

多重特征值的敏度分析及优化算法

钟万勰 程耿东

(大连工学院工程力学研究所)

摘要 本文利用瑞雷商的驻值原理导出了重特征值的二阶敏度表达式, 将它和单特征值的敏度作了比较, 利用这些敏度公式就可以构造出有效的序列二次规划算法来求解涉及重特征值的最优结构设计问题。给出了对称柱与非对称柱的最大临界力设计的数值结果, 证明了方法的有效性。

一、引言

涉及重特征值的最优结构设计问题近年来受到了广泛重视。对这类问题的最优化必要条件曾作过仔细的研究, 准则法和数学规划法都曾用来求解具有二重失稳模式和振型的各种结构的最优材料分布。特征值的敏度分析是这些研究的核心。已经指出过, 早期文献^[1]中给出的特征值对设计变量的一阶敏度公式只能适用于单特征值^[2]。如果在当前设计点要考虑的特征值是重特征值, 则只有采用和设计增量对应的一对极限向量^[3], 才能仍然按照常用公式计算一阶敏度。

本文将文献^[3]的研究加以推广, 利用瑞雷商驻值原理导出了用极限向量表示的重特征值的二阶敏度表达式, 利用这些表达式可以构造序列二次规划有效地求解涉及重特征值的最优结构设计问题。对于两端受到集中力的柱, 对称或者非对称的支承条件, 给出了使临界力最大的最优设计, 显示了本算法较已有算法为优。

二、单重特征值和特征向量的敏度分析

考虑广义特征值问题

$$K(\alpha)X = \lambda M(\alpha)X \quad (1)$$

式中 K 和 M 为依赖于设计变量 α 的对称正定矩阵, (1) 的第 i 个特征值和特征向量满足:

$$KX_i = \lambda_i MX_i \quad (2)$$

将 (2) 式两端对设计变量 α_j 求导可得:

$$\frac{\partial K}{\partial \alpha_j} X_i + K \frac{\partial X_i}{\partial \alpha_j} = \frac{\partial \lambda_i}{\partial \alpha_j} M X_i + \lambda_i \frac{\partial M}{\partial \alpha_j} X_i + \lambda_i M \frac{\partial X_i}{\partial \alpha_j} \quad (3)$$

在该式两侧前乘 X_i^T 并利用 (2) 式可导出

$$\frac{\partial \lambda_i}{\partial \alpha_j} = \frac{1}{X_i^T M X_i} \left(X_i^T \frac{\partial K}{\partial \alpha_j} X_i - \lambda_i X_i^T \frac{\partial M}{\partial \alpha_j} X_i \right) \quad (4)$$

中国科学院科学基金资助的课题

该式是常用的特征值一阶敏度表达式。为了求得特征向量的敏度，设

$$\frac{\partial X_i}{\partial \alpha_j} = \sum_{k=1}^n b_{ijik} X_k \quad (5)$$

式中 n ——矩阵 K 和 M 的阶数；

X_1, X_2, \dots, X_n ——设计向量 α 处问题 (1) 的全部特征向量；

将 (5) 代入 (3)，并在两侧前乘特征向量 X_i^T ，然后利用振型应满足的正交归一化条件，可得

$$X_i^T \frac{\partial K}{\partial \alpha_j} X_i + b_{ijil} \lambda_i = \lambda_i X_i^T \frac{\partial M}{\partial \alpha_j} X_i + \lambda_i b_{ijil} + \frac{\partial \lambda_i}{\partial \alpha_j} \delta_{ii} \quad (6)$$

由此

$$b_{ijil} = \frac{1}{(\lambda_i - \lambda_l)} X_i^T \left[\frac{\partial K}{\partial \alpha_j} - \lambda_l \frac{\partial M}{\partial \alpha_j} \right] X_l \quad i \neq l \quad (7)$$

上面提到的正交归一化条件为：

$$X_i^T M X_l = \delta_{il} \quad X_i^T K X_l = \lambda_i \delta_{il} \quad (8)$$

式中 δ_{il} ——Kronecker 符号。对于 $i=l$ ，应利用归一化条件

$$X_i^T M X_i = 1 \quad (9)$$

来推导 b_{ijil} 的公式。将 (9) 的两侧对 α_j 求导可得

$$b_{ijil} = -\frac{1}{2} X_i^T \frac{\partial M}{\partial \alpha_j} X_i \quad (10)$$

综合 (5)，(7) 和 (10) 可得

$$\frac{\partial X_i}{\partial \alpha_j} = -\frac{1}{2} X_i^T \frac{\partial M}{\partial \alpha_j} X_i X_i + \sum_{k=1, k \neq i}^n \frac{1}{(\lambda_i - \lambda_k)} X_i^T \left[\frac{\partial K}{\partial \alpha_j} - \lambda_l \frac{\partial M}{\partial \alpha_j} \right] X_l X_k \quad (11)$$

为了求得特征值的二阶敏度表达式，将 (4) 式求导并注意利用式 (9) 可得

$$\begin{aligned} \frac{\partial^2 \lambda_i}{\partial \alpha_j \partial \alpha_t} &= X_i^T \left[\frac{\partial^2 K}{\partial \alpha_j \partial \alpha_t} - \lambda_i \frac{\partial^2 M}{\partial \alpha_j \partial \alpha_t} \right] X_i + 2 \frac{\partial X_i^T}{\partial \alpha_t} \frac{\partial K}{\partial \alpha_j} X_i \\ &\quad - \frac{\partial \lambda_i}{\partial \alpha_t} X_i^T \frac{\partial M}{\partial \alpha_j} X_i - 2 \lambda_i \frac{\partial X_i^T}{\partial \alpha_t} \frac{\partial M}{\partial \alpha_j} X_i \end{aligned} \quad (12)$$

如果将 $\frac{\partial X_i}{\partial \alpha_t}$ 及 $\frac{\partial \lambda_i}{\partial \alpha_t}$ 的表达式代入 (12)，进行运算后得到：

$$\begin{aligned} \frac{\partial^2 \lambda_i}{\partial \alpha_j \partial \alpha_t} &= X_i^T \left[\frac{\partial^2 K}{\partial \alpha_j \partial \alpha_t} - \lambda_i \frac{\partial^2 M}{\partial \alpha_j \partial \alpha_t} \right] X_i + 2 \lambda_i X_i^T \frac{\partial M}{\partial \alpha_t} X_i X_i^T \frac{\partial M}{\partial \alpha_j} X_i \\ &\quad - X_i^T \frac{\partial M}{\partial \alpha_t} X_i X_i^T \frac{\partial K}{\partial \alpha_j} X_i - X_i^T \frac{\partial M}{\partial \alpha_j} X_i X_i^T \frac{\partial K}{\partial \alpha_t} X_i \\ &\quad + \sum_{k=1, k \neq i}^n \frac{2}{(\lambda_i - \lambda_k)} X_i^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_l \frac{\partial M}{\partial \alpha_j} \right) X_k X_k^T \left(\frac{\partial K}{\partial \alpha_t} - \lambda_l \frac{\partial M}{\partial \alpha_t} \right) X_l \end{aligned} \quad (13)$$

以上的敏度表达式 (4)，(11) 及 (13) 只适用于在设计 α 处特征值 λ_i 是单重的，这从 (11) 和 (13) 式来看是十分显然的，因为它们的分母上出现 $\lambda_i - \lambda_k$ ，对于重特征值则成为零；进一步，在推导 (4) 时采用了 $\frac{\partial X_i}{\partial \alpha_j}$ (见 (3) 式)，而在重特征值处这是一个不

确定的式子，因而这个推导是不合法的，（4）式也不适用于重特征值问题。下面将基于瑞雷商驻值原理给出灵敏度表达式的另一推导方法，它避免了以上的非法运算。

三、基于瑞雷商的灵敏度分析推导

瑞雷商原理指出，在设计 $\alpha + \Delta\alpha$ 处的特征值为下列驻值问题的解：

$$\begin{aligned} \text{求 } X, \quad & \text{sta. } X^T K(\alpha + \Delta\alpha) X \\ & \text{s.t. } X^T M(\alpha + \Delta\alpha) X = 1 \end{aligned} \quad (14)$$

因为在设计 α 处求得的与特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 相应的特征向量 X_1, X_2, \dots, X_n ，张成一个完备的空间，所以在设计 $\alpha + \Delta\alpha$ 处的特征向量 X 可展成

$$X = \sum_{j=1}^n C_j X_j \quad (15)$$

利用（8）式问题（14）成为

求 $c_j (j=1, 2, \dots, n)$

$$\text{sta. } \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} + \sum_{j=1}^n \lambda_j c_j^2 \quad (16)$$

$$\text{s.t. } \sum_{j=1}^n c_j^2 + \sum_{j=1}^n \sum_{k=1}^n c_j c_k T_{jk} = 1$$

$$\text{式中 } S_{jk} = X_j^T \Delta K X_k \quad T_{jk} = X_j^T \Delta M X_k \quad (17)$$

$$\Delta K = K(\alpha + \Delta\alpha) - K(\alpha), \quad \Delta M = M(\alpha + \Delta\alpha) - M(\alpha)$$

为了求解（16）可构造拉格朗日函数

$$L = \sum_{j=1}^n \lambda_j c_j^2 + \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} + \mu \left(1 - \sum_{j=1}^n c_j^2 - \sum_{j=1}^n \sum_{k=1}^n c_j c_k T_{jk} \right) \quad (18)$$

它的驻值条件为：

$$\frac{\partial L}{\partial c_t} = 2c_t \lambda_t + 2 \sum_{k=1}^n c_k S_{tk} - \mu \left(2c_t + 2 \sum_{k=1}^n c_k T_{kt} \right) = 0 \quad (19)$$

$$t = 1, 2, \dots, n$$

在它的两侧前乘 c_t ，再从 1 到 n 对 t 求和，可证明

$$\mu = \lambda |_{\alpha + \Delta\alpha} \quad (20)$$

符号 $|_{\alpha + \Delta\alpha}$ 表示该量是在设计 $\alpha + \Delta\alpha$ 处的，以下如不引起混乱将省略。条件（19）可写成矩阵形式：

$$\begin{pmatrix} \lambda_1 + S_{11} - \lambda(1 + T_{11}) & S_{21} - \lambda T_{21} & \cdots & S_{n1} - \lambda T_{n1} \\ S_{21} - \lambda T_{21} & \lambda_2 + S_{22} - \lambda(1 + T_{22}) & \cdots & S_{n2} - \lambda T_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} - \lambda T_{n1} & S_{n2} - \lambda T_{n2} & \cdots & \lambda_n + S_{nn} - \lambda(1 + T_{nn}) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (21)$$

注意到 $S_{ij} - \lambda T_{ij}$ 是和 $|\Delta\alpha|$ 同阶小量，定义小量

$$e_{ij} = S_{ij} - \lambda T_{ij} \quad (22)$$

将简化以下的叙述。特征值 λ 是下列行列式为零时的根：

$$D(\lambda) = \begin{vmatrix} \lambda_1 - \lambda + \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1n} \\ \varepsilon_{21} & \lambda_2 - \lambda + \varepsilon_{22} & \cdots & \varepsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \lambda_n - \lambda + \varepsilon_{nn} \end{vmatrix} = 0 \quad (23)$$

利用克莱姆法则, 可展开 $D(\lambda)$ 为

$$D(\lambda) = \prod_{i=1}^n (\lambda_i - \lambda + \varepsilon_{ii}) - \sum_{i=1}^{n-1} \sum_{i=i+1}^n \varepsilon_{ij}^2 \prod_{\substack{k=1 \\ k \neq i, j}}^n (\lambda_k - \lambda + \varepsilon_{kk}) + O(\varepsilon_{ij}^3) \quad (24)$$

其中 $O(\varepsilon_{ij}^3)$ 为和 ε_{ij}^3 同级的量, 为了求出 (23) 的近似根, 下面分成四种来情况讨论。

情形 A: 若所求的特征值靠近问题 (1) 的某一单根 λ_k , 则 (24) 中第一式为和 ε_{ij} 同阶的量, 准确到一阶小量, 有

$$D(\lambda) \approx \prod_{i=1}^n (\lambda_i - \lambda + \varepsilon_{ii}) \quad (25)$$

其根为 $\lambda = \lambda_k + S_{kk} - \lambda_k T_{kk}$ (26)

利用 ΔK 和 ΔM 的泰勒展式:

$$\begin{aligned} \Delta K &= \sum_{j=1}^m \frac{\partial K}{\partial \alpha_j} d\alpha_j + \frac{1}{2} \sum_{s=1}^m \sum_{t=1}^m \frac{\partial^2 K}{\partial \alpha_s \partial \alpha_t} d\alpha_s d\alpha_t + \dots \\ \Delta M &= \sum_{j=1}^m \frac{\partial M}{\partial \alpha_j} d\alpha_j + \frac{1}{2} \sum_{s=1}^m \sum_{t=1}^m \frac{\partial^2 M}{\partial \alpha_s \partial \alpha_t} d\alpha_s d\alpha_t + \dots \end{aligned} \quad (27)$$

式中 m 为设计向量 α 的维数, 将 (27) 代入 (26) 且作一阶近似, 得到

$$\Delta \lambda = \lambda - \lambda_k = S_{kk} - \lambda_k T_{kk} = \sum_{j=1}^m x_k^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_k \frac{\partial M}{\partial \alpha_j} \right) x_k d\alpha_j \quad (28)$$

将 (28) 与 (4) 式对比可见两者给出相同结果。

情形 B: 考虑二阶灵敏度的推导, 约定所求的特征根靠近 λ_k , λ_k 为问题 (1) 的单重根。注意 (24) 式的第二项中凡包含 $(\lambda_k - \lambda + \varepsilon_{kk})$ 的项是比 ε_{ij}^2 高级的小量, 舍去它们就得到 $D(\lambda)$ 的近似值:

$$D(\lambda) = \prod_{i=1}^n (\lambda_i - \lambda + \varepsilon_{ii}) - \sum_{\substack{j=1 \\ j \neq k}}^n \varepsilon_{kj}^2 \prod_{\substack{t=1 \\ t \neq j, k}}^n (\lambda_t - \lambda + \varepsilon_{tt}) \quad (29)$$

特征值应满足的方程可近似为:

$$\lambda_k - \lambda + \varepsilon_{kk} - \sum_{\substack{j=1 \\ j \neq k}}^n \frac{\varepsilon_{kj}^2}{(\lambda_j - \lambda + \varepsilon_{jj})} = 0 \quad (30)$$

注意到式 (22) 后可求得:

$$\lambda = \left[\lambda_k + S_{kk} - \sum_{\substack{j=1 \\ j \neq k}}^n \frac{\varepsilon_{kj}^2}{(\lambda_j - \lambda + \varepsilon_{jj})} \right] (1 + T_{kk})^{-1}$$

准确到二阶小量, 有

$$\lambda = \lambda_k + S_{kk} - \lambda_k T_{kk} + \lambda_k T_{kk}^2 - S_{kk} T_{kk} - \sum_{\substack{j=1 \\ j \neq k}}^n \frac{\varepsilon_{kj}^2}{(\lambda_j - \lambda_k)}$$

将 S_{kk} , T_{kk} 及 ε_{kj} 等的表达式代入, 并截断到 $d\alpha_j$ 的二阶项, 可得到用特征向量 X_j 表示的入的近似根。以靠近 λ_1 的根为例, 它和 λ_1 之差可写成:

$$\begin{aligned} \Delta\lambda_1 = & \sum_{j=1}^m \frac{\partial\lambda_1}{\partial\alpha_j} d\alpha_j + \sum_{j=1}^m \sum_{k=1}^m \frac{1}{2} X_1^T \left[\frac{\partial^2 K}{\partial\alpha_j \partial\alpha_k} - \lambda_1 \frac{\partial^2 M}{\partial\alpha_j \partial\alpha_k} \right] X_1 d\alpha_j d\alpha_k \\ & + \lambda_1 \sum_{j=1}^m \sum_{k=1}^m X_1^T \frac{\partial M}{\partial\alpha_j} X_1 X_1^T \frac{\partial M}{\partial\alpha_k} X_1 d\alpha_j d\alpha_k - \sum_{j=1}^m \sum_{k=1}^m X_1^T \frac{\partial M}{\partial\alpha_j} X_1 X_1^T \frac{\partial K}{\partial\alpha_k} X_1 d\alpha_j d\alpha_k \\ & + \sum_{j=1}^m \sum_{k=1}^m \sum_{s=3}^n \frac{X_1^T \left(\frac{\partial K}{\partial\alpha_j} - \lambda_1 \frac{\partial M}{\partial\alpha_j} \right) X_s X_1^T \left(\frac{\partial K}{\partial\alpha_k} - \lambda_1 \frac{\partial M}{\partial\alpha_k} \right) X_s}{(\lambda_1 - \lambda_s)} d\alpha_j d\alpha_k \end{aligned} \quad (31)$$

将 (31) 与 (13) 作一比较可见这儿给出的二阶敏感度表达式和用常用方法导出的是一致的。

情形 C: 设所求的特征值是靠近二重特征值 λ_1 和 λ_2 ($\lambda_1 = \lambda_2$) 的根。此时 (24) 的第一项是和 $|\Delta\alpha|^2$ 同阶的, 第二项中只有 $\varepsilon_{12} \prod_{k=3}^n (\lambda_k - \lambda + \varepsilon_{kk})$ 是和 $|\Delta\alpha|^2$ 同阶的, 其余和 $|\Delta\alpha|^3$ 同阶。保留和 $|\Delta\alpha|^2$ 同阶的项可得特征值满足的方程近似为:

$$(\lambda_1 - \lambda + \varepsilon_{11})(\lambda_1 - \lambda + \varepsilon_{22}) - \varepsilon_{12} = 0 \quad (32)$$

如果采用此式求靠近 λ_1 的两个根时可以发现得到的 $\Delta\lambda$ 的一阶近似不能表示成微分形式, 即不能表示成自变量的微分 $d\alpha_j$ 和常数 (不依赖于 $d\alpha_j$) 的乘积, 这就是重特征值不可微的特性。但是如果选择 X_1 和 X_2 使得

$$X_1^T (\Delta K - \lambda_1 \Delta M) X_2 = 0 \quad (33)$$

或近似为:
$$\sum_{j=1}^m X_1^T \left(\frac{\partial K}{\partial\alpha_j} - \lambda_1 \frac{\partial M}{\partial\alpha_j} \right) X_2 d\alpha_j = 0 \quad (33a)$$

则 (32) 成为 $(\lambda_1 - \lambda + \varepsilon_{11})(\lambda_1 - \lambda + \varepsilon_{22}) = 0$, 在准确到一阶小量的条件下可求得靠近 λ_1 的两个根:

$$\lambda' = \lambda_1 + \sum_{j=1}^m X_1^T \left(\frac{\partial K}{\partial\alpha_j} - \lambda_1 \frac{\partial M}{\partial\alpha_j} \right) X_1 d\alpha_j \quad (34)$$

$$\lambda'' = \lambda_1 + \sum_{j=1}^m X_2^T \left(\frac{\partial K}{\partial\alpha_j} - \lambda_1 \frac{\partial M}{\partial\alpha_j} \right) X_2 d\alpha_j$$

它们和 (28) 具有相同的形式, 但使用时有重要差别, 对给定的一组设计变量增量 $d\alpha_1, d\alpha_2, \dots, d\alpha_m$ 计算重特征值的一阶灵敏度时必须先求出满足条件 (33) 的一对极限向量 X_1 和 X_2 , 才能利用它们按 (34) 式计算灵敏度。

情形 D: 考虑二重特征值的二阶灵敏度。如果已经选择了满足条件 (33) 的一对特征向量 X_1 和 X_2 , 则保留到三阶小量就得到 $D(\lambda)$ 的如下近似:

$$D(\lambda) \doteq \prod_{i=1}^n (\lambda_i - \lambda + \varepsilon_{ii}) - \sum_{j=3}^m \varepsilon_{1j}^2 \prod_{\substack{k=2 \\ k \neq j}}^n (\lambda_k - \lambda + \varepsilon_{kk}) - \sum_{j=3}^m \varepsilon_{2j}^2 \prod_{\substack{k=1 \\ k \neq j}}^n (\lambda_k - \lambda + \varepsilon_{kk}) \quad (35)$$

$D(\lambda) = 0$ 可近似为:

$$\lambda_1 - \lambda + \varepsilon_{11} - \sum_{j=3}^m \frac{\varepsilon_{1j}^2}{(\lambda_j - \lambda + \varepsilon_{jj})} - \sum_{j=3}^m \frac{\varepsilon_{2j}^2 (\lambda_1 - \lambda + \varepsilon_{11})}{(\lambda_1 - \lambda + \varepsilon_{22}) (\lambda_j - \lambda + \varepsilon_{jj})} = 0 \quad (36)$$

@如果 $\varepsilon_{11} \approx \varepsilon_{22}$ 且所寻求的特征值靠近 $\lambda_1 + S_{11} - \lambda_1 T_{11}$, 则 $\lambda_1 - \lambda + \varepsilon_{11}$ 为 $O(|\Delta\alpha|^2)$ 而 $\lambda_1 - \lambda + \varepsilon_{22}$ 为 $O(|\Delta\alpha|)$, 所以与前面各项比较最后一项是高阶量, 可予舍去, 得到在准确到二阶小量的条件下靠近 $\lambda_1 + S_{11} - \lambda_1 T_{11}$ 的根满足:

$$\lambda_1 - \lambda + \varepsilon_{11} - \sum_{j=3}^n \frac{\varepsilon_{1j}^2}{(\lambda_j - \lambda + \varepsilon_{jj})} = 0 \quad (37)$$

它和(30)的唯一差别是求和号中的下标 j 从 3 开始。回到 (31) 式, 则该式最后一项的最内层对 S 求和应当从 3 开始, 换句话说, 在常用的二阶灵敏度公式 (13) 中应舍弃分母为零的项, 当然仍注意 X_1 和 X_2 应满足条件 (33)。

④如果 $\varepsilon_{11} = \varepsilon_{22}$, 则 (36) 式的最后一项也是二阶小量, 在准确到一阶小量的条件下靠近 λ_1 的特征值仍为重根, 准确到二阶时它应满足:

$$\lambda_1 - \lambda + \varepsilon_{11} - \sum_{j=3}^n \left[\frac{1}{(\lambda_j - \lambda + \varepsilon_{jj})} \sum_{i=1}^2 \varepsilon_{ij}^2 \right] = 0 \quad (38)$$

回到 (13) 式, 当前者取值为 λ_1 的特征值之一的二阶灵敏度为:

$$\begin{aligned} \frac{\partial^2 \lambda_1}{\partial \alpha_j \partial \alpha_t} = & X_1^T \left[\frac{\partial^2 K}{\partial \alpha_j \partial \alpha_t} - \lambda_1 \frac{\partial^2 M}{\partial \alpha_j \partial \alpha_t} \right] X_1 + 2\lambda_1 X_1^T \frac{\lambda M}{\partial \alpha_t} X_1 X_1^T \frac{\partial M}{\partial \alpha_j} X_1 \\ & - X_1^T \frac{\partial M}{\partial \alpha_t} X_1 X_1^T \frac{\partial K}{\partial \alpha_j} X_1 - X_1^T \frac{\partial M}{\partial \alpha_j} X_1 X_1^T \frac{\partial K}{\partial \alpha_t} X_1 \\ & + \sum_{k=3}^n \left[\frac{2}{(\lambda_1 - \lambda_k)} X_1^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_1 \frac{\partial M}{\partial \alpha_j} \right) (X_1 X_1^T + X_2 X_2^T) \left(\frac{\partial K}{\partial \alpha_t} - \lambda_1 \frac{\partial M}{\partial \alpha_t} \right) X_k \right] \quad (39) \end{aligned}$$

二重特征值中的另一个的二阶灵敏度也可用 (39) 式计算, 只要将前四项中的特征向量 X_1 换成 X_2 。

四、序列二次规划法

相当数量的寻求特征值最大的最优结构设计问题涉及重模式设计, 即最优结构具有多重模式特征值。以常用的灵敏度分析公式为基础的数值算法很难用来求出这种设计, 特别在迭代过程中求得的当前设计已经是重模式设计时, 由于没有选择正确的特征向量来计算灵敏度, 迭代记录会出现摆动。上面给出的灵敏度分析公式为较好的算法提供了一个基础。为了确定起见, 我们来考虑下列优化问题:

求设计变量 α ,

$$\begin{aligned} \max \quad & \lambda_1 \\ \text{s.t.} \quad & g_j(\alpha) \leq 0, \quad j=1, 2, \dots, J \end{aligned} \quad (40)$$

其中 λ_1 是问题 (1) 的最小特征值。

为了简单起见, 假定已经采用现存的某一算法求出了一个设计 $\alpha^{(k)}$, 上标 k 表示这是第 k 次迭代的结果。为了使得下一轮迭代求得的设计 $\alpha^{(k)} + \Delta\alpha$ 具有更高的最小特征值 λ_1 , 则当 $\alpha^{(k)}$ 是单模式设计时求解下列二次规划:

$$\begin{aligned} \text{求} \quad & \Delta\alpha \\ \max \quad & \Delta\lambda_1 \end{aligned}$$

$$s.t. \quad g_j(\alpha^{(k)}) + \nabla^T g_j(\alpha^{(k)}) \Delta \alpha \leq 0, \quad j=1, 2, \dots, J$$

其中 $\Delta \lambda_1$ 由 (31) 式给出, 但需要将 $d\alpha$ 记号作更动为 $\Delta \alpha$, $\nabla^T g_j(\alpha^{(k)})$ 是约束函数 $g_j(\alpha)$ 在设计 $\alpha^{(k)}$ 处的梯度。

当 $\alpha^{(k)}$ 是二重模式设计时, 则求解下列二次规划:

求 $\Delta \alpha$

$$\max \Delta \lambda_1$$

$$s.t. \quad \sum_{j=1}^m X_1^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_1 \frac{\partial M}{\partial \alpha_j} \right) X_1 \Delta \alpha_j = \sum_{j=1}^m X_2^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_1 \frac{\partial M}{\partial \alpha_j} \right) X_2 \Delta \alpha_j \quad (42a)$$

$$\sum_{j=1}^m X_1^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_1 \frac{\partial M}{\partial \alpha_j} \right) X_2 \Delta \alpha_j = 0 \quad (42b)$$

$$g_j(\alpha^{(k)}) + \nabla^T g_j(\alpha^{(k)}) \Delta \alpha \leq 0, \quad j=1, 2, \dots, J \quad (42c)$$

其中 $\Delta \lambda_1$ 中的二阶偏导采用 (39) 式。由 (41) 或 (42) 求得 $\Delta \alpha$ 后便可令 $\alpha^{(k+1)} = \alpha^{(k)} + \Delta \alpha$ 重复上列过程, 即重复地执行一系列的二次规划直到满足某种收敛准则为止。

上面给出的 (41)、(42) 的目标函数中亦可将二次项去掉, 只保留线性项, 这就得了序列线性规划算法。无论是序列二次规划还是序列线性规划, 均可采用运动极限或进行粗糙一维搜索^[5]来控制设计的修改量, 以保证迭代历史的稳定上升。

需要指出的是, 有的文献中^[2], 考虑到重特征值的不可微特征, 在传统的特征值优化问题的迭代算法中, 增加了一个内循环, 用来确定和设计变量增量 $\Delta \alpha$ 相应的特征向量 X_1 和 X_2 , 以满足条件 (33a), 这就使得算法变得十分复杂, 造成计算上的困难。本文建议的序列规划法并不需要这种内循环。容易证明, 如果对于任意的一对满足 M-正交、归一化要求的特征向量 X_1 和 X_2 , 求得的设计增量 $\Delta \alpha$ 满足 (42b) 和 (42c) 两式, 则可以采用任意一个归一化的特征向量 X 按 (4) 计算特征值的一阶增量, 而无需另行寻求和 $\Delta \alpha$ 相应的特征向量。事实上, 设和 $\Delta \alpha$ 相应的极限向量为 X_1 :

$$X_1 = aX_1 + bX_2 \quad a^2 + b^2 = 1$$

则利用 (42b), (42c) 可以证明

$$\sum_j X_1^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_1 \frac{\partial M}{\partial \alpha_j} \right) X_1 \Delta \alpha_j = \sum_j X_1^T \left(\frac{\partial K}{\partial \alpha_j} - \lambda_1 \frac{\partial M}{\partial \alpha_j} \right) X_1 \Delta \alpha_j$$

这一事实可以说成在设计空间中沿着重模式设计的子空间, 传统的一阶敏度计算公式 (4) 恢复了其正确性与特征向量的选择无关。看来正是由于这一特点, Olhoff 的迭代格式仍然给出了正确的结果^[6]。

数值试验已经证明了上列算法的有效性。

五、对称与非对称柱的数值结果

考虑两端受到集中力 P 的柱的稳定性。约定柱的材料体积 V 已经给定, 要求确定沿柱长的断面积分布以使失稳临界力 P 尽可能地高。在下列计算中认为柱的横断面为矩形, 它的宽度 b 已给定, h 可变, 取作设计变量, 只考虑沿高度方向的失稳 (宽度方向可认为有足够的横向支承防止其横向失稳)。考虑了两种端部条件: 夹住端-夹住端; 两端均不允许转

动,但支承在弹簧刚度分别为 k_1 和 k_2 的两个线性弹簧上,当 $k_1 \neq k_2$ 时,这个柱是非对称支承的。

采用有限元法求解这个问题时, M 阵与设计变量无关, K 阵由单元刚度阵 K_i 迭加而成,

$$K = \sum K_i h_i$$

其中 \bar{K}_i 为与 h_i 无关的比刚度阵, i 指第 i 个单元的量。以前给出的敏度表达式可以进一步具体化。按上面的方法进行数值计算时发现,特征值的二阶敏度组成的海森阵不一定正定,采用通常的三次规划求解器求解 (41)、(42) 遇到困难。我们在目标项上增加阻尼阵 CI , 其中 C 为一个适当大的数,随迭代进展而减小, I 为 $m \times m$ 阶单位阵,在求解 (41)、(42) 得到 $\Delta\alpha$ 后将它看作给出了一个搜索方向,改进的设计 α^{k+1} 是由沿 $\Delta\alpha$ 作一维搜索而求得的。

在 TQ-16 机上编制了上列算法的程序,由于计算机的条件,将全柱分割为 24 个单元,在柱长 $\frac{1}{4}$ 及 $\frac{3}{4}$ 处局部都加密,对非对称支承的柱,单元划分也是不对称的。在两端夹住柱的情况下,计算中并未强加对称性条件,但采用下列算法时,迭代得到的设计始终是对称的,在两种

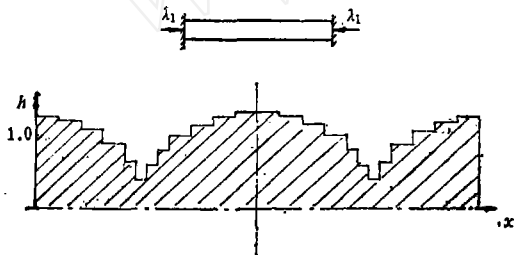


图 1. 两端夹住柱的最大临界力设计, $E=20000$, $b=1$, $l=12$, $V=12$, $P_{cr}=624.68002$, 次大特征值为 624.68014, $\lambda_1=624.68002$, $\lambda_2=624.68014$, $J=ch^3$, $A=bh$

支承条件下,柱的失稳临界力随迭代进展而稳定上升,没有遇到文献 [2] 中报导的困难,只是在当第一特征值与第二特征值相遇后,迭代进展很缓慢。鉴于这样的计算暂时局限于专题讨论,本文未仔细研究提高效率的途径。

图 1 和 2 给出了计算的两个结果。在两端夹住时,半段柱的断面积分布对柱长 $1/4$ 处不对称,这是和单模式设计的明显差别。图 2 给出的为非对称设计,图 3 为其迭代历史。值得一提,文献中迄今未见非对称结构的重模式设计,本文的算法与结构的对称性无关,显示了其优越性。

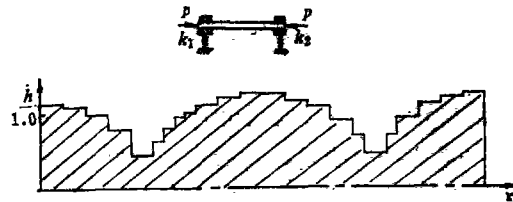


图 2. 两端弹性支柱的最大临界力设计, $K_1=2000$, $K_2=400$, 其它数据同图 1, 迭代最终结果, $\lambda_1=619.114120$, $\lambda_2=619.114335$ 单元分割不对称

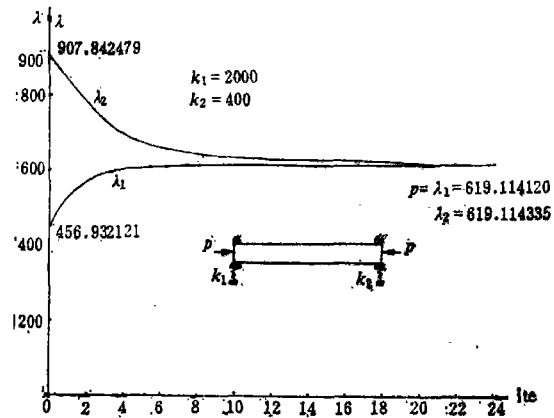


图 3. 图 2 所示例题的临界力迭代记录, 横坐标为迭代次数, 纵坐标为临界力。

六、推广到多重模式最优设计

设当前设计 $\alpha^{(k)}$ 是 n_0 重设计, 要求新设计 $\alpha^{(k)} + \Delta\alpha$ 在一阶近似范围内仍是 n_0 重模

式设计, 特征值的二阶敏度公式将为:

$$\lambda_1 T_{11}^2 - S_{11} T_{11} - \sum_{j=n_0+1}^n \sum_{i=1}^{n_0} \varepsilon_{ij}^2 / (\lambda_j - \lambda_1) \quad (43)$$

这里特征向量和设计的修改量应满足

$$\begin{aligned} \varepsilon_{ij} &= 0 \quad (i=1, 2, \dots, n_0-1; j=i+1, \dots, n) \\ \varepsilon_{11} &= \varepsilon_{22} = \dots = \varepsilon_{n_0 n_0} \end{aligned} \quad (44)$$

将 (44) 处理作约束条件, (43) 处理作目标函数便可建立求 $\Delta \alpha$ 的二次规划问题。

七、结 束 语

上面针对有限维空间, 利用瑞雷商驻值原理推导了重特征值的二阶敏度公式, 设计变量与结构自由度数均是有限个数。但是, 这里给出的基本思想可运用到设计变量或结构自由度无限多的分布参数结构优化问题中去。按上面算法给出的柱的设计是否是全局最优设计, 尚需进一步的数值试验

参 考 文 献

- [1] Wittrick, W. H., Rate of change of eigenvalues with reference to buckling and vibration problems, *Journal of the Royal Aeronautical Society*, **66**, September (1962), 590.
- [2] Choi, K. K. and Haug, E. J., Optimization of Distributed Parameter Structures, Sijthoff & Noordhoff (1981), 219—277.
- [3] 程耿东, 钟万勰, 关于两端固定柱抗失稳优化问题的注记, *力学学报*, **27** 1 (1983)
- [4] 钱令希, 钟万勰, 程耿东, 隋允康, 工程结构优化设计的一个途径——序列二次规划 SQP, *计算结构力学及其应用*, 1 (1984)。
- [5] 程耿东, 汪榴, 形状优化数值方法与粗糙的一维搜索策略, *Engineering Optimization*, (待发表) 1983 年 11 月
- [6] Olhoff, N. and Rasmussen, S. H., On Single and Bimodal Optimum Buckling Loads of Clamped Columns, *Int. J. Solids & Stru.*, **13**, (1977) 605—614

Sensitivity Analysis of Multimodal Eigenvalues and Related Optimization Techniques

Zhong Wanxie Cheng Gengdong

(Research Institute of Engineering Mechanics Dalian Institute of Technology)

Abstract

Based on the Rayleigh Quotient principle, this paper presents an analysis for the sensitivity of eigenvalues. Truncating the characteristic equation up to the different order, we have obtained the first order and the second order sensitivity expressions both for the single and multimodal eigenvalues. The difference between the new and the traditional derivation is pointed and discussed. By making use of the new sensitivity expressions, an effective Sequential Quadratic programming technique is suggested to solve the structural optimization problems, in which the double or multimodal eigenvalues are involved. The notable feature of the SQP technique is no need of any inner iteration to determine the eigenvectors associated with the undetermined design changes. In this way, the new algorithm is much more simple than some existing algorithms in literatures. Numerical results are shown for the optimal design of columns against buckling. The columns are either clamped-clamped or elastically supported at the two ends and are acted on by two concentrated forces at the two ends. Constraints are imposed on the material volume. It is shown that the new algorithm enables us to obtain bimodal design with unsymmetric thickness distribution.