关于空间复盖方法

严仲德

总体最优化中的空间复盖方法[1]，(Space covring techniques)例如Evtushenko(Евтушенко)方法[2]、Shubert方法[3]等从71年开始出现以来，得到不少人的关注。

这里将从两个方面对空间复盖方法进一步深入探讨。

(1) 在空间复盖方法中，对Lipschitz条件的常数L，假设为已知的，而实际上要知道L将是一个同等困难的问题，这里将给出一个L的统计估计法，这方法基于"总体最优化的熵理论"[4]一文中$T(x)$分布的一个深入结果。

(2) 指出空间复盖方法中这二个方法实质上是熵理论[4]中渐进收缩方法的一个特例，就是[4]中ε熵界，$\varepsilon=0$的特殊情况，注意到空间复盖方法在高维时是有缺点的，而使用ε熵界将克服这一缺点，使之在高维时能行之有效。

§1

用空间复盖方法求S维空间上有界闭区域Ω上函数$y=f(X)$的总极小值，需已知关于X的Lipschitz条件中的常数L，才能实施，而实际上要知道$y=f(X)$在Ω上的L常数是十分困难的问题。

注意到当$y=f(X)$存在各偏导数时，偏导数数值的上界，就可作为Lipschitz条件中的L。

故可运用〔4〕中Ω上$f(X)$方向导数统计分布(或绝对值)的$T_0(x)$(或$T(x)$)，其随机变量$\xi(=f'_l(X))$的上界来代替Lipschitz条件中的L。

运用$T(x)$的经验分布函数$T_n(x)$的最大子样作为ξ的上界的近似，进而可用来估计L。

$T_n(x)$的第k项子样分位数为$\xi_k^{(n)}$，$\xi_k^{(n)}$最大项为$\xi_n^{(n)}$，这时$k=n=n-l+1$，有$l=1$。

这样就可用子样边序项的极限分布来统计估计常数L。

利用斯米尔诺夫的结果〔5〕〔6〕。

设随机变量

$$U_k^{(n)}=2n\left[1-F\left(\xi_k^{(n)}\right)\right]$$

$k=n-l+1$渐近地服从一个具有自由度$2l$的χ^2一分布

$$\lim_{n\to\infty}P\left(c<U_k^{(n)}<d\right)=\int_c^d\frac{u^{l-1}e^{-\frac{u}{2}}}{2^l\Gamma(l)}du$$

本文1980年6月13日收到

这里 $\xi_k^{(n)}$ 是将子样中的数值由大到小排列后的第l个值，$F(x)$是随机变量ξ的分布函数，在这里$F(x)$就是$T_o(x)$。

特别是最大值$k=n$，$l=1$时，有

$$T_o(\xi_n^{(n)})=F\xi(_n^{(n)})=1-\frac{1}{2n}U_n^{(n)}$$

$U_n^{(n)}$的渐近分布是自由度为 2 的χ^2—分布，n充分大时有

$$P(U_n^{(n)}\leqslant\chi_\alpha^2)\geqslant 1-\alpha$$

（这里$P(\chi^2>\chi_\alpha^2)=\dfrac{1}{2^{\frac{l}{2}}\Gamma\left(\frac{l}{2}\right)}\displaystyle\int_{\chi_\alpha^2}^{+\infty}e^{-\frac{u}{2}}u^{\frac{l}{2}-1}du=\alpha$，$l$ 是自由度）

$$T_o(\xi_n^{(n)})=F(\xi_n^{(n)})\geqslant 1-\frac{1}{2n}\chi_\alpha^2$$

约有 $1-\alpha$ 的把握判定

$$T_o(\xi_n^{(n)})\geqslant 1-\frac{1}{2n}\chi_\alpha^2$$

注意到ξ的样本空间是Ω上任取点、任取方向的方向导数值，Ω是S维闭区域，而任一方向由一个S维单位球面决定，（即为一$S-1$维闭有界区域）。故实际是在一$\Omega\times\Omega_{S-1}$的 $2S-1$ 维有界闭区域中取点，设其测度为$\mu(\Omega\times\Omega_{S-1})$。

由于$T_o(\xi_n^{(n)})=P(\xi<\xi_n^{(n)})$，是 $\xi<\xi_n^{(n)}$出现的概率，而这里是在$\Omega\times\Omega_{S-1}$上随机抽取，故在$P(\xi<\xi_n^{(n)})\mu(\Omega\times\Omega_{S-1})$的测度上有 $\xi=f_l'(X)\leqslant\xi_n^{(n)}$ 即在这些测度上 $L\leqslant\xi_n^{(n)}$。

即在 $T_o(\xi_n^{(n)})\mu(\Omega\times\Omega_{S-1})$的测度范围上Lipschitz 条件常数$L$有$L\leqslant\xi_n^{(n)}$。

由于 $T_o(\xi_n^{(n)})\geqslant 1-\frac{1}{2n}\chi_\alpha^2$故有$1-\alpha$ 的把握判定，在大于$\left(1-\frac{1}{2n}\chi_\alpha^2\right)\mu(\Omega\times\Omega_{S-1})$的测度上有 $L\leqslant\xi_n^{(n)}$。这样就得定理：

定理1.1：Ω上$f(X)$的方向导数统计分布$T_o(x)$进行n次抽样，其最大子样为 $\xi_n^{(n)}$， 把 $\xi_n^{(n)}$ 作为$f(X)$关于 X的Lipschitz 条件常数L的近似，则有 $1-\alpha$ 的把握判定在大于$\left(1-\frac{1}{2n}\chi_\alpha^2\right)\mu(\Omega\times\Omega_{S-1})$的测度上$L\leqslant\xi_n^{(n)}$，(即在原区域$1-\frac{1}{2n}\chi_\alpha^2$ 范围上)当 n 越大时越正确。（这里 α，χ_α^2是从自由度为 2 的χ^2—分布表上查得）

例如 $\alpha=0.01$ 时 $\chi_\alpha^2=9.210$，若取500个子样 $T_o(\xi_n^{(n)})\geqslant 1-\frac{1}{2n}\chi_\alpha^2=0.99079$。就得到约有99%的把握判定在原区域的99.079%范围上有 $L\leqslant\xi_n^{(n)}=\xi_{500}^{(500)}$

这样就给出了一个可实现的Lipschitz条件常数L的统计估计法，一个求近似值的易于实施的方法，而且有统计的依据。

§2

容易看出空间复盖方法中Evtushenko方法及Shubert方法一定意义下是熵理论中渐进收缩方法的一个特例，是其中ε熵界在$\varepsilon=0$时的特例。

记$A(X)=|\mathrm{grad}f(X)|$，其上界$A_\Omega=\sup\limits_{X\in\Omega}A(X)$就是$\Omega$上$f(X)$的最大方向导数，偏导数的上界$A_\Omega$就可作为Lipschitz条件常数$L$。

由〔4〕引理3.1、系2：若Ω上一点X_j，有$f(X_j)>\bar{f}_{min}$，且$A(X)$几乎处处存在，在Ω上有上界A_Ω，则必存在r，在邻域$|X-X_j|<r$上，$H(X)=0$。（$\bar{f}_{min}$是目前已知最小值）。

具体定量化有：由〔4〕中定义4.1，一般地ε熵界$G_\varepsilon=T^{-1}(1-2H^{-1}(\varepsilon))$。

由引理4.1：对ε熵界G_ε，若$|X-X_j|<R=\dfrac{f(X_j)-\bar{f}_{min}}{G_\varepsilon}$，则有$H(X)\leqslant\varepsilon$。

特别当$\varepsilon=0$，$G_\varepsilon=T^{-1}(1)=A_\Omega$

$$|X-X_j|<R=\frac{f(X_j)-\bar{f}_{min}}{A_\Omega}$$

$$A_\Omega|X-X_j|<f(X_j)-\bar{f}_{min}$$

$$f(X_j)-A_\Omega|X-X_j|>\bar{f}_{min}$$

又Lipschitz条件中$L=A_\Omega$

$|f(X_j)-f(X)|\leqslant L|X_j-X|=A_\Omega|X_j-X|$，由于考虑总极小值，所以仅注意$f(X)$变小的情况。

$$f(X_j)-f(X)\leqslant A_\Omega|X_j-X|$$

就有

$$f(X)\geqslant f(X_j)-A_\Omega|X-X_j|>\bar{f}_{min}$$

在$|X-X_j|<R$范围内的点总有$f(X)>\bar{f}_{min}$，在X点完全排除了小于目前已知最小值的可能性，所以可以不再考虑这些点。

这样就将X_j点的一个邻域，以至Ω中所有这样的邻域排除出去，使搜索总极小值的区域缩小。例如Shubert方法也可逐渐缩小区域，就是将不可能的区域排除，实际上就是〔4〕中将0熵的区域排除（就是把没有总极小值信息的区域排除）。

我曾用最大方向导数A_Ω来进行试算（即用0熵界试算）发现在高维时有很大的计算量（这一实践得来的结论与〔1〕中对Evtushenko及Shubert方法评论是相一致的）它们仅适用于低维，若不强调0熵来渐进收缩，而用ε熵界来代替强调完全不可能小于总极小值的0熵界，就留下了余地，可以大大减少计算工作量，而最后的总极小值还能用熵科学地数量地进行衡量，这也从一个侧面说明ε熵界引入的必要性和合理性。

参 考 文 献

〔1〕 Dixon L. C. W, Szegö G. P. (eds), Towards Global Optimisation(1975), 2(1978)

〔2〕 Евтущенко Ю. Г., Вычисл. матем.и матем. физ, 11, 6, (1971), 1390—1403.

〔3〕 Shubert B. O., SIAM J. Numer, Anal. 9, 3. pp. 379—388(1972)

〔4〕 严仲德，总体最优化的熵理论，《自然杂志》12(1979)〔上海师范学院论文选编(1978)〕。

〔5〕 M. 费史，概率论与数理统计。

〔6〕 Смирнов Н. В. Предельные законы распределения для члеов вориационного ряда. москва(1949).