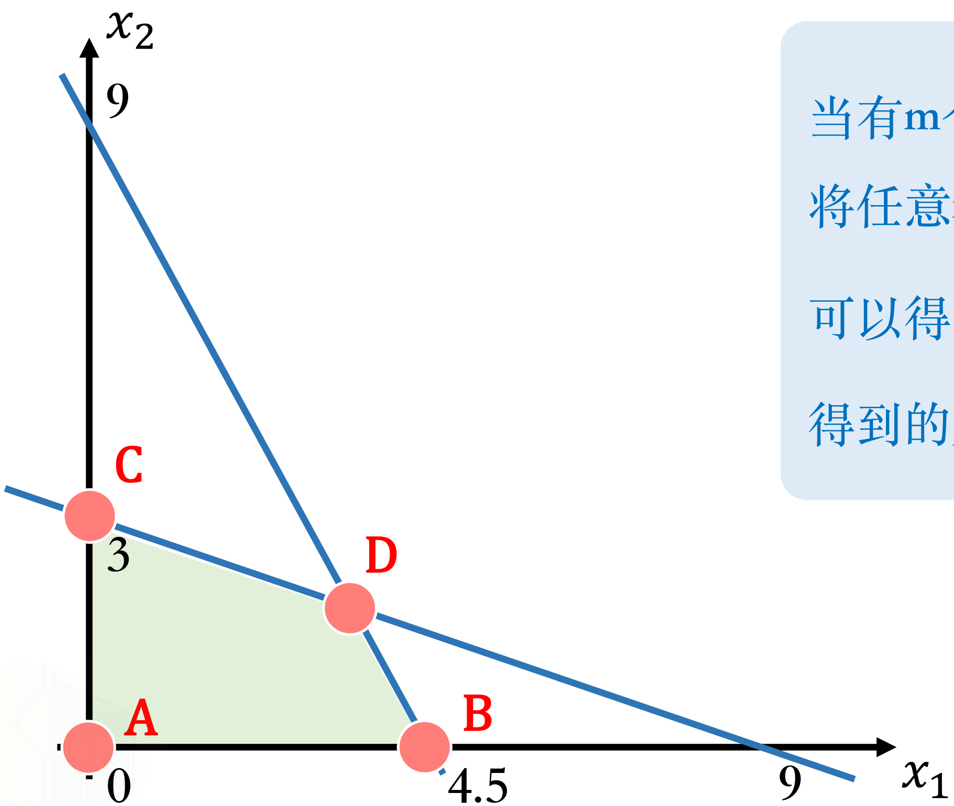
称为基可行解

重要性质2	基可行解与顶点对应

单纯形法解二元线性规划

$$
\begin{aligned}
\text{min}\quad & -3x_1-2x_2\\
\text{s.t.}\quad & 2x_1+x_2+x_3=9,\\
& x_1+3x_2+x_4=9,\\
\text{var.}\quad & x_1\ge 0, x_2\ge 0, x_3\ge 0, x_4\ge 0.
\end{aligned}
$$

如何根据标准形式找到所有顶点？



当有m个独立等式和n个变量(m⩽n)：

将任意n-m个变量设为0，计算其余m个变量取值，

可以得到共$\frac{n!}{m!(n-m)!}$个解，剔除其中非可行解，

得到的所有**基可行解**即所有**顶点**

单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

重要性质1	对任何线性规划问题，最优解若存在，可行域的顶点中一定有最优解。

如何根据标准形式找到所有顶点？

重要性质2	基可行解与顶点对应

单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

重要性质1　对任何线性规划问题，最优解若存在，可行域的顶点中一定有最优解。

如何根据标准形式找到所有顶点？

重要性质2　基可行解与顶点对应

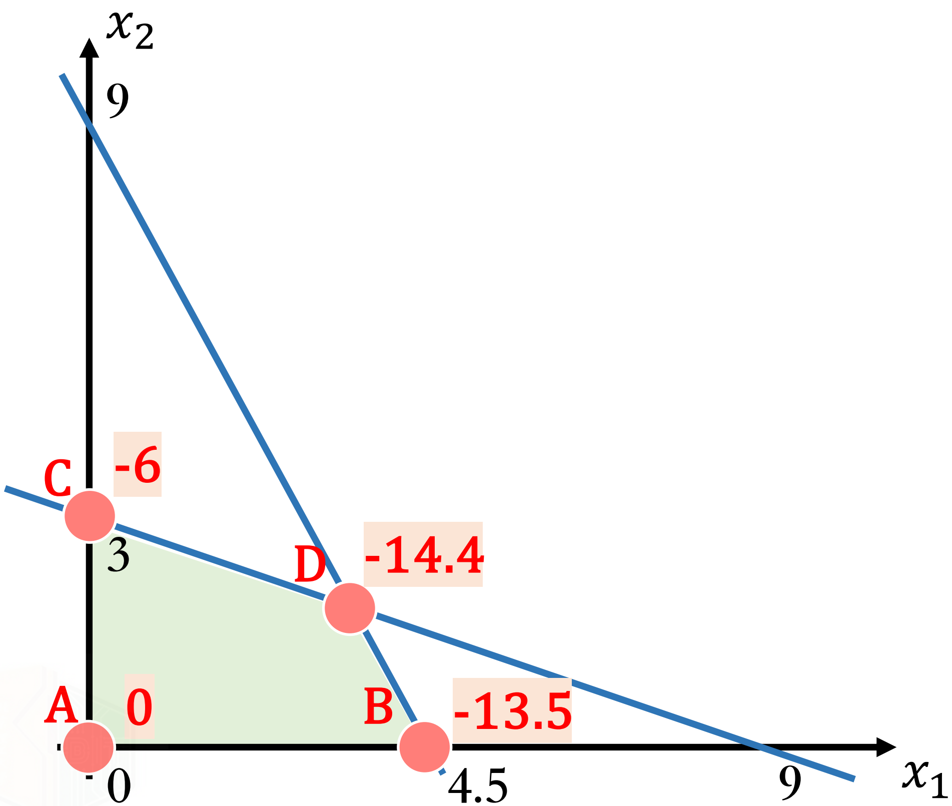
检查所有顶点（基可行解）？或按照何种顺序进行搜索？

单纯形法解二元线性规划

$$
\begin{aligned}
\text{min} \quad & -3x_1 - 2x_2 \\
\text{s.t.} \quad & 2x_1 + x_2 + x_3 = 9, \\
& x_1 + 3x_2 + x_4 = 9, \\
\text{var.} \quad & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0.
\end{aligned}
$$

检查所有顶点（基可行解）？

或按照何种顺序进行搜索？



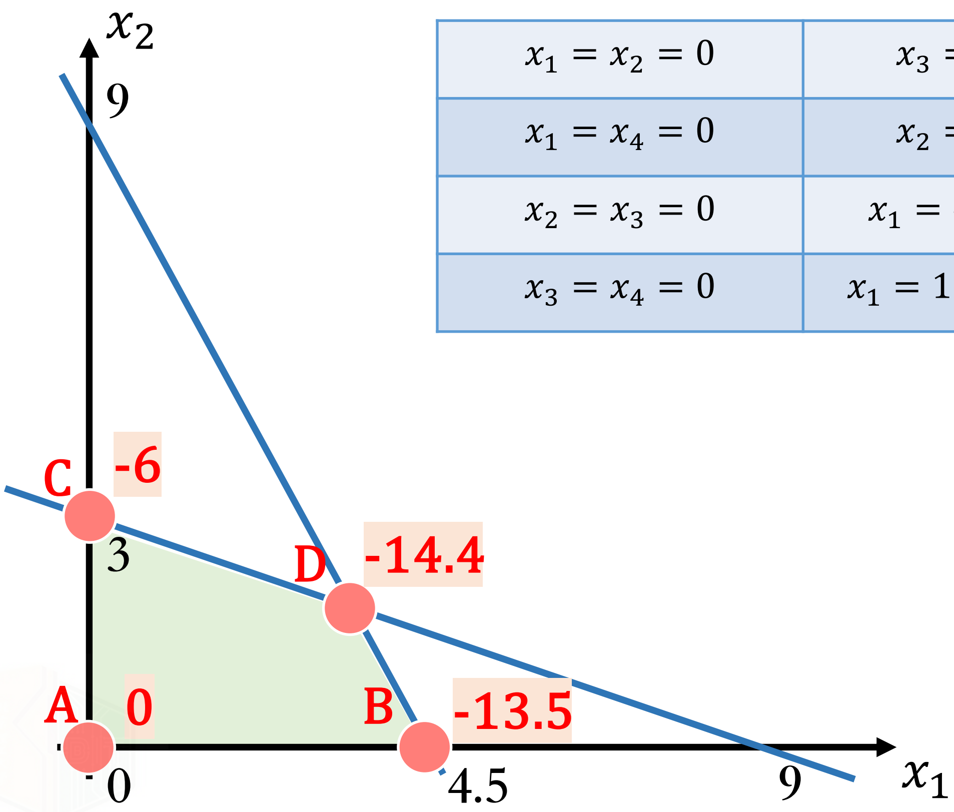
单纯形法解二元线性规划

$$
\begin{aligned}
\text{min} \quad & -3x_1 - 2x_2 \\
\text{s.t.} \quad & 2x_1 + x_2 + x_3 = 9, \\
& x_1 + 3x_2 + x_4 = 9, \\
\text{var.} \quad & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0.
\end{aligned}
$$

检查所有顶点（基可行解）？

或按照何种顺序进行搜索？

$x_1 = x_2 = 0$	$x_3 = 9, x_4 = 9$	对应顶点A	目标函数= 0
$x_1 = x_4 = 0$	$x_2 = 3, x_3 = 6$	对应顶点C	目标函数= -6
$x_2 = x_3 = 0$	$x_1 = 4.5, x_4 = 4.5$	对应顶点B	目标函数= -13.5
$x_3 = x_4 = 0$	$x_1 = 18/5, x_2 = 9/5$	对应顶点D	目标函数= -14.4

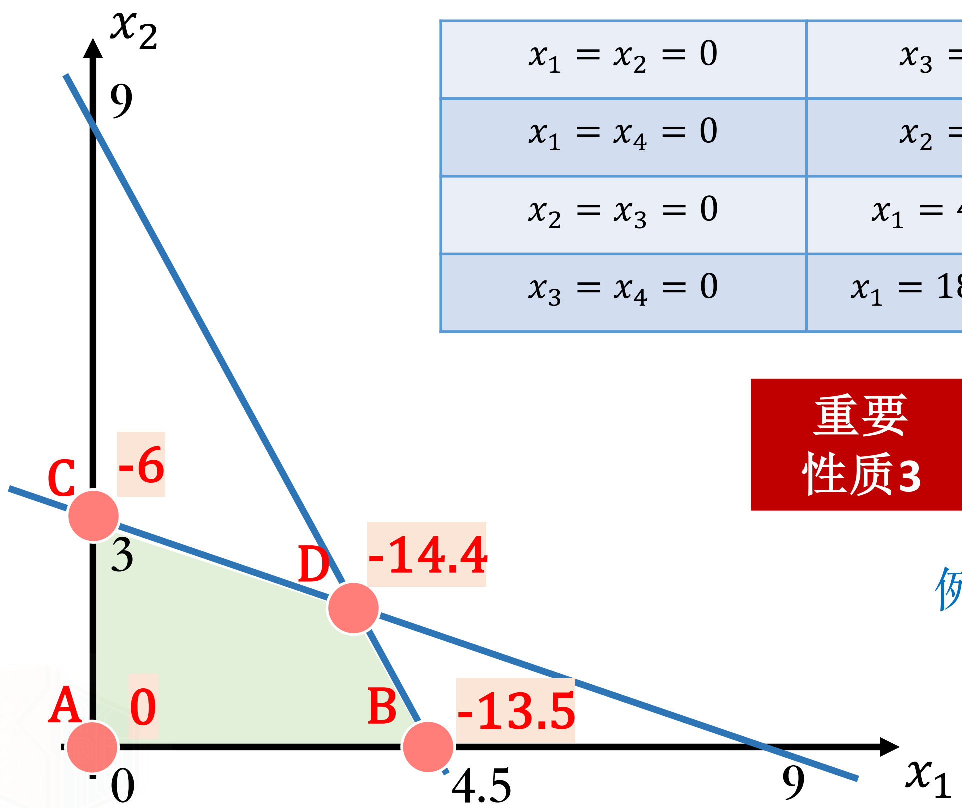


单纯形法解二元线性规划

$$\begin{aligned}
\text{min} \quad & -3x_1 - 2x_2 \\
\text{s.t.} \quad & 2x_1 + x_2 + x_3 = 9, \\
& x_1 + 3x_2 + x_4 = 9, \\
\text{var.} \quad & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0.
\end{aligned}$$

检查所有顶点（基可行解）？

或按照何种顺序进行搜索？



$x_1 = x_2 = 0$	$x_3 = 9, x_4 = 9$	对应顶点A	目标函数$= 0$
$x_1 = x_4 = 0$	$x_2 = 3, x_3 = 6$	对应顶点C	目标函数$= -6$
$x_2 = x_3 = 0$	$x_1 = 4.5, x_4 = 4.5$	对应顶点B	目标函数$= -13.5$
$x_3 = x_4 = 0$	$x_1 = 18/5, x_2 = 9/5$	对应顶点D	目标函数$= -14.4$

重要性质3	按目标函数减小的方向，沿相邻顶点搜索，可以达到最优顶点（最优解）

例如，顶点A 到 顶点B 到 顶点D

单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

重要性质1：对任何线性规划问题，最优解若存在，可行域的顶点中一定有最优解。

如何根据标准形式找到所有顶点？

重要性质2：基可行解与顶点对应

检查所有顶点（基可行解）？或按照何种顺序进行搜索？

重要性质3：按目标函数减小的方向，沿相邻顶点搜索，可以达到最优顶点（最优解）

单纯形法解二元线性规划

是否可以缩小最优解的搜索空间？

重要性质1：对任何线性规划问题，最优解若存在，可行域的顶点中一定有最优解。

如何根据标准形式找到所有顶点？

重要性质2：基可行解与顶点对应

检查所有顶点（基可行解）？或按照何种顺序进行搜索？

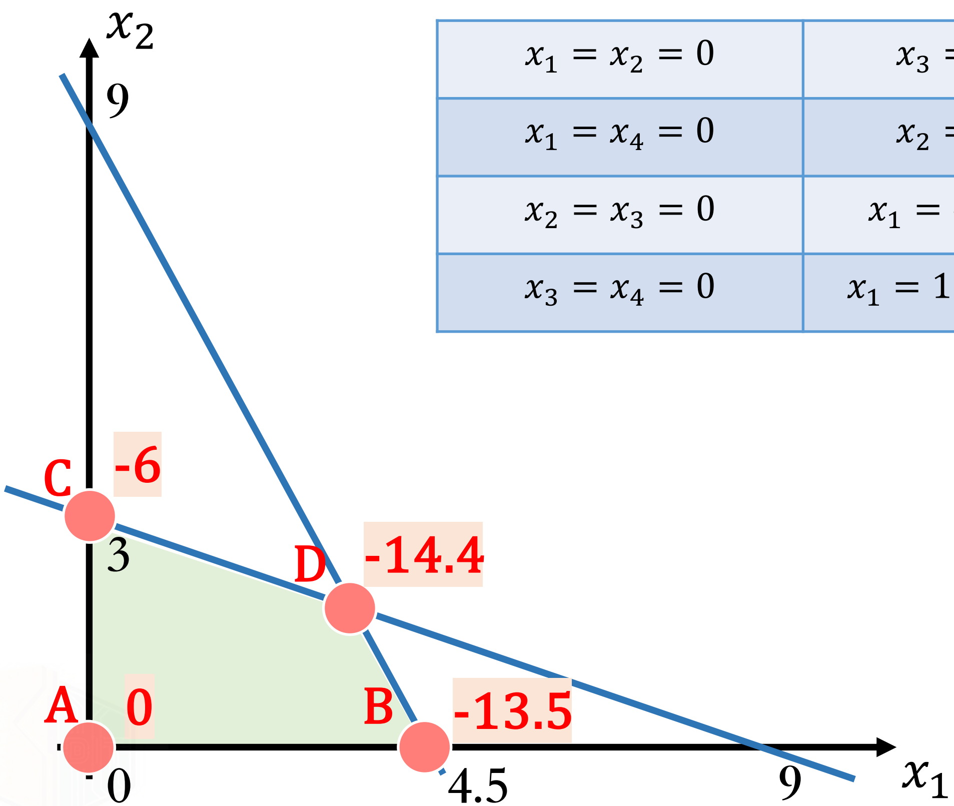
重要性质3：按目标函数减小的方向，沿相邻顶点搜索，可以达到最优顶点（最优解）

如何判断哪些基可行解对应的是相邻顶点？

单纯形法解二元线性规划

$$
\begin{aligned}
\text{min} \quad & -3x_1-2x_2 \\
\text{s.t.} \quad & 2x_1+x_2+x_3=9, \\
& x_1+3x_2+x_4=9, \\
\text{var.} \quad & x_1\ge 0, x_2\ge 0, x_3\ge 0, x_4\ge 0.
\end{aligned}
$$

如何判断哪些基可行解对应的是相邻顶点？

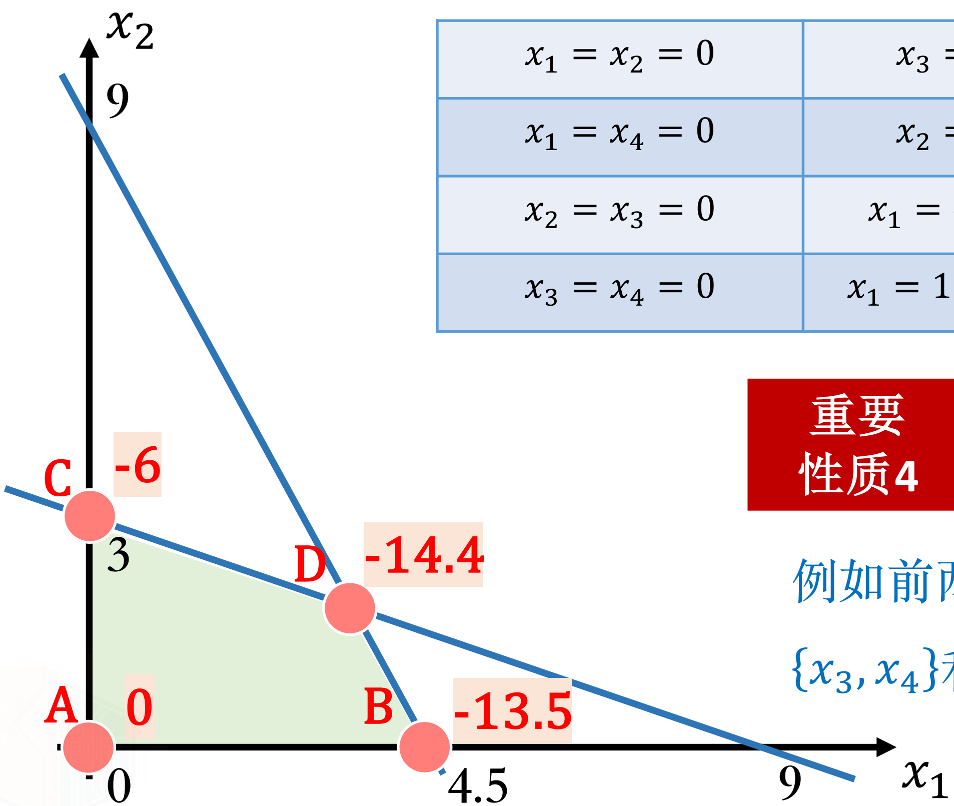


$x_1=x_2=0$	$x_3=9, x_4=9$	对应顶点A	目标函数$=0$
$x_1=x_4=0$	$x_2=3, x_3=6$	对应顶点C	目标函数$=-6$
$x_2=x_3=0$	$x_1=4.5, x_4=4.5$	对应顶点B	目标函数$=-13.5$
$x_3=x_4=0$	$x_1=18/5, x_2=9/5$	对应顶点D	目标函数$=-14.4$

单纯形法解二元线性规划

$$
\begin{aligned}
\text{min}\quad & -3x_1-2x_2\\
\text{s.t.}\quad & 2x_1+x_2+x_3=9,\\
& x_1+3x_2+x_4=9,\\
\text{var.}\quad & x_1\ge 0, x_2\ge 0, x_3\ge 0, x_4\ge 0.
\end{aligned}
$$

如何判断哪些基可行解对应的是相邻顶点？



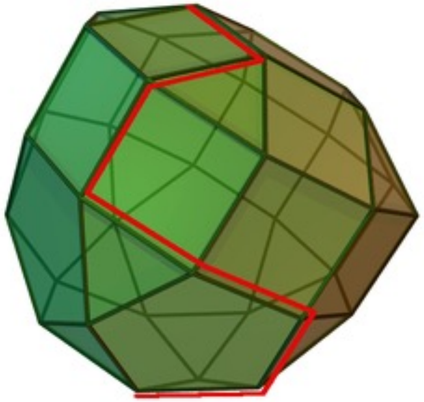
$x_1=x_2=0$	$x_3=9, x_4=9$	对应顶点A	目标函数$=0$
$x_1=x_4=0$	$x_2=3, x_3=6$	对应顶点C	目标函数$=-6$
$x_2=x_3=0$	$x_1=4.5, x_4=4.5$	对应顶点B	目标函数$=-13.5$
$x_3=x_4=0$	$x_1=18/5, x_2=9/5$	对应顶点D	目标函数$=-14.4$

重要性质4 通常，两个恰好相差一个非零变量的基可行解对应两个相邻的顶点

例如前两个基可行解，分别对应的非零变量为$\{x_3,x_4\}$和$\{x_2,x_3\}$，它们对应的顶点A和C相邻

单纯形法解二元线性规划

重要性质1	对任何线性规划问题，最优解若存在，可行域的顶点中一定有（全局）最优解。
重要性质2	基可行解与顶点对应
重要性质3	按目标函数减小的方向，沿相邻顶点搜索，可以达到最优顶点（最优解）
重要性质4	通常，两个恰好相差一个非零变量的基可行解对应两个相邻的顶点



单纯形法的核心思想:
按目标函数减小的方向，沿一系列对应相邻顶点的基可行解搜索

本讲小结

- 梯度下降法、牛顿法

- 线性规划的定义、例子

- 单纯形法

主要参考资料

Bo Jiang <CS257 Linear and Convex Optimization> Slides

www.geogebra.org

Wikipedia Taylor‘s theorem词条

Stephen Boyd <Convex Optimization> Book

《数学建模》全国赛培训课件

Andras London <Applications of Linear Programming> Book

Ron Parr <CPS 570 Linear Programming and Game Theory> Slides

《5年高考，3年模拟》

Bruno F. Santos <AE4423 Lect 6.1 Crew Scheduling Context> Video

Yong Wang <Operations Research-The Simplex Method> Video

Wen Shen <Math 484 Linear Programming> Video

谢谢！