第12卷第1期 1990年3月
重庆建筑工程学院学报
JOURNAL OF CHONGQING INSTITUTE OF ARCHITECTURE AND ENGINEERING
Vol.12 No.1 March 1990

最优化的一种新途径

李泽民
（基础科学系）

摘 要 本文用n维欧氏空间R^n中的隐函数定理研究等式约束问题的最优性必要条件，从而得出解这类问题的一种新途径。它较经典的Lagrange乘子法可减少解方程组的维数。

关键词 隐函数定理，等式约束问题，最优化，欧几里得空间，Lagrange乘子

引 言

本文在研究乘积Banach空间中的优化问题时，从抽象空间得出的结果推得关系式

$$P=(M^{-1}N)^TQ$$ [1]

现在直接用R^n中的隐函数定理来导出这个结论。

1 最优性必要条件

我们知道，对方程组

$$h(x)=0$$

有下面著名的隐函数定理，其中$x\in R^n$，$h:R^n\to R^m$，设$m<n$，记$h(x)=(h_1(x),\cdots,h_m(x))$。

引理(隐函数定理)[2]。设$x^0=(x_1^0,\cdots,x_n^0)\in R^n$满足：

（1） 在x^0的某个邻域中$h\in C^p$，$p\geq 1$；

（2） $h(x^0)=0$；

（3） $m\times m$阶 Jacobian 矩阵：

$$J=\begin{pmatrix}\dfrac{\partial h_1(x^0)}{\partial x_1} & \cdots & \dfrac{\partial h_1(x^0)}{\partial x_m}\\ \vdots & & \vdots\\ \dfrac{\partial h_m(x^0)}{\partial x_1} & \cdots & \dfrac{\partial h_m(x^0)}{\partial x_m}\end{pmatrix}$$

本文1988年7月6日收到。

是非奇异的。

则存在$\hat{x}^0=(x^0_{m+1},\cdots,x^0_n)\in R^{n-m}$的一个邻城$U$，使得$x_1,\cdots,x_m$是$\hat{x}=(x_{m+1},\cdots,x_n)\in U$的函数，记成$x_i=\phi_i(\hat{x})$，$i=1,\cdots,m$，具有下列性质：

(1) $\phi_i\in C^p$；

(2) $x^0_i=\phi_i(\hat{x}^0)$，$i=1,\cdots,m$；

(3) $h_i(\phi_1(\hat{x}),\cdots,\phi_m(\hat{x}),\hat{x})=0$，$i=1,\cdots,m$，$\hat{x}\in U$。

现在考虑等式约束问题

$$(PE)\begin{cases}\min\ f(x)\\ s.t.\ g(x)=0\end{cases}$$

其中$x\in R^n$，$f:R^n\to R$，$g:R^n\to R^m$，$m<n$。记$g(x)=(g_1(x),\cdots,g_m(x))$，$p=n-m$，令

$$P=\left(\frac{\partial f(x^0)}{\partial x_1},\cdots,\frac{\partial f(x^0)}{\partial x_p}\right)^T,$$

$$Q=\left(\frac{\partial f(x^0)}{\partial x_{p+1}},\cdots,\frac{\partial f(x^0)}{\partial x_n}\right)^T,$$

$$N=\begin{pmatrix}\frac{\partial g_1(x^0)}{\partial x_1}&\cdots&\frac{\partial g_1(x^0)}{\partial x_p}\\ \vdots&&\vdots\\ \frac{\partial g_m(x^0)}{\partial x_1}&\cdots&\frac{\partial g_m(x^0)}{\partial x_p}\end{pmatrix}.$$

$$M=\begin{pmatrix}\frac{\partial g_1(x^0)}{\partial x_{p+1}}&\cdots&\frac{\partial g_1(x^0)}{\partial x_n}\\ \vdots&&\vdots\\ \frac{\partial g_m(x^0)}{\partial x_{p+1}}&\cdots&\frac{\partial g_m(x^0)}{\partial x_n}\end{pmatrix}$$

定理 1　设$x^0\in R^n$是(PE)的最优解·$f:R^n\to R$，$g:R^n\to R^m$连续可微，矩阵M非奇异，则有

$$P=(M^{-1}N)^TQ$$

证　因$g(x^0)=0$，M非奇异，于是由隐函数定理$\exists\ \hat{x}^0=(x^0_1,\cdots,x^0_p)$的一个邻域$U$，使得$x_{p+1},\cdots,x_n$是$\hat{x}=(x_1,\cdots,x_p)\in U$的函数，即有

$$x_i=\varphi_i(\hat{x}),\quad \hat{x}\in U,\ i=p+1,\cdots,n$$

这些函数具有下列性质

(1) $\varphi_i\in C^1$；

(2) $x^0_i=\varphi_i(\hat{x}^0),i=p+1,\cdots,n$；

(3) $g_i(\hat{x},\ \varphi_{p+1}(\hat{x}),\cdots,\varphi_n(\hat{x}))=0\quad i=1,\cdots,m,\ \hat{x}\in U$。　(1)

因x^0是(PE)的最优解，即

$$f(x^0)\le f(x)\quad \forall x\in\{x\in R^n:g(x)=0\}$$

所以，有

$$f(\hat{x}^0,\ \varphi_{p+1}(\hat{x}^0),\cdots,\varphi_n(\hat{x}^0)) \le f(\hat{x},\varphi_{p+1}(\hat{x}),\cdots,\varphi_n(\hat{x}))\quad \forall x\in U \tag{2}$$

令

$$F(\hat{x})=f(\hat{x},\ \varphi_{p+1}(\hat{x}),\cdots,\varphi_n(\hat{x})),\quad \hat{x}\in U$$

由(2)式得

$$F(\hat{x}^0)\le F(\hat{x})\quad \forall \hat{x}\in U$$

故$\hat{x}^0$是$F(\hat{x})$的极小点，于是，

$$\frac{\partial F(\hat{x}^0)}{\partial x_i}=0,\ i=1,2,\cdots,p$$

由此得

$$\begin{pmatrix}\frac{\partial f(x^0)}{\partial x_1}\\ \vdots\\ \frac{\partial f(x^0)}{\partial x_p}\end{pmatrix}+\begin{pmatrix}\frac{\partial\varphi_{p+1}(\hat{x}^0)}{\partial x_1}&\cdots&\frac{\partial\varphi_n(\hat{x}^0)}{\partial x_1}\\ \vdots&&\vdots\\ \frac{\partial\varphi_{p+1}(\hat{x}^0)}{\partial x_p}&\cdots&\frac{\partial\varphi_n(\hat{x}^0)}{\partial x_p}\end{pmatrix}\begin{pmatrix}\frac{\partial f(x^0)}{\partial x_{p+1}}\\ \vdots\\ \frac{\partial f(x^0)}{\partial x_n}\end{pmatrix}=0 \tag{3}$$

由（1）式同样有

$$\begin{pmatrix}\frac{\partial g_i(x^0)}{\partial x_1}\\ \vdots\\ \frac{\partial g_i(x^0)}{\partial x_p}\end{pmatrix}+\begin{pmatrix}\frac{\partial\varphi_{p+1}(\hat{x}^0)}{\partial x_1}&\cdots&\frac{\partial\varphi_n(\hat{x}^0)}{\partial x_1}\\ \vdots&&\vdots\\ \frac{\partial\varphi_{p+1}(\hat{x}^0)}{\partial x_p}&\cdots&\frac{\partial\varphi_n(\hat{x}^0)}{\partial x_p}\end{pmatrix}\begin{pmatrix}\frac{\partial g_i(x^0)}{\partial x_{p+1}}\\ \vdots\\ \frac{\partial g_i(x^0)}{\partial x_n}\end{pmatrix}=0$$

$$i=1,2,\cdots,m \tag{4}$$

令

$$A=\begin{pmatrix}\frac{\partial\varphi_{p+1}(\hat{x}^0)}{\partial x_1}&\cdots&\frac{\partial\varphi_n(\hat{x}^0)}{\partial x_1}\\ \vdots&&\vdots\\ \frac{\partial\varphi_{p+1}(\hat{x}^0)}{\partial x_p}&\cdots&\frac{\partial\varphi_n(\hat{x}^0)}{\partial x_p}\end{pmatrix}$$

则（4）式变成

$$\begin{pmatrix}\frac{\partial g_i(x^0)}{\partial x_1}\\ \vdots\\ \frac{\partial g_i(x^0)}{\partial x_p}\end{pmatrix}=-A\begin{pmatrix}\frac{\partial g_i(x^0)}{\partial x_{p+1}}\\ \vdots\\ \frac{\partial g_i(x^0)}{\partial x_n}\end{pmatrix}\quad i=1,\cdots,m$$

从而

$$\begin{pmatrix} \frac{\partial g_1(x^0)}{\partial x_1} & \cdots & \frac{\partial g_m(x^0)}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial g_1(x^0)}{\partial x_p} & \cdots & \frac{\partial g_m(x^0)}{\partial x_p} \end{pmatrix} = -A \begin{pmatrix} \frac{\partial g_1(x^0)}{\partial x_{p+1}} & \cdots & \frac{\partial g_m(x^0)}{\partial x_{p+1}} \\ \vdots & & \vdots \\ \frac{\partial g_1(x^0)}{\partial x_n} & \cdots & \frac{\partial g_m(x^0)}{\partial x_n} \end{pmatrix}$$

即

$$N^T = -AM^T$$

因M非奇异，所以

$$A = -N^T(M^T)^{-1} \qquad (5)$$

最后，根据(3)、(5)两式便得出

$$P = (M^{-1}N)^T Q$$ 证毕

例 1

$$\begin{cases} \min\ f = x_1^2 + x_2^2 + x_3^2 \\ \text{s.t. } g_1 = x_1^2 + x_2^2 - x_3 = 0 \\ \qquad g_2 = x_1 + x_2 + x_3 - 1 = 0 \end{cases}$$

最优解是$\left(\frac{-1+\sqrt{3}}{2}, \frac{-1+\sqrt{3}}{2}, 2-\sqrt{3}\right)$. 这时，$p = 3-2 = 1$,

$$M = \begin{bmatrix} -1+\sqrt{3} & -1 \\ 1 & 1 \end{bmatrix}, \quad M^{-1} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 \\ -1 & -1+\sqrt{3} \end{bmatrix}$$

$$N = (-1+\sqrt{3},\ 1)^T$$

$$P = -1+\sqrt{3}$$

$$Q = (-1+\sqrt{3},\ 2(2-\sqrt{3}))^T$$

从而

$$M^{-1}N = (1,\ 0)^T$$

$$(M^{-1}N)^T Q = -1+\sqrt{3}$$

因此有 $P = (M^{-1}N)^T Q$.

注 从定理 1 的证明看出，矩阵 P, Q, M, N 中变量 $x_1, \cdots, x_n$ 的下标不一定要依次取，关键在于约束函数 g 在点 x^0 的 Frechet 一导数要有一个 m 阶子块非奇异， 这等价于梯度向量组 $\nabla g_1(x^0), \cdots, \nabla g_m(x^0)$线性无关。

下面，讨论等式 $P = (M^{-1}N)^T Q$ 与 Kuhn-Tucker 条件的关系。

$P = N^T(M^{-1})^T Q$ 即是

$$\begin{pmatrix} \frac{\partial f(x^0)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x^0)}{\partial x_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1(x^0)}{\partial x_1} & \cdots & \frac{\partial g_m(x^0)}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial g_1(x^0)}{\partial x_p} & \cdots & \frac{\partial g_m(x^0)}{\partial x_p} \end{pmatrix} (M^{-1})^T Q \qquad (6)$$

然而

$$\begin{pmatrix} \frac{\partial g_1(x^0)}{\partial x_{p+1}} & \cdots & \frac{\partial g_m(x^0)}{\partial x_{p+1}} \\ \vdots & & \vdots \\ \frac{\partial g_1(x^0)}{\partial x_n} & \cdots & \frac{\partial g_m(x^0)}{\partial x_n} \end{pmatrix} (M^{-1})^T Q = M^T (M^{-1})^T Q = Q \tag{7}$$

所以由(6)，(7)式有

$$\begin{pmatrix} \frac{\partial f(x^0)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x^0)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1(x^0)}{\partial x_1} & \cdots & \frac{\partial g_m(x^0)}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial g_1(x^0)}{\partial x_n} & \cdots & \frac{\partial g_m(x^0)}{\partial x_n} \end{pmatrix} (M^{-1})^T Q \tag{8}$$

即是

$$\nabla f(x^0) = \nabla g(x^0)^T (M^{-1})^T Q$$

令 $\lambda = (\lambda_1, \cdots, \lambda_m)^T = -(M^{-1})^T Q$，则由(8)式得

$$\nabla f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0 \tag{9}$$

因此，若 $P = (M^{-1}N)^T Q$，那末(9)式自然成立，故有下面定理。

定理2　若 $P = (M^{-1}N)^T Q$，则

$$\nabla f(x^0) + \lambda^T \nabla g(x^0) = 0$$

其中 $\lambda = -(M^{-1})^T Q$。

于是，根据定理2，可用求解方程组

$$\begin{cases} P = (M^{-1}N)^T Q \\ g(x) = 0 \end{cases} \tag{10}$$

来得出等式约束问题(PE)的K—T点，其中

$$P = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_p}\right)^T \qquad Q = \left(\frac{\partial f(x)}{\partial x_{p+1}}, \cdots, \frac{\partial f(x)}{\partial x_n}\right)^T$$

$$N = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial g_m(x)}{\partial x_1} & \cdots & \frac{\partial g_m(x)}{\partial x_p} \end{pmatrix}$$

$$M = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_{p+1}} & \cdots & \frac{\partial g_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m(x)}{\partial x_{p+1}} & \cdots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix}$$

使 M 非奇异的解便是（PE）的 K—T 点。

例2　$\begin{cases} \min\ f = x_1^2 + x_2^2 + x_3^2 \\ \text{s.t.}\ g = x_1 + x_2 + x_3 - 1 = 0 \end{cases}$

这时 $p=3-1=2$

$$P=(2x_1,\ 2x_2)^T$$
$$Q=2x_3$$
$$N=(1,\ 1)$$
$$M=1$$

由（10）式得

$$\begin{aligned} &x_1=x_3\\ &x_2=x_3\\ &x_1+x_2+x_3-1=0 \end{aligned} \tag{11}$$

此方程组的解为$\left(\frac{1}{3},\ \frac{1}{3},\ \frac{1}{3}\right)^T$。这不仅是$K-T$点且易验证还是最优解。

现在用经典的 Lagrange 乘子法求解例 2．令

$$L=x_1^2+x_2^2+x_3^2+\lambda(x_1+x_2+x_3-1)$$

解方程组

$$\begin{aligned} &\frac{\partial L}{\partial x_1}=2x_1+\lambda=0\\ &\frac{\partial L}{\partial x_2}=2x_2+\lambda=0\\ &\frac{\partial L}{\partial x_3}=2x_3+\lambda=0\\ &\frac{\partial L}{\partial \lambda}=x_1+x_2+x_3-1=0 \end{aligned} \tag{12}$$

得最优解$x^0=\left(\frac{1}{3},\ \frac{1}{3},\ \frac{1}{3}\right)^T$

从上述两种算法看出，前者需解方程组(11)的维数为 3，而后者需解方程组(12)的维数是 4．事实上，一般地，方程组（10）的维数$=p+m=n-m+m=n$. 对规划（PE）用经典 Lagrange 乘子法时，由 Lagrange 函数

$$L(x,\lambda)=f(x)+\sum_{i=1}^{m}\lambda_i g_i(x)$$

对各变元的偏导数构成的方程组是

$$\frac{\partial L}{\partial x_i}=0,\quad i=1,\ 2,\ \cdots n,$$

$$\frac{\partial L}{\partial \lambda_j}=0,\quad j=1,\ 2,\ \cdots,\ m$$

其维数$=n+m$. 可见，当 m 较大时，两种方法所解相应方程组的维数差别是很大的，从这个意义讲前者较后者简单。

例 3
$$\begin{cases} \min\quad f=x_1^2+x_2^2+x_3^2\\ \text{s.t.}\quad g_1=x_1^2+x_2^2-x_3=0\\ \qquad\ \ g_2=x_1+x_2+x_3-1=0 \end{cases}$$

这时$p=3-2=1$.

$$P=2x_1$$
$$Q=(2x_2,\ 2x_3)^T$$
$$N=(2x_1,\ 1)^T$$
$$M=\begin{pmatrix}2x_2 & -1\\ 1 & 1\end{pmatrix}$$

M 的行列式

$$|M|=2x_2+1 \qquad \left(x_2\neq-\frac{1}{2}\right)$$

M 的逆矩阵

$$M^{-1}=\frac{1}{2x_2+1}\begin{pmatrix}1 & 1\\ -1 & 2x_2\end{pmatrix}$$

于是，由(10)式得

$$x_1=\frac{1}{2x_2+1}(x_2(2x_1+1)-2x_3(x_1-x_2))$$
$$x_1^2+x_2^2-x_3=0$$
$$x_1+x_2+x_3-1=0$$

其实数解为$\left(\frac{-1+\sqrt{3}}{2},\ \frac{-1+\sqrt{3}}{2},2-\sqrt{3}\right)^T$, $\left(\frac{-1-\sqrt{3}}{2},\ \frac{-1-\sqrt{3}}{2},2+\sqrt{3}\right)^T$.
二者均是K—T点，但前者是最优解，后者非也。

参 考 文 献

〔1〕李泽民，乘积空间中的等式约束问题及其应用，待发表
〔2〕David G. Luenberger, Linear and Nonlinear Programming, second Edition, 1984

（编辑：姚国安）

A NEW APPROACH OF OPTIMIZATION

Li Zemin

(Department of Natural Science)

ABSTRACT In this paper, a necessary optimality condition of problems with equality constraints is investigated by the implicit function theorem in n-dimensional Euclidean space. A new approach solving these problems is obtained. The number of dimensions of corresponding system or equations is less than the classic Lagrangian multiplier method.

KEY WORDS optimality conditon, equality constraint, implicit function, Euclidean space, Lagrangian multiplier