運籌報导

（1）

中国科学院数学研究所
运筹学研究室主编
1977年9月

前　　言

运筹学是一门应用性较强的学科。它所研究的课题，无论是应用性的，抑或是理论性的，都具有明显的实际背景。研究的成果，有些可直接有助于改进生产过程或生产的计划、组织和管理，有些则逐步地接近实际，为将来的实际应用奠定基础。由于运筹学与生产实践有着广泛的、密切的联系，所以虽然它与别的学科比较，历史较短，但却发展迅速：它的一些分支有其自身的国际组织和专门杂志（如数学规划、应用概率、系统分析等），一些新的分支正在迅速成长（如组合数学、模糊集合等）。在我国，运筹学近年来越来越受到人们的注意，研究工作正逐渐开展。为了便于交流研究成果，为了将运筹学的一些新方法和动态给以迅速推广和介绍，特创办不定期的运筹报导。

一、本报着重选登：1）理论方面带有创造性的研究成果；2）实际任务中带有较普遍性（即具有一定参考意义的）工作总结；3）系统地介绍新方法的文章；4）国内外有关运筹学活动的情况介绍；5）书评。

二、由于我们对编辑工作还缺乏经验，本刊目前只刊登本室同志所写的文章，或少量外单位通过予约撰写的文章作者有权将其文章同时投交别的杂志。

三、欢迎读者对本报提出批评和意见。

目 录

带非线性约束的最小平方和（I）

§1 引言

令 $x=(x_1,\cdots,x_n)^T$，$f(x)=(f_1(x),\cdots\cdots f_m(x))^T$，$g(x)=(g_1(x),\cdots\cdots,g_l(x))^T$，作集合 $R=\{x \mid g(x)\geq 0\}$，作模函数 $\|f(x)\|^2=\sum_{i=1}^{m}f_i(x)^2$．本文对问题

$$(P) \qquad \min_{x\in R} \|f(x)\|^2$$

给出一个特殊解法．在用 Penalty 内点法求（P）的解时，当参数 γ 很小时，往々产生病态的现象，以致无法得到较满意的近似解；在用通常的方法求无条件约束问题解时，往々由于一维极值问题的计算量太大，以致得不到较好的近似解．本文企图利用模型的特殊性，给出一近似解法，以克服上述之困难．基本想法是在每点迭代时，若用性质较好的函数去近似 Penalty 函数，可期望大々减少一维搜索的计算量，加快收敛速度．

令

$$P(x,\gamma)=\frac{1}{2}\|f(x)\|^2-\gamma\sum_{j=1}^{l}\log g_j(x)$$

其中 $\gamma>0$ 为一参数．令

$$f_{ix}(x)=\left(\frac{\partial f_i(x)}{\partial x_1},\cdots\cdots,\frac{\partial f_i(x)}{\partial x_n}\right),$$

$$g_{jx}(x)=\left(\frac{\partial g_j(x)}{\partial x_1},\cdots\cdots,\frac{\partial g_j(x)}{\partial x_n}\right),$$

$$\nabla f(x)=\begin{pmatrix} f_{1x}(x) \\ \vdots \\ f_{mx}(x)\end{pmatrix},\ \nabla p(x,\gamma)=\left(\frac{\partial p(x,\gamma)}{\partial x_1},\cdots,\frac{\partial p(x,\gamma)}{\partial x_n}\right)$$

假设：$f_i(x)(i=1,\cdots,m)$，与 $g_j(x)(j=1,\cdots,l)$ 有一阶连续偏导数．在满足 $\nabla p(x^*,y)=0$ 的 x^* 处，$L(x,u)=\frac{1}{2}\|f(x)\|^2-u^Tg(x)\ (u\geqq 0)$ 对 x 为凸．在由下列迭代程序产生的 $\{x^k\}$ 的极限点 $\bar{x}$ 处，$-g_j(x)(j=1,\cdots,l)$ 严格凸．约束集合 $R=\{x\mid g(x)\geqq 0\}$ 有界，且 $R^0=\{x\mid g(x)>0\}\neq\Phi$．

记 $R_k=\{x\mid g_j(x^k)+g_{jx}(x^k)(x-x^k)\geqq 0\ (j=1,\cdots l)\}$

$R_k^0=\{x\mid g_j(x^k)+g_{jx}(x^k)(x-x^k)>0\ (j=1,\cdots,l)\}$

对任何 $x^k\in R^0$，存在 $\delta_k>0$，使

$$S_k=\{x\mid |x_i-x_i^k|\leq\delta_k\ (i=1,\cdots,n)\}\subseteq R^0$$

在 $x^k\in R^0$ 作 $p(x,\gamma)$ 的近似函数

$$Q(x,x^k,\gamma)=\frac{1}{2}\left\|\begin{pmatrix} f(x^k)+\nabla f(x^k)(x-x^k) \\ B(x-x^k)\end{pmatrix}\right\|^2-\gamma\sum_{j=1}^{l}\log\left[g_j(x^k)+g_{jx}(x^k)(x-x^k)\right],\quad x\in S_k\wedge R_k^0.$$

其中 B 为 $m\times n$ 矩阵，B^TB 正定．（例如，B^TB 为单位矩阵．）易见 $Q(x,x^k,\gamma)$ 在 $S_k\wedge R_k^0$ 中，对 x 为连续严格凸函数．

§2 迭代程序

取定 $\gamma>0$，$\delta_k=\delta_0>0$

1°．取 $x^0\in R^0$．令 $k:=0$．

2°. 设 $x^k \in R^\circ$ 已知. 若 $\nabla p(x^k, \gamma)=0$. 则 x^k 为问题 (P) 的近似解; 若 $\nabla p(x^k, \gamma) \neq 0$. 则以 x^k 为初始点. 求严格凸规划

$$\min_{y \in S_k \cap R_k^\circ} Q(y, x^k, \gamma)$$

的最优解 y^k. (在迭代时, 保持 $y \in R_k^\circ$ 的可行性, 求 $\min\limits_{y \in S_k} Q(y, x^k, \gamma)$ 的解 y^k.)

3°. 求单变量问题

$$\min_{\lambda \in \Lambda_k} p(x^k+\lambda(y^k-x^k), \gamma)$$

$$\Lambda_k=\{\lambda \geq 0 \mid x^k+\lambda(y^k-x^k) \in R\}$$

的解 λ_k. 令 $x^{k+1}=x^k+\lambda_k(y^k-x^k)$. $k:=k+1$ 转入 2°.

§3. 收敛性

定理 1. $\min\limits_{y \in S_k \cap R_k^\circ} Q(y, x^k, \gamma)$ 存在唯一解 y^k, 且 y^k-x^k 为 R 在 x^k 的可行方向.

证 $x^k \in S_k \cap R_k^\circ$. 作集合

$$W_k=\{x \mid x \in S_k \cap R_k^\circ, Q(x, x^k, \gamma) \leq Q(x^k, x^k, \gamma)\}.$$

欲证 W_k 为闭集. $W_k \subseteq S_k \cap R_k^\circ \subseteq S_k \cap R_k$ (闭集).

设 $z^p \in W_k$, $z^p \longrightarrow z$. $z \in S_k \cap R_k$.

若 $z \in S_k \cap (R_k-R_k^\circ)$. 则 $\lim\limits_{p \to \infty} Q(z^p, x^k, \gamma)=+\infty$

当 p 充分大时 $Q(z^p, x^k, \gamma)>Q(x^k, x^k, \gamma)$. 此与 $z^p \in W_k$

矛盾．故$\bar{z}\in S_k\cap R_k^o$．因$Q(x,x^k,\eta)$在$S_k\cap R_k^o$连续，故$\bar{z}\in W_k$．故W_k是有界闭集，故

$$\min_{y\in S_k\cap R_k^o} Q(y,x^k,\eta)=\min_{y\in W_k} Q(y,x^k,\eta)$$

存在解y^k．因Q为严格凸函数，故y^k唯一．

又由 $g(x^k+\lambda(y^k-x^k))=g(x^k)+\lambda g_x(x^k)(y^k-x^k)+o(\lambda)$

$$=(1-\lambda)g(x^k)+\lambda[g(x^k)+g_x(x^k)(y^k-x^k)]+o(\lambda)$$

$g(x^k)+g_x(x^k)(y^k-x^k)>0$，$g(x^k)>0$，知存在$\bar{\lambda}>0$，当$0\leq\lambda\leq\bar{\lambda}$时，$g(x^k+\lambda(y^k-x^k))\geqq 0$．故得欲证．

定理2．x^k为$\min\limits_{y\in S_k\cap R_k^o} Q(y,x^k,\eta)$的最优解，当且仅当 $\nabla P(x^k,\eta)=0$．

证 因$x^k\in S_k^o\cap R_k^o$．故x^k为$\min\limits_{y\in S_k\cap R_k^o} Q(y,x^k,\eta)$的解，当且仅当 $0=\nabla Q(x^k,x^k,\eta)=\nabla P(x^k,\eta)$。

定理3 设$\min\limits_{y\in S_k\cap R_k^o} Q(y,x^k,\eta)$的解$y^k\neq x^k$．

则 $\nabla P(x^k,\eta)(y^k-x^k)<0$．

证 因 $\nabla P(x^k,\eta)=\nabla Q(x^k,x^k,\eta)$．

$$Q(y^k,x^k,\eta)\geqq Q(x^k,x^k,\eta)+\nabla Q(x^k,x^k,\eta)(y^k-x^k)$$
$$=Q(x^k,x^k,\eta)+\nabla P(x^k,\eta)(y^k-x^k).$$

若 $\nabla P(x^k,\eta)(y^k-x^k)\geqq 0$，将得 $Q(y^k,x^k,\eta)\geqq Q(x^k,x^k,\eta)$之矛盾．

定理4. 设 $\tilde{x} \in R^{\circ}$. $\nabla p(\tilde{x}, \gamma) = 0$，函数 $L(x,u) = \frac{1}{2}\|f(x)\|^2 - u^T g(x)\ (u \geqq 0)$ 在 $\tilde{x}$ 凸. 则

$$\|f(\tilde{x})\|^2 \geqslant \min_{x \in R} \|f(x)\|^2 \geqslant \|f(\tilde{x})\|^2 - 2\gamma l.$$

证. 令 $u_j^{\circ} = \gamma / g_j(\tilde{x}) \quad (j = 1, \cdots, l)$.

$$u^{\circ} = (u_1^{\circ}, \cdots, u_l^{\circ})^T$$

则 $$0 = \nabla p(\tilde{x}, \gamma) = f(\tilde{x})^T \nabla f(\tilde{x}) - \gamma \sum_{j=1}^{l} \frac{g_{jx}(\tilde{x})}{g_j(\tilde{x})}$$

$$= f(\tilde{x})^T \nabla f(\tilde{x}) - u^{\circ T} g_x(\tilde{x})$$

$$= L_x(\tilde{x}, u^{\circ}).$$

当 $x \in R$ 时，

$$\frac{1}{2}\|f(x)\|^2 \geqslant \frac{1}{2}\|f(x)\|^2 - u^{\circ T} g(x) = L(x, u^{\circ})$$

$$\geqslant L(\tilde{x}, u^{\circ}) + L_x(\tilde{x}, u^{\circ})(x - \tilde{x})$$

$$= L(\tilde{x}, u^{\circ}) = \frac{1}{2}\|f(\tilde{x})\|^2 - \gamma l,$$

即得欲证.

定理5 设 $\lim x^k = \bar{x} \in R^{\circ}$. y^k 为

$$\min_{y \in S_k \cap R_k^{\circ}} Q(y, x^k, \gamma)$$ 的解，$\lim y^k = \bar{y}$. 若 $-g_j(x)$ $(j = 1, \cdots, l)$ 在 $\bar{y}$ 严格凸. 则 $\bar{y}$ 必为

$$\min_{y \in \bar{S} \cap \bar{R}^{\circ}} Q(y, \bar{x}, \gamma)$$

的唯一解，其中

$$\bar{S}=\{x \mid |x_i-\bar{x}_i| \leqslant \bar{\delta},\ i=1,\cdots,n\},\quad \bar{\delta}=\lim \delta_k>0.$$

$$\bar{R}^0=\{x \mid g_j(\bar{x})+g_{jx}(\bar{x})(x-\bar{x})>0,\ j=1,\cdots,l\}.$$

证．由 Kuhn-Tucker 定理，有在 u^k, v^k，使

$$(*)\begin{cases}\begin{pmatrix} f(x^k)+\nabla f(x^k)(y^k-x^k) \\ B(y^k-x^k)\end{pmatrix}\begin{pmatrix}\nabla f(x^k) \\ B\end{pmatrix}-\eta\sum_{j=1}^{l}\dfrac{g_{jx}(x^k)}{g_j(x^k)+g_{jx}(x^k)(y^k-x^k)} \\ \qquad +u^{k^T}-v^{k^T}=0. \\ u^{k^T}(\delta_k e+x^k-y^k)=0,\ \delta_k e+x^k-y^k\geqq 0,\ u^k\geqq 0,\ e=(1,\cdots,1)^T. \\ v^{k^T}(\delta_k e-x^k+y^k)=0,\ \delta_k e-x^k+y^k\geqq 0,\ v^k\geqq 0.\end{cases}$$

故

$$\delta_k u^{k^T}e=u^{k^T}(y^k-x^k)\leqslant -f(x^k)^T\nabla f(x^k)(y^k-x^k)$$

$$+\eta\sum_{j=1}^{l}\frac{g_{jx}(x^k)(y^k-x^k)}{g_j(x^k)+g_{jx}(x^k)(y^k-x^k)}$$

$$\leqslant \sup_{x,y\in S_k}\left|f(x)^T\nabla f(x)(y-x)\right|+\eta l.$$

故 $\{u^k\}$ 有界．同样，$\{v^k\}$ 有界．不妨设 $u^k\to\bar{u}, v^k\to\bar{v}$.

在 $(*)$ 中令 $k\to\infty$．因 $-g_j(x)$ 在 $\bar{x}$ 严格凸，

$g_j(\bar{x})+g_{jx}(\bar{x})(\bar{y}-\bar{x})>g_j(\bar{y})\geqq 0$（若 $\bar{y}\neq\bar{x}$）．故

$(\bar{y},\bar{u},\bar{v})$ 满足 $\displaystyle\min_{y\in\bar{S}\cap\bar{R}^0} Q(y,\bar{x},\eta)$ 的 Kuhn-Tucker

条件，故得欲证．

<u>定理 6</u>．设 $x^k\in R^0$．$\lim x^k=\bar{x}$．则

(1). $x \in R^0$.

(2). 存在与k无关的 $\alpha>0$，使

$$x^k+\lambda(y^k-x^k)\in R \quad (0\le\lambda\le\alpha).$$

证. (1). 因 $\{P(x^k,\eta)\}$ 单调下降. 故若有 $g_j(\bar{x})=0$，将得 $\lim P(x^k,\eta)=P(\bar{x},\eta)=+\infty$ 之矛盾

(3). 因 $x^k\in R^0$. 由$g(x)$连续性以及R有界，存在 $\alpha^k>0$，使 $x^k+\lambda(y^k-x^k)\in R$ $(0\le\lambda\le\alpha_k)$，而对某个 j_k $(1\le j_k\le l)$，有 $g_{j_k}(x^k+\alpha_k(y^k-x^k))=0$. 若 $\inf\limits_k \alpha_k=0$ 则存在子列 $\{\alpha_k'\}$，$\lim\alpha_k'=0$. 故存在子列 $\{k''\}\subseteq\{k'\}$，及某个与k无关的j，使

$$g_j(x^{k''}+\alpha_{k''}(y^{k''}-x^{k''}))=0,\quad \lim\alpha_k''=0.$$

故 $g_j(\bar{x})=\lim g_j(x^{k''}+\alpha_{k''}(y^{k''}-x^{k''}))=0$. 此与 $\bar{x}\in R^0$ 矛盾，故 $\inf\alpha_k=\alpha>0$.

<u>定理7</u>. 设 $x^0\in R^0$. $y^k\neq x^k$. 则 $x^k\in R^0$. 因此 $\{P(x^k,\eta)\}$ 为单调下降序列.

证. 作 $W_k=\{x^k+\lambda(y^k-x^k)\mid x+\lambda(y^k-x^k)\in R^0, P(x^k+\lambda(y^k-x^k),\eta)\le P(x^k,\eta)\}$. 与定理1同样证 W_k 有界闭.

$$\min_{\lambda\in\Lambda_k} P(x^k+\lambda(y^k-x^k),\eta)=\min_{x\in W_k} P(x^k+\lambda(y^k-x^k),\eta),$$

即 $x^{k+1}\in R^0$. 由定理3，易见 $\{P(x^k,\eta)\}$ 单调下降.

定理8.（收敛定理）. 在§1 假设下，由迭代程序所得的点列 $\{x^k\}$ 有下列性质：

(1). $\{P(x^k,\eta)\}$ 单调下降且收敛；

(2) 若 $\nabla P(x^k,\eta)=0$，则

$$\|f(x^k)\|^2 \geqslant \operatorname*{Min}_{x\in R}\|f(x)\|^2 \geqslant \|f(x^k)\|^2-2\eta l;$$

(3) 若 $\bar{x}$ 是 $\{x^k\}$ 的极限点. 则 $\nabla P(\bar{x},\eta)=0$，且

$$\|f(\bar{x})\|^2 \geqslant \operatorname*{Min}_{x\in R}\|f(x)\|^2 \geqslant \|f(\bar{x})\|^2-2\eta l.$$

证. (1)是定理3及R有界性的推论.

(2)和(3)的第二个结论是定理4和定理6的推论. 只须证明 $\nabla P(\bar{x},\eta)=0$.

不妨设 $\operatorname{Lim} x^k=\bar{x}$，$\operatorname{Lim} y^k=\bar{y}$. 若 $\nabla P(\bar{x},\eta)\neq 0$，由定理5，2，3，必有 $\nabla P(\bar{x},\eta)(\bar{y}-\bar{x})<0$.

由 Tayler 展式，

$$P(x^{k+1},\eta)\leqslant P(x^k+\lambda(y^k-x^k))=P(x^k,\eta)+\frac{\lambda}{2}\nabla P(x^k,\eta)(y^k-x^k)$$

$$+\lambda\left[\frac{1}{2}\nabla P(x^k,\eta)(y^k-x^k)+\frac{o(\lambda)}{\lambda}\right],$$

其中 $\frac{o(\lambda)}{\lambda}$ 对 $\lambda>0$ 连续. 由定理3，对一切 k，

$$\nabla P(x^k,\eta)(y^k-x^k)<0.$$

若对一切 $\lambda>0$，有

$$\frac{1}{2}\nabla P(x^k,\eta)(y^k-x^k)+\frac{o(\lambda)}{\lambda}\leqslant 0.$$

由定理6，存在与 k 无关的 $\alpha>0$，使

$$x^k+\lambda(y^k-x^k)\in R,\quad (0\le\lambda\le\alpha).$$

故 $P(x^{k+1},\eta)\le P(x^k,\eta)+\frac{\alpha}{2}\nabla P(x^k,\eta)(y^k-x^k)$. 因 $\{P(x^k,\eta)\}$ 收斂,

故得 $$\nabla P(\bar{x},\eta)(\bar{y}-\bar{x})=\lim \nabla p(x^k,\eta)(y^k-x^k)=0$$

之矛盾.

若存在 $\lambda_k'>0$, 使

$$\frac{1}{2}\nabla p(x^k,\eta)(y^k-x^k)+\frac{o(\lambda)}{\lambda}\begin{cases}<0, & 0<\lambda<\lambda_k'.\\ =0. & \lambda=\lambda_k'\end{cases},$$

则 $P(x^{k+1},\eta)\le P(x^k,\eta)+\frac{1}{2}\mathrm{Min}\{\alpha,\lambda_k'\}\nabla p(x^k,\eta)(y^k-x^k)$.

故 $\lim \mathrm{Min}\{\alpha,\lambda_k'\}\cdot\nabla p(x^k,\eta)(y^k-x^k)=0$.

$\lim \mathrm{Min}\{\alpha,\lambda_k'\}=0$.

故存在子列 $\{\lambda_{k'}'\}$, 使 $\lim\limits_{k'\to\infty}\lambda_{k'}'=0$ 由

$$\frac{1}{2}\nabla p(x^{k'},\eta)(y^{k'}-x^{k'})+\frac{o(\lambda_{k'}')}{\lambda_{k'}'}=0$$

$$\lim \frac{o(\lambda_{k'}')}{\lambda_{k'}'}=0,$$

又得 $\nabla p(\bar{x},\eta)(\bar{y}-\bar{x})=0$ 之矛盾. 故 $\nabla p(\bar{x},\eta)=0$

一个新的既约梯度法及其收敛性

一、引言

在1962年P. Wolfe提出了解决一类非线性规划问题的一种标法，称为"既约梯度法"（Reduced Gradient method）[1]。这种方法起始是用来解决具有线性约束条件和有连续微商的非线性目标函数的规划问题，在使用中它成功地解决了大量的实际问题，一般它比梯度射影法更有效。（1969年Abadie和Carpentier把它推广到具有非线性约束的情况[4]）。然而却一直不能证明Wolfe的方法所标出的点列收敛到最优解。1972年Wolfe自己发表了一个简单的数值例子[2]，在这个例子中，用这个方法所求出的点列收敛到一个非最优解。最近，P. Huard给出一个改进的既约梯度法[3]，并试图在更强的假设条件下建立迭代点列的收敛性（即收敛到最优解）。但是，他的证明是不正确的，原因是他在证明中用到两个错误的命题：1) Rosen的"梯度射影法"[5]具有这种收敛性（[3], p.44）。其实Rosen的梯度射影法和Wolfe的方法一样，不能保证收敛到最优解（甚至1969年Goldfarb改进了的梯度射影方法[6]也不能克服这种缺陷）Polak[7]引进了ε一假设，证明了Rosen-Polak方法的收敛性，Huard则未引入这一假设；2) 级数 $\frac{\alpha}{\beta}\sum_{k}|y^{k}|$ 是收敛的。Huard是由 $\sum_{k}|y^{k}_{I_k}|^2<+\infty$，$|y^{k}|\leqslant\beta^2|y^{k}_{I_k}|^2$，和 $|y^{k1}_{I_k}|\geqslant\alpha>0$，$\forall k\in N$ 推出 $\frac{\alpha}{\beta}\sum|y^{k}|<+\infty$（[3], p.48），看来他弄混了 $|y^{k}_{I_k}|$ 和

$\left|y^{k+1}_{I_k}\right|$，而$\left|y^{k}_{I_k}\right|$是趋于0的。因此 Huard 的比较繁杂的标法也是不成功的。

在本文中我们给出一个新的既约梯度方法，并且在比[3]弱的假设条件下证明了或者在有限步迭代后得到最优解，或者得到一序列，它的任意一极限点都是最优解。同时我们的方法比 Huard 的方法要简单。

二、假设条件和符号

我们在本文中研究如下的非线性规划问题：

$$(P)\quad \min\{f(x) \mid Ax=b,\ x\geq 0\}, \tag{1}$$

其中$x\in R^n$，A是$m\times n$矩阵（$m<n$），$b\in R^m$。我们假设：

(i) 函数$f: R^n\to R^1$是凸函数，且$f\in C^1$（即f有连续的一阶偏导数）；

(ii) 对$x=(x_1,\cdots,x_n)$，令$S(x)=\{j \mid x_j>0\}$，设对$\forall x\in R$，矩阵$A^{S(x)}$的秩为m（R是全体可行解的集；$A^{S(x)}$是由标点属于$S(x)$的列向量所成的子矩阵，一般A^J_L表示A中属于J列和L行的元素所成的子矩阵，$J\subset\{1,2,\cdots,n\}$，$L\subset\{1,2,\cdots,m\}$）。

我们所给的这两条假设比 Huard[3] 的假设条件要弱。他假设f是有连续的二阶偏导数的凹函数（Huard 求的是f的

最大值），並且除上面的（ii）外，　他还假定了可行集 R 有界，以及 $\exists \beta' > 0$，使得 $-\beta'|y|^2 \leq y \cdot H(x) y \leq 0$，$\forall x \in R^n$，$\forall y \in R^n$（$H(x)$ 是 f 的二阶偏导数的矩陣）。

在本文中我们仍使用〔3〕中的一些符号：

$I \subset \{1, 2, \cdots, n\}$，令 $|I|$ 表示 I 中元素的个数；若 $|I| = m$，且矩陣 A^I 的秩为 m，则称 I 是一组基；

$\bar{I}$ 表示 $\{1, 2, \cdots, n\} - I$，即 I 以外的全体标号；

$T(I)$ 表示矩陣 $(A^I)^{-1} A$，$T^{\bar{I}}(I)$ 表示 $(A^I)^{-1} A^{\bar{I}}$，

$T_r^{\bar{I}}(I)$ 表示 $T^{\bar{I}}(I)$ 的标号为 r 的行；

$t(I)$ 表示 $(A^I)^{-1} b$；

$\nabla_I f(x) = \left(\frac{\partial f(x)}{\partial x_j}, j \in I\right)$，$\nabla_{\bar{I}} f(x) = \left(\frac{\partial f(x)}{\partial x_j}, j \in \bar{I}\right)$；

显然对任意基 I，若 $x = (x_I, x_{\bar{I}}) \in R$，则有等式

$$x_I = (A^I)^{-1} b - (A^I)^{-1} A^{\bar{I}} x_{\bar{I}} = t(I) - T^{\bar{I}}(I) x_{\bar{I}} 。 \quad (2)$$

如将（2）式代入目标函数 $f(x)$，则得到 $\bar{f}(x_{\bar{I}}): R^{n-m} \to R^1$：

$$\bar{f}(x_{\bar{I}}) = f(t(I) - T^{\bar{I}}(I) x_{\bar{I}}, x_{\bar{I}}),$$

显然有

$$\nabla \bar{f}(x_{\bar{I}}) = \nabla_{\bar{I}} f(x) - \nabla_I f(x) \cdot T^{\bar{I}}(I), \quad (3)$$

我们称 $\nabla \bar{f}(x_{\bar{I}})$ 为"既约梯度"。又令

$$d(x_{\bar{I}}) = -\nabla \bar{f}(x_{\bar{I}}) 。$$

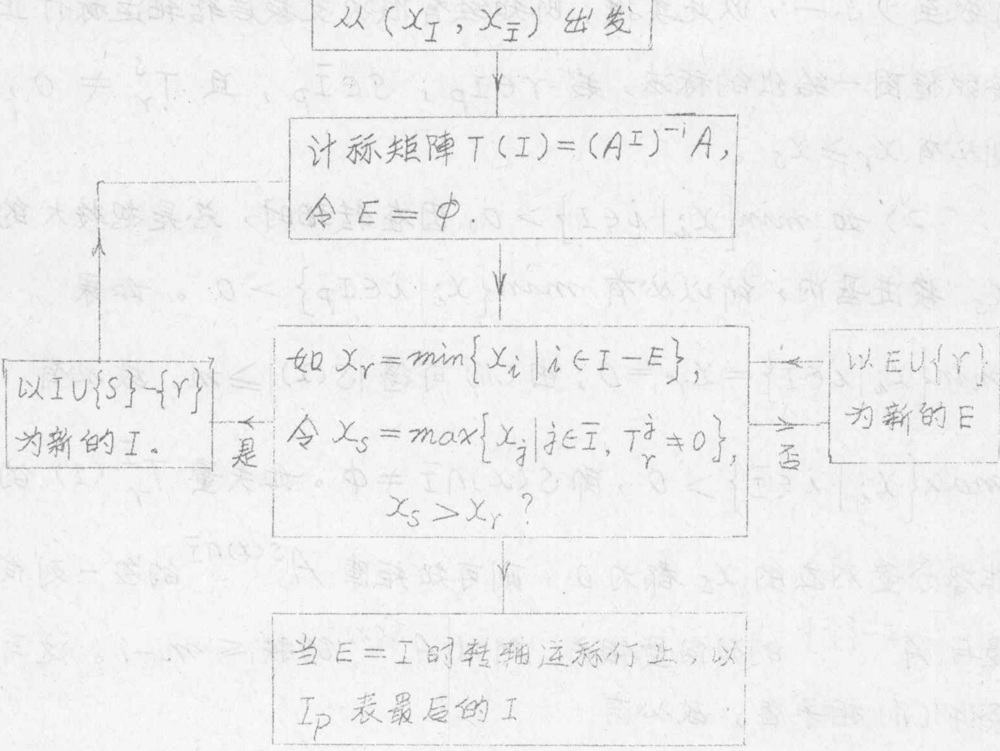
三、转轴运算

在求解问题（P）的迭代过程中，如何进行"转轴"(pivot)是既约梯度法的重要部份（转轴是基的一种变换，参见线性规划）。Wolfe[1]和Huard[3]在他们的方法中曾提出相近似的转轴方法，他们的方法都不能保证求出的序列$\{X^k\}$的极限点是最优解。下面我们给出一新的转轴方法：

设$X=(X_I, X_{\bar{I}})$为任一可行点，I为任意一基。对下标$r\in I$，相应地有矢量$T_r^{\bar{I}}(I)$，令$T_r^{\bar{I}}(I)$中不为0的分量T_r^j所对应的x_j中的最大数为x_s，即

$$x_s=max\{x_j; j\in \bar{I}, T_r^j\neq 0\}$$

如$x_s>x_r$，则将r与s交换；如$x_s\leq x_r$，则r不换。显然当r与s交换后，$I\cup\{s\}-\{r\}$是一新的基，这称为一次转轴。这种手续一直进行到基中之元素不能再被交换为止。整个转轴运算表成如下的框图



框 图 一

定理1. 在假设(ii)下， 对任意可行点 $x\in R$ 和任意基 I，必有 1) 经有限次交换后，转轴运算停止，此时对 $r\in I_p$ 和 $s\in\bar{I}_p$，如 $T_r^s\neq 0$，则有 $x_r\geq x_s$；2) $\min\{x_i \mid i\in I_p\}>0$。

证. 1) 按照定义，$r\in I$ 和 $s\in\bar{I}$ 的交换只在 $x_s>x_r$ 时才有可能发生。我们令 $Z=\{(i,j) \mid i\in I, j\in\bar{I}, x_j>x_i\}$，集合 Z 内元素的个数显然是有限的。若 r 与 s 进行交换，对于新的基 $I'=I\cup\{s\}-\{r\}$，不难推知 $Z'=\{(i,j) \mid i\in I', j\in\bar{I}', x_j>x_i\}$ 内元素的个数比 Z 内元素的