第23卷第1期 河 海 大 学 学 报 Vol. 23 No. 1
1995年1月 JOURNAL OF HOHAI UNIVERSITY Jan. 1995

结构优化设计与最大熵原理*

杨仲侯 杨海霞 金 峰

(工 程 力 学 系)

摘要 介绍了在结构最优设计中，将序列二次规划数学模式与最大熵原理结合的解法. 不仅使迭代计算实现公式化，不解一个联立方程；同时在每一次迭代计算也随之给出了相应的约束信息. 较之其它方法，免去了专设的可行性与最优性的检查. 计算表明，本文的计算格式，不仅使计算过程简捷、清晰，而且适应性也为之宽广了. 如能选择好计算参数 p，还能进一步使迭代迅速收敛，对于联合优化问题也有较高的效率.

关键词 结构设计 序列二次规划 最大熵原理 联合优化

中图法分类号 TU318

Structural Optimization Design and Maximum Entropy Theory

Yang Zhonghou Yang Haixia Jin Feng (*Dept. of Engineering Mechanics, Hohai Univ.*)

A new optimization method which combines SQP with maximum entropy theory is disscussed. With this method, there is no need to solve linear equations and the process of iteration can be described by formula. Meanwhile, the constraint in formation can be gained in every interation process. Compared with other methods, the method can avoid the checking of feasibility and optimality and has wider adaptability. Examples show that the process is concise and clear. If the parameters are chosen properly, the iteration would converge rapidly. Moreover, the method is efficient in solving combinatory optimization problems.

Keywords structure design SQP maximum entropy theory combinatory optimization

结构优化设计方法虽多，但只有将几个方法加以综合，才能解决较为复杂的工程问题. 序列线性规划有简单易行的优点，但有时计算常在最小点附近振荡，收敛缓慢；序列二次规划是一个精度高，适应性强，收敛快的方法[1]，但每一步都有大量的计算工作，还需要加以可行性与最优性检查才能说明在计算点的约束情况. 我国学者李兴斯结合桁架的优化设计提出了

* 水利部科学基金项目成果

收稿日期：1994－01－08

作者简介：杨仲侯 男 教授 结构力学专业 主要从事复杂结构分析方法及优化设计研究 已发表《拱坝应力分析的拱梁分载位移法》《桩基结构的空间分析》等论文

一种新的作法，即序列线性规划与最大熵相结合的方法[2]. 该法第一次引用了代理约束和信息熵的概念，新颖别致，使计算达到了定型化和公式化. 接着王希诚在《信息熵原理在结构优化设计的应用》中讨论了同一问题，也还是限于桁架的算例. 继之，工程博士生江建祥提出"结构优化序列熵迭代法"[3]，它是将二次规划与最大熵原理结合的一种解法，使之适应性增加了. 他讨论了梁、刚架、钢筋混凝土框架的优化实例. 但算例只给出了最后结果，无法窥其细节，影响了应用. 笔者研究了同一问题，发现序列二次规划与最大熵原理相结合的优化方法有很多优点，不仅使计算公式化、定型化，还能适应各种情况下的优化问题，而且在计算点的约束情况了如指掌，不需要另加可行性与最优性的检查. 如能对每一个实际问题选择合适的参数 p，再加上移动极限，可进一步加快计算速度. 本文将用较多的实例说明这些优点.

1　结构优化的最大熵方法

1987 年，A. B. Templemen 和我国学者李兴斯提出了以信息熵为基础的约束非线性规划方法[2]，它将约束非线性规划问题等效转化成只有一个代理约束的约束非线性规划问题，用最大熵准则确定代理约束中的约束乘子，使问题变得易于求解的约束非线性规划问题. 该方法避免了区分主动约束和被动约束带来的复杂性；同时使得优化过程公式化，概念清楚，方法简单.

1.1　代理约束概念

约束非线性规划的数学模型一般为

$$\left.\begin{aligned}&\min f(X)\\ \text{s.t.}\quad &g_j(X)\leqslant 0\end{aligned}\right\}\qquad j=1,2,\cdots,m \tag{1}$$

其中 m 个约束可以用一个"代理约束"的概念统一为一个约束，使问题(1)成为

$$\left.\begin{aligned}&\min f(X)\\ \text{s.t.}\quad &G(X)=\sum_{j=1}^{m}\lambda_j g_j(X)\leqslant 0\end{aligned}\right\} \tag{2}$$

式中　$G(X)$ ——代理约束；$\lambda_j(j=1,2,\cdots,m)$ ——代理乘子，分别为 m 个约束的权系数，它们必须满足

$$\sum_{j=1}^{m}\lambda_j=1\qquad \lambda_j\geqslant 0\qquad j=1,2,\cdots,m \tag{3}$$

在结构优化的实际问题中，在最优点至少有一个约束是主动约束，因此可处理成等式约束. 故问题(2)成为

$$\left.\begin{aligned}&\min f(X)\\ \text{s.t.}\quad &G(X)=\sum_{j=1}^{m}\lambda_j g_j(X)=0\end{aligned}\right\} \tag{4}$$

可以证明，存在一组最优乘子 λ_j^*，使问题(4)与问题(1)同解. 因此，如果我们能找到一组最优乘子 λ_j^*，就可简便地求解约束非线性规划式(4)了. 最大熵方法就是采用最大熵准则来确定代理乘子 λ_j^* 的.

1.2　信息熵与最大熵准则

设有一信源集合

$$X = \begin{Bmatrix} a_1 & a_2 & \cdots & a_n \\ p_1 & p_2 & \cdots & p_n \end{Bmatrix}$$

式中 a_i——信源中可能出现的消息；p_i——消息 a_i 出现的概率，它满足概率的一般条件

$$\sum p_i = 1 \qquad p_i \geqslant 0 \qquad j = 1,2,\cdots,n$$

对信源中可能出现的结果，存在不确定性与其概率分配有关. Shannon 对这个不确定性，作了定量的描述，并定义

$$H(X) = -K\sum_{i=1}^{n} p_i \ln p_i$$

为信源的信息量，称为信息熵. 式中 K 是一个与量度单位有关的正常数. 可以看出熵函数的极值性和上凸性，并且 $H(X)$ 越大信息源所含信息量越大，不确定性的程度也越大.

当我们对信源作分析时，比较切合实际的准则是认为信源未知部分发生的概率相同，即此时有最大的熵. 于是，Jaynes 在 1957 年提出一个最大熵准则：当根据部分信息进行推理时，必须选择一组概率分配，它应具有最大的熵，并服从一切已知的信息，这是能做出的唯一无偏分配. 这个准则的数学表达式为

$$\left.\begin{aligned} &\max H(X) = -K\sum_{i=1}^{n} p_i \ln p_i \\ \text{s.t.} \quad &\sum_{i=1}^{n} p_i a_i = E(a_i(X)) \\ &\sum_{i=1}^{n} p_i = 1 \qquad p_i \geqslant 0 \end{aligned}\right\} \quad i = 1,2,\cdots,n \tag{5}$$

式中 $[x_1,\ \cdots,\ x_n]^{\mathrm{T}}$——随机变量的离散值；$[p_1,\ \cdots,\ p_n]^{\mathrm{T}}$——该随机变量取各离散值的概率；$a_i(X)$——均值、方差等可观测值；$E(a_i(X))$——相应的观测量值.

1.3 利用最大熵准则确定代理乘子 λ_j

现在，我们可以将式(4)中的 m 个约束看作是某一信源发出的 m 个信息，每一个信息发射的比重各有不同，代理乘子 λ_j 为第 j 个约束信息的概率，一组 λ_j 可理解为约束信息在计算点的概率分配. 因此，可将上述优化问题转化为最大熵问题，就可用最大熵准则求解代理乘子 λ_j. 其数学表达式为

$$\left.\begin{aligned} &\max H(X) = -K\sum_{j=1}^{m} \lambda_j \ln \lambda_j \\ \text{s.t.} \quad &\sum_{j=1}^{m} \lambda_j g_j(X) = 0 \\ &\sum_{j=1}^{m} \lambda_j = 1 \qquad \lambda_j \geqslant 0 \end{aligned}\right\} \quad j = 1,2,\cdots,m \tag{6}$$

引入拉格朗日乘子 α 与 β，问题(6)的拉格朗日函数为

$$L(\lambda_j,\alpha,\beta) = -K\sum_{j=1}^{m} \lambda_j \ln \lambda_j + \alpha\sum_{j=1}^{m} \lambda_j g_j(X) + \beta\left(\sum_{j=1}^{m} \lambda_j - 1\right)$$

再由条件

$$\frac{\partial L}{\partial \lambda_j} = 0 \qquad \frac{\partial L}{\partial \alpha} = 0 \qquad \frac{\partial L}{\partial \beta} = 0$$

不难求得

$$\lambda_j = \frac{\exp[pg_j(X)]}{\sum_{j=1}^{m}\exp[pg_j(X)]} \qquad p = a/K > 0 \tag{7}$$

由于 X 是未知，故求 λ_j^* 的过程只能是迭代的．随迭代过程，对约束信息（松紧程度）的了解越来越多，λ_j 的不确定性越来越少，因而约束的信息熵也越来越小，可见 K 应随迭代逐步减小，p 随迭代而递增．

2　用序列二次规划解规划式(4)

序列二次规划具有精度高，收敛快，适应性强的优点．因此，当 λ_j 求出后，为配合前述迭代，本文采用序列二次规划求解规划(4)．式(4)的序列二次规划的数学模式如下：

$$\left.\begin{aligned} &\min f(X) = f(X^k) + \nabla f(X^k)^{\mathrm{T}}\delta X + \frac{1}{2}\delta X^{\mathrm{T}}H(X^k)\delta X \\ &\text{s.t.}\quad \sum_{j=1}^{m}\lambda_j[g_j(X^k) + \nabla g_j(X^k)^{\mathrm{T}}\delta X = 0 \end{aligned}\right\} \tag{8}$$

式中　$H(X^k)$ ——目标函数在 X^k 点的 Hessian 矩阵；$\nabla f(X^k)$，$\nabla g_j(X^k)$ ——目标函数和约束函数在 X^k 点的梯度向量．

设
$$A = \sum_{j=1}^{m}\lambda_j\nabla g_j(X^k) \qquad B = \sum_{j=1}^{m}\lambda_j g_j(X^k) \tag{9}$$

η 为拉格朗日乘子，于是式(8)的拉格朗日函数数成为

$$L(\delta X,\eta) = f(X^k) + \nabla f(X^k)\delta X + \frac{1}{2}\delta X^{\mathrm{T}}H(X^k)\delta X + \eta(A^{\mathrm{T}}\delta X + B) \tag{10}$$

最后由 $\dfrac{\partial L}{\partial\delta X} = 0$，$\dfrac{\partial L}{\partial\eta} = 0$，可求出

$$\eta = \frac{B - A^{\mathrm{T}}H^{-1}\nabla f(X^k)}{A^{\mathrm{T}}H^{-1}A} \qquad \delta X = -H^{-1}(\nabla f(X^k) + \eta A) \tag{11}$$

3　实 例 分 析

总结以上各节，本法的迭代格式如下：

a．给定 X^0，p^0，Δp，ε．

b．令 $k=0$，计算

$$\lambda_j^k = \frac{\exp[p^k g_j(X^k)]}{\sum_{j=1}^{m}\exp[p^k g_j(X^k)]} \qquad \nabla f(X^k) = \left[\frac{\partial f(X)}{\partial x_1}\cdots\frac{\partial f(X)}{\partial x_n}\right]^{\mathrm{T}}_{X^k}$$

$$A^k = \sum_{j=1}^{m}\lambda_j^k\nabla g_j(X^k) \qquad B^k = \sum_{j=1}^{m}\lambda_j^k g_j(X^k)$$

$$H(X^k) = \begin{bmatrix} \dfrac{\partial^2 f(X)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(X)}{\partial x_1\partial x_n} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial^2 f(X)}{\partial x_n\partial x_1} & \cdots & \dfrac{\partial^2 f(X)}{\partial x_n^2} \end{bmatrix}_{X^k} \qquad \eta^k = \frac{B^k - (A^k)^{\mathrm{T}}(H^k)^{-1}\nabla f(X^k)}{(A^k)^{\mathrm{T}}(H^k)^{-1}(A^k)}$$

$$\delta X^k = -(H^k)^{-1}[\nabla f(X^k) + \eta(A^k)] \qquad X^{k+1} = X^k + \delta\Delta^k \qquad p^{k+1} = p^k + \Delta p$$

c. 令 $k=1, 2, \cdots$，重复第二步至前后两轮结果接近，满足 $\varepsilon^k \leqslant \varepsilon$ 为止.

例1 求 $X=[x_1 \quad x_2]^{\mathrm{T}}$[5]

$$\min f(X) = x_1^2 + x_2^2$$

$$\text{s.t.} \quad g_1(X) = \frac{x_1^2}{20} - x_2 + 1 \leqslant 0$$

$$g_2(X) = \frac{x_2^2}{20} - x_1 + 1 \leqslant 0$$

$$x_1, x_2 \geqslant 0$$

表1 例1结果

方法		$f(X)$	x	x_2	λ_1	λ_2
本文	1	0.5000	0.5000	0.5000	0.48	0.48
	4	2.2182	1.0556	1.0556	0.50	0.50
	5	2.2291	1.0557	1.0557	0.50	0.50
可行方向法		2.2230	1.0560	1.0560		
序列线性化法		2.2260	1.0550	1.0550		

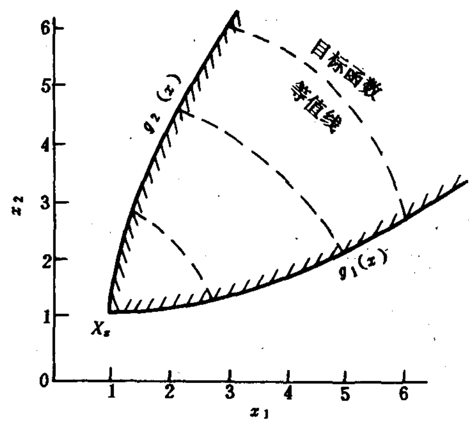

图1 例1图解

因对称取 $x_1^0=0, x_2^0=0, p^0=0, \Delta p=3, \varepsilon=0.001$，经五次迭代收敛于最优解 $X^*=[1.0557 \quad 1.0557]^{\mathrm{T}}$，$f(X^*)=2.2291$. $\lambda_1=\lambda_2=0.5$ 说明最优点在两个约束的交点（见图1、表1）. *

例2 为各教材引用的三杆桁架最轻设计的考题，其优化的数学模式如下*：

求 $X=[x_1 \quad x_2 \quad x_1]^{\mathrm{T}}$

$$\min W(X) = 2\sqrt{2}x_1 + x_2$$

$$\text{s.t.} \quad g_1(X) = x_2 + \sqrt{2}x_1 - \sqrt{2}x_1^2 - 2x_1x_2 \leqslant 0$$

$$g_2(X) = \sqrt{2} - \sqrt{2}x_1 - 2x_2 \leqslant 0$$

$$g_3(X) = 4x_2 - 3\sqrt{2}x_1^2 - 6x_1x_2 \leqslant 0$$

$$x_1, x_2 \geqslant 0$$

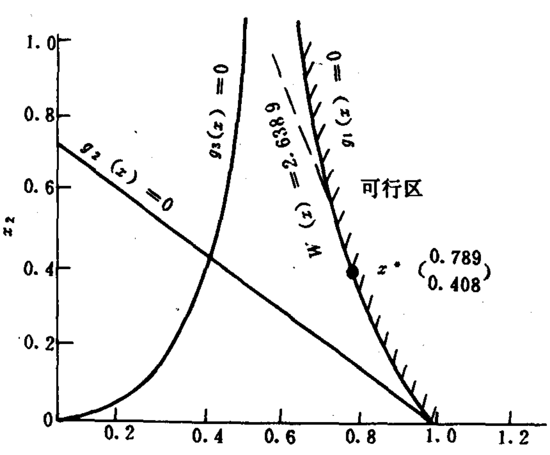

图2 例2图解

优化时为保证Hessian矩阵正定和提高精度，先将上规划转换到倒设计变量空间运行，但每一迭代结果仍由原设计变量输出. 取 $x_1^0=1, x_2^0=1, p^0=0, \Delta p=5, \varepsilon=0.001$，并采用限步长迭代运算，经11次迭代得最优解：$x_1^*=$

* 杨仲侯. 工程结构的优化设计方法. 南京：河海大学，1991

0.789 0，$x_2^*=0.407\ 3$，$W(X^*)=2.638\ 6$. $\lambda_1=1$，$\lambda_2=\lambda_3=\lambda_4=\lambda_5=0$，说明最优点在 $g_1(x)=0$ 的约束曲线上（见图 2、表 2）.

表2　例 2 结果

方法		$W(x)$	x_1	x_2	λ_1	λ_2	λ_3	λ_4	λ_5
本	1	2.614 6	0.665 2	0.674 9	0.164 9	0.012 1	0.823 0	0	0
	10	2.638 6	0.789 1	0.406 9	1	0	0	0	0
文	11	2.638 6	0.789 0	0.407 3	1	0	0	0	
齿行法		2.642 9	0.768 0	0.408 0					
图解法		2.638 9	0.789 0	0.408 0					

例 3　图 3 示一工字型组合梁，截面系数有统计关系 $F=0.8I^{0.5}$，$S=0.78I^{0.25}$，设 σ，E，I 及 q 为已知，求梁最轻下的的截面模数 S_1，S_2[4].

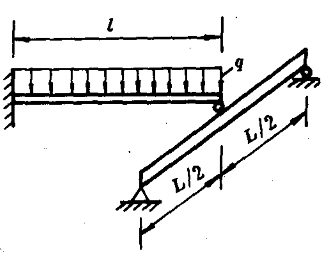

图 3　组合型钢梁

该题经简化后的优化模型为

求 $X=[x_1\ x_2]^T$

$$\min W(X)=x_1^{2/3}+x_2^{2/3}$$

$$\text{s.t.}\quad g_1(X)=16(1-2\beta)-x_1\leqslant 0$$

$$g_2(X)=8\beta-x_2\leqslant 0$$

$$x_1,x_2\geqslant 0$$

由于本例约束方程及目标函数非线性程度高，极小点较多（图 4），为使迭代向最小点逼近，采用变步长的限步长法，取 $x_1^0=2\sqrt{2}$，$x_2^0==\sqrt{3}$，$p^0=0$，$\Delta p=7$，$\varepsilon=0.001$，步长 $FK=0.6$，变步长系数 $CK=0.8$，使计算经 38 次迭代稳步走向全局最小点，见表 3.

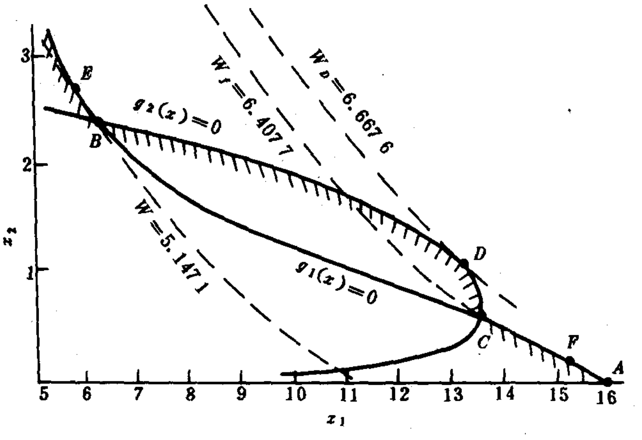

图 4　例 3 图解

例4 图5示一两杆桁架，在结点C受有100 kN的垂直荷载，以截面面积x_1,x_2及C点的铅直坐标y为设计变量，$\bar{\sigma}$=100 000 kPa，y的变化区间为1.0～3.0 m，x_1与x_2必须为正.

表3 例3结果

方法		$W(X)$	x_1	x_2	λ_1	λ_2
本	1	3.140	1.661 1	2.291 0	1	0
文	38	5.147	5.924 3	2.562 6	1	0
复形法		5.148	5.841 0	2.629 1		
解析解		5.147	5.925 8	2.562 1		

该问题是一个联合优化问题，它的最小体积的优化数学模式可以叙述为

求 $$X = [x_1, x_2, y]^T$$

$$\min W(X) = x_1\sqrt{16+y^2} + x_2\sqrt{1+y^2}$$

s.t. $$g_1(X) = 20\sqrt{16+y^2}/(yx_1) \leqslant 100\ 000$$

$$g_2(X) = 80\sqrt{1+y^2}/(yx_2) \leqslant 100\ 000$$

$$1 \leqslant y \leqslant 3 \qquad x_1, x_2 \geqslant 0$$

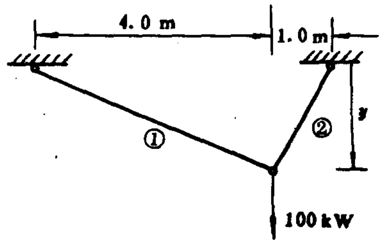

图5 两杆桁架

取x_1^0=1.000 00，x_2^0=1.000 00，y^0=1.000 00，p^0=0.0，Δp=0.000 1，e=0.001. 对各变量分别采用变步长限制迭代计算，经21次迭代收敛于分层模拟坐标法的结果（见图6、表4).

表4 例4结果

方法		x_1 $\times10^{-3}$	x_2 $\times10^{-3}$	y	$W(X)$ $\times10^{-3}$	λ_1	λ_2	λ_3	λ_4
本	1	5.69	3.33	0.903	27.80	0.25	0.25	0.25	0.25
文	21	0.45	0.89	2.000	4.00	0.25	0.25	0.25	0.25
模拟坐标法		0.448	0.896	2.000	4.00				

4 结 束 语

a. 本文讨论的方法具有精度高，收敛快，适应广，计算简便，每一例都迭代到底，不受收敛条件的限制，每一迭代点的约束情况了如指掌，不需要作特殊检查.

b. 在收敛极小点的$\lambda_j \geqslant 0$，可知：如$\lambda_j > 0$，对应的约束是紧约束；如$\lambda_j = 0$对应的约束不是紧约束. 在每一计算点的$\lambda_j \geqslant 0$，只能说明约束对该点的松紧程度.

c. 为了使计算稳步走向极小点，用移动极限控制迭代步长是可取的，如例3.

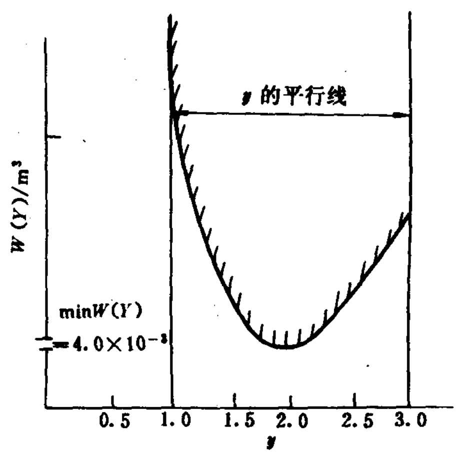

图6 例4图解

d. 从例4的计算表中还见到 λ_j 都大于0这可解释为，对静定结构每走一步应力都达到了满应力；y 是协调变量，只要在允许范围内变化，故在每计算点也是紧约束之故．至于 λ_j 都相同，又由于在每一计算点 $e^{y f_j(X)}$ 的指数都近于零之故．

e. 本法收敛的快慢还与计算参数 X^0，y^0，ΔP 有关．例1参数选取对计算的影响，见表5．

表5　参数 X^0，Δp 的影响

x_1^0	x_2^0	ΔP	N	x_1^*	x_2^*	$W(X^*)$
−1	−1	1.5	6	1.055 3	1.055 3	2.227 3
1	1	1.5	6	1.055 3	1.055 3	2.227 3
−1	−1	3	4	1.055 6	1.055 6	2.228 7
1	1	3	4	1.055 6	1.055 6	2.228 7
−1	−1	5	4	1.055 7	1.055 7	2.229 1
1	1	5	4	1.055 7	1.055 7	2.229 1

参 考 文 献

1 杨仲侯，顾　浩，潘正初等．序列二次规划法及其在水工结构优化设计中的应用．华东水利学院学报，1984，12(4)：54～65

2 李兴斯．结构优化设计的最大熵方法．计算结构力学及其应用，1988，6(1)：36～45

3 江建祥．结构优化的序列熵迭代法．基建优化，1991，12(3)：24～28

4 王光远，董明耀．结构优化设计．北京：高等教育出版社，1987．41～47