阻尼最小二乘法中权的选择原理

王大麒　朱匀华　陈志恬
（数学力学系）

自从1944年 k.Levenberg[1]发表阻尼最小二乘法后，有许多应用和改进形式，但其效果取决于权因子$w^{(k)}$、阻尼系数$q^{(k)}$和阻尼因子s的选取。本文建立了一套"权"的选择原理，并导出一系列新的结果：(一)对$w^{(k)}$权。得到$w^{(k)}$的构造公式(2.7)，导出按精度比例齐步下降权(2.11)和(2.13)等，比目前普遍使用的权(2.9)更为优越。(二)对$q^{(k)}$权。得到$q^{(k)}$权的构造公式(2.18)，并导出敛速较快的$q^{(k)}$权(2.20)和(2.22)等，敛速快于目前普遍使用的$q^{(k)}$权(2.27)。(三)对s权。提出用逼近衡度代替 Wynne 线性度控制s的新方法，结果优于 Wynne 线性度。(四)对边界条件。提出用$q^{(k)}$权控制边界条件的新方法，克服了目前常用方法中存在的许多缺点。(五)§4〔定理4〕的(5)和〔定理5〕得到了阻尼最小二乘法得出的迭代点使加权函数值下降的构造性证法；并从构造性证明中诱导出加权函数下降值大小的控制方法。

本文提出的方法已很好地应用于光学和电网络自动设计中，最近又开始应用于化学方面的计算中，对收敛问题和敛速均获得了良好的效果。参加本文原理的程序试验和计算的还有林康有、汤爱珍、庞道初、鄢克、罗卫国、周展新、闫学玉等。

§1 阻尼最小二乘法及其核心问题

有一类最优化设计问题，可归结为求不等式组

$$-\varepsilon_i \leqslant F_i(x) \leqslant \varepsilon_i \quad (i=1,2,\cdots,m) \tag{1.1}$$

的解，其中$x \in D \subset R^n$，$\varepsilon_i \geqslant 0$为精度，$F_i(x)$为偏差函数。

(1.1)可用阻尼最小二乘法求其数值解。当得到第k个迭代点$x^{(k)}$以后，在D上定义**加权函数**（也称评价函数或价值函数）

$$\phi_k(x)=\sum_{i=1}^{m} w_i^{(k)} F_i^2(x), \tag{1.2}$$

并在R^n中定义**阻尼函数**

$$\Psi_k(x)=\sum_{i=1}^{m} w_i^{(k)} \widetilde{F}_i^2(x)+s\sum_{j=1}^{n} q_j^{(k)}\left(x_j-x_j^{(k)}\right)^2, \tag{1.3}$$

其中

$$\widetilde{F}_i(x)=F_i(x^{(k)})+\sum_{j=1}^{n}\frac{\partial F_i(x^{(k)})}{\partial x_j}(x_j-x_j^{(k)}) \quad (i=1,2,\cdots m), \tag{1.4}$$

常数$w_i^{(k)}\geqslant 0$称为**权因子**，常数$q_j^{(k)}>0$称为**阻尼系数**，常数 $s\geqslant 0$称为**阻尼因子**。因权因子$w^{(k)}=(w_1^{(k)}, w_2^{(k)},\cdots,w_m^{(k)})$，阻尼系数$q^{(k)}=(q_1^{(k)},q_2^{(k)},\cdots,q_n^{(k)})$和阻尼因子$s$均起权的作用，我们统称为“权”。以阻尼函数$\Psi_k(x)$的最小值点$x^{(k+1)}$作为紧接于$x^{(k)}$的次一个迭代点的迭代方法就叫**阻尼最小二乘法**。

1°　阻尼最小二乘法迭代程序计算公式

阻尼最小二乘法迭代程序计算公式可写成矩阵形式

$$X^{(k+1)}=X^{(k)}-[SQ+A^TWA]^{-1}A^TWF, \tag{1.5}$$

或

$$X^{(k+1)}=X^{(k)}-\frac{1}{s}\left[I+\frac{1}{s}Q^{-1}A^TWA\right]^{-1}Q^{-1}A^TWF, \tag{1.6}$$

其中

$$A=(a_{ij})_{m\times n},\quad a_{ij}=\frac{\partial F_i(x^{(k)})}{\partial x_j}\begin{pmatrix} i=1,2,\cdots,m\\ j=1,2,\cdots,n\end{pmatrix},$$

$$W=\begin{pmatrix} w_1^{(k)} & & \bigcirc\\ & w_2^{(k)} \ddots & \\ \bigcirc & & w_m^{(k)}\end{pmatrix}_{m\times m},\quad Q=\begin{pmatrix} q_1^{(k)} & & \bigcirc\\ & q_2^{(k)} \ddots & \\ \bigcirc & & q_n^{(k)}\end{pmatrix}_{n\times n},$$

$$X^{(k+1)}=\begin{pmatrix} x_1^{(k+1)}\\ x_2^{(k+1)}\\ \vdots\\ x_n^{(k+1)}\end{pmatrix}_{n\times 1},\quad X^{(k)}=\begin{pmatrix} x_1^{(k)}\\ x_2^{(k)}\\ \vdots\\ x_n^{(k)}\end{pmatrix}_{n\times 1},\quad F=\begin{pmatrix} F_1(x^{(k)})\\ F_2(x^{(k)})\\ \vdots\\ F_m(x^{(k)})\end{pmatrix}_{m\times 1},$$

A^T是A的转置矩阵，Q^{-1}是Q的逆矩阵。

2°　阻尼最小二乘法核心问题

由迭代公式(1.5)(或(1.6))可知，迭代点$x^{(k+1)}$是依赖于权$w^{(k)}$、$q^{(k)}$和s的，它们一起决定了$x^{(k+1)}$的位置。用阻尼最小二乘法得到的迭代点列$\{x^{(k)}\}$能否收敛于(1.1)的解，收敛速度如何，就取决于这些权的选取。因此权的选取问题是阻尼最小二乘法的核心问题。

§2　权的构造原理与构造方法

一、$w^{(k)}$权

$w^{(k)}$权决定$\phi_k(x)$在点$x^{(k)}$邻域的“形状”和$\phi_k(x)$在点$x^{(k)}$的下降方向。我们利用它来控制各$F_i^2(x)$从$x^{(k)}$到下一个迭代点的下降值的大小。

1°. $w^{(k)}$权的构造公式

由§4〔定理4〕的(5)知当$s\in[s_4,+\infty)$时，紧接于$x^{(k)}$的次一个迭代点$x_{(s)}^{(k+1)}$位于$\phi_k(x)$在点$x^{(k)}$的下降方向上，§4〔定理5〕又告诉我们，当$s\in[s_5,+\infty)$时，$\phi_k(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$有该定理所估计的下降值。因此，$F_i^2(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值大小，就取决于$F_i^2(x)$在点$x^{(k)}$的下降方向与$\phi_k(x)$在点$x^{(k)}$的下降方向相同程度如何和$F_i^2(x)$与$\phi_k(x)$在点$x^{(k)}$邻域的函数超曲面相象的程度如何。因$F_i^2(x)$与$\phi_k(x)$在点$x^{(k)}$的下降方向相同程度可用夹角

$$\theta_{ki}=\angle\left(-grad\phi_k(x^{(k)}),\ -grad w_i^{(k)}F_i^2(x^{(k)})\right)$$
$$=\angle\left(grad\phi_k(x^{(k)}),grad w_i^{(k)}F_i^2(x^{(k)})\right)\tag{2.1}$$

的大小作主要特征，一般θ_{ki}越小相同程度越好。但

$$grad\phi_k(x^{(k)})=\sum_{i=1}^{m}w_i^{(k)}grad\,F_i^2(x^{(k)}),\tag{2.2}$$

因而由§4〔定理1〕有

$$\theta_{ki}\xrightarrow[\left[w_i^{(k)}\|grad F_i^2(x^{(k)})\|\left(\sum_{i=1}^{m}w_i^{(k)}\|grad\,F_i^2(x^{(k)})\|\right)^{-1}\right]\to 1]{}0\tag{2.3}$$

故量

$$u_{ki}=w_i^{(k)}\|grad\,F_i^2(x^{(k)})\|\left(\sum_{i=1}^{m}w_i^{(k)}\|grad\,F_i^2(x^{(k)})\|\right)^{-1}\tag{2.4}$$

可作为夹角θ_{ki}大小的主要特征，u_{ki}越大，一般θ_{ki}越小。因量u_{ki}同时也是$F_i^2(x)$与$\phi_k(x)$在点$x^{(k)}$邻域的函数超曲面相象程度的主要特征，u_{ki}越大，一般越相象。因此，量u_{ki}越大，一般地$F_i^2(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值越大。但

$$w_i^{(k)}\|grad\,F_i^2(x^{(k)})\|=2w_i^{(k)}|F_i(x^{(k)})|\,\|grad\,F_i(x^{(k)})\|$$
$$(i=1,2,\cdots,m),\tag{2.5}$$

故量

$$v_i^{(k)}=w_i^{(k)}|F_i(x^{(k)})|\,\|grad\,F_i(x^{(k)})\|\quad(i=1,2,\cdots,m)\tag{2.6}$$

是$F_i^2(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值大小的**主要标志量**。一般地，$v_i^{(k)}$大$F_i^2(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值就大。由(2.6)得到$w^{(k)}$权的构造公式

$$w_i^{(k)}=v_i^{(k)}\left[|F_i(x^{(k)})|\,\|grad\,F_i(x^{(k)})\|\right]^{-1}\quad(i=1,2,\cdots,m).\tag{2.7}$$

2^0　$w^{(k)}$权的构造方法

$w^{(k)}$权是按收敛方式的要求来构造的，因$v_i^{(k)}$是$F_i^2(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值大小的主要标志量，赋给标志量$v^{(k)}=(v_i^{(k)},v_2^{(k)}\cdots,v_m^{(k)})$的不同变化规律，就得到不同的收敛方式。例如

1）如果希望在迭代点函数值大的偏差函数下降快些，可取

$$v_i^{(k)}=|F_i(x^{(k)})|\quad(i=1,2,\cdots,m). \tag{2.8}$$

于是由(2.7)有

$$w_i^{(k)}=\|grad\,F_i(x^{(k)})\|^{-1}\quad(i=1,2,\cdots,m). \tag{2.9}$$

这是目前国内外普遍使用的一个权。

2）如果希望偏差函数在迭代点按精度比例齐步下降，可取

$$v_i^{(k)}=\varepsilon_i^{-1}|F_i(x^{(k)})|\quad(i=1,2,\cdots,m). \tag{2.10}$$

于是由(2.7)有

$$w_i^{(k)}=\left[\varepsilon_i\|grad\,F_i(x^{(k)})\|\right]^{-1}\quad(i=1,2,\cdots,m). \tag{2.11}$$

由(2.10)知这个权能促使各偏差函数$F_i^2(x)$在各迭代点的值朝着按精度比例大致相等的方向下降。我们在设计中普遍使用这个权。大量计算结果表明，这个权比(2.9)更为优越，是促使各偏差函数按精度比例齐步下降的权。

3）和2)类似的思想，可取

$$v_i^{(k)}=\left[\varepsilon_i^{-1}|F_i(x^{(k)})|\right]^{\alpha}\quad(i=1,2,\cdots,m), \tag{2.12}$$

使

$$w_i^{(k)}=\varepsilon_i^{-\alpha}|F_i(x^{(k)})|^{\alpha-1}\|gradF_i(x^{(k)})\|^{-1}\quad(i=1,2,\cdots,m), \tag{2.13}$$

其中$\alpha\geqslant0$，表示按精度比例的幂次。

二、$q^{(k)}$权

由§4〔定理2〕知，当$w^{(k)}$和$q^{(k)}$权固定以后，$x^{(k+1)}_{((s_0,+\infty))}$是一条连续超曲线，这条连续超曲线按该定理的(2),(3),(4)所描写的方式摆置在点$x^{(k)}$的邻域中．按照这种摆置，迭代点$x_{(s)}^{(k+1)}$的增量$\Delta x(s)=x_{(s)}^{(k+1)}-x^{(k)}$与向量

$$G=-\left(\frac{1}{q_1^{(k)}}\frac{\partial\phi_k(x^{(k)})}{\partial x_1},\quad\frac{1}{q_2^{(k)}}\frac{\partial\phi_k(x^{(k)})}{\partial x_2},\cdots,\right.$$
$$\left.\frac{1}{q_n^{(k)}}\frac{\partial\phi_k(x^{(k)})}{\partial x_n}\right) \tag{2.14}$$

的夹角

$$\angle(G,\Delta x(s))\xrightarrow[s\longrightarrow\infty]{}0. \tag{2.15}$$

这指明了迭代点$x_{(s)}^{(k+1)}$的近似方位。由此可知，当$w^{(k)}$权固定以后，$q^{(k)}$权可以用来控制连续超曲线$x_{([s_0,+\infty))}^{(k+1)}$在点$x^{(k)}$邻域中的摆置方式，进而控制在点$x^{(k)}$的迭代增量$\Delta x(s)$的方向。

1° $q^{(k)}$权的构造公式

若要求在点$x^{(k)}$的迭代增量的方向近似于

$$\Delta x^{(k)}=\left(\Delta x_1^{(k)},\Delta x_2^{(k)},\cdots,\Delta x_n^{(k)}\right), \tag{2.16}$$

则由(2.15)式知可令

$$\Delta x^{(k)}=G. \tag{2.17}$$

从而得到$q^{(k)}$权的构造公式为

$$q_j^{(k)}=-\left(\Delta x_j^{(k)}\right)^{-1}\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\quad(j=1,2,\cdots,n). \tag{2.18}$$

2° $q^{(k)}$权的构造方法

$q^{(k)}$权的构造就是按照要求设计迭代增量的近似方向$\Delta x^{(k)}$。例如

1) 对灵敏参数进行“阻尼”的$q^{(k)}$权设计

一般地，在迭代过程中越灵敏的参数改变量越大，这将导致灵敏参数易超界，或不易得到好的逼近解。通常用“阻尼”的思想来解决这一问题，阻挡迭代点迅速逼近灵敏参数的边界。据此，迭代增量近似方向可设计为

$$\Delta x_j^{(k)}=sgn\left[-\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\right]\left|\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\right|^{-\alpha}\ (j=1,2,\cdots,n), \tag{2.19}$$

其中$\alpha>0$，它使迭代过程中越灵敏的参数变化越小，实现“阻尼”的思想。于是由(2.18)有

$$q_j^{(k)}=\left|\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\right|^{1+\alpha}\quad(j=1,2,\cdots,n). \tag{2.20}$$

2) 快速下降的$q^{(k)}$权设计

由(2.16)和(2.18)知在点$x^{(k)}$的增量的近似方向取为

$$\Delta x_j^{(k)}=sgn\left[-\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\right]\left|\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\right|^{\alpha}\quad(j=1,2,\cdots,n), \tag{2.21}$$

则
$$q_j^{(k)}=\left|\frac{\partial\phi_k(x^{(k)})}{\partial x_j}\right|^{1-\alpha}\quad(j=1,2,\cdots,n), \tag{2.22}$$

其中$\alpha>0$。而由后面(2.30)式知，当$\theta_k(x_{(s)}^{(k+1)})$的值固定以后，加权函数$\phi_k(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值就由值

$$-\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}\left\|\Delta x(s)\right\| \tag{2.23}$$

决定。因此，对(2.22)选取适当的α，让它给出的迭代点使(2.23)有较大的值，则这个$q^{(k)}$权就能促使加权函数快速下降，但此权使灵敏参数容易出界，要与边界条件结合在

一起使用，对边界条件严加控制。

3）与阻挡出界相结合的混合型$q^{(k)}$权设计

当边界条件表现为

$$a_j \leqslant x_j \leqslant b_j \quad (j=1,2,\cdots,n) \tag{2.24}$$

时，则由(2.16)和(2.18)知权

$$q_j^{(k)}=\begin{cases}\left(x_j^{(k)}-a_j\right)^{-\lambda}\left|\dfrac{\partial\phi_k(x^{(k)})}{\partial x_j}\right| & \text{当}a_j<x_j^{(k)}<a_j+\mu_j\text{时，}\\ \left|\dfrac{\partial\phi_k(x^{(k)})}{\partial x_j}\right|^{1+\alpha} & \text{当}a_j+\mu_j\leqslant x_j\leqslant b_j-\mu_j\text{时，}\\ \left(b_j-x_j^{(k)}\right)^{-\lambda}\left|\dfrac{\partial\phi_k(x^{(k)})}{\partial x_j}\right| & \text{当}b_j-\mu_j<x_j<b_j\text{时，}\\ & (j=1,2,\cdots,n),\end{cases} \tag{2.25}$$

是阻尼与阻挡出界相结合的混合型$q^{(k)}$权，其中 λ,α,μ_j 为正的经验常数，μ_j不超过$\dfrac{b_j-a_j}{2}$，而权

$$q_j^{(k)}=\begin{cases}\left(x_j^{(k)}-a_j\right)^{-\lambda}\left|\dfrac{\partial\phi_k(x^{(k)})}{\partial x_j}\right| & \text{当}a_j<x_j^{(k)}<a_j+\mu_j\text{时，}\\ \left|\dfrac{\partial\phi_k(x^{(k)})}{\partial x_j}\right|^{1-\alpha} & \text{当}a_j+\mu_j\leqslant x_j^{(k)}\leqslant b_j-\mu_j\text{时，}\\ \left(b_j-x_j^{(k)}\right)^{-\lambda}\left|\dfrac{\partial\phi_k(x^{(k)})}{\partial x_j}\right| & \text{当}b_j-\mu_j<x_j^{(k)}<b_j\text{时}\\ & (j=1,2,\cdots,n),\end{cases} \tag{2.26}$$

是快速下降与阻挡出界相结合的混合型$q^{(k)}$权，其中λ,μ_j与(2.25)同，α是正的经验常数。

计算实践表明，(2.20)的α取$0\leqslant\alpha\leqslant1$，(2.22)的$\alpha$取为$1.1\leqslant\alpha\leqslant1.6$，这些权与(2.11)的$w^{(k)}$权相配合能使偏差函数得到较佳的齐步下降速度。特别是(2.22)更为突出，它比目前普遍使用的阻尼系数

$$q_j^{(k)}=\sum_{i=1}^{m}w_i^{(k)}a_{ij}^2 \quad (j=1,2,\cdots,n) \tag{2.27}$$

更为优越，获得更快的敛速。

在实际计算中通常把$q_j^{(k)}$分解为

$$q_j^{(k)}=q_0^{(k)}\ q_{1j}^{(k)}\ q_{2j}^{(k)} \quad (j=1,2,\cdots n) \tag{2.28}$$

其中$q_{2j}^{(k)}$是自动权，取(2.26)，(2.25)，(2.22)，(2.20)等形式，$q_{1j}^{(k)}$作人工权和特殊用途使用，$q_0^{(k)}$是规格化常数．

三、S 权

由§4〔定理2〕知s的作用是控制迭代点$x_{(s)}^{(k+1)}$位于连续超曲线$x_{([s_0,+\infty))}^{(k+1)}$的什么位置上，使$\phi_k(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$获得较大的下降值。

国内外普遍使用W_{ynne}线性度来控制s，但此方法有时不能获得较佳的敛速．对此我们提出用逼近衡度来控制s。

函数

$$\theta_k(x)=[\phi_k(x)-\widetilde{\phi}_k(x)][\phi_k(x^{(k)})-\widetilde{\phi}_k(x)]^{-1} \tag{2.29}$$

称为$\phi_k(x)$的逼近衡度函数，$\widetilde{\phi}_k(x)$是$\phi_k(x)$在点$x^{(k)}$的台劳展开的线性部分．$\theta_k(x)$可用来衡量$\widetilde{\phi}_k(x)$对$\phi_k(x)$的逼近程度，$\theta_k(x)$越接近于0，表示$\widetilde{\phi}_k(x)$越接近于$\phi_k(x)$；当$\theta_k(x)=0$时，就有$\widetilde{\phi}_k(x)=\phi_k(x)$．因此，由于$\widetilde{\phi}_k(x)$是一次函数，就可以很简单地通过$\theta_k(x)$来了解$\phi_k(x)$的性态。

特别地，由于(2.29)可等价地写成

$$\phi_k(x^{(k)})-\phi_k(x)=[1-\theta_k(x)]\left[-\frac{\partial\phi_k(x^{(k)})}{\partial(x-x^{(k)})}\right]\left\|x-x^{(k)}\right\|, \tag{2.30}$$

其中$\dfrac{\partial\phi_k(x^{(k)})}{\partial(x-x^{(k)})}$是方向导数．故可通过$\theta_k(x(t))$来计算$\phi_k(x)$从$x^{(k)}$到$\phi_k(x)$在点$x^{(k)}$的下降方向射线

$$L_x=\left\{x(t)\,|\,x(t)=x^{(k)}+t(x-x^{(k)}),0\leqslant t<+\infty,\frac{\partial\phi_k(x^{(k)})}{\partial(x-x^{(k)})}<0\right\} \tag{2.31}$$

上的任一点$x(t)$的下降值$[\phi_k(x^{(k)})-\phi_k(x(t))]$。不仅如此，还可以由$\theta_k(x)$的定义和(2.30)推出：

1°　若存在点$x^{(k)}$的某一凸邻域使$\phi_k(x)$在其上是凸函数时，则必存在区间

$$[\theta_0,\theta^0]\qquad(0<\theta_0<\theta^0<1), \tag{2.32}$$

使$\theta_k(x)$在任一(2.31)的射线L_x上的值，只要满足

$$\theta_0\leqslant\theta_k(x(t_0))\leqslant\theta^0$$

$x(t_0)$就是$\phi_k(x)$在L_x靠近$x^{(k)}$的一段上函数值较小的点。

2°　若存在点$x^{(k)}$的某一凸邻域使$\phi_k(x)$在其上是凹函数时，则必存在区间

$$(-\infty,\theta_-]\qquad(\theta_-<0), \tag{2.33}$$

使$\theta_k(x)$在任一(2.31)的射线L_x上的值只要满足

$$\theta_k(x(t_0))\leqslant\theta_-,$$

$x(t_0)$就是$\phi_k(x)$在L_x靠近$x^{(k)}$的一段上函数值较小的点。

而由§4〔定理4〕的(5)和〔定理2〕的(1)知连续超曲线$x^{(k+1)}_{([s_0,+\infty))}$是被包含在$\phi_k(x)$在点$x^{(k)}$的下降方向区域中的．故只要$\theta_k(x^{(k+1)}_{(s)})$落入(2.32)(或(2.33))的区间，$\phi_k(x)$从$x^{(k)}$到$x^{(k+1)}_{(s)}$就有较大的下降值。因此，我们把(2.32)和(2.33)称为较优敛速区间。

这就导出了由控制$\theta_k(x^{(k+1)}_{(s)})$落入较优敛速区间而使$\phi_k(x)$从$x^{(k)}$到$x^{(k+1)}_{(s)}$有较大下降值的控制方法。这就是$s$权的选择准则。

§4〔定理4〕的(4)和$\theta_k(x)$在2°情形中函数值分布状态导出了如何增减s控制$\theta_k(x_{(s)}^{(k+1)})$落入较优敛速区间的控制方法。即(1)当$\theta_k(x_{(s)}^{(k+1)})>\theta^0$时，增大$s$；(2)当$0<\theta_k(x_{(s)}^{(k+1)})<\theta_0$时，缩小$s$；(3)当$\theta_-<\theta_k(x_{(s)}^{(k+1)})\leqslant 0$时，先增大$s$，直至出现$\theta_k$的值反而上升时，就改为缩小$s$.

经实践验证，$\theta_-,\theta_0,\theta^0$取如下值

$$-0.5\leqslant\theta_-\leqslant-0.3,\quad 0.4\leqslant\theta_0\leqslant 0.45,\quad 0.55\leqslant\theta^0\leqslant 0.6 \tag{2.34}$$

时有较优的敛速。

大量计算实践证明，用逼近衡度来控制s，比Wynne线性度能获得更佳的收敛速度。

§3 用$q^{(k)}$权控制边界条件的方法

我们根据$q^{(k)}$权的作用，提出专门用(2.28)中的$q_{1j}^{(k)}$来控制边界条件表现为(2.24)的控制方法。即令

$$q_{1j}^{(k)}=\begin{cases} c_j(x_j^{(k)}-a_j)^{-\lambda} & \text{当 } a_j<x_j^{(k)}<a_j+\mu_j\text{时}, \\ c_j & \text{当 } a_j+\mu_j\leqslant x_j^{(k)}\leqslant b_j-\mu_j\text{时}, \\ c_j(b_j-x_j^{(k)})^{-\lambda} & \text{当 } b_j-\mu_j<x_j^{(x)}<b_j\text{时}, \end{cases} \tag{3.1}$$
$$(j=1,2,\cdots,n),$$

其中c_j,λ,μ_j,均为正的经验常数.而(2.28)中的$q_{2j}^{(k)}$取(2.22)(或(2.20))等形式。由(2.16)和(2.18)知，这时当$x_j^{(k)}$越靠近边界时迭代增量$|\Delta x_j(s)|$越小，起到阻挡迭代点出界的作用，使超界问题得到控制。而$x_j^{(k)}$离边界远者其增量受$q_{2j}^{(k)}$权所控制。

计算经验得出当$q_{1j}^{(k)}$取(3.1)时，可取$3\leqslant\lambda\leqslant 5$，而$c_j$用来使各参数变化时保持大致相同的比例。

(3.1)已包括了单边约束条件，这只要从(3.1)中取$a_j=-\infty$(或$b_j=+\infty$)就可得到。

〔注〕 权(3.1)和(2.25)及(2.26)的处理思想相同，但处理方法有差异，因而常数λ也有差异。

§4 主要定理证明

这一节将证明前面论述的权的构造原理与构造方法中所用到的若干主要定理：

〔定理1〕 设$e_1,e_2,\cdots e_m$为R^n中的m个向量，变数$w_i\geqslant 0(i=1,2,\cdots,m)$，$\sigma=\sum\limits_{i=1}^{m}w_ie_i$，则向量$\sigma$与$w_{i_0}e_{i_0}$的夹角

$$\angle(\sigma, w_{i_0}e_{i_0}) \xrightarrow[\left[w_{i_0}\|e_{i_0}\|\left(\sum\limits_{i=1}^{m} w_i\|e_i\|\right)^{-1}\right] \longrightarrow 1]{} 0,$$

其中$i_0 \in \{1,2,\cdots,m\}$。

证 因为

$$\left[w_{i_0}\|e_{i_0}\|\left(\sum_{i=1}^{m} w_i\|e_i\|\right)^{-1}\right] \to 1 \Longleftrightarrow \left[(w_i\|e_i\|)(w_{i_0}\|e_{i_0}\|)^{-1}\right] \to 0 \quad (i \neq i_0),$$

而

$$\cos_\angle(\sigma, w_{i_0}e_{i_0}) = \cos_\angle(\sigma, e_{i_0}) = \frac{(\sigma, e_{i_0})}{\|e_{i_0}\|\|\sigma\|}$$

$$= \frac{1 + \sum\limits_{i \neq i_0} \dfrac{w_i\|e_i\|}{w_{i_0}\|e_{i_0}\|} \cos_\angle(e_i, e_{i_0})}{\left[1 + \sum\limits_{i,i'}{}' \left(\dfrac{w_i\|e_i\|}{w_{i_0}\|e_{i_0}\|}\right)\left(\dfrac{w_{i'}\|e_{i'}\|}{w_{i_0}\|e_{i_0}\|} \cos_\angle(e_i, e_{i'})\right)\right]^{1/2}}$$

其中$\sum\limits_{i,i'}{}'$表i和i'不同时等于i_0的那些项之和。从而有

$$\cos_\angle(\sigma, w_{i_0}e_{i_0}) \xrightarrow[\left[w_{i_0}\|e_{i_0}\|\left(\sum\limits_{i=1}^{m} w_i\|e_i\|\right)^{-1}\right] \longrightarrow 1]{} 1.$$

故夹角

$$\angle(\sigma, w_{i_0}e_{i_0}) \xrightarrow[\left[w_{i_0}\|e_{i_0}\|\left(\sum\limits_{i=1}^{m} w_i\|e_i\|\right)^{-1}\right] \longrightarrow 1]{} 0.$$

〔引理 1〕

$$grad\phi_k(x^{(k)}) = 2(A^T W F)^T.$$

这由

$$\frac{\partial\phi_k(x^{(k)})}{\partial x_j} = 2\sum_{i=1}^{m} w_i^{(k)} F_i(x^{(k)}) \frac{\partial F_i(x^{(k)})}{\partial x_j} \quad (j = 1,2,\cdots,n)$$

直接推得。

〔定理 2〕 设$F_i(x)(i=1,2,\cdots,m)$定义在$D(\subset R^n)$上，且在D中有一阶连续偏导数，$x^{(k)} \in D^0$（D^0是D的全部内点的集合），常数$q_j^{(k)} > 0(j=1,2\cdots,n)$，常数$w_i^{(k)} \geqslant 0(i=1,2,\cdots,m)$，$x_{(s)}^{(k+1)}$是阻尼函数(1.3)的最小值点。

$$\Delta x(s) = x_{(s)}^{(k+1)} - x^{(k)},$$

$$G = -\left(\frac{1}{q_1^{(k)}} \frac{\partial\phi_k(x^{(k)})}{\partial x_1}, \frac{1}{q_2^{(k)}} \frac{\partial\phi_k(x^{(k)})}{\partial x_2},\right.$$

$$\cdots \frac{1}{q_n^{(k)}} \quad \frac{\partial \phi_k(x^{(k)})}{\partial x_n} \Bigg),$$

则

（1）$\exists s_0>0$，使$x_{(s)}^{(k+1)}$是从$[s_0, +\infty)$到R^n中的连续向量函数；

（2）$x_{(s)}^{(k+1)} \xrightarrow[s\longrightarrow\infty]{} x^{(k)}$；

（3）$s\Delta x(s) \xrightarrow[s\longrightarrow\infty]{} -(Q^{-1}A^TWF)^T = \frac{1}{2}G$；

（4）向量G与$\Delta x(s)$的夹角

$$\angle(G, \Delta x(s)) \xrightarrow[s\longrightarrow\infty]{} 0.$$

证　(1),(2)和(3)由(1.6)和〔引理1〕直接推出。

现证(4)　令

$$H = Q^{-1}A^TWF, \tag{4.1}$$

由〔引理1〕有

$$G = -[Q^{-1}(grad\phi_k(x^{(k)}))^T]^T = -2[Q^{-1}A^TWF]^T = -2H^T, \tag{4.2}$$

从而

$$\|G\| = [G^TG]^{1/2} = 2[H^TH]^{1/2}. \tag{4.3}$$

而由(1.6)，有

$$\|\Delta x(s)\| = [(\Delta X(s))^T(\Delta X(s))]^{1/2}$$
$$= \frac{1}{s}\left\{\left[\left(I+\frac{1}{s}Q^{-1}A^TWA\right)^{-1}H\right]^T\left[\left(I+\frac{1}{s}Q^{-1}A^TWA\right)^{-1}H\right]\right\}^{1/2}. \tag{4.4}$$

但内积

$$(G, \Delta x(s)) = G[\Delta X(s)]$$
$$= \frac{2}{s}\left\{ H^T\left[I+\frac{1}{s}Q^{-1}A^TWA\right]^{-1}H\right\}. \tag{4.5}$$

从而当$\|G\|$和$\|\Delta x(s)\|$均不为0时，由(4.5)、(4.4)和(4.3)，有

$$Cos_{\angle}(G, \Delta x(s)) = \frac{(G, \Delta x(s))}{\|G\|\,\|\Delta x(s)\|}$$

$$= \frac{H^T\left[I+\frac{1}{s}Q^{-1}A^TWA\right]^{-1}H}{[H^TH]^{1/2}\left\{\left[\left(I+\frac{1}{s}Q^{-1}A^TWA\right)^{-1}H\right]^T\left[\left(I+\frac{1}{s}Q^{-1}A^TWA\right)^{-1}H\right]\right\}^{1/2}}$$

$$\xrightarrow[s\longrightarrow\infty]{} \frac{H^TH}{[H^TH]^{1/2}[H^TH]^{1/2}} = 1.$$

因而夹角

$$\angle(G, \Delta x(s)) \xrightarrow[s\longrightarrow\infty]{} 0.$$

而其余情况上式自然成立。

〔**引理2**〕 设 $w_i^{(k)}(i=1,2,\cdots,m)$, $q_j^{(k)}>0(j=1,2,\cdots,n)$, $s>0$ 使(1.5)的逆矩阵 $(sQ+A^TWA)^{-1}$存在, $\|grad\phi_k(x^{(k)})\|\neq 0$, $x^{(k+1)}$是阻尼函数$\Psi_k(x)$的最小值点, $\Delta x=x^{(k+1)}-x^{(k)}$, $F_i(x)$ $(i=1,2,\cdots,m)$在D中有一阶连续偏导数, $x^{(k)}\epsilon D°$, 则

(1) $\Psi_k(x)$有唯一的最小值点$x^{(k+1)}$, 且$x^{(k+1)}\neq x^{(k)}$;

(2) $$\sum_{i=1}^m w_i^{(k)}\left[\sum_{j=1}^n \frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j\right]^2\neq 0.$$

证 因$\|grad\phi_k(x^{(k)})\|\neq 0$,由〔引理1〕和(1.5)直接推得(1).

现用反证法证(2).因若(2)不成立, 有

$$\sum_{i=1}^m w_i^{(k)}\left[\sum_{j=1}^n \frac{\partial F_i(x^{(k)})}{\partial x_i}\Delta x_j\right]^2=0,$$

则必有

$$w_i^{(k)}\left[\sum_{j=1}^n \frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j\right]=0,\quad (i=1,2,\cdots,m).$$

再注意到$x^{(k+1)}\neq x^{(k)}$,则知阻尼函数在点$x^{(k+1)}$的值

$$\begin{aligned}\Psi_k(x^{(k+1)}) &= \sum_{i=1}^m w_i^{(k)}\widetilde{F}_i^2(x^{(k+1)})+s\sum_{j=1}^n q_j^{(k)}\Delta x_j^2\\ &=\sum_{i=1}^m w_i^{(k)}F_i^2(x^{(k)})+2\sum_{i=1}^m w_i^{(k)}F_i(x^{(k)})\left[\sum_{j=1}^n \frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j\right]+\\ &+\sum_{i=1}^m w_i^{(k)}\left[\sum_{j=1}^n \frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j\right]^2+s\sum_{j=1}^n q_j^{(k)}\Delta x_j^2\\ &=\phi_k(x^{(k)})+s\sum_{j=1}^n q_j^{(k)}\Delta x_j^2>\phi_k(x^{(k)})=\Psi_k(x^{(k)}).\end{aligned}$$

这与$x^{(k+1)}$是$\Psi_k(x)$的最小值点的结论矛盾. 证毕.

〔**定理3**〕 在〔引理2〕的假设下, 再令

$$\widetilde{\phi}_k(x)=\phi_k(x^{(k)})+\sum_{j=1}^n \frac{\partial\phi_k(x^{(k)})}{\partial x_j}(x_i-x_j^{(k)}),$$

$$\approx\!\!\!\!\phi_k(x)=\sum_{i=1}^m w_i^{(k)}\widetilde{F}_i^2(x),$$

则 (1) $\widetilde{\phi}_k(x^{(k+1)})<\overset{\approx}{\phi}_k(x^{(k+1)})<\Psi_k(x^{(k+1)})<\phi_k(x^{(k)})$;

(2) 在Δx方向上的方向导数

$$\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x}<-\left[\sum_{i=1}^m w_i^{(k)}\left(\frac{\partial F_i(x^{(k)})}{\partial\Delta x}\right)^2+s\sum_{j=1}^n q_j^{(k)}\cos^2\alpha_j\right]\|\Delta x\|$$

$$\leqslant -\left[\sum_{i=1}^{m} w_i^{(k)}\left(\frac{\partial F_i(x^{(k)})}{\partial \Delta x}\right)^2 + sq_0\right]\|\Delta x\| < -sq_0\|\Delta x\|,$$

其中$\alpha_1,\alpha_2\cdots,\alpha_n$是$\Delta x$的方向角，$q_0 = min\left\{q_1^{(k)},q_2^{(k)},\cdots,q_n^{(k)}\right\}$.

证　因

$$\widetilde{F}_i^2(x^{(k+1)}) = F_i^2(x^{(k)}) + 2F_i(x^{(k)})\sum_{j=1}^{n}\frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j +$$

$$+\left[\sum_{j=1}^{n}\frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j\right]^2 \quad (i=1,2,\cdots,m),$$

$$\phi_k(x) = \sum_{i=1}^{m} w_i^{(k)} F_i^2(x),$$

直接推得

$$\approx\!\!\!\!\phi_k(x^{(k+1)}) = \widetilde{\phi}_k(x^{(k+1)}) + \sum_{i=1}^{m} w_i^{(k)}\left[\sum_{j=1}^{n}\frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j\right]^2. \tag{4.6}$$

从而由〔引理2〕，并注意到$\Psi_k(x^{(k)}) = \phi_k(x^{(k)})$，则由$\widetilde{\phi}_k(x)$，$\approx\!\!\!\!\phi_k(x)$和$\Psi_k^{(x)}$的定义，就得到(1).

现证(2)因

$$\widetilde{\phi}_k(x^{(k+1)}) = \phi_k(x^{(k)}) + \sum_{j=1}^{n}\frac{\partial \phi_k(x^{(k)})}{\partial x_j}\Delta x_j$$

$$= \phi_k(x^{(k)}) + \frac{\partial \phi_k(x^{(k)})}{\partial \Delta x}\|\Delta x\|,$$

$$\sum_{j=1}^{m}\frac{\partial F_i(x^{(k)})}{\partial x_j}\Delta x_j = \frac{\partial F_i(x^{(k)})}{\partial \Delta x}\|\Delta x\| \quad (i=1,2,\cdots,m),$$

又由(4.6)和(1)，有

$$\phi_k(x^{(k)}) + \frac{\partial \phi_k(x^{(k)})}{\partial \Delta x}\|\Delta x\| + \sum_{i=1}^{m} w_i^{(k)}\left[\frac{\partial F_i(x^{(k)})}{\partial \Delta x}\|\Delta x\|\right]^2 +$$

$$+\sum_{j=1}^{n} q_j^{(k)}\Delta x_j^2 = \approx\!\!\!\!\phi_k(x^{(k+1)}) + s\sum_{j=1}^{n} q_j^{(k)}\Delta x_j^2$$

$$= \Psi_k(x^{(k+1)}) < \phi_k(x^{(k)}),$$

从而

$$\frac{\partial \phi_k(x^{(k)})}{\partial \Delta x} < -\left[\sum_{i=1}^{m} w_i^{(k)}\left(\frac{\partial F_i(x^{(k)})}{\partial \Delta x}\right)^2 + s\sum_{j=1}^{n} q_j^{(k)}\left(\frac{\Delta x_j}{\|\Delta x\|}\right)^2\right]\|\Delta x\|$$

$$= -\left[\sum_{i=1}^{m} w_i^{(k)}\left(\frac{\partial F_i(x^{(k)})}{\partial \Delta x}\right)^2 + s\sum_{j=1}^{n} q_j^{(k)}\cos^2\alpha_j\right]\|\Delta x\|.$$

(2)得证.

〔**定理4**〕 在〔定理2〕的假设下，又设

$$0<q_0\leqslant q_j^{(k)}\leqslant q^0\ (j=1,2,\cdots,n),\quad \|grad\phi_k(x^{(k)})\|=\varepsilon_0>0, \tag{4.7}$$

$$\theta_k(x)=[\phi_k(x)-\widetilde{\phi_k}(x)]\,[\phi_k(x^{(k)})-\widetilde{\phi_k}(x)]^{-1}, \tag{4.8}$$

$\widetilde{\phi_k}(x)$是$\phi_k(x)$在点$x^{(k)}$的台劳展开的线性部分，

则 （1） $\phi_k(x_{(s)}^{(k+1)})<\phi_k(x^{(k)})\Longleftrightarrow\theta_k(x_{(s)}^{(k+1)})<1$；

（2） $\theta_k(x)$在集

$$D_1=\left\{x\,|\,x\in[D-\{x^{(k)}\}],\ \text{方向导数}\left|\frac{\partial\phi_k(x^{(k)})}{\partial(x-x^{(k)})}\right|\geqslant l_1>0\right\}$$

中连续，且

$$\theta_k(x)\xrightarrow[x\to x^{(k)}]{}0\quad(x\in D_1),$$

并∃$\delta>0$，使$x\in D_1\cap S(x^{(k)},\delta)$时，有

$$\theta_k(x)<1;$$

（3） ∃$s_1\geqslant0$，使$s\in[s_1,+\infty)$时，有方向导数

$$\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}<-\frac{\varepsilon_0q_0}{4q^0};$$

（4） ∃$s_2\geqslant0$，使$\theta_k^*(s)\equiv\theta_k(x_{(s)}^{(k+1)})$在$[s_2,+\infty)$中连续，且

$$\theta_k^*(s)\xrightarrow[s\to\infty]{}0,$$

又∃$s_3\geqslant0$，使$s\in[s_3,+\infty)$时有

$$\theta_k^*(s)<1;$$

（5） ∃$s_4\geqslant0$，使$s\in[s_4,+\infty)$时，有

$$\phi_k(x_{(s)}^{(k+1)})<\phi_k(x^{(k)}),$$

且 $$\phi_k(x(t))<\phi_k(x^{(k)}),$$

其中 $$x(t)=x^{(k)}+t\left[x_{(s)}^{(k+1)}-x^{(k)}\right]\quad(0<t\leqslant1).$$

证（1） 因由〔定理3〕的（1）有$\phi_k(x^{(k)})-\widetilde{\phi_k}(x_{(s)}^{(k+1)})>0$，故由(2.30)知，有

$$\phi_k(x_{(s)}^{(k+1)})<\phi_k(x^{(k)})\Longleftrightarrow\theta_k(x_{(s)}^{(k+1)})<1.$$

证（2） 因$F_i(x)(i=1,2,\cdots,m)$，从而是$\phi_k(x)$在D中一阶偏导数连续，故$\phi_k(x)$在D中连续；而$[\phi_k(x^{(k)})-\widetilde{\phi_k}(x)]$显然在$D_1$中连续且不等于0；故由$\theta_k(x)$的定义知$\theta_k(x)$在$D_1$中连续。又由$\phi_k(x)$在$D$中一阶偏导数连续，且$x^{(k)}\in D^\circ$，推出$\phi_k(x)$在点$x^{(k)}$可微。且

$$\phi_k(x^{(k)})-\widetilde{\phi}_k(x)=-\frac{\partial\phi_k(x^{(k)})}{\partial(x-x^{(k)})}\left\|x-x^{(k)}\right\|.$$

这样，由$x\in D_1$和$\theta_k(x)$的定义，就有

$$\left|\theta_k(x)\right|=\left|\varepsilon\left\|x-x^{(k)}\right\|\left[-\frac{\partial\phi_k(x^{(k)})}{\partial(x-x^{(k)})}\left\|x-x^{(k)}\right\|\right]^{-1}\right|$$

$$\leqslant|\varepsilon|\,l_1^{-1}\xrightarrow[\|x-x^{(k)}\|\longrightarrow 0]{}0. \tag{4.9}$$

从而$\exists\ \delta>0$，使$x\in D_1\cap S(x^{(k)},\delta)$时，有

$$\theta_k(x)<1.$$

证(3)　由〔定理2〕知当$s\to\infty$时，$s\Delta x(s)\to\frac{1}{2}G$，故$\exists s_1\geqslant 0$，使$s\in[s_1,+\infty)$时，有

$$\left\|\frac{1}{2}G-s\Delta x(s)\right\|\leqslant\frac{\varepsilon_0}{4q^0}.$$

又
$$\frac{1}{2}\|G\|-s\|\Delta x(s)\|\leqslant\left\|\frac{1}{2}G-s\Delta x(s)\right\|\leqslant\frac{\varepsilon_0}{4q^0},$$

故
$$-s\|\Delta x(s)\|\leqslant\frac{\varepsilon_0}{4q^0}-\frac{1}{2}\|G\|\leqslant\frac{\varepsilon_0}{4q^0}-\frac{1}{2q^0}\|grad\phi_k(x^{(k)})\|$$

$$=\frac{\varepsilon_0}{4q^0}-\frac{\varepsilon_0}{2q^0}=-\frac{\varepsilon_0}{4q^0}. \tag{4.10}$$

这样，当$s\in[s_1,+\infty)$时，根据〔定理3〕的(2)就有

$$\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}<-sq_0\|\Delta x(s)\|\leqslant-\frac{\varepsilon_0 q_0}{4q^0}.$$

证(4)　由〔定理2〕的(1)和(2)，本定理的(2)、(3)和(4.9)，可直接推得(4)成立。

证(5)　由〔定理2〕的(2)，知必存在$s_4\geqslant 0$，使$s\in[s_4,+\infty)$时，有

$$\left\|x_{(s)}^{(k+1)}-x^{(k)}\right\|<\delta\ 和\ \frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}<-\frac{\varepsilon_0 q_0}{4q^0}.$$

取$l_1=\frac{\varepsilon_0 q_0}{4q^0}$，则由(2)有

$$\theta_k(x_{(s)}^{(k+1)})<1\ (s\in[s_4,+\infty)).$$

又由(1)有

$$\phi_k(x_{(s)}^{(k+1)})<\phi_k(x^{(k)})\ (s\in[s_4,+\infty)).$$

而当$s\in[s_4,+\infty)$时，由

$$x(t)=x^{(k)}+t\left[x_{(s)}^{(k+1)}-x^{(k)}\right]\qquad(0<t\leqslant 1),$$

有
$$\left\|x(t)-x^{(k)}\right\|=t\left\|x_{(s)}^{(k+1)}-x^{(k)}\right\|\leqslant\left\|x_{(s)}^{(k+1)}-x^{(k)}\right\|<\delta$$

和
$$\frac{\partial\phi_k(x^{(k)})}{\partial(x(t)-x^{(k)})}=\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}<-l_1.$$

从而由(2)有

$$\theta_k(x(t))<1 \qquad (0<t\leqslant 1). \tag{4.11}$$

又由〔定理3〕的(1)，有

$$\sum_{j=1}^{n}\frac{\partial\phi_k(x^{(k)})}{\partial xj}\left[x_{j(s)}^{(k+1)}-x_j^{(k)}\right]<0.$$

因而当$0<t\leqslant 1$时，有

$$\sum_{j=1}^{n}\frac{\partial\phi_k(x^{(k)})}{\partial xj}\left[x_j(t)-x_j^{(k)}\right]=t\sum_{j=1}^{n}\frac{\partial\phi_k(x^{(k)})}{\partial xj}\left[x_{j(s)}^{(k+1)}-x_j^{(k)}\right]<0.$$

进而有

$$\widetilde{\phi_k}(x(t))=\phi_k(x^{(k)})+\sum_{j=1}^{n}\frac{\partial\phi_k(x^{(k)})}{\partial xj}\left[x_j(t)-x_j^{(k)}\right]<\phi_k(x^{(k)}).$$

即

$$\phi_k(x^{(k)})-\widetilde{\phi}_k(x(t))>0. \tag{4.12}$$

由(4.12)、(4.11)和(2.30)，有

$$\phi_k(x(t))<\phi_k(x^{(k)}).$$

本定理证毕。

〔**定理5**〕 在〔定理4〕的假设下，$\exists s_5=max\{s_1,s_3\}$，使$s\in[s_5+\infty)$时，加权函数$\phi_k(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$的下降值

$$\phi_k(x^{(k)})-\phi_k(x_{(s)}^{(k+1)})>\left[1-\theta_k(x_{(s)}^{(k+1)})\right]$$

$$\cdot\left[\sum_{i=1}^{m}w_i^{(k)}\left(\frac{\partial F_i(x^{(k)})}{\partial\Delta x(s)}\right)^2+s\sum_{j=1}^{n}q_j^{(k)}cos^2\alpha_j\right]\left\|\Delta x(s)\right\|^2$$

$$\geqslant\left[1-\theta_k(x_{(s)}^{(k+1)})\right]\left[\sum_{i=1}^{m}w_i^{(k)}\left(\frac{\partial F_i(x^{(k)})}{\partial\Delta x(s)}\right)^2+sq_0\right]\left\|\Delta x(s)\right\|^2$$

$$>sq_0\left[1-\theta_k(x_{(s)}^{(k+1)})\right]\left\|\Delta x(s)\right\|^2$$

$$\geqslant\frac{\varepsilon_0q_0}{4q^0}\left[1-\theta_k(x_{(s)}^{(k+1)})\right]\left\|\Delta x(s)\right\|.$$

证　因〔定理4〕的假设满足〔定理3〕的条件，故由(2.30)，〔定理3〕的(2)和(4.10)可直接推得本定理的结果。

〔注1〕 上面各定理中，关于函数在定义域D中一阶偏导数连续的假设可分别减弱。

〔注2〕 (2.30)式的方程

$$\phi_k(x^{(k)})-\phi_k(x_{(s)}^{(k+1)})=\left[1-\theta_k(x_{(s)}^{(k+1)})\right]\left[-\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}\right]\left\|\Delta x(s)\right\|$$

是获得〔定理 4〕的(5)和〔定理 5〕的结果的构造性表达式。不仅如此，它还诱导出用选择$q^{(k)}$权，使其对应的$\theta_k(x_{(s)}^{(k+1)})$落入较优敛速区间时，

$$\left[-\frac{\partial\phi_k(x^{(k)})}{\partial\Delta x(s)}\right]\left\|\Delta x(s)\right\|$$

有较大值的方法，使加权函数$\phi_k(x)$从$x^{(k)}$到$x_{(s)}^{(k+1)}$获得较大的下降值。

参 考 文 献

〔1〕K. Levenberg, *Quart. Appl. Math.*, 2 (1944), 164.

〔2〕A. Girard, *Rev. Opt.*, 37 (1958), 225, 397.

〔3〕C. G. Wynne, *Proc. Phys. Soc.*, London 73 (1959), 777.

〔8〕常群，光学設計文集，科学出版社，1976.

〔5〕南京大学数学系計算数学專业編，光学系統自动設計中的数值方法，国防工业出版社，1976.

The Selective Principle of Weights in the Method of Damped Least Squares

Wang Daqi Zhu Yunhua Chen Zhitian

Abstract

This article illuminates the selective principle of weights, and this principle inducts into new dealing method of damped least squares.