组合优化与凸优化

第2章 线性规划

刘绍辉

计算机科学与技术学院 哈尔滨工业大学

shliu@hit.edu.cn

2023年春季

第2章 线性规划

- **2.0凸性简单介绍**
- **2.1线性规划的基本概念**
  - 线性规划问题及其数学模型
  - 线性规划模型的标准形式及解的概念
  - 图解法
  - 线性规划问题的基本理论
- **2.2 单纯形法**
  - 单纯形法基本原理
  - 单纯形表
  - 大M法与两阶段法
- **2.3 对偶理论**
  - 对偶单纯形法
- **2.4 灵敏度分析**
  - 灵敏度分析

第2章 线性规划

◆ 问题的一般描述

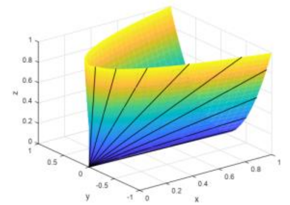
- $\boldsymbol{min\ f_0(x), s.t. f_j(x) \le (\ge \setminus =)0, j = 1 \cdots m, x \in Q}$,$\boldsymbol{Q}$表示可行集
- 约束优化问题
- 无约束优化问题
- 光滑优化问题：可微
- 非光滑优化问题：有些不可微
- 线性优化问题：约束是仿射的（**Ax+b**）
  - ✓ 如果$\boldsymbol{f_0}$也是仿射的，称为线性规划问题
  - ✓ 如果$\boldsymbol{f_0}$是二次函数，则为二次规划问题
  - ✓ 如果所有函数$\boldsymbol{f_0, \cdots, f_m}$都是二次函数，称为二次约束的二次规划问题
- 若决策变量必需是整数，则可以通过如下约束描述
  - ✓ $\boldsymbol{sin\left(\pi x^{(i)}\right) = 0, i = 1, \cdots, n}$，整数优化问题

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- 矩阵A的秩表示为$rank(A)$,非空集合$C \subseteq R^n$的维数$dimC$定义为：
  - $n - max\{rank(A): A \in R^{n\times n}, Ax = Ay$,对所有$x, y \in C\}$
- 凸组合(convex combination):点$x_1, x_2, \cdots, x_n$的凸组合是指点$x = \Sigma_{i=1}^{n} \alpha_i x_i, \alpha_i \geq 0$, 且$\Sigma \alpha_i = 1$
- 仿射组合：点$x_1, x_2, \cdots, x_n$的仿射组合是指点$x = \Sigma_{i=1}^{n} \alpha_i x_i$, 且$\Sigma \alpha_i = 1$
- 锥组合：点x_1, x_2的锥组合是指形如$x = \alpha_1 x_1 + \alpha_2 x_2, \alpha_1 > 0, \alpha_2 > 0$
- 凸集(convex set):令$A \subseteq R^n$是凸的，如果$\forall x, y \in A$,则$\alpha x + (1-\alpha) y \in A$,$\alpha \in (0, 1)$
- 凸包（convex Hull)
  - 集合A的凸包$conv(A)$定义为A中点的所有凸组合.
  - 凸包是包含集合的最小凸集
  - 离散点的凸包示例，扇形凸包示例
- 仿射包（affine Hull)
  - 集合$A \subseteq R^n$的仿射包为A中点的组合：affine$A$ $:= \{x | x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k, x_1, x_2, \cdots, x_k \in A, \sum_i \alpha_i = 1\}$
  - 一般情况下，一个集合的仿射包实际上是包含该集合 的最小的仿射集

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- **极点(extreme point):点$x \in A$如果不能表示为非空凸集A中两个任何其它点的凸组合的形式，则称为极点**
  - 集合A的极点是这样的点$x \in A$，使得$x \notin conv(A\backslash\{x\})$
  - 集合A为凸，当且仅当A中点的所有凸组合仍在A中
  - 集合A的凸包是包含A的最小凸集
  - 凸集的交是否是凸集？凸集的并集是否是凸集？
- **凸锥(convex cone)**
  - 集合$A \subseteq \mathbb{R}^n$中任意元素$x, y \in A, \lambda, \mu \geq 0$,有$\lambda x + \mu y \in A$,则称集合$A$为一个凸锥。若$x_1, \cdots, x_k \in A$,对任意$x \in A$,存在数$\lambda_1, \cdots, \lambda_k \geq 0$,使得$x = \sum_{i=1}^{k} \lambda_i x_i$,则称凸锥由$x_1, \cdots, x_k$生成
  - 若锥是由有限向量集生成的，则称该锥是有限生成的
- **半正定锥**
  - S^n为$n \times n$的对称矩阵集合，$S_+^n = \{x \in S^n | x \geq 0\}$为$n \times n$的半正定矩阵集合，$S_{++}^n = \{x \in S^n | x > 0\}$为$n \times n$的正定矩阵集合
  - 则S_+^n是凸锥，也成为半正定锥
  - 如$(x\ y; y\ z) \in S_+^2$,点(x, y, z)构成一个半正定锥



- **保凸运算：任意多个凸集的交和仿射变换都是保凸的，缩放、平移和投影都是仿射变换，从而证明集合是凸集，或者直接按照定义证明**

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- **顶点(Vertex):如果点x是两个或多个线段、边缘、表面的交点，则称为顶点**
  - 注意，凸集的顶点都是极点
- **如果两个极点由一条边(线，或面)连接，则称为相邻的极点**
- **超平面(Hyperplane):集合$H=\{x|Ax=b\}, a\neq 0$:超平面将R^n空间分成两个半空间$H^+=\{x|Ax\geq b\}, H^-=\{x|Ax\leq b\}$**
- **多胞形(polyhedron)$P\subseteq R^n$:有限个半空间的交**
  - 一个多面锥（polyhedral cone)是形如$\{x: Ax\leq 0\}$的多胞形
  - 多面锥是锥，且是有限生成锥
  - 若$dimP=n$，则称P是满维的；满维等价于它有一内点
- **多面体(polytope):有界的多胞形**
- **引理（Minkowski,1896)：设$C=\{x\in R^n, Ax\leq 0\}$是一多面锥，则C是肯定由方程组$My=b'$的解集的一个子集生成，这里M是由$\binom{A}{I}$的n个线性无关行所组成，$b'=\pm e_j$,其中e_j是某一单位向量。**

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- 证明：令$A \in R^{m\times n}$,考虑方程组$My = b'$,这里M是由$\begin{pmatrix} A \\ I \end{pmatrix}$的$n$个线性无关行所组成，$b' = \pm e_j$,其中$e_j$是某一单位向量.令$y_1, \cdots, y_t$是这些方程组中属于$C$的解，将证$C$是由$y_1, \cdots, y_t$所生成。
- 首先设$C = \{x: Ax=0\}$,即C是线性子空间，记$C = \{x: A'x = 0\}$,这里A'是由A的极大线性无关行所组成，令I'是I中的某些行，满足$\begin{pmatrix} A' \\ I' \end{pmatrix}$是满秩方阵，则$C$是被方程组$\begin{pmatrix} A' \\ I' \end{pmatrix} x = \begin{pmatrix} 0 \\ b \end{pmatrix}$,对$b = \pm e_j, j = 1, \cdots, dimC$的解所生成。
- 对于一般情况，则用数学归纳法来证明。若C不是线性子空间，选择A的一行a和A的子矩阵A'使得$\begin{pmatrix} A' \\ a \end{pmatrix}$的行是线性无关的且$\{x: A'x = 0, ax \le 0\} \subseteq C$,由此可得，存在下标$s \in \{1, \cdots, t\}$使得$A'y_s = 0$且$ay_s = -1$。
- 现任取$z \in C$,令$a_1, \cdots, a_m$是A的行，且$\mu := min\{\frac{a_i z}{a_i y_s}: i = 1, \cdots, m, a_i y_s < 0\}$，则有$\mu \ge 0$，令$k$是取到最大值的一个下标，考虑$z' := z - \mu y_s$，由$\mu$的定义，我们有$a_j z' = a_j z - \frac{a_k z}{a_k y_s} a_j y_s (j = 1, \cdots, m)$，因此$z' \in C' := \{x \in C: a_k x = 0\}$,因为$C'$是锥，它的维数比$C$少1（由于$a_k y_s < 0$和$y_s \in C$）,由归纳法，$C'$是由$y_1, \cdots, y_t$的子集所生成，故存在$\lambda_1, \cdots, \lambda_t \ge 0$使得$z' = \sum_{i=1}^{t} \lambda_i y_i$，令$\lambda'_s := \lambda_s + \mu$（注意到$\mu \ge 0$）和$\lambda'_i := \lambda_i$（$i \ne s$）,得到$z = z' + \mu y_s = \sum_{i=1}^{t} \lambda'_i y_i$

◆ 任何多面锥都是有限生成的。

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

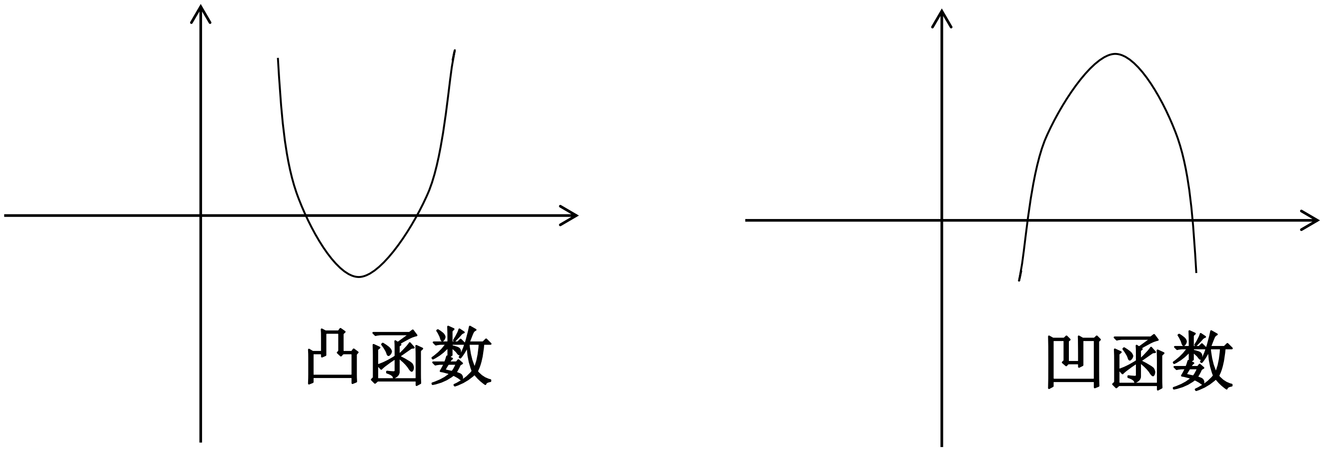
◆ **定理：集合P是多面体⇔集合P是有限点集的凸包**

➢ 证明：设$P=\{x\in R^n\colon Ax\le b\}$是一非空多面体，构造$C=\left\{\begin{pmatrix}x\\ \lambda\end{pmatrix}\in R^{n+1}\colon \lambda\ge 0, Ax-\lambda b\le 0\right\}$,则$P=\left\{x\colon \begin{pmatrix}x\\ 1\end{pmatrix}\in C\right\}$。集合$C$是一个多面锥，由前述定理可知，它由有限个非零向量所生成，记为$\begin{pmatrix}x_1\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}x_k\\ \lambda_k\end{pmatrix}$，因为$P$是有界的，所有$\lambda_i$非零，不失一般性，可令所有$\lambda_i$为1，故$x\in P\Leftrightarrow\begin{pmatrix}x\\ 1\end{pmatrix}=\mu_1\begin{pmatrix}x_1\\ 1\end{pmatrix}+\cdots+\mu_k\begin{pmatrix}x_k\\ 1\end{pmatrix}$对某些$\mu_1,\cdots,\mu_k\ge 0$成立，换言之，$P$是$x_1,\cdots,x_k$的凸包。

➢ ⇐:设P是$x_1,\cdots,x_k\in R^n$的凸包，则$x\in P\Leftrightarrow\begin{pmatrix}x\\ 1\end{pmatrix}\in C$，这里$C$是由$\begin{pmatrix}x_1\\ 1\end{pmatrix},\cdots,\begin{pmatrix}x_k\\ 1\end{pmatrix}$所生成的锥，因此由前述定理知$C$是多面锥，故有$C=\left\{\begin{pmatrix}x\\ \lambda\end{pmatrix}\colon Ax+\lambda b\le 0\right\}$，从而推得$P=\{x\in R^n\colon Ax+b\le 0$

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- **单纯形(simplex):一个R^n空间中的凸多面体，由$n+1$个不在同一超平面上的点所生成。**
  - 二维时为三角形，三维时为四面体
- **凸函数(convex function)和凹函数(concave function)**



- **函数f是凸(严格凸)的，对任意$0<\lambda<1, f(\lambda x+(1-\lambda)y) \le (<) \lambda f(x)+(1-\lambda)f(y)$**

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- **对于函数$f(x)$，$x \in R^n$,其Hessian矩阵$H\big(f(x)\big) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)$,H是对称矩阵，梯度为向量:$G\big(f(x)\big) = \big(\partial f(x)/\partial x\big)$**
  - 若一阶梯度或二阶Hessian矩阵在定义域的每个点上都存在，则称函数为一阶或二阶可微，若其还连续，则为连续可微（此时Hessian对称）
  - 若函数$f: \mathbb{R}^n \to \mathbb{R}^m$为向量值函数，此时有Jacobi矩阵$J(x) \in R^{m \times n}$，其第$i$行是分量$f_i(x)$梯度的转置:$J(x) = \left(\frac{\partial f_i(x)}{\partial x_j}\right)$,梯度的Jacobi矩阵就是Hessian矩阵
- **函数$f(x)$是凸的，如果其海塞矩阵H是半正定的**
- **拐点(infection point)或者鞍点(saddle point)**
  - 鞍点：曲线或曲面导数为0，但不是极值点
  - 拐点：曲线的曲率符号发生变化，凸凹性发生变化的点，不一定导数为0
- **函数$f(x)$是凸的，则满足$\nabla f(x^*) = 0$的点为极小值点，如果$f(x)$是凹的，则为极大值点，**
- **由$\nabla f(x) = 0$可以找到临界点x^*,**
  - 如果$H(f(x))$是半正定的，则为极小点
  - 如果$H(f(x))$是半负定的，则为极大点
  - 如果$H(f(x))$是不定的，则为拐点

2.0简单回顾一下线性代数里面的概念凸性(Convexity)

- **梯度的定义**
  - 给定函数$f \in R^n \to R$,f在点x的一个邻域内有意义，若存在向量$g \in R^n$，满足$lim_{p \to 0} \frac{f(x+p)-f(x)-g^T p}{||p||} = 0$,其中范数是任意的范数，则称函数$f$在点$x$处可微，而$g$为函数$f$在点$x$处的梯度
- **梯度李普希兹连续**
  - 给定可微函数f,若存在$L > 0$,对任意的$x, y \in domf$,有$||\nabla f(x) - \nabla f(y)|| \leq L||x - y||$,则称之为梯度李普希兹连续的，有时也称之为$L$ –光滑
  - 表明梯度的变化被自变量x的变化所控制
  - 设可微函数$f(x)$的定义域$domf = R^n$,且为梯度李普希兹连续的，则函数$f(x)$有二次上界：
  - $f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{L}{2}||y - x||^2, \forall x, y \in domf$
  - 构造辅助函数$g(t) = f\big(x + t(y - x)\big), t \in [0, 1], g(0) = f(x), g(1) = f(y), g'(t) = \nabla f\big(x + t(y - x)\big)^T (y - x)$,由等式$g(1) - g(0) = \int_0^1 g'(t)dt \Rightarrow f(y) - f(x) - \nabla f(x)^T (y - x) = \int_0^1 (g'(t) - g'(0)dt \leq \frac{L}{2}||y - x||^2$
- **可微函数存在一个全局极小点x^*,$f(x)$为L –光滑，则对任意的x有：**
- $\frac{1}{2L}||\nabla f(x)||^2 \leq f(x) - f(x^*)$
- **矩阵变量函数的导数**
  - 梯度的推广，称为$Frechet$可微

2.1 线性规划的基本概念-例

例：某工厂拥有A、B、C 三种类型的设备，生产甲、乙两种产品。每件产品在生产中需要占用的设备机时数，每件产品可以获得的利润以及三种设备可利用的时数如下表所示：

	产品甲	产品乙	设备能力（h）
设备A	3	2	65
设备B	2	1	40
设备C	0	3	75
利润（元/件）	1500	2500	

2.1 线性规划的基本概念-例

- 问题：工厂应如何安排生产可获得最大的总利润？
- 解：设变量x_i为第i种（甲、乙）产品的生产件数（$i=1$，2）。根据题意，我们知道两种产品的生产受到设备能力（机时数）的限制。对设备A，两种产品生产所占用的机时数不能超过65，于是我们可以得到不等式：$3x_1+2x_2 \leqslant 65$；
- 对设备B，两种产品生产所占用的机时数不能超过40，于是我们可以得到不等式：$2x_1+x_2 \leqslant 40$；

2.1 线性规划的基本概念–例(续1)

- 对设备C，两种产品生产所占用的机时数不能超过75，于是我们可以得到不等式：$3x_2 \leqslant 75$；另外，产品数不可能为负，即 $x_1, x_2 \geqslant 0$。同时，我们有一个追求目标，即获取最大利润
- 于是可写出目标函数z为相应的生产计划可以获得的总利润：$z=1500x_1+2500x_2$
- 综合上述讨论，在加工时间以及利润与产品产量成线性关系的假设下，把目标函数和约束条件放在一起，可以建立如下的线性规划模型：

2.1 线性规划的基本概念-例(续2)

◆目标函数(Objective Function)

$$Max\ z = 1500x_1 + 2500x_2$$

◆约束条件(Constrain conditions)

◆**Subject to：s.t.**
$$\begin{cases} 3x_1 + 2x_2 \le 65 \\ 2x_1 + x2 \le 40 \\ 3x_2 \le 75 \\ x_1, x_2 \ge 0 \end{cases}$$

2.1 线性规划的基本概念-例(续3)

- 例某铁器加工厂制作100套钢架，每套要用长为2.9米，2.1米和1.5米的圆钢各一根，已知原料长为7.4米，问应如何下料，可使所用材料最省？
- 解：7.4米有很多种切割方法，先计算出比较省料的几种方法，然后将问题转化为使剩余料最少的方案：I:2.9+1.5*3=7.4, II:2*2.9+1.5=7.3，III:2*2.1+2*1.5 = 7.2,IV:2.9+2*2.1=7.1，V:2.1+3*1.5=6.6
- 假设按第i种方案下料的原料根数为$x_i(i=1,2,\ldots,5)$,则要求：$min\ z=0x_1+0.1x_2+0.2x_3+0.3x_4+0.8x_5$,且满足$x_1+2x_2+x_4=100, 2x_3+2x_4+x_5=100, 3x_1+x_2+2x_3+3x_5=100$,同时要求$x_i\geq 0$,且为整数

2.1 线性规划的基本概念-例(续4)

- 这些都是典型的利润最大化的生产计划问题。
- 其中， “Max”是英文单词“Maximize”的缩写，含义为“最大化”
- “s.t.”是“subject to”的缩写，表示“满足于……”
- 上述模型的含义是：在给定条件限制下，求使目标函数z达到最大的x_1，x_2 的取值
- 它是一个目标函数、约束函数都是线性函数的最优化问题，即线性规划问题(LP)

2.1 线性规划的基本概念-线性规划-模型的一般形式

◆目标函数

$$Min(Max)Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

◆约束条件

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n \le (=, \ge) b_1 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots a_{in}x_n \le (=, \ge) b_i \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n \le (=, \ge) b_m \\ x_1, x_2, x_3, \cdots, x_n \ge 0 \end{cases}$$

2.1 线性规划的基本概念-模型的标准形式

◆**目标函数**

$$Max(Min)\ z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

◆**约束条件**

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1 \\ \qquad\vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots a_{in}x_n = b_i \\ \qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = b_m \\ x_1, x_2, x_3, \cdots, x_n \geq 0 \end{cases}$$

把一般的LP化成标准型的过程称为线性规划问题的标准化

方法：

1 目标标准化

max Z 等价于 min (- Z)

min $Z' = -\sum c_j x_j$

2 化约束为等式

加松弛变量、减剩余变量

3 变量非负化做变换

或 $x'_j = -x_j \qquad x'_j \geq 0$

$x_j = x'_j - x''_j \qquad x''_j \geq 0$

4 右端非负

注意：右端项要求非负

松弛变量(Slack Variable)：化不等式为等式约束

2.1 线性规划的基本概念-模型的标准形式

◆将下述问题转化为标准形式

- $Min\ \ f = -3.6\,x_1 + 5.2\,x_2 - 1.8\,x_3$
- $s.t.\ 2.3\,x_1 + 5.2\,x_2 - 6.1\,x_3 \le 15.7$
- $4.1\,x_1 + 3.3\,x_3 \ge 8.9$
- $x_1 + x_2 + x_3 = 38$
- $x_1, x_2, x_3 \ge 0$

◆将下述问题转化为标准形式

- $Min\ f = 3\,x_1 - 5\,x_2 - 8\,x_3 + 7\,x_4$
- $s.t.\ 2\,x_1 - 3\,x_2 + 5\,x_3 + 6\,x_4 \le 28$
- $4\,x_1 + 2\,x_2 + 3\,x_3 - 9\,x_4 \ge 39$
- $6\,x_2 + 2\,x_3 + 3\,x_4 \le -58$
- $x_1, x_3, x_4 \ge 0$

2.1 线性规划的基本概念-模型的标准形式

◆**例子**

$$\min Z = -2x_1 + 3x_2 - x_3$$

$$s.t.\begin{cases} x_1 - x_2 + x_3 \le 10 \\ 3x_1 + 2x_2 - x_3 \ge 8 \\ x_1 - 3x_2 + x_3 = -1 \\ x_1, x_2 \ge 0, x_3\text{符号不受限制} \end{cases}$$

$$\max Z' = 2x_1 - 3x_2 + (x_3' - x_4) + 0\cdot x_5 + 0\cdot x_6$$

$$s.t.\begin{cases} x_1 - x_2 + (x_3' - x_4) + x_5 = 10 \\ 3x_1 + 2x_2 - (x_3' - x_4) - x_6 = 8 \\ -x_1 + 3x_2 - (x_3' - x_4) = 1 \\ x_1, x_2, x_3', x_4, x_5, x_6 \ge 0 \end{cases}$$

线性规划的标准形式：

$$
\text{(LP)} \quad \begin{cases} \boldsymbol{Max(Min)} \ c^T x \\ s.t. \quad Ax = b \\ \qquad\qquad x \geq 0 \end{cases}
$$

其中, c , $x \in R^n$； $b \in R^m$； A为 $m \times n$ 矩阵，且一般假设 $\boldsymbol{rank(A) = m}$

2.1 线性规划的基本概念-解的基本概念

- **可行解(Feasible solution)有时也称为可行域**
  - $\{x \mid Ax = b, x \geq 0\}$
- **最优解:使目标函数取最优的可行解**
- **基(basis),基变量(basic Var.),非基变量(non-basic var.)**
  - 若$B = [P_1\ P_2\ \ldots\ P_m]$为$A$的一个$m$阶可逆矩阵，$A = [B\ N], \boldsymbol{x} = [\boldsymbol{x}_B\ \boldsymbol{x}_N]^T$,则称$B$为一个基或基矩阵，对应$\boldsymbol{x}_B$：$x_1, \cdots, x_m$称为基变量，剩余的$\boldsymbol{x}_N$：$n - m$个变量称为非基变量。
- **基本解:非基变量为0时，满足约束$Ax = b$的解**
  - 基本解至少有$n - m$个分量为0，至多有m个非零分量
  - 非零分量的个数少于m时，称为退化的基本解
  - 基本解的个数最多有$C(n, m) = n!/(m!\,(n - m)!)$

2.1 线性规划的基本概念-解的基本概念

◆**基本可行解**(basic feasible solution:BFS)

- 还满足非负约束x≥0的基本解。基本可行解的个数至多为C(n,m)=n!/(m!(n-m)!)

◆**最优基本可行解**

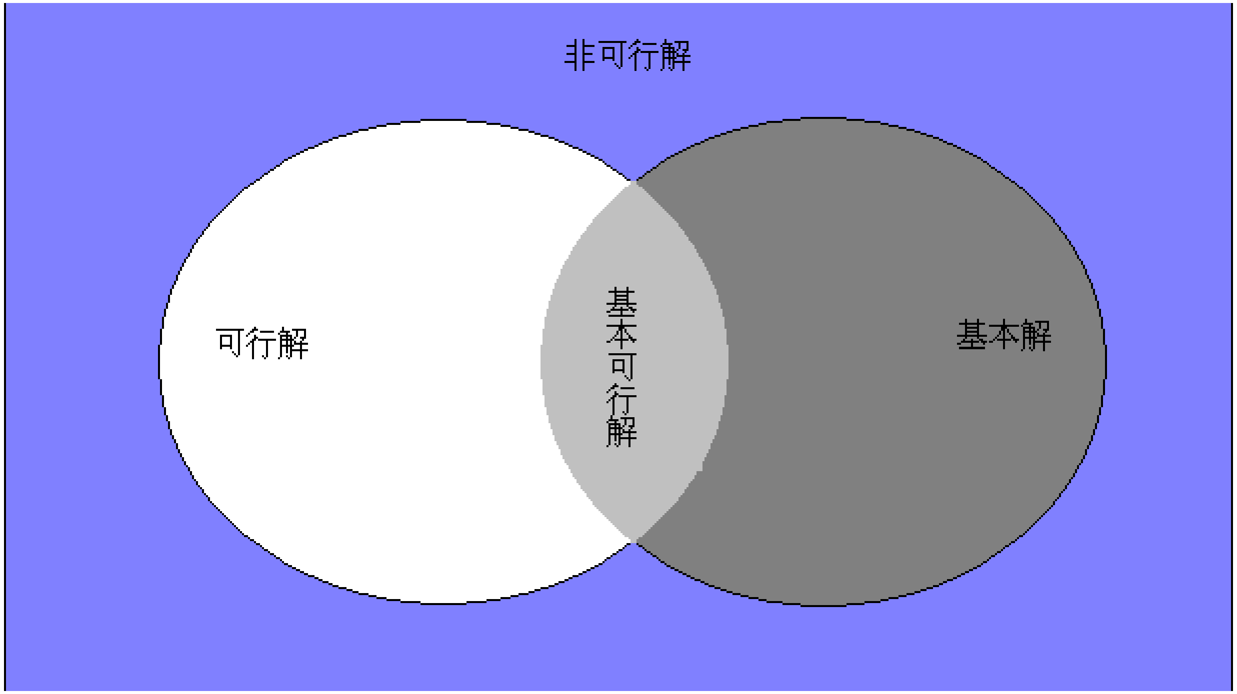
- 使目标函数取最优的基本可行解
- 注：BFS=>$Ax_B=b, x_B\geq 0$;x^*是最优解，则是一个BFS，且$c^Tx^*\geq c^Tx$, 对任意BFS x

◆$Ax=b\Rightarrow(B\ \ N)\begin{pmatrix}x_B\\x_N\end{pmatrix}=Bx_B+Nx_N=b$, 这时进一步可写为$Bx_B=b-Nx_N\Rightarrow x_B=B^{-1}(b-Nx_N)$，$x=\begin{pmatrix}x_B\\x_N\end{pmatrix}$

- 如果$x_N=0\Rightarrow x_B=B^{-1}b$, 这$x=\begin{pmatrix}x_B\\x_N\end{pmatrix}=\begin{pmatrix}B^{-1}b\\0\end{pmatrix}$称为与基$B$对应的基本解，若$x_B\geq 0$，则称之为基本可行解

2.1 线性规划的基本概念-解的基本概念

◆可行解、基本解



2.1 线性规划的基本概念-解的基本情况

◆**解的存在性：若(LP)的可行域(带约束的多面体)非空，则可行域是个凸集，且(LP)一定存在有限最优解或无界最优解**

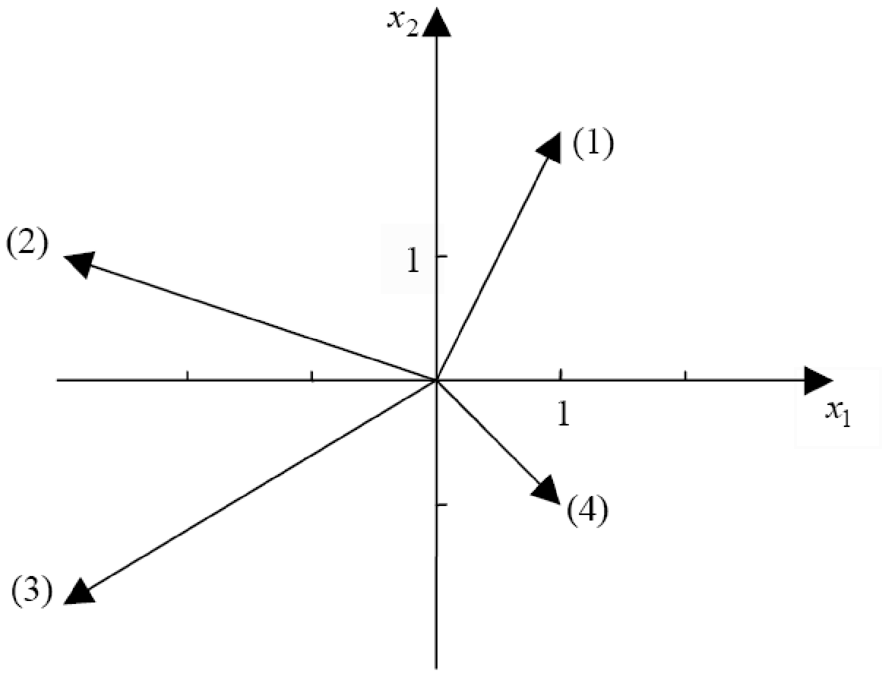
◆**解在顶点的可达性：若(LP)存在有限最优解，则最优解可在某个顶点处达到**

◆**顶点与基本可行解的关系：x_0是(LP)的可行域顶点的充分必要条件是x_0是(LP)的基本可行解**

→可通过求基本可行解得到有限最优解

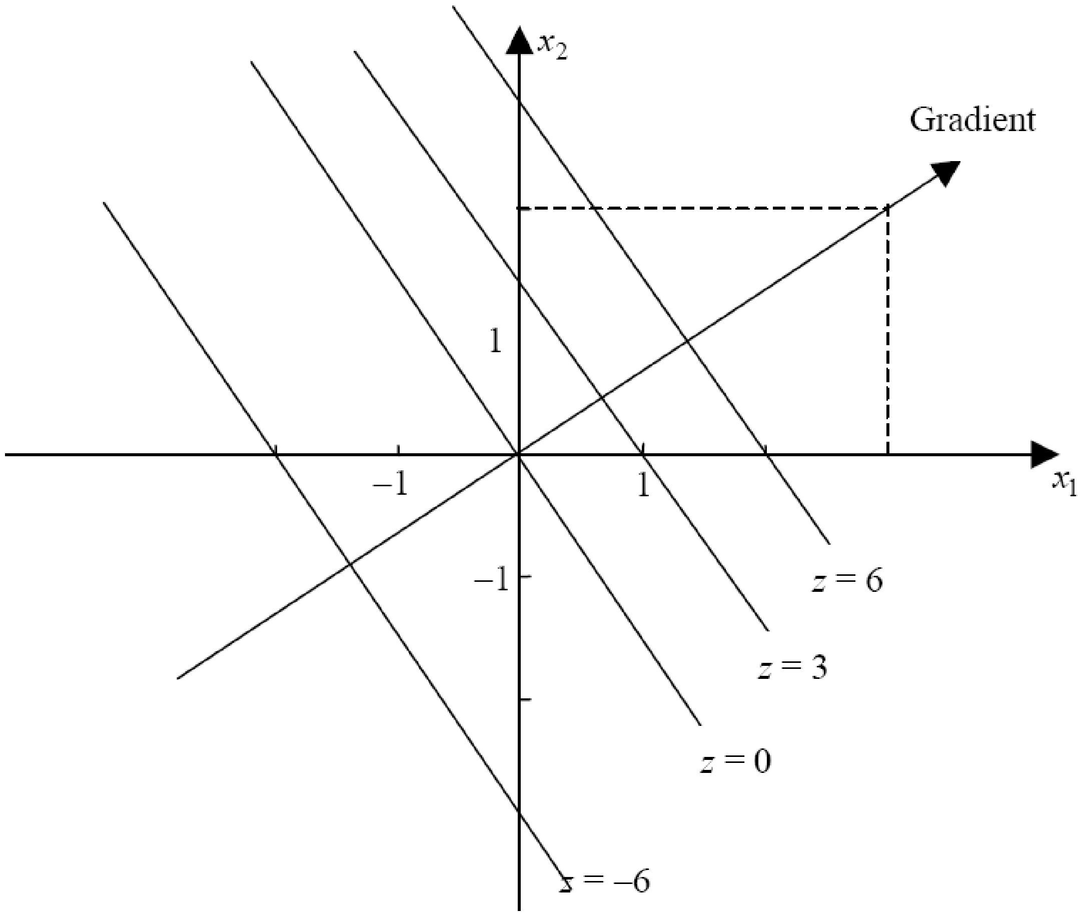
◆考虑下列目标函数的梯度(Gradient)情况

- $Max\ z = x_1 + 2x_2$
- $Max\ z = -3x_1 + x_2$
- $Max\ z = -3x_1 - 2x_2$
- $Max\ z = x_1 - x_2$

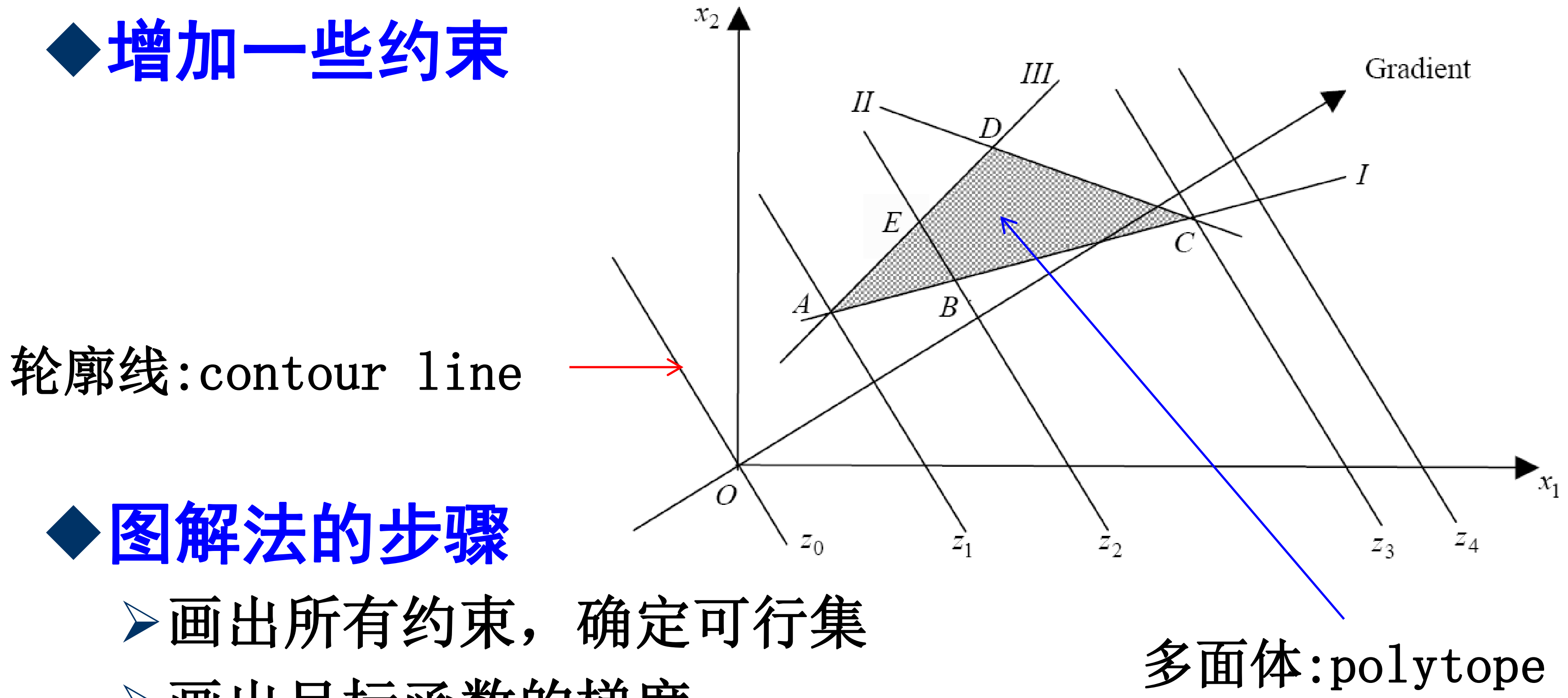


◆考虑下列目标函数的梯度情况

- $Max\ z = 3x_1 + 2x_2$
- $z = -6, 0, 3, 6$相应不同的等值线



◆**增加一些约束**



◆**图解法的步骤**

- 画出所有约束，确定可行集
- 画出目标函数的梯度
- 按目标函数的梯度方向平移轮廓线，直至最后一个可行点，这个点即为最优点$\bar{X}$
- 超平面约束集合相交于该点$\bar{X}$，解方程得到最优解的坐标，然后确定最终目标函数值$\bar{z}=c\bar{X}$

- 那这意味着最优点肯定在边界上，这么做是否合理呢？
- **定理**(Dantzig,1951)假设可行集(Feasible set)是非空且有界的。则至少有一个最优解位于极点上。
- **证明**：假设X_o是唯一的最优解，且不是极点，则：$\exists\lambda \in (0,1), X_o = \lambda X_1 + (1-\lambda)X_2$,且$X_0, X_1$不是最优解，即 $cX_0 > cX_1, cX_0 > cX_2 \Rightarrow c(\lambda X_1 - \lambda X_2 + X_2 - X_1) > 0 \Rightarrow cX_1 < cX_2$，类似的，$cX_2 < cX_1$，矛盾。对所有的最优点重复上述过程，则得证！

◆例子，考虑下列LP问题

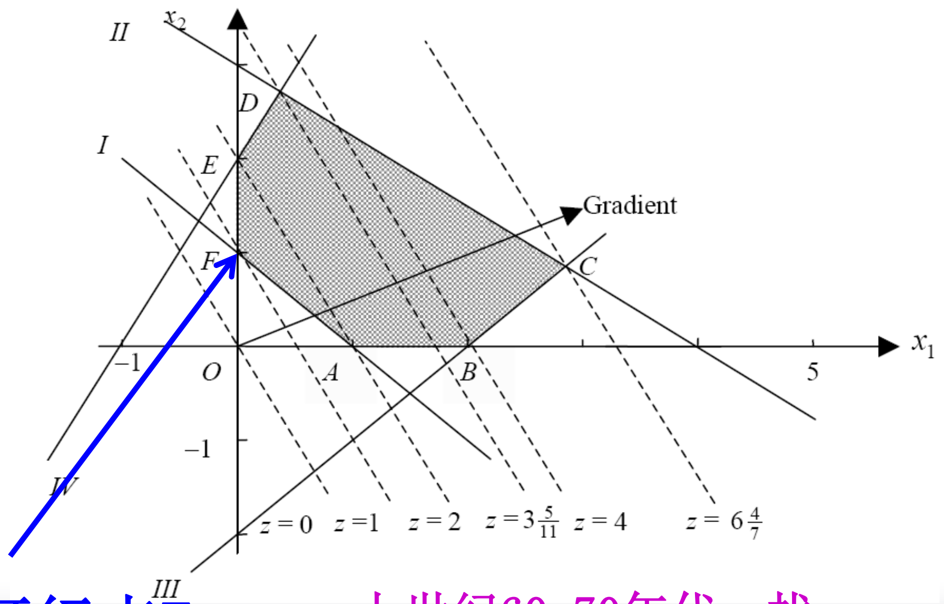
◆$Max\ z = 2x_1 + x_2$,

Optimal soln： $X = (\frac{20}{7}, \frac{6}{7})$

$$Z = \frac{46}{7}$$

◆$s.t.$

$$\begin{cases} x_1 + x_2 \geq 1 & I \\ 3x_1 + 4x_2 \geq 12 & II \\ x_1 - x_2 \leq 2 & III \\ -2x_1 + x_2 \leq 2 & IV \\ x_1 \geq 0 & V \\ x_2 \geq 0 & VI \end{cases}$$



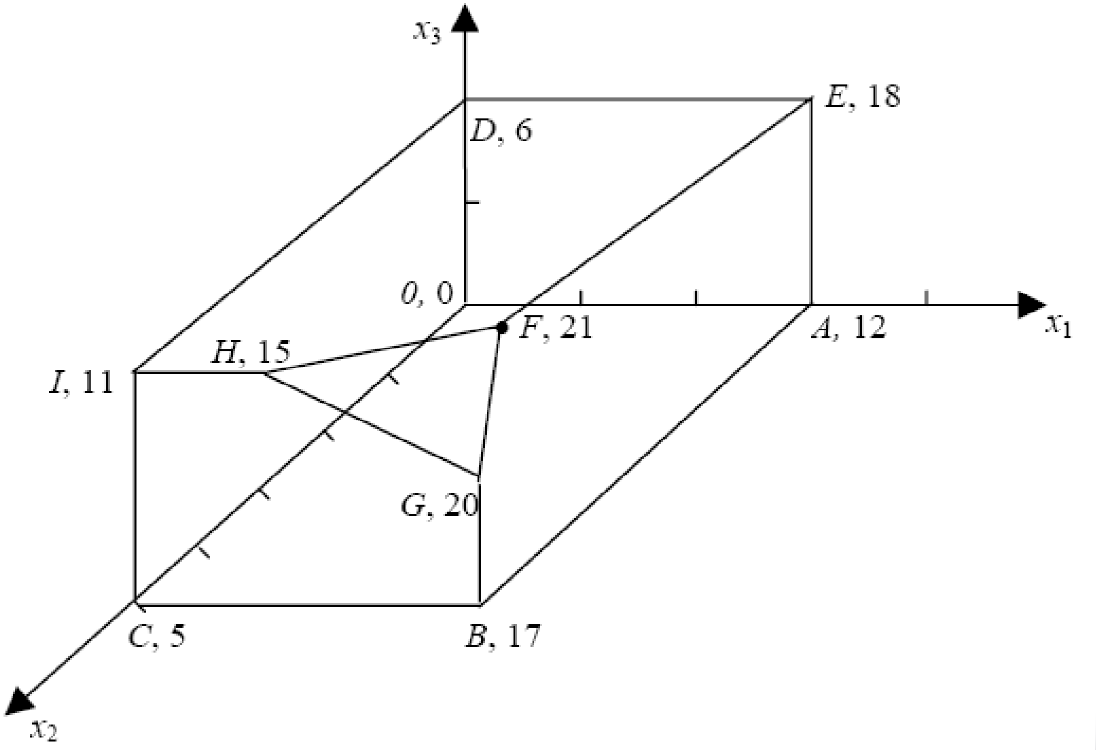
从任意初始可行点出发，定义一个方向为改进的可行方向：移动一小步会保持其可行性，且目标函数的值会得到改进

初始可行点F

上世纪60. 70年代，找最短路径

- **例子**
- $\boldsymbol{Max\ z = 4x_1 + x_2 + 3x_3}$
- $$\begin{cases} x_1 \le 3 \\ x_2 \le 5 \\ x_3 \le 2 \\ x_1 + x_2 + 2x_3 \le 10 \\ x_1, x_2, x_3 \ge 0 \end{cases}$$

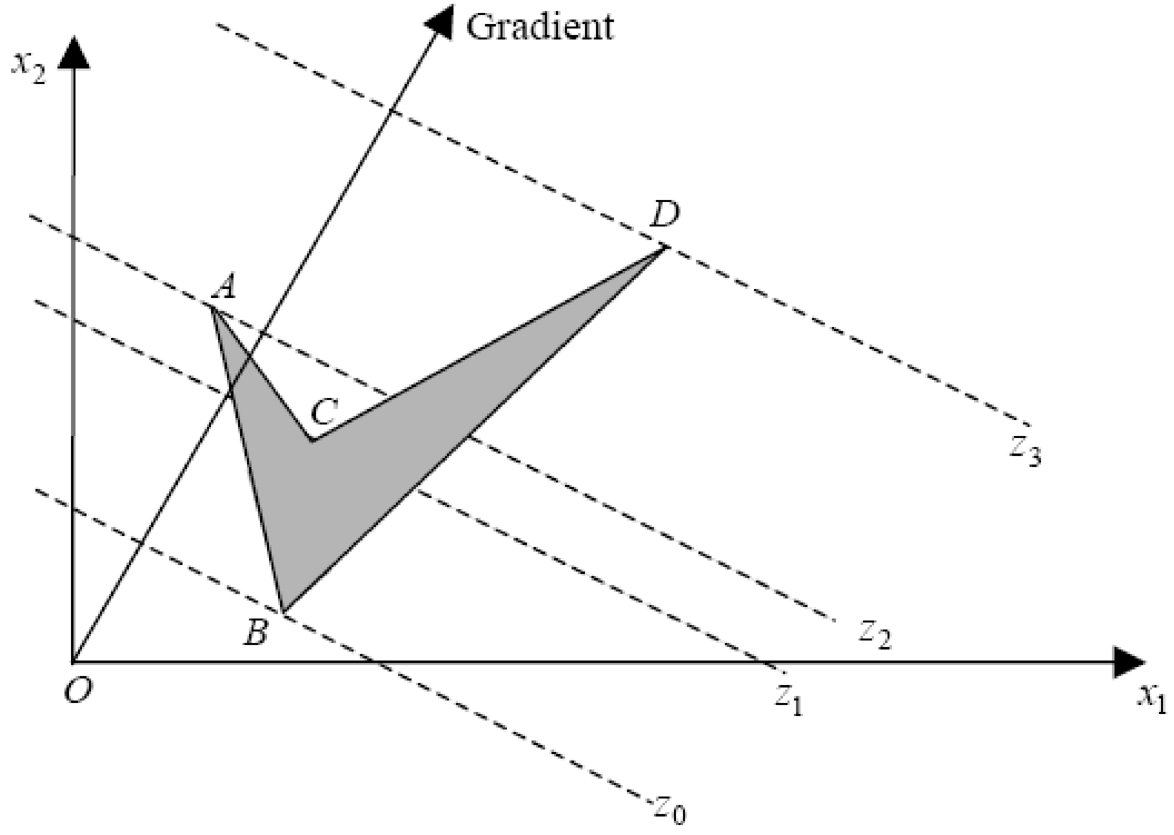
如果能提前知道哪条路径包含最少的中间极点，这将有助于减少计算量！



单纯形路径:0->F,至少有8条合法的路径:

OABGF, OAEF, OCBGF, OCIHGF, OCIHF, ODEF, ODIHGF, ODIHF

◆例子，如何找可行方向，B为起始点



◆显然，A不是全局最优点，而是一个局部最优点

- 从这些例子中可以看到，可能会找到局部最优点
  - 对于集合S上的Max问题，局部最优点$\hat{X} \in S$,对某个任意小的$\epsilon > 0$,所有$\tilde{X} \in S$, $\left\|\hat{X} - \tilde{X}\right\| \le \epsilon$，都有$c^T\tilde{X} \le c^T\hat{X}$,全局最优点$\bar{X}$则表明,任意$\tilde{X} \in S$，都有$c^T\bar{X} \ge c^T\tilde{X}$
- 但对于凸可行集和线性目标函数来说，每个局部最优点都是全局最优点
- 在封闭多面体的相邻极点之间按改进可行方向搜索可找到局部最优点
- 由于LP问题是凸的，因此按照这种搜索方式找到的局部最优点就是LP问题的全局最优解
- 1957年,Hirsch提出了一个猜想：从任意一个极点移动到任何其它的极点，最多只需要m步，m是约束的数目

Hirsch WM (1957) Hirsch conjecture, verbal communication to Dantzig

◆针对Hirsch的猜想，Klee和Walkup(1967)证明了$n-m\le 5$,n为变量数目，并证明猜想对无界多面体不成立

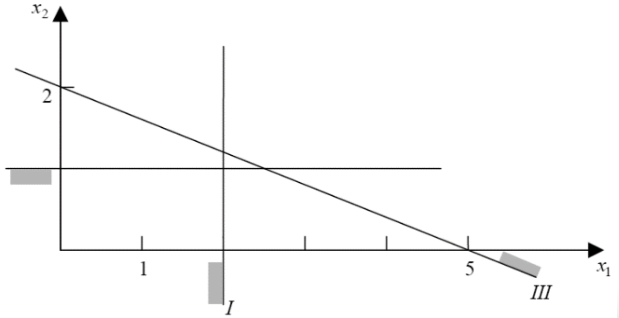
◆图解法时，可能会碰到以下四种特殊情况

➢1.没有可行解：考虑如下约束
$$\begin{cases} x_1 \le 2 & I \\ x_2 \le 1 & II \\ 2x_1+5x_2 \ge 10 & III \\ x_1,\ x_2 \ge 0 \end{cases}$$

➢显然，约束$I\times 2+II\times 5\le 9$，与约束$III$矛盾，因此图解法的时候这三个超平面没有围成一个多面体

➢说明什么问题呢？如何解决？

✓修改原来问题的建模方法，例如调整约束的右边项，一般是每个约束都适当调整，而不要单独调整某一个约束

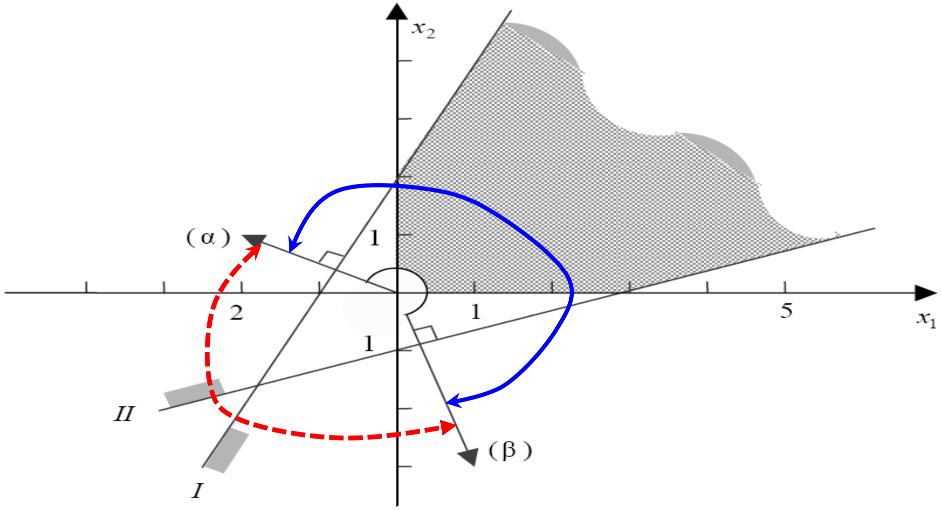


Klee VL, Walkup DW (1967) The d-step conjecture for polyhedra of dimension $d<6$. Acta Mathematica 117: 53-78

2.1 线性规划的基本概念-图解法(Graphical Solution)(续9)

- **2.存在无界最优解：** 多面体是无界的，并且目标函数的梯度指向分离超平面正交的射线所形成的圆锥体的方向，考虑如下约束

$$\begin{cases} -2x_1 + x_2 \le 2 \ \ I \\ x_1 \ \ -3x_2 \le 3 \ \ II \\ x_1 \ \ , \ \ x_2 \ge 0 \end{cases}$$



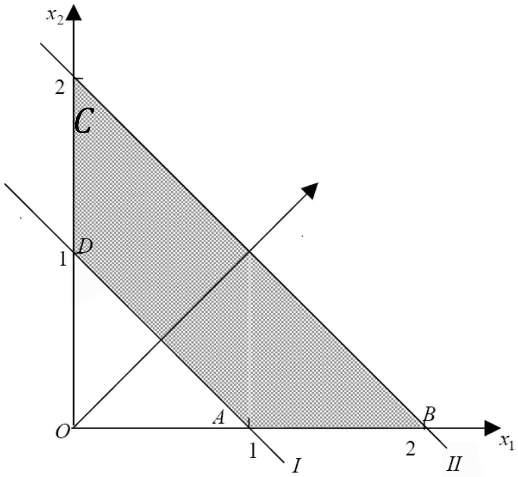
- 目标函数
  - ✓ (α): $Max\ z = -2x_1 + x_2$
  - ✓ (β): $Max\ z = x_1 - 3x_2$

- 限定蓝色实线区间之内的任何目标函数

(指目标函数的梯度方向在蓝色曲线围成的区间之内时)都有无界最优解

- 限定红色虚线区间之内的任何目标函数(指目标函数的梯度方向在红色虚线围成的区间之内时)都有有限最优解，只可能在三个极点处取得最优解
- 如果出现无界最优解的情况，说明形式化约束和目标函数的时候有问题，可能忘记某些约束条件，导致已有的形式化太松了

- 3.在极点处出现对偶退化的情况(Dual degeneracy):相邻极点中至少有一个具有相同的目标函数值，如果在最优极点上出现对偶退化，则说明至少存在一个可替换的最优解
- 对偶退化出现的充要条件目标函数的梯度与不可约约束形成的超平面正交
- 考虑如下LP问题：$Max\ z = x_1 + x_2, s$



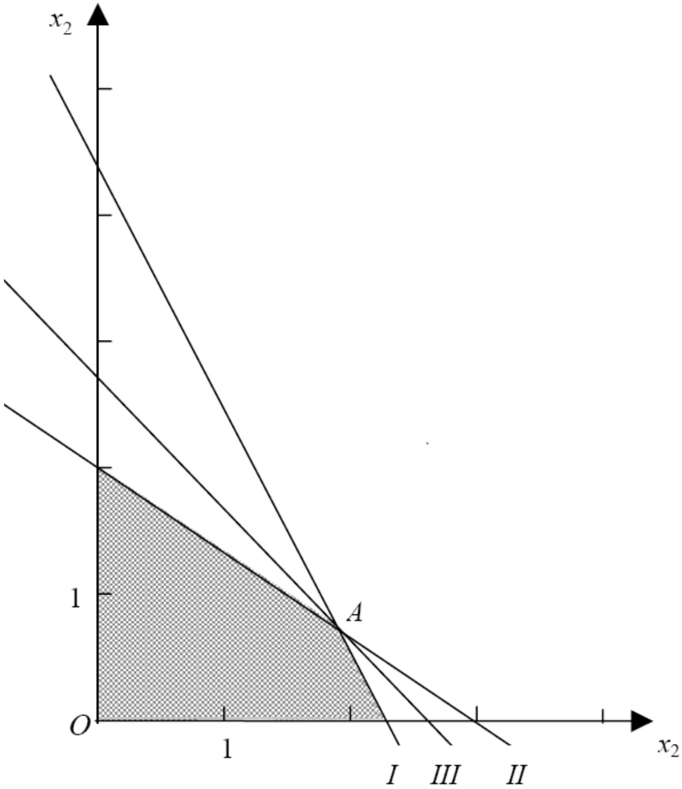
- A和D，B和C都是对偶退化点，B和C 明存在可替换的最优解，实际上，BC线段上任何点都是最优解

◆**4.原始退化(Primal degeneracy)：如果超过n个超平面相交于R^n空间中的一个点(这样，该点就被超定了,overdetermined)， 这种情况仅仅依赖于约束条件，与目标函数无关**

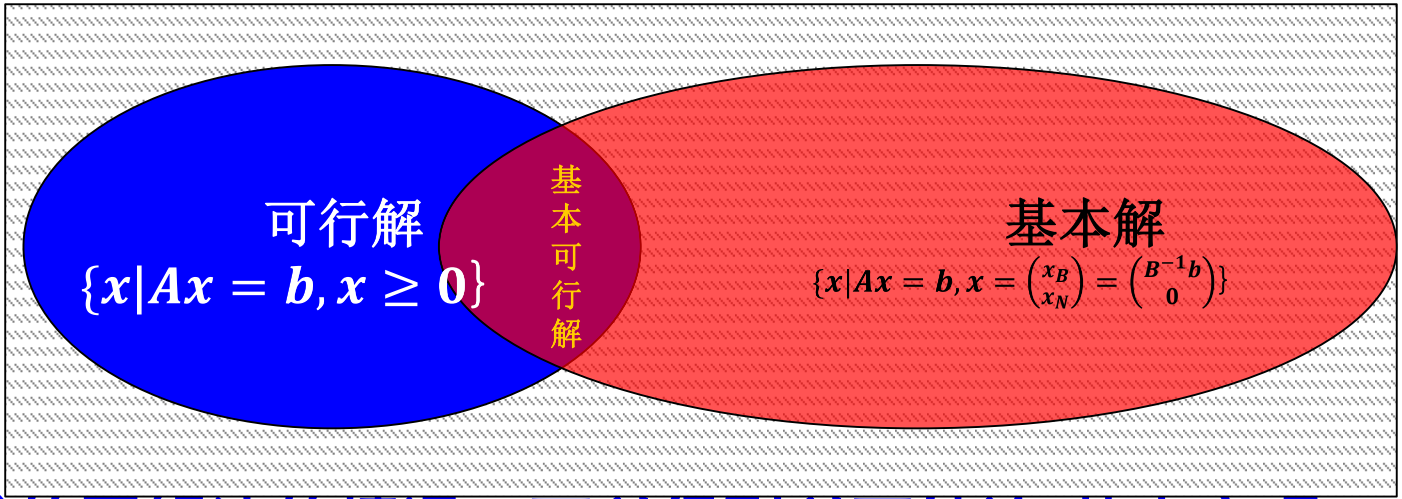
◆**例如约束：**

$$\begin{cases} 4x_1 + 2x_2 \le 9 & I \\ 2x_1 + 3x_2 \le 6 & II \\ 6x_1 + 5x_2 \le 15 & III \\ x_1, \; x_2 \ge 0 \end{cases}$$

◆**原始退化点$A(1\frac{7}{8},\frac{3}{4})$**



- $Ax=b \Rightarrow (B\ \ N)\begin{pmatrix}x_B\\x_N\end{pmatrix}=Bx_B+Nx_N=b$, 这时进一步可写为$Bx_B=b-Nx_N \Rightarrow x_B=B^{-1}(b-Nx_N)$，$x=\begin{pmatrix}x_B\\x_N\end{pmatrix}$
  - 如果$x_N=0 \Rightarrow x_B=B^{-1}b$, 这$x=\begin{pmatrix}x_B\\x_N\end{pmatrix}=\begin{pmatrix}B^{-1}b\\0\end{pmatrix}$称为与基$B$对应的基本解，若$x_B\geq 0$，则称之为基本可行解



- **从上述的图解法的情况，可以得到以下结论(基本定理)**
  - 1.LP问题可行解x为基本可行解的充要条件是x中正分量对应的系数列向量线性无关（如何证明？）
  - 2.LP问题的每一个基本可行解x对应可行域(就是约束多面体)的一个极点

◆ 线性规划-基本定理

➢ 2.LP问题的每一个基本可行解x对应可行域(就是约束多面体)的一个极点

证明:不失一般性，设基本可行解x的前m个分量为基变量(注意$x_{m+1} \dots x_n$均为0)

$$x = [x_1\ x_2 \cdots x_m\ x_{m+1} \cdots x_n]^T$$

(1)设x是基本可行解$\Rightarrow$ x是极点

反证法: 若x不是极点，则x可表示为可行域S中的两个不同点$x^{(1)}, x^{(2)}$的严格凸组合,令$[p_1\ p_2 \cdots p_m]$为基变量对应的基矩阵，其中p_i为A中的列向量

则由$x = \lambda x^{(1)} + (1-\lambda)x^{(2)} \Rightarrow x^{(1)}, x^{(2)}$的后$n-m$个成份必为0，而且$x^{(1)}, x^{(2)}$前$m$个分量与$[p_1\ p_2 \cdots p_m]$相乘后都为$b$，相减，可知$\Sigma_{i=1}^{m}\left(x_i^{(1)} - x_i^{(2)}\right)p_i = 0 \Rightarrow x_i^{(1)} = x_i^{(2)}$，进而$x^{(1)} = x^{(2)}$，与假设矛盾，证毕.

(2) x是极点$\Rightarrow$ x是基本可行解

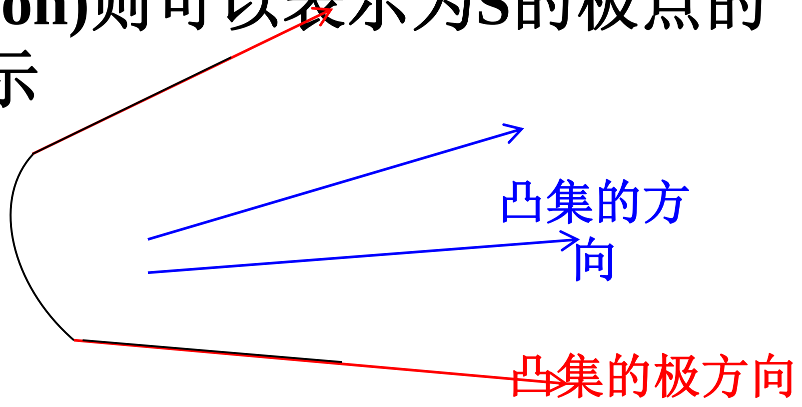
反证法:若x不是一个基本可行解(BFS).x是可行域上的点，因此必定是可行解。从而其非零分量必为正分量，不妨设$x = [x_1 \cdots x_k 0 \cdots 0]^T$,非零分量对应的$p_i$必线性相关:

$\Sigma_{\mathrm{i=1}}^{\mathrm{k}}\delta_i p_i = 0$;另$\Sigma_{\mathrm{i=1}}^{\mathrm{k}} x_i p_i = b$,取$\lambda = min_{\delta_j \neq 0, j=1\ to\ k}\left(\frac{x_j}{|\delta_j|}\right)$,从而有$\Sigma_{\mathrm{i=1}}^{\mathrm{k}}(x_i - \lambda\delta_i)p_i = b$,

$\Sigma_{\mathrm{i=1}}^{\mathrm{k}}(x_i + \lambda\delta_i)p_i = b \Rightarrow x$可由两个不等的可行域上的点的凸组合表示，产生矛盾，得证!

◆线性规划-基本定理

- **3.**有界凸集S(polytope)上的任意一点都可以表示为S的极点的凸组合,如果是无界凸集(polyhedron)则可以表示为S的极点的凸组合加上极方向的正组合来表示



- **4.**LP问题若可行域有界，且存在最优解，则目标函数必可在其可行域的某个顶点达到
  - ✓证明：假设不在顶点达到，利用基本定理3，可以得出顶点也是最优解
- 从这些基本定理可以得到以下结论
  - ✓LP问题的可行域是凸集，可能无界，但顶点数有限
  - ✓LP问题每个基本可行解对应可行域的一个极点
  - ✓LP问题有最优解，则必可在某些极点上达到最优值

◆LP中的基本定理

- 5.对于LP问题，若存在可行解，则必存在基本可行解，其中约束矩阵$A_{m\times n}$的秩为m

- 证明：设$A=[p_1\, p_2 \cdots p_n]$,且$x_0=\left[x_0^1, x_0^2, \cdots, x_0^n\right]^T$为LP的一个可行解，显然$x_0^i \geq 0$，不妨令前$k$个分量为正，则有$\sum_{i=1}^{k} x_0^i p_i = b$,此时有两种情况

- a)$p_1, \ldots, p_k$线性无关，显然存在基本可行解(BFS);

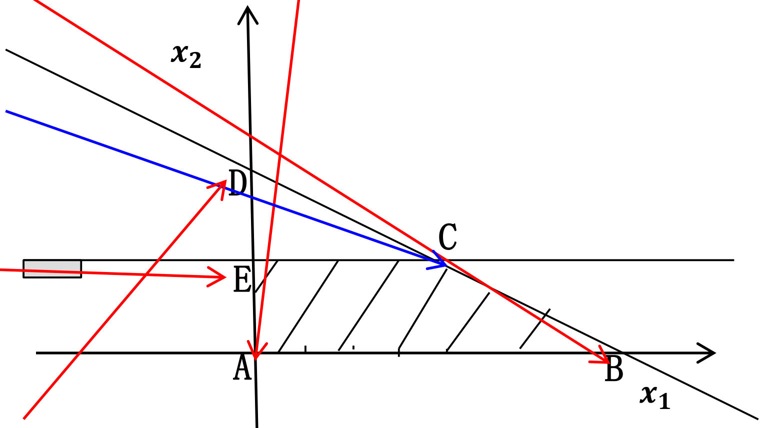
- b) $p_1, \ldots, p_k$线性相关，则存在不全为0的数δ_i,使得$\sum_{i=1}^{k}\delta_i p_i = 0$,至少有一个$\delta_i > 0$, $\sum_{i=1}^{k}(x_0^i - \lambda\delta_i)p_i = b$， 记$x_1=[x_0^1 - \lambda\delta_1, \cdots, x_0^k - \lambda\delta_k, 0, \cdots, 0]^T$,显然，若令$\lambda = min_{\delta_j \neq 0, j=1\ to\ k}\left(\frac{x_0^j}{|\delta_j|}\right)$,则$x_1$仍然是可行解，但正分量的个数最多为$k-1$，如果此时对应的$p_i$线性无关，证毕，否则，重复进行下去.

- 上述结论表明：可以用代数的办法寻找最优解：先找基本可行解($A_{m\times n}: rank(A)=m$基本解个数$\le C_n^m$),最多为C_n^m个

- 例子:LP问题 $max\ z=x_1+4x_2, s.t.\begin{cases} x_1+2x_2\le 8 \\ \qquad\quad x_2\le 2 \\ x_1\ ,\ x_2\ge 0 \end{cases}$
  - 先化为标准形式，然后再求其基本解，在判断是否是基本可行解
  - 基本可行解为$x_1=(4,2,0,0)^T, x_2=(8,0,0,2)^T, x_3=(0,2,4,0)^T, x_4=(0,0,8,2)^T$
  - 然后代入目标函数分别的目标函数值为12,8,8,0，可见最优解为x_1，最优目标值为12

2.1 线性规划的基本概念-线性规划问题的基本理论(续4)(注1)

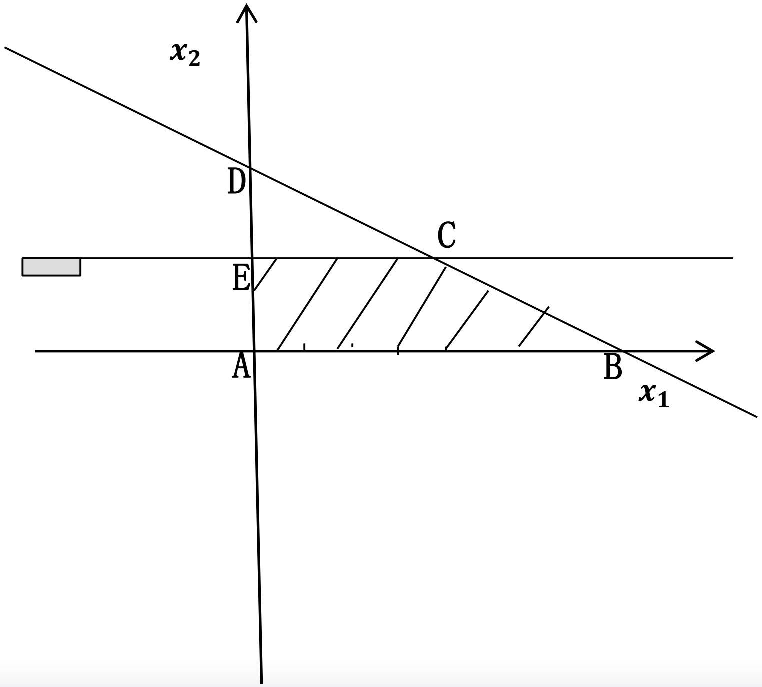
- $max\, z = x_1 + 4x_2, s.t.\begin{cases} x_1 + 2x_2 \le 8 \\ \quad\quad x_2 \le 2 \\ x_1,\ x_2 \ge 0 \end{cases}$,化为标准型
- $\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$,选$B = (p_3\ p_4) \Rightarrow x^{(0)} = \begin{pmatrix} x_N^{(0)} \\ x_B^{(0)} \end{pmatrix} = (0,0,8,2)^T, z = 0$
- 选$B = (p_1\ p_4) \Rightarrow x^{(1)} = \begin{pmatrix} x_N^{(1)} \\ x_B^{(1)} \end{pmatrix} = (8,0,0,2)^T, z = 8$
- 选$B = (p_1\ p_2) \Rightarrow x^{(2)} = \begin{pmatrix} x_N^{(2)} \\ x_B^{(2)} \end{pmatrix} = (4,2,0,0)^T, z = 12$
- 选$B = (p_2\ p_3) \Rightarrow x^{(3)} = \begin{pmatrix} x_N^{(3)} \\ x_B^{(3)} \end{pmatrix} = (0,2,4,0)^T, z = 8$
- 选$B = (p_2\ p_4) \Rightarrow x^{(4)} = \begin{pmatrix} x_N^{(4)} \\ x_B^{(4)} \end{pmatrix} = (0,4,0,-2)^T, z = 8$



- 单纯形法:从A点出发，此时用x_N表示x_B, $z \Rightarrow x_B = \begin{pmatrix} 8-x_1-2x_2 \\ 2-x_2 \end{pmatrix}, z = x_1 + 4x_2$,因此$x_2$进基，由于$\min\left(\frac{8}{2},\frac{2}{1}\right) = 2$，因此$x_4$离基，因此当前基为$(p_2, p_3)$,即E点

2.1 线性规划的基本概念-线性规划问题的基本理论(续4)(注2)

- 因此当前基为(p_2,p_3),即E点，$x_B=\begin{pmatrix}x_2\\x_3\end{pmatrix}=\begin{pmatrix}2-x_4\\4-x_1+2x_4\end{pmatrix}$, $z=8+x_1-4x_4$,此时最优值$z=8$
- 此时x_1系数为正，选为入基变量，由于增加x_1,先使x_3为0,因此x_3离基，此时基为(p_1,p_2)，对应C点，$x_B=\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}4-x_3+2x_4\\2-x_4\end{pmatrix}$, $z=12-x_3-2x_4$,此时最优值$z=12$,变量系数均为负，即为最优值



- 可以看出，采用代数法，先找到A点，然后通过入基变量和离基变量，分别遍历其E点，然后迭代至C点，确认为最优点，显然D点并没有出现在步骤中。
- 下面看看单纯形表格

◆单纯形表

c_j			1	4	0	0	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	θ
0	x_3	8	1	2	1	0	$\frac{8}{2}$
0	x_4	2	0	**1**	0	1	$\frac{2}{1}$
$-z$		0	1	4	0	0	

c_j			1	4	0	0	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	θ
0	x_3	4	1	0	1	-2	∞
4	x_2	2	0	**1**	0	1	$\frac{2}{1}$
$-z$		-8	1	0	0	-4	

c_j			1	4	0	0	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	θ
0	x_3	4	**1**	0	1	-2	4
4	x_2	2	0	**1**	0	1	∞
$-z$		-8	1	0	0	-4	

c_j			1	4	0	0	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	θ
1	x_1	4	**1**	0	1	-2	
4	x_2	2	0	**1**	0	1	
$-z$		-12	0	0	-1	-2	

2.2 单纯形法(simplex)

- **把所有基本可行解找出来，这么做是否可行？**
  - $C_n^m: C_{40}^{20} \approx 1.3 \times 10^{11}$
- **是否可以从一个基本可行解，去寻找目标函数值更优的基本可行解呢？(只往好的方向去寻找)**
  - 如何找到一个初始的BFS
  - 如何判断BFS已经是最优解
  - 如何去寻找一个更优的BFS
- **丹齐格(G.B.Dantzig)提出的单纯形法(Simplex)解决了这个问题**
- $Ax = b \Rightarrow (B \ \ N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} = Bx_B + Nx_N = b$, 这时进一步可写为$Bx_B = b - Nx_N \Rightarrow$ $x_B = B^{-1}(b - Nx_N)$, $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$, $A = [p_1\, p_2 \cdots p_k \cdots p_n]$

2.2 单纯形法(simplex)(续1)

◆ 例子，LP 问题$Max\ z = 1500x_1 + 2500x_2$； s.t.$\begin{cases} 3x_1 + 2x_2 \le 65 \\ 2x_1 + \ \ x_2 \le 40 \\ \qquad\ \ 3x_2 \le 75 \\ x_1, x_2 \ge 0 \end{cases}$

化为标准型:

$$Max\ z = 1500x_1 + 2500x_2 + 0x_3 + 0x_4 + 0x_5;\ \text{s.t.}\begin{cases} 3x_1 + 2x_2 + x_3 = 65 \\ 2x_1 + \ \ x_2 + x_4 = 40 \\ \quad 3x_2 + x_5 = 75 \\ x_1, x_2, x_3, x_4, x_5 \ge 0 \end{cases}$$

第一步:找初始的基本可行解，$A = \begin{pmatrix} 3 & 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}$，选$A$中三列线性无关的，这里显然$p_3, p_4, p_5$正好是三阶单位阵，然后用非基变量来表示基变量和目标函数:

$$\begin{cases} x_3 = 65 - 3x_1 - 2x_2 \\ x_4 = 40 - 2x_1 - x_2 \\ x_5 = 75 - 3x_2 \end{cases},z = 1500x_1 + 2500x_2$$

然后令非基变量为0，可知$x_B = (x_3, x_4, x_5) = (65, 40, 75), z = 0$

2.2 单纯形法(simplex)(续2)

- $x_N=(x_1,x_2)^T=(0,0)^T$,这时函数值$z=1500x_1+2500x_2=0$,显然不是最优的，增加x_1或x_2的值都可令目标函数值增加，但x_2增加更快
  $\begin{cases} x_3=65-3x_1-2x_2 \\ x_4=40-2x_1-x_2 \\ x_5=75-3x_2 \end{cases}$，此时为确保$(x_3,x_4,x_5)\geq 0$,而$x_1$为0，则需要将$(x_3,x_4,x_5)$中某个变量移出基变量(称为Leaving variable)，而x_2进入基变量中(称为Entering variable)，如何选取离基变量呢？$min\{\frac{65}{2},\frac{40}{1},\frac{75}{3}\}$对应的变量，这时增加$x_2$会导致该变量$x_5$最先达到约束边界。这时新的基变量为$x_B=\{x_2,x_3,x_4\}$, $x_N=\{x_1,x_5\}$,可行基为$\{p_2,p_3,p_4\}$,再次用非基变量表示基变量和目标函数:
- $\begin{cases} x_2= 25-1/3x_5 \\ x_3= 15-3x_1+2/3x_5 \\ x_4=15-2x_1+1/3x_5 \end{cases}$, $z=62500+1500x_1-\frac{2500}{3}x_5$,**这时$x_N=(x_1,x_2)^T=$** $(0,0)^T$,这时函数值$z=62500$,但是如果增加x_1的值，目标函数值仍然会增加，因此将x_1作为入基变量，$min\{25,\frac{15}{3},\frac{15}{2}\}$对应的$x_3$作为出基变量，再次用$x_N$表示$x_B$和目标函数

2.2 单纯形法(simplex)(续3)

- $$\begin{cases} x_1 = 5 - 1/3x_3 + 2/9x_5 \\ x_2 = 25 - 1/3x_5, \\ x_4 = 5 + 2/3x_3 - 1/9x_5 \end{cases}$$
- 目标函数$z = 70000 - 500x_3 - 500x_5$
- 此时目标函数中的非基变量系数均为负，目标函数已经最优，此时最优解为$x^* = (5, 25, 0, 5, 0)^T$

2.2 单纯形法(simplex)(续3)

◆单纯形法的基本步骤

- 1)构造一个初始基本可行解：对已经标准化的模型，设法从约束矩阵中构造一个m阶的单位阵
- 2)判断当前基本可行解是否为最优解：求出用非基变量x_N来表示基变量x_B和目标函数z，这个称为LP问题的典式(canonical form规范式)，将目标函数的典式中非基变量前的系数称为检验数，最大问题所有检验数为非正数，则当前解即为最优解
- 3)若当前解不是最优解，则进行基变换，迭代到下一个基本可行解：进基变量从目标函数典式中选取正的最大检验数对应的非基变量；再从当前基变量中选取一个离基变量，选取准则为除进基变量外，令其余非基变量为0，再按最小比值准则确立离基变量，然后回到2)循环至找到最优解或者判断无最优解结束。

2.2 单纯形法(simplex)(续4)

◆单纯形的基本步骤

➢1)确定初始基本可行解

✓约束都是≤，直接添加松弛变量，构成单位阵

✓约束包含≥，减去松弛变量，添加一个人工变量，构成单位阵

✓约束包含=，直接添加人工变量，构成单位阵

➢例如确定 $\begin{cases} 3x_1+2x_2-4x_3\le 50 \\ 2x_1-x_2+3x_3\ge 30 \\ x_1+x_2+x_3=20 \\ x_1,\ x_2,\ x_3\ge 0 \end{cases}$ 的初始基本可行解

➢经过引入4个变量后，其基矩阵为$[p_4\ p_6\ p_7]$,对应基变量为$[x_4, x_6, x_7]$

2.2 单纯形法(simplex)(续5)

◆单纯形的基本步骤

- 2)最优性检验(判断当前的基本可行解是否为最优解)
- 假设初始基本可行解为$x^{(0)}=(b_1,b_2,\cdots,b_m,0,\cdots,0)^T$则用非基变量表示基变量可得:$x_i=b_i-\sum_{j=m+1}^{n}a_{ij}x_j$, $i=1\cdots m$,注意这里的b_i,a_{ij}经过多次迭代后与原始的不一样了，后续均用′来表示，然后代入目标函数，整理后可得:
- $z=\sum_{i=1}^{m}c_ix_i+\sum_{j=m+1}^{n}c_jx_j=\sum_{i=1}^{m}c_ib_i'+\sum_{j=m+1}^{n}\left(c_j-\sum_{i=1}^{m}c_ia_{ij}'\right)x_j$, 蓝色部分记为$\sigma_j$,称为非基变量的检验数，$\sum_{i=1}^{m}c_ib_i'=z_0$为当前基本可行解的目标函数值
- 定理：如果关于非基变量的所有检验数$\sigma_j\le 0$,则当前基本可行解就是最优解(注意对Max问题)

2.2 单纯形法(simplex)(续6)

◆单纯形的基本步骤

- 基变换(进基变量和离基变量的选取)
  - ✓如果有多个非基变量的检验数为正，则选取最大的检验数对应的非基变量作为进基变量(Entering variable)：$\sigma_{m+t} = max_{j\in J_N}\sigma_j(\sigma_j > 0)$, 注意这里用$J_N$，$J_B$分别表示非基变量和基变量的下标集合
  - ✓离基变量选取为：当进基变量逐渐增大时，当前基变量值最先下降到0值的变量为出基变量
  - ✓如果选取的基均为单位基向量，则其操作方便，因此可组成单纯形表

◆无穷多最优解及无界解的情况

- 若当前基本可行解中，非基变量x_j对应的检验数$\sigma_j \le 0, j \in J_N$，且其中有一个非基变量$x_{m+t}$对应的检验数$\sigma_{m+t} = 0$，则该LP问题有无穷多个最优解
- 若当前基本可行解中，非基变量x_{m+t}对应的检验数$\sigma_{m+t} > 0$，且其对应的系数列向量p'_{m+t}中所有分量$a'_{i,m+t} \le 0, i = 1 \cdots m$

则该LP问题有无界解(或称无最优解)

2.2 单纯形法(simplex)(续7)

◆单纯形表

$$\max z = 2x_1 + 5x_2$$
$$\begin{aligned} s.t.\quad & x_1 + 2x_2 \le 8 \\ & x_1 \le 4 \\ & x_2 \le 3 \\ & x_1, x_2 \ge 0 \end{aligned}$$

◆化为标准型

$$\max\quad z = 2x_1 + 5x_2 + 0x_3 + 0x_4 + 0x_5$$
$$\begin{aligned} s.t.\quad & x_1 + 2x_2 + x_3 + 0x_4 + 0x_5 = 8 \\ & x_1 + 0x_2 + 0x_3 + x_4 + 0x_5 = 4 \\ & 0x_1 + x_2 + 0x_3 + 0x_4 + x_5 = 3 \\ & x_1, x_2, x_3, x_4, x_5 \ge 0 \end{aligned}$$

◆矩阵形式

$$\max c^T x$$
$$s.t.\quad Ax = b,\ x \ge 0$$
$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} = [p_1\ p_2\ p_3\ p_4\ p_5]$$
$$b = [8\ 4\ 3]^T,\ c = [2\ 5\ 0\ 0\ 0]^T$$

$c_j \rightarrow$			c_1	c_2	c_3	c_4	c_5	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	x_5	θ
c_1	x_1	b_1	1	0	0	a_{14}	a_{15}	
c_2	x_2	b_2	0	1	0	a_{24}	a_{25}	
c_3	x_3	b_3	0	0	1	a_{34}	a_{35}	
	$-z$	$-z_0$	0	0	0	σ_4	σ_5	

◆单纯形表

$c_j \rightarrow$			c_1	c_2	c_3	c_4	c_5	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	x_5	θ
c_1	x_1	b_1	1	0	0	a_{14}	a_{15}	
c_2	x_2	b_2	0	1	0	a_{24}	a_{25}	
c_3	x_3	b_3	0	0	1	a_{34}	a_{35}	
	$-z$	$-z_0$	0	0	0	σ_4	σ_5	

◆**注意:$z_0 = \sum_{i=1}^{m} c_i b_i$,$\sigma_i = c_i - \sum_{k=1}^{m} c_k a_{ki}$,取最大值作为进基变量，不妨设为$k$,即$x_k$为进基变量，最右边的一列为判断离基变量的值:$b_i/a_{ik}$,如果$a_{ik} = 0$,该值设为无穷大，选择最小的作为离基变量**

2.2 单纯形法(simplex)(续7)(注2)

◆用单纯形表计算

$$\max c^T x$$

$$s.t.\ \ Ax = b,\ x \geq 0$$

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} = [p_1\ p_2\ p_3\ p_4\ p_5]$$

$$b = [8\ 4\ 3]^T,\ \ c = [2\ 5\ \ 0\ \ 0\ \ 0]^T$$

◆练习

◆最优值为19，最优解为$x^* = (2,3,0,2,0)^T$

◆注意：如果计算过程中出现检验数最大的正列k对应的列$a_{ik} \leq 0$，则无有限最优解，如果对应的检验数为0，则有无穷多组最优解

2.2 单纯形法(simplex)(续7)(注3)

注意：这几个例子对应PPT29-31页图解法的情况

◆ 例 $Max\ z = 1500x_1 + 2500x_2$， s.t. $\begin{cases} 3x_1 + 2x_2 \le 65 \\ 2x_1 + x_2 \le 40 \\ 3x_2 \le 75 \\ x_1, x_2 \ge 0 \end{cases}$,如何用单纯形方法求解？

◆ 最优解: $x^* = (5, 25, 0, 5, 0)^T, z^* = 7000$,单纯形表最后的检验数为 $(0, 0, -500, 0, -500)$

◆ 例 $Max\ z = 1500x_1 + 1000x_2$， s.t. $\begin{cases} 3x_1 + 2x_2 \le 65 \\ 2x_1 + x_2 \le 40 \\ 3x_2 \le 75 \\ x_1, x_2 \ge 0 \end{cases}$,

◆ 最优解: $x^* = (15, 10, 0, 0, 45)^T, z^* = 32500$,单纯形表最后的检验数为 $(0, 0, -500, 0, 0)$,注意此时有无限多最优解

2.2 单纯形法(simplex)(续7)(注4)

◆ 单纯形表格如下，体会里面的运算过程

C_B	X_B	b	1500	2500	0	0	0	θ_i
			x_1	x_2	x_3	x_4	x_5	
0	x_3	65	3	2	1	0	0	32.5
0	x_4	40	2	1	0	1	0	40
0	x_5	75	0	(3)	0	0	1	25
$-z$		0	1500	2500*	0	0	0	
0	x_3	15	(3)	0	1	0	-2/3	5
0	x_4	15	2	0	0	1	-1/3	7. 5
2500	x_2	25	0	1	0	0	1/3	
$-z$		-62500	1500*	0	0	0	-2500/3	
1500	x_1	5	1	0	1/3	0	-2/ 9	
0	x_4	5	0	0	-2/3	1	1/9	
2500	x_2	25	0	1	0	0	1/3	
$-z$		-70000	0	0	-500	0	-500	

◆ **注意，**画红框的列要保持非负，因为对应资源约束，有几何意义；画蓝框的检验数非正后，达到最优，如果此时非基变量检验数有0值，则有无限多最优解，确保进基变量为单位向量，这通过初等行变换来完成，注意Ax=b参与

2.2 单纯形法(simplex)(续7)(注5)

- 例$Max\ z = 7x_1 + x_2 - 4x_3 + 2x_4$，s.t. $\begin{cases} x_1 + 4x_2 - 2x_3 + x_4 + x_5 = 5 \\ 2x_1 - x_2 - 5x_3 + 3x_4 + x_6 = 11 \\ -x_1 + 3x_2 + x_3 + x_7 = 15 \\ x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0 \end{cases}$
- 使用单纯形表格计算时，第二个单纯形表格出现第3个变量x_3检验数大于0，但对应的基向量均为非正值，因而无有限最优解，计算终止

- 例$Max\ z = \frac{3}{4}x_4 - 20x_5 + \frac{1}{2}x_6 - 6x_7$，s.t. $\begin{cases} x_1 + \frac{1}{4}x_4 - 8x_5 - x_6 + 9x_7 = 0 \\ x_2 + \frac{1}{2}x_4 - 12x_5 - \frac{1}{2}x_6 + 3x_7 = 0 \\ x_3 + x_6 = 1 \\ x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0 \end{cases}$
- 使用单纯形法，$x_B = (x_1, x_2, x_3)^T, B = (p_1, p_2, p_3)$,经过六次迭代又回到$B$,因此求不到最优解，但实际上$x^* = \left(\frac{3}{4}, 0, 0, 1, 0, 1, 0\right)^T, z^* = 5/4$
- 如果一个基本可行解存在取0值的基变量，则称为退化的基本可行解，对应的基称为退化基，存在退化的情况，有可能求不到最优解
- 采用摄动法，字典序法，最简单的方法：出基和入基变量选取下标最小的变量
- 可以具体分析PPT第29-31页的例子，进一步理解

2.2 单纯形法-标准解决办法

◆当初始基本可行解不能通过观察法很容易得到时，如何确定初始基本可行解？

$$\begin{cases} a_{11}x_1+\cdots a_{1n}x_n=b_1 \\ a_{21}x_1+\cdots a_{2n}x_n=b_2 \\ \quad\vdots \qquad\qquad \vdots \\ a_{m1}x_1+\cdots a_{mn}x_n=b_m \\ x_i,b_i\ge 0 \end{cases} \xRightarrow{\text{人工变量法}} \begin{cases} a_{11}x_1+\cdots a_{1n}x_n+x_{n+1}=b_1 \\ a_{21}x_1+\cdots a_{2n}x_n+x_{n+2}=b_2 \\ \quad\vdots \qquad\qquad \vdots \\ a_{m1}x_1+\cdots a_{mn}x_n+x_{n+m}=b_m \\ x_i,b_i\ge 0,x_{n+i}\ge 0 \end{cases}$$

$x_{n+i}(i=1,\cdots,m)$称为人工变量. $(\overbrace{0,\cdots,0}^{n},b_1,\cdots,b_m)^T$即为初始基本可行解.

- **带来的问题：人工变量的引入，改变了原问题的约束条件，得到的是与原问题不同的新问题，而新问题的最优解不一定是原问题的最优解（除非新问题的最优解正好人工变量都为零）.（人工变量是“非法”的变量，松弛变量是“合法”的变量）**
- **解决方法：**
  - 大M法：惩罚法，引入$M>0$为一个充分大的正数，在原目标函数中加入$-M$乘上每个变量

$$max\ z=c_1x_1+\cdots+c_nx_n-Mx_{n+1}-\cdots-Mx_{n+m}$$

因此求解这个新问题就是从最大化角度迫使人工变量为0，从而达到求解目的。

$x^*=(x_1^*,\cdots,x_n^*,x_{n+1}^*,\cdots,x_{n+m}^*)$为最优解，则令人工变量为0，剩余的变量即为原问题的最优解，人工变量不全为0，说明原问题无可行解

  - 两阶段法：分为两步，第一步约束不变，目标函数改为：$max\ z'=-x_{n+1}-\cdots-x_{n+m}$,尽量迫使变量为0，达到求解原问题的一个基本可行解的目的；第二步，求解原问题，以第一步得到的基本可行解求解原问题。
  - 第一阶段$x^*=(x_1^*,\cdots,x_n^*,x_{n+1}^*,\cdots,x_{n+m}^*)$为最优解，则令人工变量为0，剩余的变量即为原问题的一个基本可行解，这时目标函数值为0；否则人工变量不全为0，说明原问题无可行解

2.2 单纯形法-大M法

$$LP:\begin{cases}\max\ c^T x\\ s.t.\quad Ax=b,\\ \qquad\quad x\ge 0.\end{cases}\Rightarrow DMLP:\begin{cases}\max\ c^T x-M\vec{1}^T x_a\\ s.t.\quad Ax+x_a=b,\\ \qquad\quad x\ge 0, x_a\ge 0.\end{cases}$$

其中，A为$m\times n$矩阵,$b\ge 0$,M是个很大的正数,x_a是人工变量

- **若$(x,0)^T$是DMLP的有限最优解，则x是LP的最优解**
- **若$(x,x_a)^T(x_a\neq 0)$是DMLP的有限最优解，则LP无可行解**
- **若DMLP无有限最优解，则LP或者无有限最优解或者无可行解**

2.2 单纯形法-大M法(续1)

- 例：大M法求解$LP: max\ z=3x_1-x_2-x_3, s.t.\begin{cases} x_1-2x_2+x_3\le 11 \\ -4x_1+x_2+2x_3\ge 3 \\ -2x_1+x_3=1 \\ x_1,x_2,x_3\ge 0 \end{cases}$
- 引入人工变量并标准化:$max\ z=3x_1-x_2-x_3+0x_4+0x_5-Mx_6-Mx_7, s.t. Ax=b, A=\begin{pmatrix} 1 & -2 & 1 & 1 & 0 & 0 & 0 \\ -4 & 1 & 2 & 0 & -1 & 1 & 0 \\ -2 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}, b=\begin{pmatrix} 11 \\ 3 \\ 1 \end{pmatrix}$
- 使用单纯形法，$B^{(0)}=(p_4\ p_6\ p_7), B^{(1)}=(p_4,p_6,p_3), B^{(2)}=(p_4,p_2,p_3)$, $B^{(3)}=(p_1,p_2,p_3)$，最终$(\sigma_i)=(0\ 0\ 0\ -\frac{1}{3}-\frac{1}{3}\left(-M+\frac{1}{3}\right)\ \left(-M+\frac{2}{3}\right))$,已到最优，此时$x^*=(4,1,9,0,0,0,0)^T, z^*=2$,从而得到最优解

2.2 单纯形法-大M法(续2)

表格

c_B	x_B	b	3	-1	-1	0	0	$-M$	$-M$	
			x_1	x_2	x_3	x_4	x_5	x_6	x_7	θ
0	x_4	11	1	-2	1	1	0	0	0	11
$-M$	x_6	3	-4	1	2	0	-1	1	0	$3/2$
$-M$	x_7	1	-2	0	1	0	0	0	1	$1/1$
$-z$		$4M$	$3-6M$	$-1+M$	$-1+3M$	0	$-M$	0	0	
0	x_4	10	3	-2	0	1	0	0	-1	---
$-M$	x_6	1	0	1	0	0	-1	1	-2	1
-1	x_3	1	-2	0	1	0	0	0	1	∞
-z		$-M-1$	1	$-1+M$	0	0	$-M$	0	-$3M+1$	
0	x_4	12	3	0	0	1	-2	2	-5	4
-1	x_2	1	0	1	0	0	-1	1	-2	∞
-1	x_3	1	-2	0	1	0	0	0	1	---
-z		2	1	0	0	0	-1	$-M+1$	$-M-1$	
3	x_1	4	1	0	0	$1/3$	$-2/3$	$2/3$	$-5/3$	
-1	x_2	1	0	1	0	0	-1	1	$-2/3$	
-1	x_3	9	0	0	1	$2/3$	$-4/3$	$4/3$	$-7/3$	

2.2 单纯形法-两阶段法

◆ **Two-Phase method: First Phase, Second Phase**

$$\text{FPLP}:\begin{cases}\max\ -\vec{1}^{T}x_a\\ s.t.\quad Ax+x_a=b,\\ \qquad x\ge 0, x_a\ge 0.\end{cases}\qquad \text{SPLP}:\begin{cases}\max\ c^{T}x\\ s.t.\quad Ax=b,\\ \qquad x\ge 0.\end{cases}$$

其中A为$m\times n$矩阵，$b\ge 0$，x_a是人工变量

◆ **若$(x,x_a)^T(x_a=0)$是FPLP的有限最优解，且x_a的分量都是非基变量，则x是SPLP的一个基本可行解**

◆ **若$(x,x_a)^T(x_a>0)$是FPLP的有限最优解，则SPLP无可行解**

◆ **若$(x,x_a)^T(x_a=0)$是FPLP的有限最优解，且x_a的某些分量是基变量，则可通过主元素消去法，用原变量中的非基变量，替换出基变量中的人工变量，得到SPLP的一个基本可行解**

2.2 单纯形法-两阶段法-例子

◆ 用两阶段法求解$max\ z = 3x_1 - x_2 - x_3, s.t. \begin{cases} x_1 - 2x_2 + x_3 \le 11 \\ -4x_1 + x_2 + 2x_3 \ge 3 \\ -2x_1 + x_3 = 1 \\ x_1, x_2, x_3 \ge 0 \end{cases}$

◆ 第一步，标准化并引入人工变量，建立辅助问题，$max\ z' = -x_6 - x_7, s.t. \begin{cases} x_1 - 2x_2 + x_3 + x_4 = 11 \\ -4x_1 + x_2 + 2x_3 - x_5 + x_6 = 3 \\ -2x_1 + x_3 + x_7 = 1 \\ x_i \ge 0 \end{cases}$

◆ 用单纯形表，从$x_B^{(0)} = (x_4, x_6, x_7)^T \Rightarrow x_B^{(1)} = (x_4, x_6, x_3)^T \Rightarrow x_B^{(2)} = (x_4, x_2, x_3)^T$,此时检验数$(\sigma_i) = (0,0,0,0,0,-1,-1)$, 得最优解$x^* = (0,1,1,12,0,0,0)^T$,最优值$z'^* = 0$,由于人工变量的值均为0，故得标准化后的原问题的基本可行解为$x = (0,1,1,12,0)^T$

◆ 第二步，产生初始单纯形：a)由第一步的最优单纯形表删除第$x_{n+1}, \cdots, x_{n+m}$列；b)把第一行的目标函数系数行换为原问题目标函数的系数；c)重新计算检验数行；

➢ 得到最优解$x^* = (4,1,9,0,0)^T$,最优值$z^* = 2$,从而返回原问题的最优解和最优值

2.2 单纯形法-两阶段法-例子

- $Max\ z = 5x_1 + 2x_2 + 3x_3 - x_4$

$$s.t.\ \begin{cases} x_1 + 2x_2 + 3x_3 = 15 \\ 2x_1 + x_2 + 5x_3 = 20 \\ x_1 + 2x_2 + 4x_3 + x_4 = 26 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

- **第一阶段(FPLP):** $Max\ z = -x_5 - x_6$

$$s.t.\ \begin{cases} x_1 + 2x_2 + 3x_3 + x_5 = 15 \\ 2x_1 + x_2 + 5x_3 + x_6 = 20 \\ x_1 + 2x_2 + 4x_3 + x_4 = 26 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{cases}$$

- **采用单纯形法解FPLP**

2.2 单纯形法-两阶段法-例子

◆ FPLP单纯形表格如下

			0	0	0	0	-1	-1	
C_B	X_B		x_1	x_2	x_3	x_4	x_5	x_6	θ_i
-1	x_5	15	1	2	3	0	1	0	5
-1	x_6	20	2	1	(5)	0	0	1	4
0	x_4	26	1	2	4	1	0	0	6.5
-z		35	3	3	8	0	0	0	
-1	x_5	3	-1/5	(7/5)	0	0	1	-3/5	15/7
0	x_3	4	2/5	1/5	1	0	0	1/5	20
0	x_4	10	-3/5	6/5	0	1	0	-4/5	25/3
-z		3	-1/5	7/5	0	0	0	-8/5	
0	x_2	15/7	-1/7	1	0	0	5/7	-3/7	
0	x_3	25/7	3/7	0	1	0	-1/7	2/7	25/3
0	x_4	52/7	-3/7	0	0	1	-6/7	-2/7	
-z		0	0	0	0	0	-1	-1	

◆ 由PPT第60页的结论，得到原问题的基本可行解：$(0,\ 15/7,\ 25/7,\ 52/7)^T$

◆第二阶段(SPLP)，先按照PPT61页下面的方法构造单纯形

			5	2	3	-1	
C_B	X_B		x_1	x_2	x_3	x_4	θ_i
2	x_2	15/7	-1/7	1	0	0	
3	x_3	25/7	(3/7)	0	1	0	25/3
-1	x_4	52/7	-3/7	0	0	1	
$-z$		-53/7	25/7	0	0	0	
2	x_2	10/3	0	1	1/3	0	
5	x_1	25/3	1	0	7/3	0	
-1	x_4	11	0	0	1	1	
$-z$		-112/3	0	0	-25/3	0	

◆从而得到原问题的最优解: $\left(\frac{25}{3}, \ \frac{10}{3}, \ 0, \ 11\right)^T$,最优目标值: $\frac{112}{3}$

2.3 对偶问题(Duality)

	产品甲	产品乙	设备能力(h)
设备A	3	2	65
设备B	2	1	40
设备C	0	3	75
利润（元/件）	1500	2500	

◆再看本章开始的例子

- 假设工厂考虑不安排生产，而准备将设备出租，收取租费。于是需要为每种设备的台时进行估价
- 假设A,B,C台时租费估价为y_1, y_2, y_3，由上面的表格可知，生产一件产品甲，需要台时3,2,0，如果将其不用于生产，而是出租，得到租费：$3y_1+2y_2+0y_3$,为了不低于生产利润，因此要求$3y_1+2y_2+0y_3 \geq 1500$
- 同理，对产品乙:$2y_1+1y_2+3y_3 \geq 2500$,且$y_i \geq 0$
- 企业现在总台时数为65,40,75，都用于出租，则总收入为：$f(\mathrm{y})=65y_1+40y_2+75y_3$,但出租费用应该尽可能低，才能是出租计划成交，因此要求$\min f(y)$
- 这样就得到了对偶规划问题
  - ✓对减少管理工作的盲目性提供了更多的科学依据，与原规划问题互相对应，从不同的角度对企业的经营管理进行分析研究

- 例子
- 某家具厂木器车间生产木门与木窗两种产品，木门利润56元/扇，木窗30元/扇。生产木门需要木工4小时，油漆工2小时；生产木窗木工3小时，油漆工1小时。车间每日可用木工总工时120小时，油漆工50小时。问该如何安排生产才能是每日利润对大？
  - $\mathrm{Max}\ z=56x_1+30x_2, s.t. 4x_1+3x_2\le 120, 2x_1+x_2\le 50, x_1,x_2\ge 0$
  - 最优解$x^*=(x_1,x_2)^T=(15,20)^T=1440$（元）

2.3 对偶问题(Duality)

- 某家具厂木器车间生产木门与木窗两种产品，木门利润56元/扇，木窗30元/扇。生产木门需要木工4小时，油漆工2小时；生产木窗木工3小时，油漆工1小时。车间每日可用木工总工时120小时，油漆工50小时。问该如何安排生产才能是每日利润对大？
- 从另一个角度思考：假若有一个个体经营者，手中有一批木器家具生产订单，他想利用该木器车间的木工与油漆工来加工完成他的订单。就需要考虑付给该车间每个工时的价格。他可以构造一个数学模型来研究如何定价才能既使木器车间觉得有利可图从而愿意替他加工这批订单，又使自己所付的工时费用总数最少呢？

2.3 对偶问题(Duality)

- 设w_1为付给木工每个工时的价格，w_2为付给油漆工每个工时的价格，则该个体经营者的目标函数为每日所付工时费用最小：
- $\min f = 120w_1 + 50w_2$
- 注意：该个体经营者所付的价格不能太低，至少不能低于该车间生产木门、木窗所得到的收入，否则该车间觉得无利可图就不会给他加工这批订单，因此：
- $4w_1 + 2w_2 \geq 56; 3w_1 + w_2 \geq 30, w_1, w_2 \geq 0$
- 此时最优解：$w^* = (w_1^*, w_2^*)^T = (2, 24)^T, f^* = 1440$

2.3 对偶问题(续1)

◆ 对称形式的对偶问题

◆ $LP:\begin{cases} max\ z = c^T x \\ s.t. Ax \le b \\ x \ge 0_{n\times1} \end{cases}$ $DP:\begin{cases} min\ f = b^T y \\ A^T y \ge c \\ y \ge 0_{1\times m} \end{cases}$

◆ 对称性定理：对偶问题的对偶是原问题

◆ 练习，求$LP: max\ z = 2x_1 + 2x_2, s.t.\begin{cases} 2x_1 + 4x_2 \le 1 \\ x_1 + 2x_2 \le 1 \\ 2x_1 + x_2 \le 1 \\ x_1, x_2 \ge 0 \end{cases}$的对称形式的对偶问题

2.3 对偶问题(续2)

◆如果原问题是标准形式，如何定义其对偶问题

◆$LP: \max c^T x, s.t. \begin{cases} Ax = b \\ x \geq 0 \end{cases} \Rightarrow \begin{cases} Ax \leq b \\ Ax \geq b \\ x \geq 0 \end{cases} \Rightarrow$

$$\begin{cases} Ax \leq b \\ -Ax \leq -b \\ x \geq 0 \end{cases} \Rightarrow \begin{cases} \begin{pmatrix} A \\ -A \end{pmatrix} x \leq \begin{pmatrix} b \\ -b \end{pmatrix} \\ x \geq 0 \end{cases}$$

◆$DP: \min b^T w; s.t. \begin{cases} A^T w \geq c \\ w\text{无正负限制} \end{cases}$

◆如何化为非对称的对偶规划？

2.3 对偶问题(续3)

- **如何化为非对称的对偶规划**
  - a)将模型统一为max, ≤或者$\boldsymbol{min}$, ≥，对其中的等式或无非负约束变量执行下面步骤
  - b)对等式约束，则在对偶规划中与此约束对应的那个变量没有非负限制
  - c)若某个变量没有非负限制，则对偶问题于此对应的的变量对应的约束为等式

- 例$LP: max\ z = x_1 - x_2 + 5x_3 - 7x_4, s.t. \begin{cases} x_1 + 3x_2 - 2x_3 + x_4 = 25 \\ 2x_1 + 7x_3 + 2x_4 \geq -60 \\ 2x_1 + 2x_2 - 4x_3 \leq 30 \\ -5 \leq x_4 \leq 10, x_1, x_2 \geq 0, x_3\text{没有非负限制} \end{cases}$

- 先将约束变为≤ 的形式$s.t. \begin{cases} x_1 + 3x_2 - 2x_3 + x_4 = 25 \\ -2x_1 - 7x_3 - 2x_4 \leq 60 \\ 2x_1 + 2x_2 - 4x_3 \leq 30 \\ x_4 \leq 10 \\ -x_4 \leq 5 \\ x_1, x_2 \geq 0, x_3, x_4\text{没有非负限制} \end{cases}$ =>直接写出DP问题

- $min\ f = 25y_1 + 60y_2 + 30y_3 + 10y_4 + 5y_5, s.t. \begin{cases} y_1 - 2y_2 + 2y_3 \geq 1 \\ 3y_1 + 2y_3 \geq -1 \\ -2y_1 - 7y_2 - 4y_3 = 5 \\ y_1 - 2y_2 + y_4 - y_5 = -7 \\ y_2, y_3, y_4, y_5 \geq 0, y_1\text{没有非负限制} \end{cases}$

2.3 对偶问题(续3注)

◆变量约束的对应关系

PRIMAL – DUAL TABLE	
primal (maximize)	*dual* (minimize)
$\underline{A}$ (coefficient matrix)	$\underline{A}^T$ (transpose of the coefficient matrix)
$\underline{b}$ (right-hand side vector)	$\underline{b}$ (cost vector)
$\underline{c}$ (price vector)	$\underline{c}$ (right hand side vector)
i^{th} constraint is $=$ type	the dual variable y_i is unrestricted in sign
i^{th} constraint is $\leq$ type	the dual variable $y_i \geq 0$
i^{th} constraint is $\geq$ type	the dual variable $y_i \leq 0$
x_j is unrestricted	j^{th} dual constraint is $=$ type
$x_j \geq 0$	j^{th} dual constraint is $\geq$ type
$x_j \leq 0$	j^{th} dual constraint is $\leq$ type

2.3 对偶问题(续4)

◆ 对偶定理LP: $\begin{cases} \max z = c^T x \\ s.t.\ Ax \le b \\ x \ge 0 \end{cases}$ DP: $\begin{cases} \min f = b^T y \\ A^T y \ge c \\ y \ge 0 \end{cases}$

◆ 定理(弱对偶定理)：若x, y分别是LP, DP问题的可行解，则$c^T x \le b^T y$

- 推论(最优性准则定理)：若x^0, y^0分别是LP, DP问题的可行解，当$c^T x^0 = b^T y^0$时,若x^0, y^0分别是LP, DP问题的最优解
- 推论：若LP有可行解，则LP有最优解的充要条件是DP有可行解
- 推论：若DP有可行解，则DP有最优解的充要条件是LP有可行解

◆ 从而利用对偶理论容易判断LP问题是否存在最优解

- 若LP存在可行解，而其DP问题没有可行解，则LP问题无最优解
- 若LP存在可行解，而其DP问题也存在可行解，则两个问题都有最优解

◆ 定理(主对偶定理)：若原规划LP问题有最优解，则对偶规划问题DP也有最优解，反之亦然，并且两者的目标函数值相等

2.3 对偶问题(续5)

◆ 定理：若原规划LP问题有最优解，则对偶规划问题DP也有最优解，反之亦然，并且两者的目标函数值相等

◆ 证明： $LP:\begin{cases} max\ z = c^T x \\ s.t. Ax \le b \\ x \ge 0 \end{cases}$，引入松弛变量$x_s$,约束化为$Ax + Ex_s = b$,令$A^s = (A, E), x^s = \left(x^T, x_s^T\right)^T, c^s = \left(c^T, 0^T\right)^T$,从而的标准形式LP1: $\begin{cases} max\ z = c^T x \\ s.t. A^s x^s \le b \\ x^s \ge 0 \end{cases}$,设$B$为LP1的最优基，现在证明DP问题也有最优解，由单纯形法可知：检验数$\sigma^T = c_B^{sT} - c_B^{sT} B^{-1} A^s \le 0$,即$c_B^{sT} \le c_B^{sT} B^{-1} A^s \Rightarrow \left(\mathrm{c}^{\mathrm{T}}, 0^{\mathrm{T}}\right) \le c_B^{sT} B^{-1}(A, E) \Rightarrow c^T \le c_B^{sT} B^{-1} A, c_B^{sT} B^{-1} \ge 0^T$,令$y^T = c_B^{sT} B^{-1}$,则有$y^T A \ge c^T, y \ge 0$,即为DP问题的可行解，即有$c^T x = b^T y$,其中$x$为原规划问题的最优解，从而$y$即为DP问题的最优解

◆ 注意，DP问题的最优解可以从检验数得到，为什么？

2.3 对偶问题(续6)

◆**检验数**$\sigma^T = c_B^{sT} - c_B^{sT} B^{-1} A^s \le 0$，DP问题的最优解$y^T = c_B^{sT} B^{-1}$，**注意**$A^s$**的后**$m$**列为单位矩阵，**$c^s$**的后**$m$**个分量为0，所以**$\sigma^T$**的展开式为**$\sigma^T = (\sigma_1, \cdots, \sigma_n, \cdots, \sigma_{n+m}) = (c_1, \cdots, c_n, 0, \cdots, 0) - (y_1, \cdots, y_m)\begin{pmatrix} * & * & \cdots & * & 1 & 0 & \cdots & 0 \\ * & * & \cdots & * & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * & 0 & 0 & \cdots & 1 \end{pmatrix} = (*, *, \cdots, *, -y_1, -y_2, \cdots, -y_m)$,**即**$\sigma$**的后**$m$**个分量的负值，为DP问题的最优解**

◆**看下面的例子**

2. 3 对偶问题(续7)

对偶问题的最优解

◆ 例$Max\ z = 1500x_1 + 2500x_2$，s.t. $\begin{cases} 3x_1 + 2x_2 \le 65 \\ 2x_1 + x_2 \le 40 \\ 3x_2 \le 75 \\ x_1, x_2 \ge 0 \end{cases}$

			1500	2500	0	0	0	
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	θ_i
0	x_3	65	3	2	1	0	0	32.5
0	x_4	40	2	1	0	1	0	40
0	x_5	75	0	(3)	0	0	1	25
-z		0	1500	2500*	0	0	0	
0	x_3	15	(3)	0	1	0	-2/3	5
0	x_4	15	2	0	0	1	-1/3	7. 5
2500	x_2	25	0	1	0	0	1/3	
-z		-62500	1500*	0	0	0	-2500/3	
1500	x_1	5	1	0	1/3	0	-2/ 9	
0	x_4	5	0	0	-2/3	1	1/9	
2500	x_2	25	0	1	0	0	1/3	
-z		-70000	0	0	-500	0	-500	

练习:写出上述问题的对偶问题，并求解，印证这个结论

2.3 对偶问题(续7)(注)

◆影子价格

- 是一个向量，它的分量表示最优目标值随相应资源数量变化的变化率
- 若x^*, y^*分别为（LP）和（DP）的最优解，那么，$c^T x^* = b^T y^*$，根据
- $f = b^T y^* = b_1 y_1^* + b_2 y_2^* + \cdots + b_m y_m^*$ 可知 $\partial f / \partial b_i = y_i^*$，表示 b_i 变化1个单位对目标f产生的影响，称 y_i^*为b_i的影子价格
- *注意*：若$\boldsymbol{B}$是最优基，$\boldsymbol{y}^* = \left(\boldsymbol{B}^T\right)^{-1} \boldsymbol{c}_B$为影子价格向量

2.3对偶问题(续7)

◆对偶单纯形法(The Dual Simplex Method)

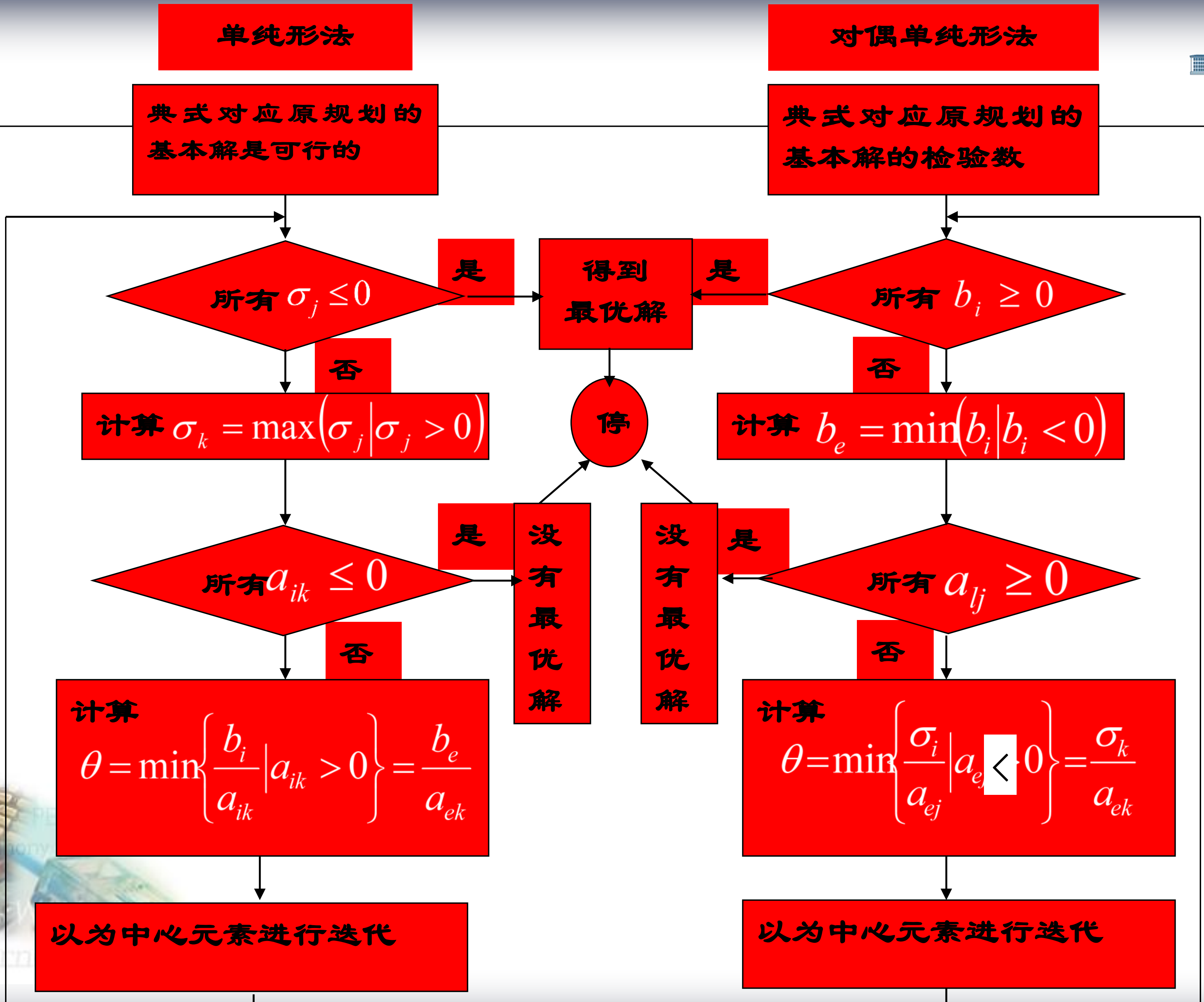
- 求解原规划问题的一种方法，采用单纯形法的思想和对偶的思想
- 从原规划问题的一个基本解出发，此基本解不一定可行，但它对应着一个对偶可行解(检验数均非正)，因此，也可以说从一个对偶可行解出发，然后检验原规划问题的基本解是否可行，即是否有负的分量，如果有小于0的分量，则进行迭代，求另一个基本解，此基本解对应另一个对偶可行解(检验数均非正)，如果得到的基本解的分量皆为非负，则该基本解为最优解。也就是说对偶单纯性法在迭代过程中始终保持对偶解的可行性(检验数均非正),是原规划问题的基本解由不可行逐渐变为可行

2.3对偶问题(续8)

- **对偶单纯形法(The Dual Simplex Method)**
- **例子：用对偶单纯形法求解LP：** $\min f=2x_1+3x_2+4x_3,\ s.t.\begin{cases}x_1+2x_2+x_3\ge 3\\ 2x_1-x_2+x_3\ge 4\\ x_1,x_2,x_3\ge 0\end{cases}$
- **标准化:** $\max z=-2x_1-3x_2-4x_3,\ s.t.\begin{cases}-x_1-2x_2-x_3+x_4=-3\\ -2x_1+x_2-3x_3+x_5=-4\\ x_1,x_2,x_3,x_4,x_5\ge 0\end{cases}$

◆对偶单纯形法：注意与单纯形法的区别

c_i			-2	-3	-4	0	0
c_b	x_b	b	x_1	x_2	x_3	x_4	x_5
0	x_4	-3	-1	-2	-1	1	0
0	x_5	-4	[-2]	1	-3	0	1
	σ_j		-2	-3	-4	0	0
0	x_4	-1	0	[-5/2]	1/2	1	-1/2
-2	x_1	2	1	-1/2	3/2	0	-1/2
	σ_j		0	-4	-1	0	-1
-3	x_2	2/5	0	1	-1/5	-2/5	1/5
-2	x_1	11/5	1	0	7/5	-1/5	-2/5
	σ_j		0	0	-9/5	-8/5	-1/5



单纯形法和对偶单纯形法步骤

2.4 灵敏度分析

- **Sensitivity analysis**
  - 在线性规划中，模型的参数(输入数据)可能在一定限度内改变不会造成最优解发生变化
- **Post-optimal 分析(优化后分析)**
  - 如果输入数据改变造成目标发生变化，然后重新求解该最优性问题，称为**Post-optimal**分析
- **这是因为在LP中，参数一般都不是精确的。采用灵敏度分析，可以探查这种不确定性对最优解质量的影响**
- **例如，对某种产品估计的单位利润，如果灵敏度分析表明最优解对于单位利润浮动±10%，都不变化，则断定该解比仅对浮动±1%不变的最优解更鲁棒**

2.4 灵敏度分析(续1)

◆灵敏度分析

➢右端项变化对最优解的影响

➢目标函数的系数变化(单位代价或利润)对最优解的影响

◆价值系数c发生变化(基变量系数还是非基变量系数)

➢考虑检验数$\sigma_j = c_j - \Sigma_{i=1}^{m} c_{r_i} a_{r_i j}, j = 1, 2, \cdots, n$

◆1. 若c_k是非基变量的系数

设c_k变化为 $c_k + \Delta c_k$，$\sigma_k' = c_k + \Delta c_k - \sum cri\ a_{rik} = \sigma_k + \Delta c_k$

只要 $\sigma_k' \le 0$，即 $\Delta c_k \le -\sigma_k$，则最优解不变；否则，将最优单纯形表中的检验数 σ_k 用 σ_k'取代，继续单纯形法的表格计算

2.4 灵敏度分析-例子

◆例 $Max\ z = -2x_1 - 3x_2 - 4x_3$

$$
\begin{aligned}
s.t.\ \ & -x_1 - 2x_2 - x_3 + x_4 = -3 \\
& -2x_1 + x_2 - 3x_3 + x_5 = -4 \\
& x_1, x_2, x_3, x_4, x_5 \ge 0
\end{aligned}
$$

◆练习

第2章线性规划-例子(续1)

◆单纯形表

C_I			-2	-3	-4	0	0
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5
-3	X_2	2/5	0	1	-1/5	-2/5	1/5
-2	X_1	11/5	1	0	7/5	-1/5	-2/5
	σ_j		0	0	-9/5	-8/5	-1/5

C_I			-2	-3	$-4+\Delta c_3$	0	0
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5
-3	X_2	2/5	0	1	-1/5	-2/5	1/5
-2	X_1	11/5	1	0	7/5	-1/5	-2/5
	σ_j		0	0	$-9/5+\Delta c_3$	-8/5	-1/5

◆从表中看到$\sigma_3 = c_3 + \Delta c_3 - (c_2 \times a_{13} + c_1 \times a_{23})$，可得到$\Delta c_3 \le 9/5$时，原最优解不变。

2.4 灵敏度分析（续2）

➢考虑检验数$\boldsymbol{\sigma_j = c_j - \Sigma_{i=1}^{m} c_{r_i} a_{r_i j}, j = 1, 2, \cdots, n}$

◆**2. 若c_k是基变量的系数**

设 c_s 变化为 $c_s + \Delta c_s$，那么$\boldsymbol{\sigma_j' = c_j - \Sigma_{i \neq s} c_{r_i} a_{r_i j} - (c_s + \Delta c_s) a_{sj} = \sigma_j - \Delta c_s a_{sj}}$,

◆**对所有非基变量，只要对所有非基变量$\sigma_j' \leq 0$，即 $\sigma_j' \leq \Delta c_s a_{sj}$，则最优解不变；否则，将最优单纯形表中的检验数$\sigma_j$ 用σ_j'取代，继续单纯形法的表格计算**

$$Max\{\boldsymbol{\sigma}_j / a_{sj} \mid a_{sj} > 0\} \leq \Delta c_s \leq Min\{\sigma_j / a_{sj} \mid a_{sj} < 0\}$$

2.4 灵敏度分析-例子

◆ **例子** $\mathbf{Max}\ z = 2x_1 + 3x_2 + 0x_3 + 0x_4 + 0x_5$

$$
s.t.\quad
\begin{cases}
x_1 + 2x_2 + x_3 = 8 \\
4x_1 + x_4 = 16 \\
4x_2 + x_5 = 12 \\
x_1, x_2, x_3, x_4, x_5 \geq 0
\end{cases}
$$

◆ **练习**

2.4 灵敏度分析-例子

◆最优单纯形如下

C_i			2	3	0	0	0
C_B	X_B	B	x_1	x_2	x_3	x_4	x_5
2	x_1	4	1	0	0	1/4	0
0	x_5	4	0	0	-2	1/2	1
3	x_2	2	0	1	1/2	-1/8	0
	σ_j		0	0	-1.5	-1/8	0

C_i			2	$3+\Delta C_2$	0	0	0
C_B	X_B	B	x_1	x_2	x_3	x_4	x_5
2	x_1	4	1	0	0	1/4	0
0	x_5	4	0	0	-2	1/2	1
$3+\Delta C_2$	x_2	2	0	1	1/2	-1/8	0
	σ_j		0	0	$-1.5-\Delta C_2/2$	$-1/8+\Delta C_2/8$	0

◆从表中看到$\sigma_j = c_j - (c_1 \times a_{1j} + c_5 \times a_{5j} + (c_2 + \Delta c_2) \times a_{2j})j = 3,4$，可得到 $-3 \leqslant \Delta c_2 \leqslant 1$时，原最优解不变

2.4 灵敏度分析

◆右端项变化

➢设分量 b_r 变化为 $b_r + \Delta b_r$，根据前面的讨论，最优解的基变量 $x_B = B^{-1}b$，那么只要保持 $B^{-1}(b + \Delta b) \geq 0$，则最优基不变，即基变量保持，只有值的变化；否则，需要利用对偶单纯形法继续计算

◆对LP: $\boldsymbol{Max\ z = c^T x, s.t. Ax \leq b, x \geq 0}$**,其最优单纯形表中含有** $\boldsymbol{B^{-1}} = (a_{ij})\ i = 1, \ldots, m;\ j = n+1, \ldots, n+m$ **,那么新的**$x_i = (B^{-1}b)_i + \Delta b_r a_{ir}\ i = 1, \ldots, m$ **,由此可得，最优基不变的条件是**

$$Max\{-bi/a_{ir} \mid a_{ir} > 0\} \leq \Delta b_r \leq Min\{-bi/a_{ir} \mid a_{ir} < 0\}$$

2.4 灵敏度分析-右端项变化例子

◆前面例子的最优单纯形表如下

C_i			2	3	0	0	0
C_B	X_B	B	x_1	x_2	x_3	x_4	x_5
2	x_1	4	1	0	0	1/4	0
0	x_5	4	0	0	-2	1/2	1
3	x_2	2	0	1	1/2	-1/8	0
	σ_j		0	0	-1.5	-1/8	0

◆这里$B^{-1}=[0\ \frac{1}{4}\ 0;-2\ \frac{1}{2}\ 1;\frac{1}{2}-\frac{1}{8}\ 0]$,分别对应$b_1,b_2,b_3$的单列变化，设$b_1$增加4,则$x_1,x_5,x_2$分别变为:$4+0\times4=4, 4+(-2)\times4=-4<0, 2+\frac{1}{2}\times 4=4$

◆用对偶单纯形求解，可得$x^*=(4,3,2,0,0)^T, f^*=17$

2.4 灵敏度分析-增加变量

- 增加变量x_{n+1}则有相应的p_{n+1}, c_{n+1}，那么计算出$B^{-1}p_{n+1}$, $\sigma_{n+1} = c_{n+1} - \sum c_{ri} a_{ri\,n+1}$，填入最优单纯形表，
  若 $\sigma_{n+1} \leqslant 0$ 则 最优解不变；
  否则，进一步用单纯形法求解
- 沿用前面的例子，但增加一列约束，$x_6, p_6 = (2,6,3)^T, c_6 = 5$，最终可得如下单纯形表

C_i			2	3	0	0	0	5
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6
2	x_1	4	1	0	0	1/4	0	1.5
0	x_5	4	0	0	-2	1/2	1	[2]
3	x_2	2	0	1	1/2	-1/8	0	0.25
	σ_j		0	0	-1.5	-1/8	0	1.25

- 用单纯形法进一步求解，可得：$x* = (1,1.5,0,0,0,2)^T$ $f* = 16.5$

2.4 灵敏度分析-增加约束

◆若增加一个约束

➢增加约束一个之后，应把最优解代入入新的约束，若满足则最优解不变，否则填入最优单纯形表作为新的一行，引入一个新的非负变量（原约束若是小于等于形式可引入非负松弛变量，否则引入非负人工变量），并通过矩阵行变换把对应基变量的元素变为0，进一步用单纯形法或对偶单纯形法求解。

2.4 灵敏度分析–增加约束

◆仍考虑前面的例子，增加$3x_1+2x_2\leqslant 15$，原最优解不满足这个约束

C_i			2	3	0	0	0	0
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6
2	x_1	4	1	0	0	1/4	0	0
0	x_5	4	0	0	-2	1/2	1	0
3	x_2	2	0	1	1/2	-1/8	0	0
0	x_6	-1	0	0	-1	-1/2	0	1
	σ_j		0	0	-1.5	-1/8	0	0

◆经对偶单纯形法一步，可得最优解为$(3.5, 2.25, 0, 0, 3, 2)^T$，最优值为 13.75

第2章线性规划-总结

- $\boldsymbol{LP: Max\ z = c^T x, Ax = b, x \geq 0}$
- 如果$\boldsymbol{z = \frac{1}{2}x^T Hx + c^T x}$, $\boldsymbol{H}$为$\boldsymbol{n}$阶的对称矩阵，其余不变，则成为二次规划(Quadratic Programming)
- 如果目标和约束都允许非线性，称为非线性规划(nonlinear program)
- 如果变量为整数，则成为整数规划(integer program)
- 线性规划最先在经济规划(economic planning)中使用，后在数值分析，逼近论，模式识别，机器学习中都有很多应用

第2章线性规划-总结(续1)

◆应用问题简述

- 饮食问题-The Diet Problem: 花最少的钱为个人制定适合营养计划要求的食物规划：给定食物品种，确保个人营养要求
- 线性曲面拟合(Linear surface fitting): 一系列观测量$(A_i, b_i), i = 1, 2, \cdots, m$, A_i为n个元素的行向量，b_i为实数，目的值找到向量x,常数γ: $A_i x + \gamma \approx b_i, i = 1 \cdots m$, 为了测量这种拟合度，可以使用$\sum_{i=1}^{m} |A_i x + \gamma - b_i|$来度量，该问题就可以形式化为一个LP问题：$min_{x,\gamma,y}\ z = e'y, s.t. -y \le Ax + \gamma e - b \le y, e' = (1\ 1\ \cdots 1)$,如果$n = 1$,可体会其在二维空间$(A_i, b_i)$中的几何意义

◆负载均衡(Load balancing problem): 考虑n个多核处理器来执行某计算任务，这些处理器可能已经运行了别的工作，这时需要在已运行任务轻的处理器上多运行一些任务，假设p_i: 处理器当前负载, L: 代分配的总负载,x_i:剩余负载分配到第i个处理器的比例，$x_i \geq 0, \Sigma x_i = 1, \gamma$: 任务分配到$L$个处理器后,其剩余负载的最小值，此时问题可形式化为：$max_{x,\gamma}\ \gamma, s.t.\ \gamma e \leq p + xL, e'x = 1, x \geq 0$,$e$为全1向量

◆这是少数几个有close-form解的LP问题之一，当$p_i \leq \frac{L}{n}, i = 1,2 \cdots n$,最优解$\gamma = (e'p + L)/n$,即所有处理器有同样的负载。否则，需要对$p_i$排序

第2章线性规划-总结(续3)

- **资源分配(Resource Allocation):某公司需要确定如何来在某个时间段安排其资源(原材料，劳力，租用设备的时间等)来生产某些产品**
- **分类(Classification)**
  - 给定空间$\boldsymbol{R}^n$中的两组点，目标：找到该空间中的一个超平面，将其尽可能精确的分离开。然后可以使用这个超平面来分类任何新获得的点
  - **LP问题:**超平面可用向量$\boldsymbol{w} \in \boldsymbol{R}^n$,和标量$\gamma$来确定，点分类根据$\boldsymbol{w'}\boldsymbol{t} \geq \gamma$或者$\boldsymbol{w'}\boldsymbol{t} \leq \gamma$来确认，为了排除超平面退化的情况，可以加强一下条件$\boldsymbol{w'}\boldsymbol{t} \geq \gamma + \mathbf{1}, \boldsymbol{w'}\boldsymbol{t} \leq \gamma - \mathbf{1}$来分类，目标函数为：每个集合上平均分类错误的点数

第2章线性规划-总结(续4)

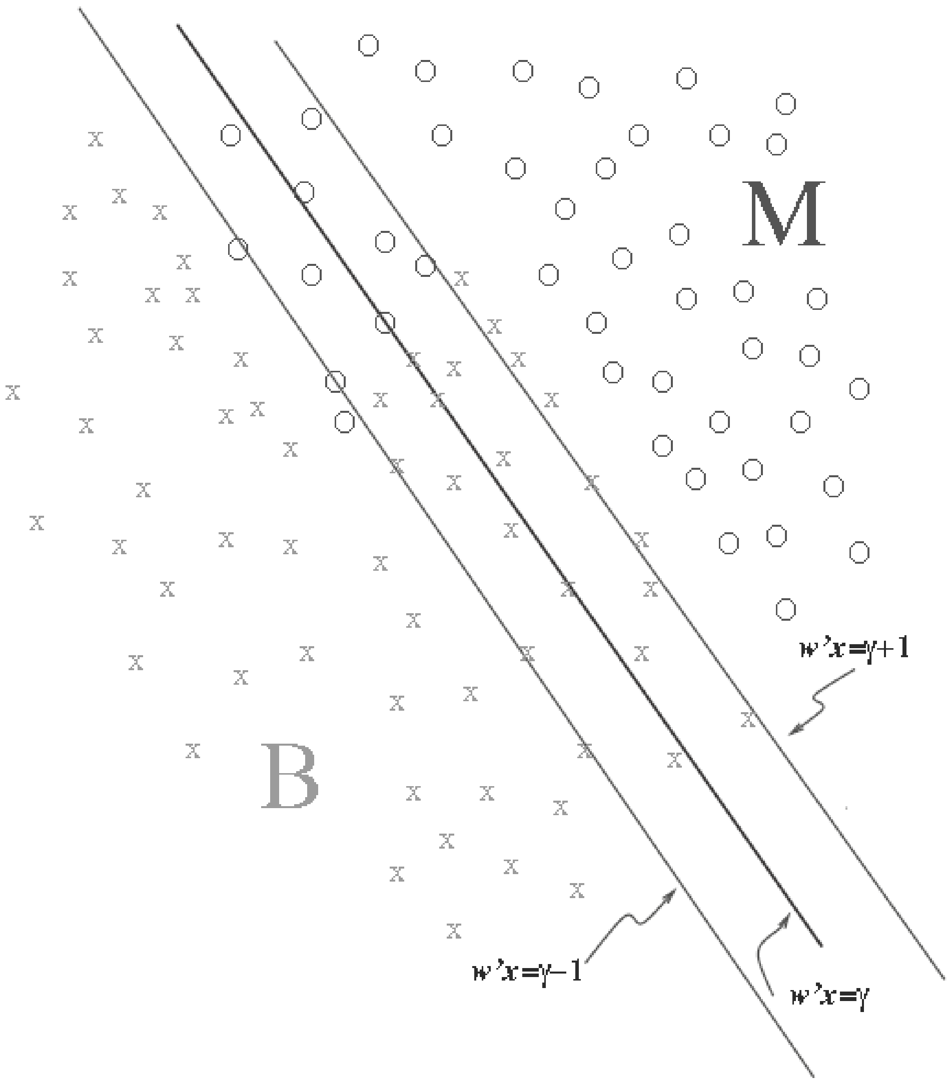
◆建立LP问题：$M_{m\times n}$:第i行包含第一个集合中第i个点的n个分量，$B_{k\times n}$：意义跟前一样，不过是第二个点集里面的点

◆违反第一类分类条件的点：$w't \geq \gamma + 1$，可由向量y通过不等式来测度:$y \geq -(Mw - \gamma e) + e, y \geq 0$,e为全1向量；类似对于第二类违反分类的点可由向量$z, z \geq (Bw - \gamma e) + e, z \geq 0$来测度

◆第一和第二个集合上平均违反分类准则的点分别为:$\frac{e'y}{m}, \frac{e'z}{k}$，因此LP问题为: $min_{w,\gamma,y,z} \frac{1}{m}e'y + \frac{1}{k}e'z, s.t.\, y \geq -(Mw - \gamma e) + e, z \geq (Bw - \gamma e) + e, (y, z) \geq 0$

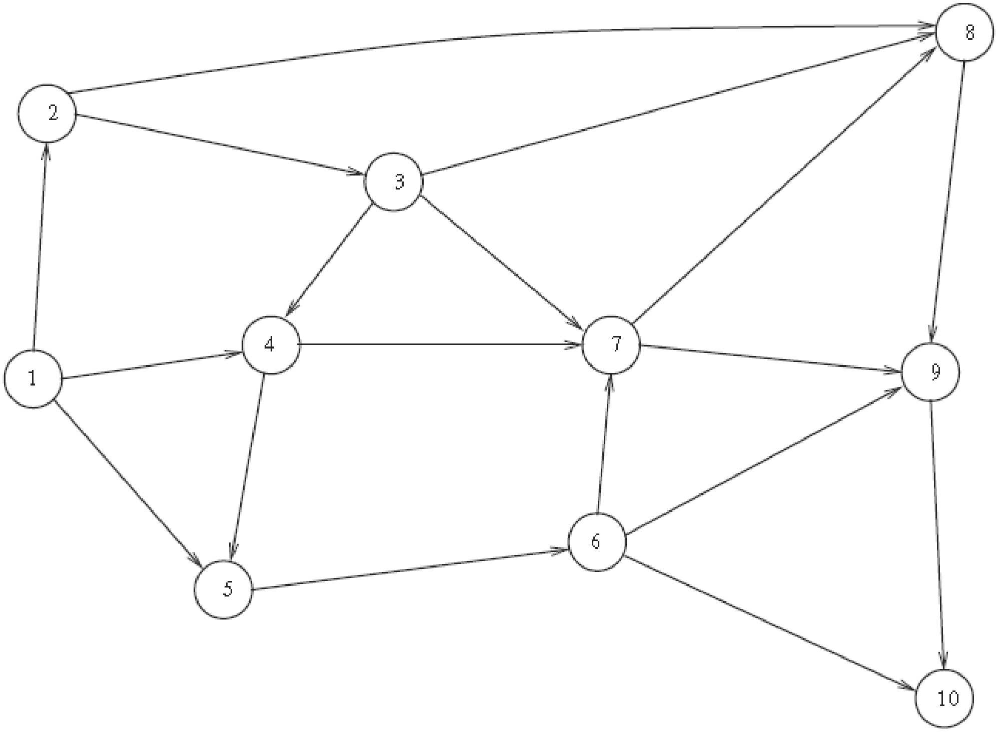
第2章线性规划-总结(续4)

- 乳腺癌分类
  - M(malignant tumors]
  - B(Benign tumors)
- 1998, Bosch & Smith 提出用LP解决美国12联邦文件是由James Madison,还是Alexande Hamilton起草的，使用某些单词的使用频率来建立超平面



第2章线性规划-总结(续5)

◆最小代价网络流(Minimum-cost network flow)



◆例如运输产品到某些城市中，如何代价最小

◆例如如何分割图像

第2章线性规划-总结(续6)

- Simplex 方法在大部分实际问题中运行良好
- 但有些问题迭代次数按变量的指数级递增，后续1979年，Khachiyan发明了椭圆方法(Ellipsoid method),运行时间以存储数据和变量的比特数的多项式函数为界，但该算法很难实现，并且实际运行时令人失望
- 1984年，Karmarkar提出了一个具有类似多项式时间界的算法，声称优于已有的所有单纯形代码，但没人验证
- 但推动了内点法(interior-point method)的研究大潮:迭代不是在顶点和边界上进行，而是可行域内部进行移动，其性能在很多问题上显著优于单纯形法

第2章线性规划-总结(续7)

- 单词"Programming":是指用于求解优化问题的逐步的数学过程，而不是具体指计算机上的实现
- Linear programming出现在1940's, 很久之前单词programming就与计算机有强烈的关联
- 从线性规划中了解到一些基本的概念和观点
  - 提出问题
  - 数学形式化描述问题
  - 提出解决办法
    - ✓找到一个初始解
    - ✓判断是否是最优解
    - ✓寻找更优的解
      - 搜索方向

第二章 概念总结

- **一些基本概念，凸性，梯度，Hessian矩阵，拐点**
- **线性规划建模**
  - LP的基本概念
  - 标准型LP问题
- **LP求解**
  - 图解法，理解其基本的思想
  - 代数法：单纯形法
  - 单纯形表，易于操作
- **LP的对偶问题**
  - 理解LP与DP的关系
- **灵敏度分析**
  - 右端项，价值系数，约束，变量

附录

- 附录：单纯形表例子
- 附录：退化循环例子

第2章线性规划-单纯形法(simplex)(续7)

◆单纯形表

$$\begin{aligned} &\max\ z = 2x_1 + 5x_2 \\ &s.t.\quad x_1 + 2x_2 \le 8 \\ &\qquad\quad x_1 \le 4 \\ &\qquad\quad x_2 \le 3 \\ &\qquad x_1, x_2 \ge 0 \end{aligned}$$

◆化为标准型

$$\begin{aligned} &\max\quad z = 2x_1 + 5x_2 + 0x_3 + 0x_4 + 0x_5 \\ &s.t.\quad x_1 + 2x_2 + x_3 + 0x_4 + 0x_5 = 8 \\ &\qquad\ \ x_1 + 0x_2 + 0x_3 + x_4 + 0x_5 = 4 \\ &\qquad\ \ 0x_1 + x_2 + 0x_3 + 0x_4 + x_5 = 3 \\ &\qquad\ \ x_1, x_2, x_3, x_4, x_5 \ge 0 \end{aligned}$$

◆矩阵形式

$$\begin{aligned} &\max\ c^T x \\ &s.t.\ \ Ax = b,\ x \ge 0 \end{aligned}$$

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} = [p_1\ p_2\ p_3\ p_4\ p_5]$$

$$b = [8\ 4\ 3]^T,\ \ c = [2\ 5\ 0\ 0\ 0]^T$$

$c_j \rightarrow$			c_1	c_2	c_3	c_4	c_5	
c_B	x_B	$\bar{b}$	x_1	x_2	x_3	x_4	x_5	θ
c_1	x_1	b_1	1	0	0	a_{14}	a_{15}	
c_2	x_2	b_2	0	1	0	a_{24}	a_{25}	
c_3	x_3	b_3	0	0	1	a_{34}	a_{35}	
	$-z$	$-z_0$	0	0	0	σ_4	σ_5	

◆找一个初始基本可行解(BFS)

$$B=[P_3\ P_4\ P_5]=I_3,\ x^{(0)}=\begin{bmatrix}x_N^{(0)}\\ x_B^{(0)}\end{bmatrix}=\begin{bmatrix}0\\ B^{-1}b\end{bmatrix}=[0\ \ 0\ \ 8\ \ 4\ \ 3]^T$$

将基变量$x_B^{(0)}$表示成非基变量$x_N^{(0)}$的组合形式：

$$\because\ Ax^{(0)}=[N\quad B]\begin{bmatrix}x_N^{(0)}\\ x_B^{(0)}\end{bmatrix}=Nx_N^{(0)}+Bx_B^{(0)}=b$$

$$\therefore\ x_B^{(0)}=B^{-1}b-B^{-1}Nx_N^{(0)}$$

此时可将目标函数转化为非基变量$x_N^{(0)}$的组合形式：

$$\begin{aligned}f(x^{(0)})&=c^Tx^{(0)}=c_B^Tx_B^{(0)}+c_N^Tx_N^{(0)}\\&=c_B^TB^{-1}b-(c_B^TB^{-1}N-c_N^T)x_N^{(0)}\\&=c_B^TB^{-1}b-\sum_{j\in NI}(c_B^TB^{-1}p_j-c_j)x_j^{(0)}\\&=0-(2x_1^{(0)}+5x_2^{(0)})=0\end{aligned}$$

- **找一个改进的初始基本可行解(BFS)**
  - $f\left(x^{(0)}\right)=c_B^TB^{-1}b+\sum_{j\in J_N}\left(c_j-c_B^TB^{-1}p_j\right)x_j^{(0)}=0+\left(2x_1^{(0)}+5x_2^{(0)}\right)=0$
- $\lambda_j=c_j-c_B^TB^{-1}p_j, j\in J_N$为检验数，根据前面的定理，检验数都为负，说明已经是最优解，否则需要将决策变量进行入基变量和出基变量轮换来求得新的基本可行解
- 显然，选取x_k,使得$\lambda_k=max\{\lambda_i, i\in J_N\}$此时入基变量选取$2=max_i\{\lambda_i\}$,即$x_2$
- 此时，$x_B^{(1)}=B^{-1}b-B^{-1}p_kx_k=x_B^{(0)}-x_kB^{-1}p_k=\bar{b}-y_kx_k=(8,4,3)^T-(2,0,1)^T\,x_2$;$x_N^{(1)}=(0,\cdots,0,x_k,0,\cdots,0)^T=(0,x_2)^T$;
- 取$d=\left(\left(B^{-1}p_k\right)^T(0,\cdots,0,-1,0,\cdots,0)^T\right)^T$,则$f\left(x^{(1)}\right)=f\left(x^{(0)}+x_kd\right)=f\left(x^{(0)}\right)+x_k\left(c_k-c_B^TB^{-1}p_k\right)=f\left(x^{(0)}\right)+\lambda_kx_k=5x_2$

◆**具体如下**

此时$x_B^{(1)}=B^{-1}b-B^{-1}p_kx_k=x_B^{(0)}-x_kB^{-1}p_k\quad(\because x_B^{(0)}=B^{-1}b-B^{-1}Nx_N)$

$$=\bar{b}-y_kx_k=(\bar{b}_1,\bar{b}_2,\cdots,\bar{b}_m)^T-(y_{1k},y_{2k},\cdots,y_{mk})^Tx_k$$

$$=(8,4,3)^T-(2,0,1)^Tx_2$$

$$x_N^{(1)}=(0,\cdots,0,x_k,0,\cdots,0)^T=(0,x_2)^T,$$

取$d=((B^{-1}p_k)^T\ (0,\cdots,0,-1,0,\cdots,0)^T)^T$，则

$$f(x^{(1)})=f(x^{(0)}+x_kd)=f(x^{(0)})+x_k(c_k-c_B^TB^{-1}p_k)=f(x^{(0)})+\lambda_kx_k=5x_2$$

为使目标函数值增加越多, 显然$x_k(x_2)$越大越好, 但$x_k(x_2)$的取值受到$x^{(1)}$必须为可行解的约束, 即

$$\bar{b}_i-y_{ik}x_k\geq 0,i=1,2,\cdots,m\Rightarrow x_k\leq\bar{b}_i/y_{ik},i=1,2,\cdots,m$$

因此$x_k=\min\{\bar{b}_i/y_{ik}\mid y_{ik}>0,i=1,2,\cdots,m\}=\bar{b}_r/y_{rk}$

$$(x_2=\min\{8/2,3/1\}=3/1=\bar{b}_3/y_{3k})$$

此时原来的基变量$x_{B_3}^{(0)}=0\ (x_5=0)$, 成为了非基变量, 新的基本可行解为

$$x^{(1)}=(x_{B_1}^{(1)},\cdots,x_{B_{r-1}}^{(1)},0,x_{B_{r+1}}^{(1)},\cdots,x_{B_m}^{(1)},0,\cdots,0,x_k,0,\cdots,0)^T\ (x^{(1)}=(0,3,2,4,0)^T)$$

第2章线性规划-单纯形法(simplex)(续11)

◆继续上述过程

x_1换入为基变量, x_3换出为非基变量

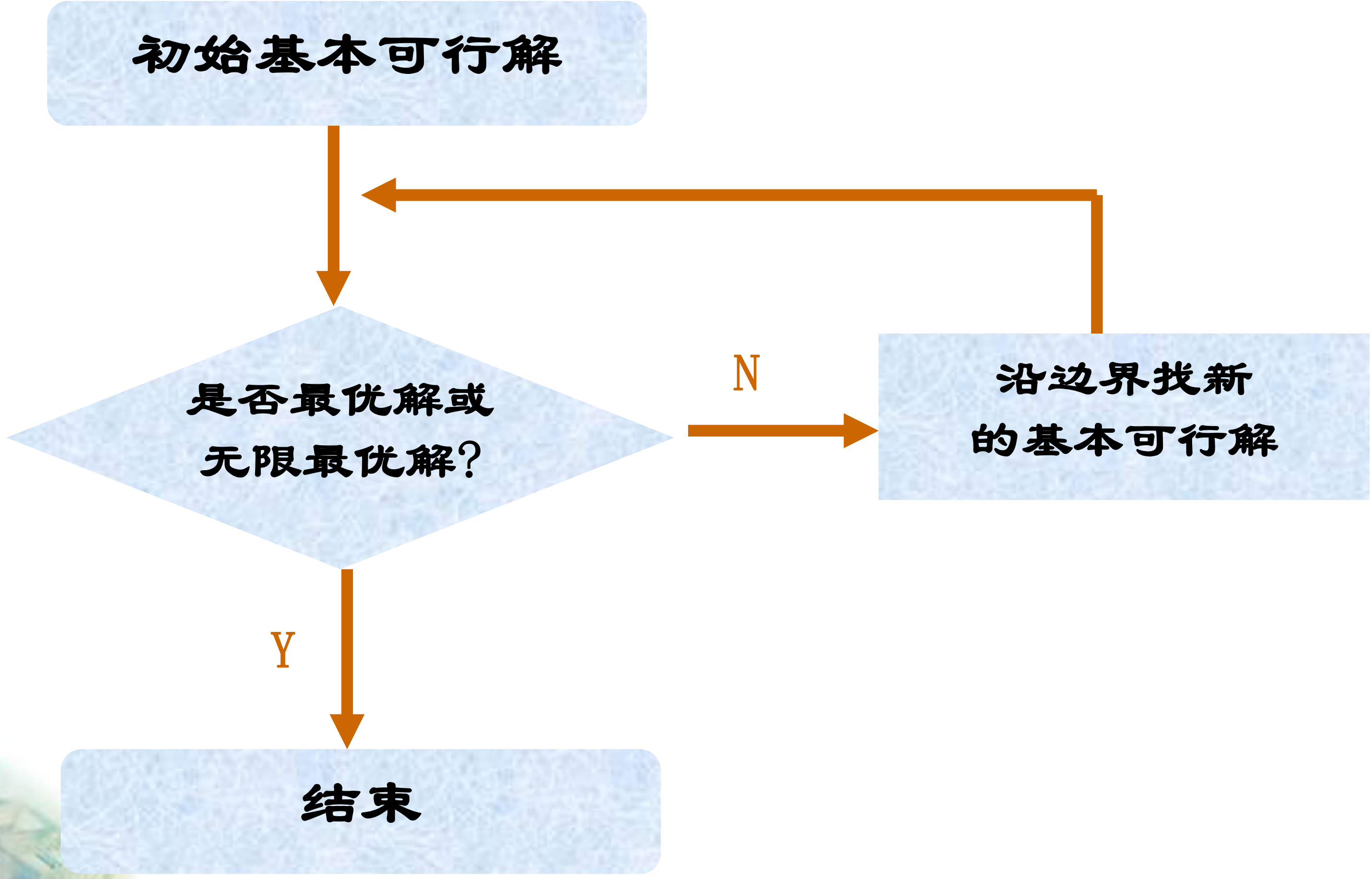
$$\begin{cases} x_1 = 2 - x_3 + 2x_5 \\ x_2 = 3 - x_5 \\ x_4 = 2 + x_3 - 2x_5 \end{cases}$$

$$x^{(2)} = (2,3,0,2,0)^T,\ f(x^{(2)}) = 19 + (-2x_3 - x_5) = 19$$

由于非基变量x_3, x_5的检验数分别为$-2, -1$,目标函数无法继续下降,因此$x^{(2)} = (2,3,0,2,0)^T$即为所求的最优解.

如何用单纯形表完成上述过程?

◆单纯形的一般过程



（1）**确定一个初始基本可行解:** $A = [B, N]$， $x^T = [x_B{}^T, x_N{}^T]$, $c^T = [c_B{}^T, c_N{}^T]$, $f = c_B{}^T x_B$. **这里** $x_B = \boldsymbol{B}^{-1}b \geq 0$, $x_N = 0$.

（2）**计算非基变量检验数**$□ = c_N{}^T - \boldsymbol{c_B^T B}^{-1}N$，**根据最优解判别定理，判断 x 是否为最优解. 若是，则停止，否则转下一步.**

（3）**求一改进的基本可行解**

1) 确定换入变量□**:** $□_k = \max\{□_j \mid j \in J_N\}$,**相应的$x_k$为换入变量**

2) 确定换出变量: 按θ规则计算

$$\theta = \min\{\ {}_i / a_{ik} \mid a_{ik} > 0\} = □_r / a_{rk}$$

以x_r为换出变量.这里$□ = \boldsymbol{B}^{-1}b, a_k = \boldsymbol{B}^{-1}p_k$

3) 用p_k代替p_{Br} ,得到新的基矩阵B, 并将B变换为单位矩阵:

以a_{rk}为主元素进行迭代(即用高斯消去法)把$p_k=(a_{1k}\ldots a_{rk}\ldots a_{mk})^T$变为 $(0\ldots1\ldots0)^T$.

（4）**重复（2）、（3）直到得到最优解.**

	z	x_B^{T}	x_N^{T}	右端项	
目标行	1	$\boldsymbol{c}_B^{\mathrm{T}}$	$\boldsymbol{c}_N^{\mathrm{T}}$	0	1行
约束行	0	$\boldsymbol{B}$	$\boldsymbol{N}$	$\boldsymbol{b}$	M行
	1列	m列	n-m列	1列	

作变换，使前m+1列对应的m+1阶矩阵变为单位矩阵，得：

检验数

	z	x_B^{T}	x_N^{T}	右端项	
目标行	1	0^{T}	$\boldsymbol{c}_N^{\mathrm{T}}-\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}$	$\boldsymbol{c}_B^{T}\boldsymbol{B}^{-1}\boldsymbol{b}$	1行
x_B	0	I	$\boldsymbol{B}^{-1}\boldsymbol{N}$	$\boldsymbol{B}^{-1}\boldsymbol{b}$	M行
	1列	m列	n-m列	1列	

$$\max \quad 2x_1 + 5x_2 + 0x_3 + 0x_4 + 0x_5$$
$$s.t. \quad x_1 + 2x_2 + x_3 + 0x_4 + 0x_5 = 8$$
$$x_1 + 0x_2 + 0x_3 + x_4 + 0x_5 = 4$$
$$0x_1 + x_2 + 0x_3 + 0x_4 + x_5 = 3$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

- 换入变量：x_2，因为max{2,5} = 5
- 换出变量：x_5，因为 min{8/2,3/1} = 3

	z	x_1	x_2	x_3	x_4	x_5	RHS
x_B	1	2	5	0	0	0	0
x_3	0	1	2	1	0	0	8
x_4	0	1	0	0	1	0	4
x_5	0	0	$\underline{1}$	0	0	1	3

$x = (0,3,2,4,0)^T$

	z	x_1	x_2	x_3	x_4	x_5	
x_B	1	2	0	0	0	-5	15
x_3	0	$\underline{1}$	0	1	0	-2	2
x_4	0	1	0	0	1	0	4
x_2	0	0	1	0	0	1	3

	z	x_1	x_2	x_3	x_4	x_5	
x_B	1	0	0	-2	0	-1	19
x_1	0	1	0	1	0	-2	2
x_4	0	0	0	-1	1	2	2
x_2	0	0	1	0	0	1	3

最优解$x^* = (2,3,0,2,0)^T, f^* = 19$

第2章线性规划-单纯形法-退化循环的例子

$$
\begin{aligned}
\min\ & -\frac{3}{4}x_4+20x_5-\frac{1}{2}x_6+6x_7 \\
s.t.\ & x_1 + \frac{1}{4}x_4-8x_5-x_6+9x_7=0, \\
& x_2+\frac{1}{2}x_4-12x_5-\frac{1}{2}x_6+3x_7=0, \\
& x_3+x_6=1, \\
& x_i\ge 0, i=1,\cdots,7.
\end{aligned}
$$

第一次迭代

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
f	0	0	0	3/4	-20	1/2	-6	0
x_1	1	0	0	<u>1/4</u>	-8	-1	9	0
x_2	0	1	0	1/2	-12	-1/2	3	0
x_3	0	0	1	0	0	1	0	1

第六次迭代

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
f	0	0	0	3/4	-20	1/2	-6	0
x_1	1	0	0	1/4	-8	-1	9	0
x_2	0	1	0	1/2	-12	-1/2	3	0
x_3	0	0	1	0	0	1	0	1

第2章线性规划-单纯形法-解决方法：摄动法

- 解决方法：摄动法，具体请看陈宝林的“最优化理论与算法（第2版）”
- 退化情形不用摄动法也不一定出现循环，事实上，退化情形是常见的，但在迭代中发生循环现象的概率很小，因此关于退化和循环的研究，主要是理论上的意义，实际计算中并不那么重要

1. 确定初始基本可行解：$A=[B\ N], B$可逆,且$B^{-1}b\geq 0$,则一个初始基本可行解为$x=\begin{bmatrix} x_B \\ x_N \end{bmatrix}=\begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}$

2. 判断初始基本可行解是否为最优解或无有限最优解，如是则停止，否则转下一步：计算非基变量检验数$\lambda_j=c_j-c_B^T B^{-1}p_j, j\in J_N$，其中$w=c_B^T B^{-1}$称为单纯形乘子. 若所有$\lambda_j\leq 0$，则已得到最优解，算法停止；若所有$\lambda_j$>0，且$B^{-1}p_j\leq 0$, 则无有限最优解,算法亦停止；否则转下一步,求一个改进的基本可行解:

3.求一个改进的基本可行解:根据$\max\{\lambda_j, j\in J_N\}=\lambda_k$,确定$x_k$为换入变量;根据$\theta=\min\{(B^{-1}b)_i/(B^{-1}p_k)_i \mid (B^{-1}p_k)_i>0\}=(B^{-1}b)_r/(B^{-1}p_k)_r$,确定$x_r$为换出变量;得到新的基矩阵$B_1$;

4. 令$B_1=B$,返回第2步

启示：

(1) 关键是计算B^{-1},由于B_1与B仅有1列不同,因此B_1^{-1}可方便地通过B^{-1}得到.

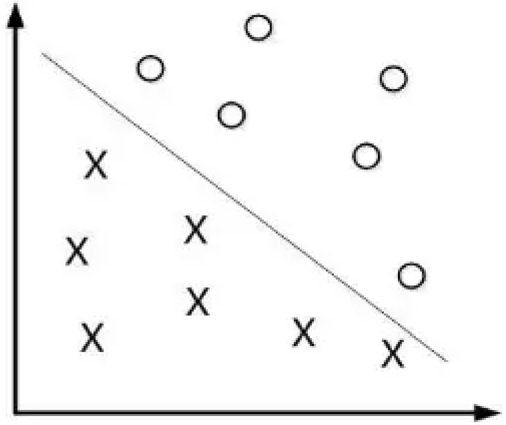
(2) 单纯形法仅需存储原始数据和每次迭代的$w=c_B^T B^{-1}, \bar{b}=B^{-1}b, f=c_B^T\bar{b}$,以及换入变量的$\lambda_k=c_k-w^T p_k, y_k=B^{-1}p_k$,无关的非基变量的数据可以不存储

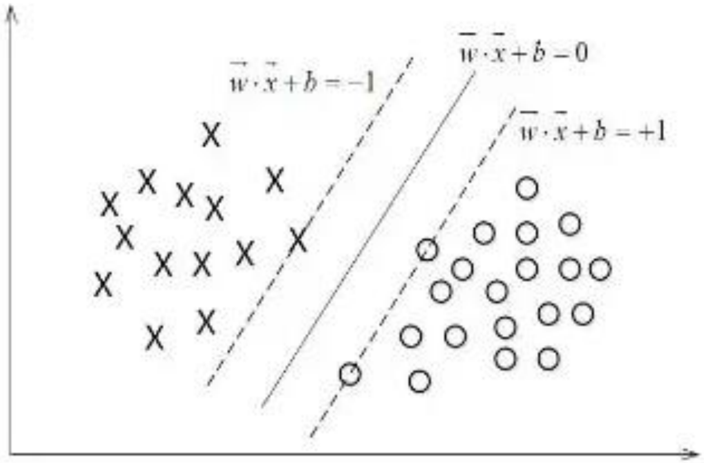
B_1^{-1}的计算：设$B_1^{-1}=(\bar{b}_{ij})_{m\times m}, B^{-1}=(b_{ij})_{m\times m}, y_k=B^{-1}p_k$,

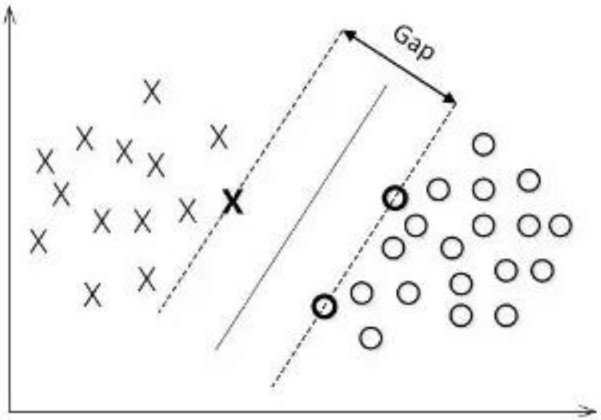
则对B_1^{-1}的每一列,有

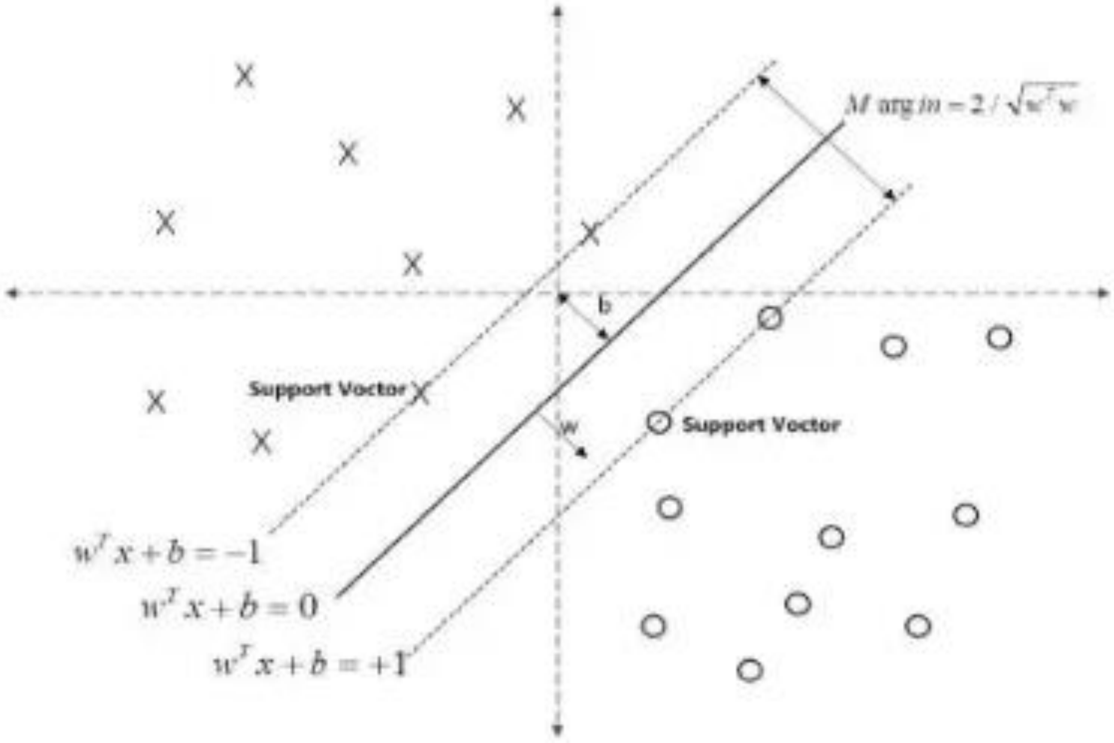
$$\begin{cases} \bar{b}_{ij}=b_{ij}-\dfrac{y_{ik}}{y_{rk}}b_{rj}, i\neq r \\ \bar{b}_{rj}=\dfrac{b_{rj}}{y_{rk}} \end{cases}$$

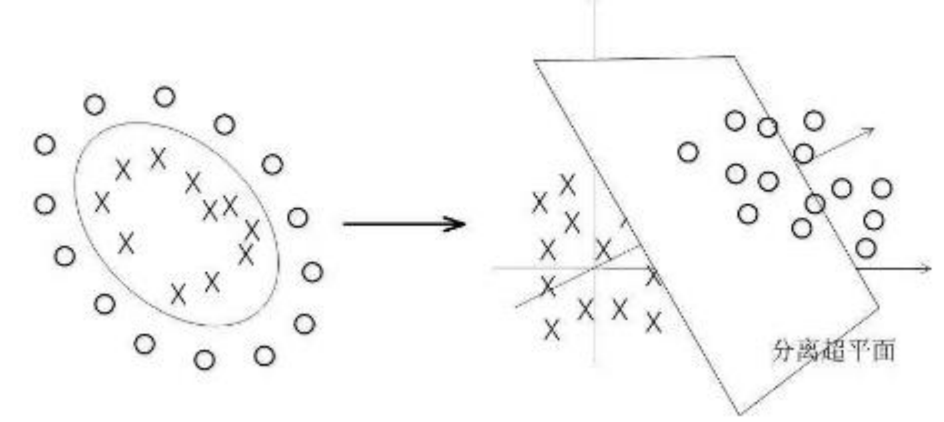
优点：主要是节省了存储空间

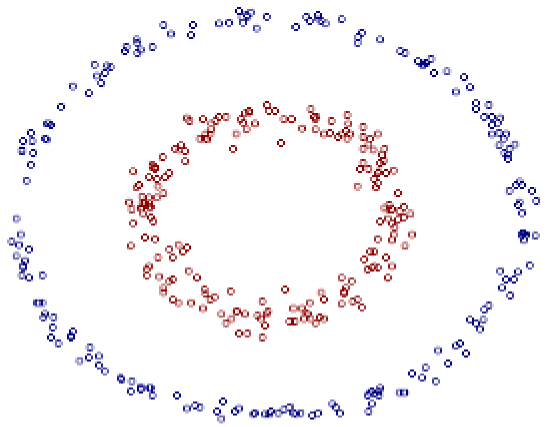


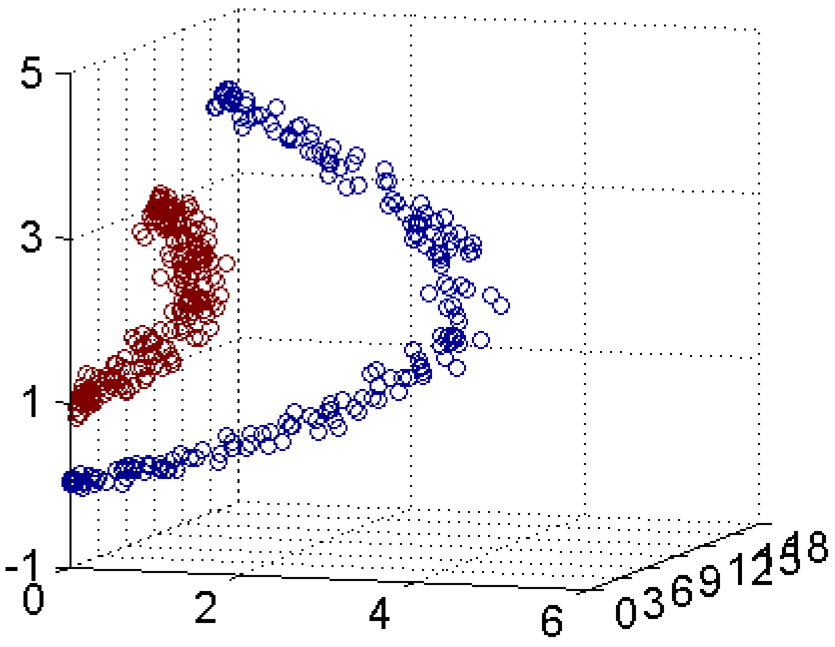

$\vec{w}\cdot\vec{x}+b=-1$
$\vec{w}\cdot\vec{x}+b=0$
$\vec{w}\cdot\vec{x}+b=+1$

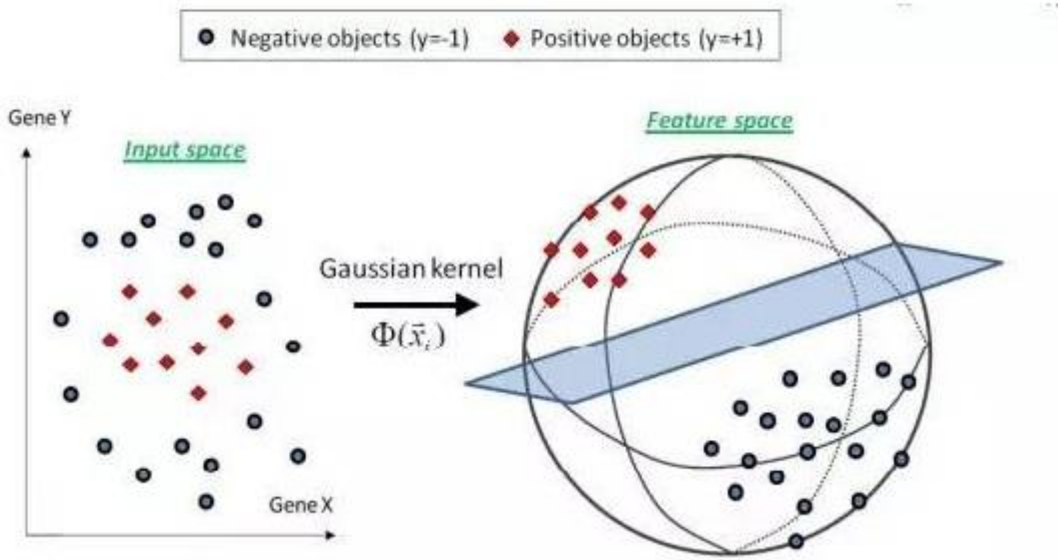

Gap


$Margin=2/\sqrt{w^T w}$
b
Support Voctor
w
Support Voctor
$w^T x+b=-1$
$w^T x+b=0$
$w^T x+b=+1$


分离超平面




5
3
1
-1
0
2
4
6
036912158


Negative objects (y=-1)
Positive objects (y=+1)
Gene Y
Input space
Gene X
Gaussian kernel
$\Phi(\vec{x}_i)$
Feature space