

求解大规模 Hamilton 矩阵特征问题的辛 Lanczos 算法的误差分析*

闫庆友^{1,2}, 魏小鹏³

(1. 山东财政学院经济统计系, 山东 济南 250014; 2. 华北电力大学工商管理学院, 北京 102206;
3. 大连大学先进设计技术中心, 辽宁 大连 116622)

摘 要: 对求解大规模稀疏 Hamilton 矩阵特征问题的辛 Lanczos 算法给出了舍入误差分析. 分析表明辛 Lanczos 算法在无中断时, 保 Hamilton 结构的限制没有破坏非对称 Lanczos 算法的本质特性. 本文还讨论了辛 Lanczos 算法计算出的辛 Lanczos 向量的 J-正交性的损失与 Ritz 值收敛的关系. 结论正如所料, 当某些 Ritz 值开始收敛时, 计算出的辛 Lanczos 向量的 J-正交性损失是必然的. 以上结果对辛 Lanczos 算法的改进具有理论指导意义.

关键词: 辛 Lanczos 算法; Hamilton 矩阵; 特征值; 误差分析; Ritz 值; Ritz 向量.

分类号: AMS(2000) 65G05, 65F15, 65F50/CLC number: O241.6

文献标识码: A **文章编号:** 1000-341X(2004)01-0091-16

1 引言

1950 年, Lanczos 提出的 Lanczos 算法^[13]是一个逐渐把一个一般矩阵 $A \in \mathbb{R}^{n \times n}$ 化为一个三对角矩阵 T 的过程. 在 Lanczos 过程执行第 j 步后, Lanczos 算法产生了两个 $n \times j$ 阶矩阵 Q_j, P_j . $Q_j = [q_1, q_2, \dots, q_j]$, $P_j = [p_1, p_2, \dots, p_j]$ 满足 $P_j^T Q_j = I$, 且

$$AQ_j = Q_j T_j + \beta_{j+1} q_{j+1} e_j^T, \quad A^T P_j = P_j T_j^T + \gamma_{j+1} p_{j+1} e_j^T. \quad (1)$$

这里 $e_j = [0, 0, \dots, 1]^T \in \mathbb{R}^j$, T_j 为三对角矩阵, $T_j = \begin{bmatrix} \alpha_1 & \gamma_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \gamma_j \\ & & \beta_j & \alpha_j \end{bmatrix}$.

小矩阵 T_j 的某些特征值随 j 的增大可以是矩阵 A 的某些特征值的好的近似. 在整个的迭代过程中, 只用到了矩阵和向量的积 Ax 和 $A^T x$, 因此此算法特别适合求大型稀疏矩阵的部分特征值. 从 1960 年至今出现了大量的研究 Lanczos 算法的文章, 如[1, 11, 17].

最近二三十年, 出现了一些研究保结构的计算 Hamilton 矩阵特征值的文章^[5, 6, 7]. 对于大

* 收稿日期: 2002-01-20

基金项目: 国家自然科学基金资助(50275013, G60174037)

作者简介: 闫庆友(1963-), 博士, 山东财政学院教授.

型稀疏 Hamilton 矩阵特征问题的求解 P. Benner 和 H. FaBbender 在文[3,4]中提出了一个隐式重新开始的辛 Lanczos 算法. 大型稀疏 Hamilton 矩阵的特征问题求解有广泛的实际意义. 例如, 求解连续时间的代数 Riccati 方程; 计算矩阵的 H_∞ 范数; 分析计算化学中的相应理论^[15]; 求解最优控制问题. 很多情况下产生的 Hamilton 矩阵为大型的稀疏的. 而且应用中只需要部分特征值及其对应的不变子空间.

保结构求解 Hamilton 矩阵问题的 Lanczos 算法(辛 Lanczos 算法)是 P. Benner 和 H. FaBbender 于 1995 提出的^[4]. 此算法逐次地把 Hamilton 矩阵约化为 Hamilton J-三对角矩阵. 文[3,4]中作者利用了隐式重新开始的思路解决辛 Lanczos 算法面临的数值问题. 对辛 Lanczos 算法的中断问题, R. Freund 和 V. Mehrmann 提出了 Look-ahead 辛 Lanczos 算法^[10].

本文对于求解 Hamilton 矩阵问题的简单的辛 Lanczos 算法在无中断的情况下进行了误差分析. 分析表明保结构的辛 Lanczos 具有非对称 Lanczos 算法某些的数值特性. 整个分析是按照 Bai^[1], Paige^[16]及 H. Fassbender^[8]的思路进行的.

2 求解 Hamilton 矩阵特征问题的辛 Lanczos 算法及其性质

给定 Hamilton 矩阵 $H \in \mathbb{H}^{2n \times 2n}$, 初始向量 $s_1 \in \mathbb{R}^{2n}$, 辛 Lanczos 算法产生一系列 Hamilton J-三对角矩阵 $\tilde{H}^{2k}$ (如果无中断发生), 满足

$$HS^{2k} = S^{2k}\tilde{H}^{2k} + \xi_{k+1}v_{k+1}e_{2k}^T, \quad (2)$$

其中 $S^{2k} \in \mathbb{R}^{2n \times 2k}$, $S^{2k}e_1 = s_1$, $(S^{2k})^T J^{2n} S^{2k} = J^{2k}$ (J-正交). 当 $k=n$ 时, 该算法生成一个辛矩阵 S , 满足 $S^{-1}HS = \tilde{H}^{2n}$, 其中 $\tilde{H}^{2n}$ 为下列 Hamilton J-三对角矩阵.

$$\tilde{H}^{2n} = S^{-1}HS = \begin{bmatrix} \delta_1 & & & & \beta_1 & \xi_2 & & & \\ & \delta_2 & & & \xi_2 & \beta_2 & \xi_3 & & \\ & & \delta_3 & & \xi_3 & \ddots & \ddots & & \\ & & & \ddots & & \ddots & \ddots & \xi_n & \\ & & & & \delta_n & & \xi_n & \beta_n & \\ \gamma_1 & & & & -\delta_1 & & & & \\ & \gamma_2 & & & & -\delta_2 & & & \\ & & \gamma_3 & & & & -\delta_3 & & \\ & & & \ddots & & & & \ddots & \\ & & & & \gamma_n & & & & -\delta_n \end{bmatrix}. \quad (3)$$

为了叙述方便, 定义如下置换矩阵. 设 $e_1, e_2, \dots, e_{2n-1}, e_{2n}$ 为 $\mathbb{R}^{2n}$ 的坐标基, 定义置换矩阵 $P^{2n} = [e_1, e_3, \dots, e_{2n-1}, e_2, e_4, \dots, e_{2n}] \in \mathbb{R}^{2n \times 2n}$. 称 $H_P = P^{2n}H(P^{2n})^T$, $S_P^{2n} = P^{2n}S^{2n}(P^{2n})^T$, $\tilde{H}_P^{2n} = P^{2n}\tilde{H}^{2n}(P^{2n})^T$ 分别为置换版本的 Hamilton 矩阵和置换版本的辛矩阵 (J-正交阵) 和置换版的 Hamilton J-三对角矩阵.

特别值得指出的是

$$J_P^{2n} = P^{2n}J^{2n}(P^{2n})^T = \text{diag}\left(\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right). \quad (4)$$

$$\tilde{H}_P^{2n} = S_P^{-1} H_P S_P = \begin{bmatrix} \delta_1 & \beta_1 & 0 & \xi_2 & & & & & \\ \gamma_1 & -\delta_1 & 0 & 0 & 0 & & & & \\ & \xi_2 & \delta_2 & \beta_2 & 0 & \xi_3 & & & \\ & & \gamma_2 & -\delta_2 & 0 & 0 & \ddots & & \\ & & & \xi_3 & \ddots & \ddots & \ddots & \xi_{n-1} & \\ & & & & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & & & \ddots & \ddots & \beta_{n-1} & 0 & \xi_n \\ & & & & & & \ddots & \ddots & 0 & 0 \\ & & & & & & & \xi_n & \delta_n & \beta_n \\ & & & & & & & & \gamma_n & -\delta_n \end{bmatrix} \quad (5)$$

为一特殊的 Hessenberg 矩阵. 因为 $S^{-1}HS = \tilde{H}$, 所以有

$$H_P S_P = S_P \tilde{H}_P = S_P \begin{bmatrix} \delta_1 & \beta_1 & 0 & \xi_2 & & & & & \\ \gamma_1 & -\delta_1 & 0 & 0 & 0 & & & & \\ & \xi_2 & \delta_2 & \beta_2 & 0 & \xi_3 & & & \\ & & \gamma_2 & -\delta_2 & 0 & 0 & \ddots & & \\ & & & \xi_3 & \ddots & \ddots & \ddots & \xi_{n-1} & \\ & & & & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & & & \ddots & \ddots & \beta_{n-1} & 0 & \xi_n \\ & & & & & & \ddots & \ddots & 0 & 0 \\ & & & & & & & \xi_n & \delta_n & \beta_n \\ & & & & & & & & \gamma_n & -\delta_n \end{bmatrix}. \quad (6)$$

保结构的辛 Lanczos 算法^[3,4]生成一系列矩阵 $S_P^{2k} = [v_1, w_1, v_2, w_2, \dots, v_k, w_k] \in \mathbb{R}^{2n \times 2k}$, 满足

$$H_P S_P^{2k} = S_P^{2k} \tilde{H}_P^{2k} + \xi_{k+1} v_{k+1} e_{2k}^T, \quad (7)$$

其中 $\tilde{H}_P^{2k} = P^{2k} \tilde{H}^{2k} (P^{2k})^T$ 为 $2k \times 2k$ 阶 Hamilton J-三对角矩阵. 因为 $(S_P^{2k})^T J_P^{2k} S_P^{2k} = J_P^{2k}$, 容易证明 $S^{2k} = (P^{2k})^T S_P^{2k}$ 的列向量张成的空间是辛空间.

由于这一约化过程非常强地依赖于约化矩阵的第一列, 可以预见可能会发生的中断和近似中断. 假设无中断发生, $S_P = [v_1, w_1, v_2, w_2, \dots, v_n, w_n]$. 给定初始向量 v_1 , 辛 Lanczos 算法通过下列方程逐列地形成

$$H_P S_P e_j = S_P \tilde{H}_P e_j, \quad j = 1, 2, \dots. \quad (8)$$

由此得到下列辛 Lanczos 算法^[3]

算法: 辛 Lanczos 算法

1. 开始: 选择初始向量 $\tilde{v}_1 \in \mathbb{R}^{2n}$, $\tilde{v}_1 \neq 0$, 令 $v_0 = 0 \in \mathbb{R}^{2n}$, $\xi_1 = \|\tilde{v}_1\|_2$, $v_1 = \frac{1}{\xi_1} \tilde{v}_1$.

2. 迭代: 对 $j = 1, 2, \dots$

(a) $\gamma_j = v_j^T J_P H_P v_j$,

(b) $\tilde{w}_j = H_P v_j - \delta_j v_j$,

(c) $w_j = \frac{\tilde{w}_j}{\gamma_j}$,

- (d) $\beta_j = -w_j^T J_P H_P w_j$,
 (e) $\tilde{v}_{j+1} = H_P w_j - \xi_j v_{j-1} - \beta_j v_j + \delta_j w_j$,
 (f) $\xi_{j+1} = \|\tilde{v}_{j+1}\|_2$,
 (g) $v_{j+1} = \frac{1}{\xi_{j+1}} \tilde{v}_{j+1}$.

算法中的 v_j 及 w_j 称为辛 Lanczos 向量, 在计算辛 Lanczos 向量 v_j 及 w_j 的过程中只用到了矩阵向量的积. 有效的执行这一算法的运算量为 $6n + (4nz + 32)k$, 其中的 nz 为 H_P 中非零元的个数, $2k$ 为计算出的辛 Lanczos 向量的个数. 算法中的参数的选取有一定自由度^[3]. 特别地 δ_j 可以自由选取. 本文的讨论中, δ_j 始终取为 1. 类似地 ξ_j 和 γ_j 也有其它选取方式^[3].

上述算法有可能发生中断和严重失稳现象, 因为 γ_j 和 ξ_{j+1} 都可能等于 0. 如果 $\tilde{w}_j = 0$ 或接近 0, 容易证明此时 $\gamma_j = 0$ 或接近 0, 将得到 H_P 的一个 $2j-1$ 维的不变空间 (或一个 $2j-1$ 维不变子空间的好的逼近). 而且, 非常容易的从 $\tilde{H}^{2j-2}$ 得到两个特征值和该特征值的左、右特征向量. 如果因为 $\tilde{v}_{j+1} = 0$ 或接近 0 导致 $\xi_{j+1} = 0$ 或接近 0, 将得到 H_P 的一个 $2j$ 维的辛不变空间 (或一个 $2j$ 维辛不变子空间的好的逼近). 此时只要任取非 0 向量 $\tilde{v}_{j+1}$ 满足

$$v_i^T J_P \tilde{v}_{j+1} = 0, \quad w_i^T J_P \tilde{v}_{j+1} = 0, \quad i = 1, 2, \dots, j, \quad (9)$$

算法将继续下去. 不过得到的 J-三对角 Hamilton 矩阵将是可约的. 从而原始问题可转化为两个小问题来处理. 不论是 $\tilde{w}_j = 0$ 还是 $\tilde{v}_{j+1} = 0$, 中断都是良性的. 如果 $\tilde{w}_{j+1}$ 和 $\tilde{v}_{j+1}$ 都不为 0, 但 $\gamma_{j+1} = 0$, 则中断是恶性的. 此时以 v_1 为第 1 列的把 H 约化为 J-三对角 Hamilton 矩阵的辛变换是不存在的. 实际上, 数值上稳定的把 H 约化为 J-三对角 Hamilton 矩阵的辛变换的是否存在非常强烈地依赖于 v_1 ^[3]但在本文的讨论中我们总是在无中断的假设下进行的.

在执行了 k 步辛 Lanczos 算法后有

$$H_P S_P^{2k} = S_P^{2k} \tilde{H}^{2k} + \xi_{k+1} v_{k+1} e_{2k}^T. \quad (10)$$

置换还原后为 $HS^{2k} = S^{2k} \tilde{H}^{2k} + P^T \xi_{k+1} v_{k+1} e_{2k}^T P = S^{2k} \tilde{H}^{2k} + \xi_{k+1} \tilde{v}_{k+1} e_{2k}^T$. 如果剩余向量的模小, 则 $\tilde{H}^{2k}$ 的 $2k$ 个特征值可能是 H 的特征值的近似. 数值实验表明剩余向量的范数小是很少出现的. 然而 $\tilde{H}^{2k}$ 某些特征值仍然可能是 H 的某些特征值 (常常是依模最大的) 的好的近似. 设 λ 是 $\tilde{H}$ 的一个特征值, y 为其对应的特征向量, 则向量 $x = S^{2k} y$ 满足

$$\begin{aligned} \|H - \lambda x\|_2 &= \|(HS^{2k} - S^{2k} \tilde{H}^{2k})y\|_2 = |\xi_{k+1}| \|e_{2k}^T y\| \|\tilde{v}_{k+1}\|_2 \\ &= |\xi_{k+1}| \|e_{2k}^T y\| \|v_{k+1}\|_2. \end{aligned} \quad (11)$$

向量 x 称为 H 的 Ritz 向量, λ 称为 H 的 Ritz 值. 如果特征向量向量 y 的最后一个分量足够小, 以至于使 (11) 的右端小, 则 $\{\lambda, x\}$ 是 H 的一个特征对的好的逼近. $\{\lambda, x\}$ 是下一列邻近问题的精确特征解:

$$(H + E)x = \lambda x, \quad (12)$$

其中 $E = -\xi_{k+1} \tilde{v}_{k+1} e_{2k}^T (S^{2k})^T J^{2n}$.

在实际计算中, 为了确定近似特征对的精度, 通常把 $|\xi_{k+1}| \|e_{2k}^T y\| \|v_{k+1}\|_2$ 作为 Ritz 估计, 此估计避免了显式计算 $(HS^{2k} - S^{2k} \tilde{H}^{2k})y$.

对非正规矩阵近似特征对的剩余模不足以介定特征值的误差, 小的 Ritz 估计不足以说明 Ritz 对 $\{\lambda, x\}$ 是 H 的一个特征对的一个好的近似. 即此时 $\min |\lambda - \mu_j|$, 其中 $\mu_j \in \sigma(H)$ 未必小^[12]. 然而它可以介定一个邻近问题, 使 Ritz 三元组 $\{\lambda, x, y\}$ 是此邻近问题的精确解^[12] (这里

y 为 H 的对应 Ritz 值 λ 的左 Ritz 向量).

下面我们给出此误差的可计算的表达式. 假设 $\tilde{H}^{2k}$ 是可对角化的, 即存在 $y \in \mathbb{C}^{2k \times 2k}$, 满足

$$Y^{-1} \tilde{H}^{2k} Y = \text{diag}(\lambda_1, \dots, \lambda_k, -\lambda_1, \dots, -\lambda_k) = \Lambda. \quad (13)$$

设 $X = S^{2k} Y = [x_1, x_2, \dots, x_{2k}]$. 则由 $HS^{2k} = S^{2k} \tilde{H}^{2k} + \xi_{k+1} \tilde{v}_{k+1} e_{2k}^T$ 可推得

$$HX = X\Lambda + \xi_{k+1} \tilde{v}_{k+1} e_{2k}^T Y. \quad (14)$$

即

$$\begin{cases} Hx_i = \lambda_i x_i + y_{2k,i} \xi_{k+1} \tilde{v}_{k+1}, \\ Hx_{k+i} = -\lambda_i x_{k+i} + y_{2k,k+i} \xi_{k+1} \tilde{v}_{k+1}, \end{cases} \quad i = 1, 2, \dots, k. \quad (15)$$

最后一个方程可改写为 $(Jx_{k+i})^T H = \lambda_i (Jx_{k+i})^T + y_{2k,k+i} \xi_{k+1} \tilde{v}_{k+1}^T J$.

利用文^[12]中的定理 2', 可知 $\{\lambda_i, x_i, (Jx_{k+i})^T\}$ 是 $M - F_{\lambda_i}$ 的特征三元组, 其中

$$\|F_{\lambda_i}\|_2 = \|\xi_{k+1}\|_2 \|v_{k+1}\|_2 \max_i \left\{ \frac{|y_{2k,i}|}{\|x_i\|_2}, \frac{|y_{2k,k+i}|}{\|x_{k+i}\|_2} \right\}. \quad (16)$$

而且, 如果 $\|F_{\lambda_i}\|_2$ 足够小, 则 $\{\lambda_i, x_i, (Jx_{k+i})^T\}$ 是原始问题的一个小的扰动问题 $H - F_{\lambda_i}$ 的精确特征三元组. 为了给出 Ritz 值是 H 的某个单特征值 θ_i 的绝对误差, 由 Wilkinson 在文^[18]的论证知, 当 $\|F_{\lambda_i}\|_2$ 是够小时, 有

$$|\theta_i - \lambda_i| \leq \text{Cond}(\lambda_i) \|F_{\lambda_i}\|_2 + O(\|F_{\lambda_i}\|_2^2), \quad (17)$$

其中 $\text{Cond}(\lambda_i)$ 是 H 的 Ritz 值 λ_i 的条件数, $\text{Cond}(\lambda_i) = \frac{\|x_i\|_2 \|Jx_{k+i}\|_2}{|x_{k+i}^T J x_i|} = \|x_i\|_2 \|x_{k+i}\|_2$.

因此辛 Lanczos 过程应在 $\|F_{\lambda_i}\|_2$ 小, 同时 $\text{Cond}(\lambda_i) \|F_{\lambda_i}\|_2$ 小到我们精度标准中止. 然而, 实际执行中所有迭代只要求残量小到给定值, 从不要求 $\text{Cond}(\lambda_i) \|F_{\lambda_i}\|_2$ 充分小, 因为这很难做到, 且经常不可能, 比如当 $\text{Cond}(\lambda_i)$ 很大时. 值得注意的是和 Ritz 估计类似, 中止的准则中, ξ_{k+1} 和 $\tilde{H}^{2k}$ 的特征向量的最后一个分量的大小 $|y_{2k,i}|$ 和 $|y_{2k,k+i}|$ 的大小仍是非常重要的量.

3 有限精度算法下的辛 Lanczos 算法

在本部分, 我们给出在有限精度算法下辛 Lanczos 算法的舍入误差分析.

为了便于分析, 首先叙述一下文献^[11]中的浮点运算模型:

$$\begin{cases} fl(ax + y) = ax + y + e, & |e| \leq u(2|ax| + |y|) + O(u^2); \\ fl(x^T y) = x^T y + e, & |e| \leq ku|x|^T |y| + O(u^2); \\ fl(Ax) = Ax + e, & |e| \leq mu|A||x| + O(u^2). \end{cases} \quad (18)$$

其中 k 为向量 x 和 y 的重叠的非零元的个数, m 为矩阵 A 的列和行中非零元个数的最大值. 如果 $x = [x_1, x_2, \dots, x_n]$, 则 $|x| = [|x_1|, |x_2|, \dots, |x_n|]$; $A = [a_{ij}]_{i=1}^n$, 则 $|A| = [|a_{ij}|]_{i=1}^n$. 在下面的分析中的另一个约定是 $\hat{\alpha}$ 表示 α 的有限精度算法下的计算值.

我们只需分析一步辛 Lanczos 算法即可. 在第 $j-1$ 步执行完以后, 即已经计算出了 $\hat{\gamma}_{j-1}$, $\hat{w}_{j-1}$, $\hat{\beta}_{j-1}$, $\hat{\xi}_j$ 和 $\hat{v}_j$. 我们要计算 $\hat{\gamma}_j$, $\hat{w}_j$, $\hat{\beta}_j$, $\hat{\xi}_{j+1}$ 和 $\hat{v}_{j+1}$. 注意 $\delta_j = 1$, 因此 δ_j 不需要计算.

先分析 $\gamma_j = v_j^T J_P H_P v_j$ 的计算值 $\hat{\gamma}_j$. 由于 J_P 的特殊结构, J_P 乘以任何向量和矩阵都不会产生任何舍入误差. 因此其不影响我们的误差分析. 设 H_P 每列或每行最多含有 m 个非零元素. 计算 $J_P H_P \hat{v}_j$ 有

$$\hat{s}_1 = fl(J_P H_P \hat{v}_j) = J_P H_P \hat{v}_j + e_1, \quad |e_1| \leq mu |J_P H_P| |\hat{v}_j| + O(u^2). \quad (19)$$

计算 γ_j 有

$$\hat{s}_2 = fl(\hat{v}_j^T \hat{s}_1) = \hat{v}_j^T \hat{s}_1 + e_2, \quad |e_2| \leq 2nu |\hat{v}_j|^T |\hat{s}_1| + O(u^2). \quad (20)$$

综合(19),(20)式有

$$\hat{\gamma}_j = \hat{v}_j^T J_P H_P \hat{v}_j + f_j, \quad f_j = \hat{v}_j^T e_1 + e_2, \quad (21)$$

$$|f_j| \leq (m + 2n)u |\hat{v}_j|^T |J_P H_P| |\hat{v}_j| + O(u^2), \quad (22)$$

下面分析 w_j 的计算值

$$\hat{s}_3 = fl(H_P \hat{v}_j) = H_P \hat{v}_j + e_3, \quad |e_3| \leq mu |H_P| |\hat{v}_j| + O(u^2). \quad (23)$$

$$\hat{w}_j = \hat{s}_4 = fl(\hat{s}_3 - \hat{v}_j) = \hat{s}_3 - \hat{v}_j + e_4, \quad |e_4| \leq u(|\hat{v}_j| + |\hat{s}_3|) + O(u^2), \quad (24)$$

$$\hat{w}_j = \hat{s}_5 = fl(\frac{\hat{s}_4}{\hat{\gamma}_j}) = \frac{\hat{s}_4}{\hat{\gamma}_j} + e_5, \quad |e_5| \leq u |\frac{\hat{s}_4}{\hat{\gamma}_j}| + O(u^2). \quad (25)$$

综合以上三式得

$$\hat{\gamma}_j \hat{w}_j = H_P \hat{v}_j - \hat{v}_j + e_3 + e_4 + \hat{\gamma}_j e_5 = H_P \hat{v}_j - \hat{v}_j + h_j, \quad (26)$$

$$|h_j| \leq (m + 2)u |H_P| |\hat{v}_j| + 2u |\hat{v}_j| + O(u^2). \quad (27)$$

β_j 的计算结果分析和 γ_j 的计算分析是一样的,有

$$\hat{\beta}_j = \hat{w}_j^T J_P H_P \hat{w}_j + e_6, \quad |e_6| \leq (m + 2n)u |\hat{w}_j|^T |J_P H_P| |\hat{w}_j| + O(u^2). \quad (28)$$

下面分析 v_{j+1} 的计算值

$$\hat{s}_7 = fl(H_P \hat{w}_j) = H_P \hat{w}_j + e_7, \quad |e_7| \leq mu |H_P| |\hat{w}_j| + O(u^2); \quad (29)$$

$$\hat{s}_8 = fl(\hat{s}_7 - \hat{\xi}_j \hat{v}_{j-1}) = \hat{s}_7 - \hat{\xi}_j \hat{v}_{j-1} + e_8, \quad |e_8| \leq u(2|\hat{\xi}_j \hat{v}_{j-1}| + |\hat{s}_7|) + O(u^2); \quad (30)$$

$$\hat{s}_9 = fl(\hat{s}_8 - \hat{\beta}_j \hat{v}_j) = \hat{s}_8 - \hat{\beta}_j \hat{v}_j + e_9, \quad |e_9| \leq u(2|\hat{\beta}_j \hat{v}_j| + |\hat{s}_8|) + O(u^2); \quad (31)$$

$$\hat{v}_{j+1} = fl(\hat{s}_9 + \hat{w}_j) = \hat{s}_9 + \hat{w}_j + e_{10}, \quad |e_{10}| \leq u(|\hat{s}_9| + |\hat{w}_j|) + O(u^2). \quad (32)$$

综合(29)–(32)式得

$$\hat{v}_{j+1} = H_P \hat{w}_j - \hat{\xi}_j \hat{v}_{j-1} - \hat{\beta}_j \hat{v}_j + \hat{w}_j + e_{11}, \quad (33)$$

其中 $|e_{11}| \leq (m + 3)u |H_P| |\hat{w}_j| + 4u |\hat{\xi}_j| |\hat{v}_{j-1}| + 3u |\hat{\beta}_j| |\hat{v}_j| + u |\hat{w}_j| + O(u^2)$.

$$\hat{s}_{12} = fl(\hat{v}_{j+1}^T \hat{v}_{j+1}) = \hat{v}_{j+1}^T \hat{v}_{j+1} + e_{12}, \quad |e_{12}| \leq nu |\hat{v}_{j+1}|^T |\hat{v}_{j+1}| + O(u^2); \quad (34)$$

$$\hat{\xi}_{j+1} = fl(\sqrt{\hat{s}_{12}}) = \sqrt{\hat{s}_{12}} + e_{13}, \quad |e_{13}| \leq u \sqrt{\hat{s}_{12}} = u(|\hat{v}_{j+1}|^T |\hat{v}_{j+1}|)^{\frac{1}{2}} + O(u^2); \quad (35)$$

$$\hat{v}_{j+1} = fl(\frac{\hat{v}_{j+1}}{\hat{\xi}_{j+1}}) = \frac{\hat{v}_{j+1}}{\hat{\xi}_{j+1}} + e_{14}, \quad |e_{14}| \leq u |\frac{\hat{v}_{j+1}}{\hat{\xi}_{j+1}}| + O(u^2). \quad (36)$$

综合(33),(36)有

$$\hat{\xi}_{j+1} \hat{v}_{j+1} = \hat{v}_{j+1} + \hat{\xi}_{j+1} e_{14} = H_P \hat{w}_j - \hat{\xi}_j \hat{v}_{j-1} - \hat{\beta}_j \hat{v}_j + \hat{w}_j + g_{j+1}. \quad (37)$$

其中 $|g_{j+1}| \leq (m + 4)u |H_P| |\hat{w}_j| + 5u |\hat{\xi}_j| |\hat{v}_{j-1}| + 4u |\hat{\beta}_j| |\hat{v}_j| + 2u |\hat{w}_j| + O(u^2)$.

总结以上讨论及(26), (27), (37)的结果得以下定理.

定理 1 设 $H \in \mathbb{R}^{2n \times 2n}$ 为一 Hamiton 矩阵, H 的每行和每列最多有 m 个非 0 元素. 如果 k 步辛 Lanczos 过程没有中断发生, 则计算出的 Lanczos 向量满足

$$\begin{cases} \hat{\gamma}_j \hat{w}_j = H_P \hat{v}_j - \hat{v}_j + h_j, \\ \hat{\xi}_{j+1} \hat{v}_{j+1} = H_P \hat{w}_j - \hat{\xi}_j \hat{v}_{j-1} - \hat{\beta}_j \hat{v}_j + \hat{w}_j + g_{j+1}, \end{cases} \quad (38)$$

$$\text{其中} \begin{cases} |h_j| \leq (m+2)u|H_P||\hat{v}_j| + 2u|\hat{v}_j| + O(u^2), \\ |g_{j+1}| \leq (m+4)u|H_P||\hat{w}_j| + 5u|\hat{\xi}_j||\hat{v}_{j-1}| + 4u|\hat{\beta}_j||\hat{v}_j| + 2u|\hat{w}_j| + O(u^2). \end{cases}$$

计算出的矩阵 $\hat{S}_P^{2k}, \hat{H}_P^{2k}$ 满足

$$H_P \hat{S}_P^{2k} = \hat{S}_P^{2k} \hat{H}_P^{2k} + \hat{\xi}_{k+1} \hat{v}_{k+1} e_{2k}^T + E_k, \quad |E_k| \leq (m+4)u|H_P||\hat{S}_P^{2k}| + 5u|\hat{S}_P^{2k}||\hat{H}_P^{2k}| + O(u^2). \quad (39)$$

如果 $\|H_P\|_F, \|\hat{S}_P^{2k}\|_F, \|\hat{H}_P^{2k}\|_F$ 的大小适中, (39) 说明递推方程

$$H_P \hat{S}_P^{2k} = \hat{S}_P^{2k} \hat{H}_P^{2k} + \hat{\xi}_{k+1} \hat{v}_{k+1} e_{2k}^T. \quad (40)$$

在有限精度算法下满足机器精度. 但是不幸的是, 在接近发生中断时, $\|\hat{S}_P^{2k}\|_F$ 可能无界定.

对于非对称 Lanczos 算法, Bai 在 [1] 中得到了一个相似的结果. 与我们的 (38) 相对应的有如下结果

$$\begin{cases} \hat{\beta}_{j+1} \hat{q}_{j+1} = A \hat{q}_j - \hat{\alpha}_j \hat{q}_j - \hat{\gamma}_j \hat{q}_{j-1} + h_j^{\text{nonsymLan}}, \\ \hat{\gamma}_{j+1} \hat{p}_{j+1} = A^T \hat{p}_j - \hat{\alpha}_j \hat{p}_j - \hat{\beta}_j \hat{p}_{j-1} + g_{j+1}^{\text{nonsymLan}}, \\ |h_j^{\text{nonsymLan}}| \leq (3+m)u|A||\hat{q}_j| + 4u|\hat{\alpha}_j||\hat{q}_j| + 3u|\hat{\gamma}_j||\hat{q}_{j-1}| + O(u^2), \\ |g_{j+1}^{\text{nonsymLan}}| \leq (3+m)u|A^T||\hat{p}_j| + 4u|\hat{\alpha}_j||\hat{p}_j| + 3u|\hat{\beta}_j||\hat{p}_{j-1}| + O(u^2). \end{cases} \quad (41)$$

因此, 辛 Lanczos 算法和非对称 Lanczos 算法的误差性态是一样的. 保结构的限制并没有给误差分析带出任何附加问题, 误差分析的结果在本质上是相同的.

众所周知, 在有限精度下, 对于对称 Lanczos 过程, 计算出的 Lanczos 向量之间正交性会失去^[17]. 对于非对称 Lanczos 算法 Bai^[1] 给出了类似的结论. 正交性的损失不是由于误差的逐渐积累造成的, 而是由于相消造成的. 那么对于我们的辛 Lanczos 算法又怎样呢? 明显地, 某些计算出的 Lanczos 向量一旦产生了舍入误差, 舍入误差会传递到后面的 Lanczos 步. 非对称 Lanczos 算法中的误差传递 Bai 在文 [1] 中给出了分析, 下面我们给出关于计算出的辛 Lanczos 向量的误差传递分析.

定义

$$L = [\hat{v}_1, \hat{w}_1, \dots, \hat{v}_k, \hat{w}_k]^T J_P [\hat{v}_1, \hat{w}_1, \dots, \hat{v}_k, \hat{w}_k], \quad (42)$$

即 $l_{2j-1, 2m-1} = \hat{v}_j^T J_P \hat{v}_m, l_{2j-1, 2m} = \hat{v}_j^T J_P \hat{w}_m, l_{2j, 2m-1} = \hat{w}_j^T J_P \hat{v}_m, l_{2j, 2m} = \hat{w}_j^T J_P \hat{w}_m$.

在精确算法中, $L = J_P^{2k}$. 易证 $x J_P x = 0$ 对任何向量都成立. 所以 $l_{2j, 2j} = l_{2j-1, 2j-1} = 0$ 总是成立的, 与 $\hat{S}_P^{2k}$ 正交性的损失无关. 分析 L 的其它元素, 有以下引理.

引理 2 L 的元素满足以下方程

$$l_{jj} = 0, \quad j = 1, \dots, 2k, \quad l_{j+1, j} = -l_{j, j+1} = -1 + \kappa_j, \quad j = 1, \dots, 2k-1,$$

其中 $|\kappa_j| \leq \frac{2(m+n+2)u|\hat{v}_j|^T |J_P| |H_P| |\hat{v}_j| + 5u|\hat{v}_j|^T |J_P| |\hat{v}_j|}{|\hat{w}_j J_P \hat{v}_j|} + O(u^2)$. 且有

$$\hat{\gamma}_m l_{2j, 2m} = -\hat{\xi}_{j+1} l_{2j+1, 2m-1} - \hat{\xi}_j l_{2j-3, 2m-1} - \hat{\beta}_j l_{2j-1, 2m-1} + g_{j+1}^T J_P \hat{v}_m + \hat{w}_j^T J_P h_m. \quad (43)$$

关于 $l_{2j, 2m-1}$ 及 $l_{2j-1, 2m-1}$ 有类似的表达式.

证明 显然 $l_{2j, 2j} = l_{2j-1, 2j-1} = 0$ ($j = 1, 2, \dots, k$), 因为对任何向量 x 有 $x J_P x = 0$. 由于 $l_{2m, 2j-1} = -l_{2j-1, 2m}$, 所以我们只需推导 $l_{2j, 2j-1}$ ($j = 1, 2, \dots, k$), $l_{2j, 2m-1}$, $l_{2j-1, 2m-1}$, 和 $l_{2j, 2m}$ ($j, m = 1, 2, \dots, k, j < m$) 即可.

首先考虑 $l_{2j,2j-1}$, 利用 $\hat{w}_j$ 和 $\hat{v}_j$ 的误差分析的结果

$$l_{2j,2j-1} = \hat{w}_j^T J_P \hat{v}_j = \frac{\hat{w}_j^T J_P \hat{v}_j + \zeta_1}{\hat{w}_j^T \hat{v}_j} + O(u^2), \quad (44)$$

其中 $|\zeta_1| = |\hat{y}_j e_s^T J_P \hat{v}_j + \hat{\xi}_j \hat{w}_j^T J_P e_{14}| \leq 2u |\hat{v}_j|^T |J_P| (|H_P| |\hat{v}_j| + |\hat{v}_j|) + O(u^2)$.

另一方面

$$\begin{aligned} \hat{\xi}_j \hat{y}_j &= \hat{\xi}_j \hat{v}_j^T J_P H_P \hat{v}_j = \hat{v}_j^T J_P H_P (\hat{v}_j + \hat{\xi}_j \delta_1) \quad (\text{其中 } |\delta_1| \leq u \hat{v}_j / \hat{\xi}_j + O(u^2)) \\ &= -(\hat{w}_j + \hat{v}_j + \delta_2)^T J_P (\hat{v}_j + \hat{\xi}_j \delta_1) \quad (\text{其中 } |\delta_2| \leq mu |H_P| |\hat{v}_j| + u |\hat{v}_j| + O(u^2)) \\ &= -\hat{w}_j^T J_P \hat{v}_j + \zeta_2 + O(u^2), \end{aligned} \quad (45)$$

$$\text{其中} \begin{cases} \zeta_2 = -(H_P \hat{v}_j + \hat{v}_j)^T J_P \delta_1 - \hat{v}_j^T J_P \delta_1 - \delta_2^T J_P \hat{v}_j, \\ |\zeta_2| \leq (2m+2n+2)u |\hat{v}_j|^T |J_P| |H_P| |\hat{v}_j| + 3u |\hat{v}_j|^T |J_P| |\hat{v}_j| + O(u^2). \end{cases}$$

由以上分析知

$$\begin{aligned} l_{2j,2j-1} &= \frac{\hat{w}_j^T J_P \hat{v}_j + \zeta_1}{-\hat{w}_j^T J_P \hat{v}_j + \zeta_2} + O(u^2) \\ &= -1 + \frac{\zeta_1 + \zeta_2}{-\hat{w}_j^T J_P \hat{v}_j + \zeta_2} + O(u^2) = -1 + \kappa_j + O(u^2). \end{aligned} \quad (46)$$

利用 Taylor 展开式将 $f(x) = \frac{\zeta_1 + \zeta_2}{x + \zeta_2}$ 在 $t = x - \zeta_2$ 点展开得

$$|\kappa_j| \leq \frac{2(m+n+2)u |\hat{v}_j|^T |J_P| |H_P| |\hat{v}_j| + 5u |\hat{v}_j|^T |J_P| |\hat{v}_j|}{|\hat{w}_j^T J_P \hat{v}_j|} + O(u^2). \quad (47)$$

假设 $l_{i,l}, i=1,2,\dots,2m-1, l=1,2,\dots,2m-1; l_{2m,l}, l=1,2,\dots,2j-1, l_{i,2m}, i=1,2,\dots,2j-1$ 已经计算出. 下面计算 $l_{2j,2m}$. 由于 $j=m$ 时, $l_{2m,2m}=0$. 因此只推导 $j < m$ 的情况即可.

$$\begin{aligned} \hat{y}_m l_{2j,2m} &= \hat{y}_m \hat{w}_j^T J_P \hat{w}_m^T = \hat{w}_j^T J_P (H_P \hat{v}_m - \hat{v}_m + h_m) \\ &= (-\hat{\xi}_{j+1} \hat{v}_{j+1} - \hat{\xi}_j \hat{v}_{j-1} - \hat{\beta}_j \hat{v}_j + \hat{w}_j + g_{j+1})^T J_P \hat{v}_m - l_{2j,2m-1} + \hat{w}_j^T J_P h_m \\ &= -\hat{\xi}_{j+1} l_{2j+1,2m-1} - \hat{\xi}_j l_{2j-1,2m-1} + \hat{\beta}_j l_{2j-1,2m-1} + g_{j+1}^T J_P \hat{v}_m + \hat{w}_j^T J_P h_m. \end{aligned} \quad (48)$$

关于 $l_{2j,2m-1}$ 及 $l_{2j-1,2m-1}$ 有相似的分析.

上述引理描述了计算出的辛 Lanczos 向量的 J_P 正交性是如何损失的. 特别的, (43) 表达了误差是如何向前传递的.

例 3.1 为了说明引理 2 的结果, 给出如下数值例子

$$H = \text{diag}(D, -D^T), \quad D = \text{diag}\left(200, 100, 50, 47, \dots, 4, 3, \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}\right).$$

所有计算是在 Inter Pentium II 120/MMX 微机 (内存 48M) 使用 Matlab5.3 软件包完成的, 机器精度 $u \approx 2.22 \times 10^{-16}$. 算法中的辛 Lanczos 过程产生一系列 Hamilton J-三对角矩阵, 这些 J-三对角矩阵的特征值逐渐逼近 H 的最大特征值 $\lambda_1 = 200$. 当初始向量 v_1 随机选取时, 表 3.1 列出了 $l_{m+1,m}$ 偏离 1 或 $l_{m,m+1}$ 偏离 -1 上界 κ_m . 可见理论上的 κ_m 比计算出的 κ_m 一般情况下要大两个数量级.

表 3.1: 引理 2 中 κ_m 的绝对值的上界, 随机初始向量

m	$ \kappa_m $ 的界	$ L_{m+1,m}+1 $	$ L_{m,m+1}-1 $
1	$2.402096358720039e-14$	$4.440892098500626e-16$	$6.661338147750939e-16$
2	$2.400521241470975e-14$	$4.440892098500626e-16$	$6.661338147750939e-16$
3	$2.399110496291188e-14$	$6.661338147750939e-16$	$6.661338147750939e-16$
4	$2.404570093107262e-14$	$9.992007221626409e-16$	$1.110223024625157e-15$
5	$2.406194364766614e-14$	$3.330669073875470e-16$	$4.440892098500626e-16$
6	$2.406541474265487e-14$	$7.771561172376096e-16$	$7.771561172376096e-16$
7	$2.407097769875368e-14$	$9.992007221626409e-16$	$9.992007221626409e-16$
8	$2.413653957852442e-14$	0	0
9	$2.414536441243320e-14$	$2.220446049250313e-16$	$2.220446049250313e-16$
10	$2.415331725795096e-14$	$6.661338147750939e-16$	$6.661338147750939e-16$
11	$2.412262285835516e-14$	$8.881784197001252e-16$	$6.661338147750939e-16$
12	$2.411371182337842e-14$	$4.440892098500626e-16$	$6.661338147750939e-16$
13	$2.399037381446854e-14$	$7.771561172376096e-16$	$7.771561172376096e-16$

由描述的 L 中元素的舍入误差的传递, 可以明显地从图 3.1 看出. 为了跟踪误差的传递, 计算 $Z=L-J_P$, 每步 Lanczos 过程产生两个 Lanczos 向量, 因此 Z 的顺序主子阵将增加两行两列, 图 3.1 给出了 $2m$ 循环 ($m=3, 4, 5, 6$) 的元素的绝对值. $m=3$ 时, Z 的元素的阶为 10^{-14} , $m=4$ 时, Z 的元素的阶为 10^{-13} , $m=5$ 时, Z 的元素的阶为 10^{-12} , $m=6$ 时, Z 的元素的阶为 10^{-10} , 即大约每次增加一阶.

表 3.2: 引理 2 中 κ_m 的绝对值的上界, 特殊的初始向量

m	$ \kappa_m $ 的界	$ L_{m+1,m}+1 $	$ L_{m,m+1}-1 $
1	$2.398636844702651e-14$	0	0
2	$2.406633209694901e-14$	$2.220446049250313e-16$	0
3	$2.402808800167200e-14$	0	0
4	$2.415620760572344e-14$	0	0
5	$2.423704534528227e-14$	$2.220446049250313e-16$	0
6	$2.400760689022892e-14$	$2.220446049250313e-16$	$2.220446049250313e-16$
7	$2.428618469893587e-14$	$6.661338147750939e-16$	$6.661338147750939e-16$
8	$2.435552450510698e-14$	0	$2.220446049250313e-16$
9	$2.440784317651194e-14$	$3.330669073875470e-16$	$4.440892098500626e-16$
10	$2.443942157857217e-14$	$1.110223024625157e-15$	$1.110223024625157e-15$
11	$2.444928341933028e-14$	$6.661338147750939e-16$	$4.440892098500626e-16$
12	$2.443967298296119e-14$	$8.881784197001252e-16$	$8.881784197001252e-16$
13	$2.441302108512050e-14$	$6.661338147750939e-16$	$6.661338147750939e-16$

当初始向量选为 $v_1=[1, 1, 10^{-11}, \dots, 10^{-11}]^T \in \mathbb{R}^{100}$ 时, 由于此时初始向量近似地等于 H_P 的最大和最小特征值所对应的的特征向量的和. 因此可以预见依最大的特大的特征值对应的

不变子空间会很快收敛,表 3.2 和图 3.2 的意义与表 3.1 和图 3.1 相同.由图 3.2 可以明显地看出,两步之后辛 Lanczos 向量的 J-正交性就已失去,但表 3.2 知此几乎不影响计算出的 κ_m .

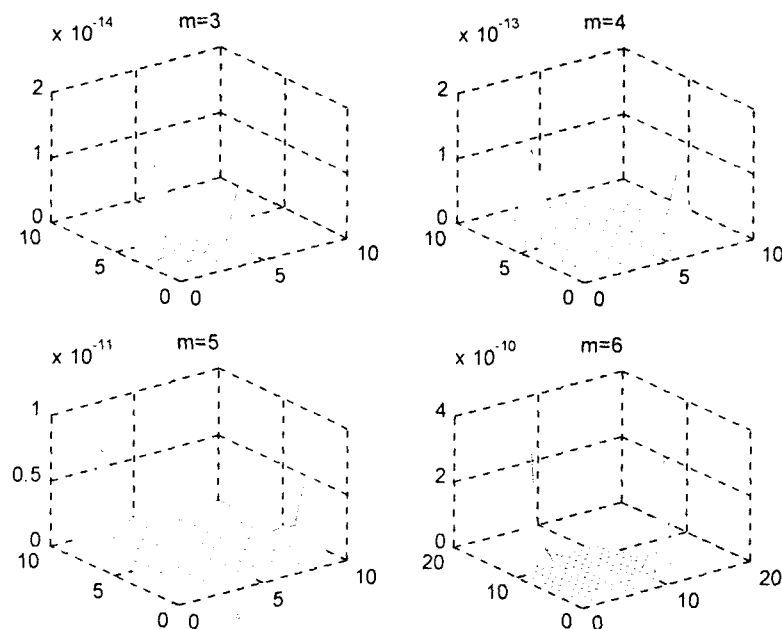


图 3.1: 误差的传递, 随机初始向量

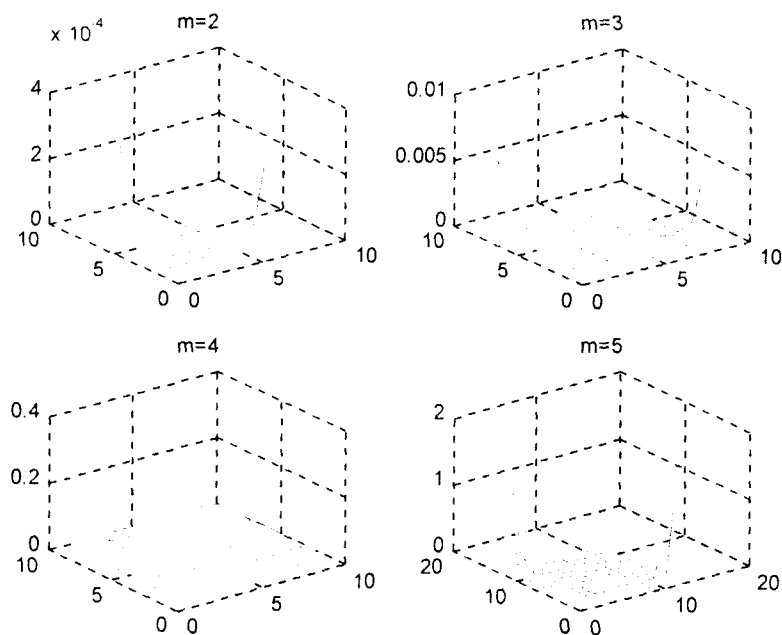


图 3.2: 误差的传递, 特殊的初始向量

下一节我们讨论 J-正交性的损失和收敛和关系,与上例显示的结果一样,J-正交性的失去伴随着 Ritz 的收敛.

4 收敛与 J-正交性的失去

众所周知,在对称 Lanczos 过程中,计算出的 Lanczos 向量正交性的失去意味着一个 Ritz 对的收敛^[17].下面我们分析辛 Lanczos 算法的收敛和 J-正交性的失去的关系,整个讨论与 Bai 在非对称 Lanczos 算法中的讨论^[1]和 F. Fassbender 对求解辛矩阵特征问题的辛 Lanczos 算法中的讨论^[8]类似.

由以前的推导知计算出的 Lanczos 向量满足 $H_P \hat{S}_P^{2k} = \hat{S}_P^{2k} \hat{H}_P^{2k} + \xi_{k+1} \hat{v}_{k+1} e_{2k}^T + E_k$. 设

$$(\hat{S}_P^{2k})^T J_P^{2k} \hat{S}_P^{2k} = J_P^{2k} + C_k + \Delta_k - C_k^T, \quad (49)$$

其中 $\Delta_k = \text{diag} \left[\begin{bmatrix} 0 & \kappa_1 \\ -\kappa_1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & \kappa_k \\ -\kappa_k & 0 \end{bmatrix} \right]$ 为严格下三角块阵,且块的大小为 2.

为了简化讨论,作两个假设.第一个假设为局部 J-正交性,即算出的辛 Lanczos 向量满足

$$\begin{bmatrix} \hat{v}_{j-1}^T \\ \hat{w}_{j-1}^T \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \hat{v}_j \\ \hat{w}_j \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \quad (50)$$

此意味着 C_k 的次对角线上的 2 阶块为 0,即 C_k 具有如下结构

$$C_k = \begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ X & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ X & X & 0 & \cdots & \cdots & 0 & 0 & 0 \\ X & X & X & \ddots & & 0 & 0 & 0 \\ & \vdots & & \ddots & \ddots & & \vdots & \\ X & X & X & \cdots & \cdots & X & 0 & 0 \\ X & X & X & \cdots & \cdots & X & X & 0 \end{bmatrix}, \text{其中 } X \text{ 为 } 2 \times 2 \text{ 的块}. \quad (51)$$

第二个假设为 $\hat{H}_P^{2k}$ 的特征问题可以精确求解,即存在矩阵 $Y_k \in \mathbb{C}^{2k \times 2k}$,使得

$$Y_k^{-1} \hat{H}_P^{2k} Y_k = \text{diag}(\lambda_1, -\lambda_1, \dots, \lambda_k, -\lambda_k). \quad (52)$$

此意味着计算出的与 λ_j 对应的 Ritz 向量为 $z_j = \hat{S}_P^{2k} y_{2j-1}$,而与 $-\lambda_j$ 对应的 Ritz 向量为 $x_j = \hat{S}_P^{2k} y_{2j}$.

定理 3 假设有限精度下的辛 Lanczos 过程满足

$$H_P \hat{S}_P^{2k} = \hat{S}_P^{2k} \hat{H}_P^{2k} + \xi_{k+1} \hat{v}_{k+1} e_{2k}^T + E_k. \quad (53)$$

等价地

$$H_P^T \hat{W}_P^{2k} = \hat{W}_P^{2k} (\hat{H}_P^{2k})^T + \xi_{k+1} J_P \hat{v}_{k+1} e_{2k-1}^T + F_k, \quad (54)$$

$$\hat{W}_P^{2k} = J_P \hat{S}_P^{2k} J_P^{2k}, \quad F_k = -J_P E_k J_P^{2k}.$$

令

$$\begin{cases} L_k^{(2)} + U_k^{(2)} = \hat{H}_P^{2k} J_P^{2k} \Delta_k - J_P^{2k} \Delta_k \hat{H}_P^{2k}, \\ L_k^{(4)} + U_k^{(4)} = (\hat{W}_P^{2k})^T E_k - F_k \hat{S}_P^{2k}, \end{cases} \quad (55)$$

其中 $L^{(2)}, L^{(4)}$ 为严格下三角块阵, $U^{(2)}, U^{(4)}$ 为严格上三角块阵, 且块的大小为 2. 则计算出的 Ritz 向量 $x_j = \hat{S}_P^{2k} y_{2j}$ 及 $z_j = \hat{S}_P^{2k} y_{2j-1}$ 满足

$$x_j^T J_P^{2k} \hat{v}_{k+1} = \frac{y_{2j}^T J_P^{2k} [U_k^{(4)} - U_k^{(2)}] y_{2j-1}}{\xi_{k+1} (e_{2k}^T y_{2j-1})} =: \frac{\Psi_1}{\xi_{k+1} (e_{2k}^T y_{2j-1})}, \quad (56)$$

$$z_j^T J_P^{2k} \hat{v}_{k+1} = \frac{y_{2j-1}^T J_P^{2k} [U_k^{(4)} - U_k^{(2)}] y_{2j}}{\xi_{k+1} (e_{2k}^T y_{2j})} =: \frac{\Psi_2}{\xi_{k+1} (e_{2k}^T y_{2j})}. \quad (57)$$

证明 我们的目的是以 J-正交失去的形式描述 $z_j^T J_P^{2k} \hat{v}_{k+1}$ 及 $x_j^T J_P^{2k} \hat{v}_{k+1}$. 在精确算法中它们的值都应该为 0. (54) 左乘 $(\hat{S}_P^{2k})^T$ 然后转置及 (53) 左乘 $(\hat{W}_P^{2k})^T$ 得

$$(\hat{W}_P^{2k})^T H_P \hat{S}_P^{2k} = \hat{H}_P^{2k} (\hat{W}_P^{2k})^T \hat{S}_P^{2k} + (\xi_{k+1} J_P \hat{v}_{k+1} e_{2k-1}^T)^T \hat{S}_P^{2k} + F_k^T \hat{S}_P^{2k}. \quad (58)$$

$$(\hat{W}_P^{2k})^T H_P \hat{S}_P^{2k} = (\hat{W}_P^{2k})^T \hat{S}_P^{2k} \hat{H}_P^{2k} + (\hat{W}_P^{2k})^T \xi_{k+1} \hat{v}_{k+1} e_{2k}^T + (\hat{W}_P^{2k})^T E_k. \quad (59)$$

(58) 与 (59) 两边相减得

$$\begin{aligned} & (\hat{W}_P^{2k})^T \hat{S}_P^{2k} \hat{H}_P^{2k} - \hat{H}_P^{2k} (\hat{W}_P^{2k})^T \hat{S}_P^{2k} \\ &= \xi_{k+1} e_{2k-1}^T \hat{v}_{k+1} J_P^T \hat{S}_P^{2k} - \xi_{k+1} (\hat{W}_P^{2k})^T \hat{v}_{k+1} e_{2k}^T + F_k^T \hat{S}_P^{2k} - (\hat{W}_P^{2k})^T E_k. \end{aligned} \quad (60)$$

上式具有很特殊的结构 $(\hat{S}_P^{2k})^T J_P^{2k} \hat{S}_P^{2k} = K = J_P^{2k} + C_k + \Delta_k - C_k^T$, 这意味着

$$(\hat{W}_P^{2k})^T \hat{S}_P^{2k} = -I^{2k} + J_P^{2k} C_k + J_P^{2k} \Delta_k - J_P^{2k} C_k^T,$$

这里 $J_P^{2k} C_k$ 及 $(J_P^{2k} C_k)^T$ 与 C_k 具有相同的形式. $J_P^{2k} \Delta_k$ 为对角矩阵

$$J_P^{2k} \Delta_k = \text{diag} \left[\begin{bmatrix} -\kappa_1 & 0 \\ 0 & -\kappa_1 \end{bmatrix}, \dots, \begin{bmatrix} -\kappa_k & 0 \\ 0 & -\kappa_k \end{bmatrix} \right]. \quad (61)$$

因此 (62) 的左端可改写为

$$\begin{aligned} & (\hat{W}_P^{2k})^T \hat{S}_P^{2k} \hat{H}_P^{2k} - \hat{H}_P^{2k} (\hat{W}_P^{2k})^T \hat{S}_P^{2k} \\ &= [J_P^{2k} C_k \hat{H}_P^{2k} - \hat{H}_P^{2k} J_P^{2k} C_k] + [J_P^{2k} \Delta_k \hat{H}_P^{2k} - \hat{H}_P^{2k} J_P^{2k} \Delta_k] + \\ & \quad [\hat{H}_P^{2k} J_P^{2k} C_k^T - J_P^{2k} C_k^T \hat{H}_P^{2k}]. \end{aligned} \quad (62)$$

由局部 J-正交性, 知 $J_P^{2k} C_k$ 具有 C_k 的特殊形式 (C_k 的次对角 2 阶块为 0). 因此有 $L^{(1)} = J_P^{2k} C_k \hat{H}_P^{2k} - \hat{H}_P^{2k} J_P^{2k} C_k$ 为严格下三角块阵, 块的大小为 2 阶.

同样分析有 $U^{(1)} = \hat{H}_P^{2k} J_P^{2k} C_k^T - J_P^{2k} C_k^T \hat{H}_P^{2k}$ 为严格上三角块阵, 且块的大小为 2 阶.

由于 $J_P^{2k} \Delta_k \hat{H}_P^{2k} - \hat{H}_P^{2k} J_P^{2k} \Delta_k$ 的二阶对角块为 0. 所以可设

$$J_P^{2k} \Delta_k \hat{H}_P^{2k} - \hat{H}_P^{2k} J_P^{2k} \Delta_k = L_k^{(2)} + U_k^{(2)}, \quad (63)$$

其中 $L_k^{(2)}$ 为严格下三角块阵, 而 $U_k^{(2)}$ 为严格上三角块阵. 因此

$$(\hat{W}_P^{2k})^T \hat{S}_P^{2k} \hat{H}_P^{2k} - \hat{H}_P^{2k} (\hat{W}_P^{2k})^T \hat{S}_P^{2k} = L_k^{(1)} + L_k^{(2)} + U_k^{(1)} + U_k^{(2)}.$$

下面分析 (60) 的右端项. 由于行向量 $\hat{v}_{k+1}^T J_P^T \hat{S}_P^{2k} = [\ast, \ast, \dots, \ast, 0, 0]$, 即只在前 $2n-2$ 位置上非零元素, 最后两个分量为 0 (局部 J-正交性). 故 $L_k^{(3)} = -\xi_{k+1} e_{2k-1} \hat{v}_{k+1}^T J_P^T \hat{S}_P^{2k}$ 为严格下三角块阵, 块的大小为 2 阶. 类似地分析可知 $U_k^{(3)} = \xi_{k+1} (\hat{W}_P^{2k})^T \hat{v}_{k+1} e_{2k}^T$ 为严格上三角块阵, 块的大小为 2. 因此

$$L_k^{(1)} + L_k^{(2)} + L_k^{(3)} + U_k^{(1)} + U_k^{(2)} + U_k^{(3)} = F_k^T \hat{S}_P^{2k} - (\hat{W}_P^{2k})^T E_k. \quad (64)$$

此式意味着 $F_k^T \hat{S}_P^{2k} - (\hat{W}_P^{2k})^T E_k$ 的对角上的块阵为 0. 因此可设

$$F_k^T \hat{S}_P^{2k} - (\hat{W}_P^{2k})^T E_k = L_k^{(4)} + U_k^{(4)},$$

其中 $L_k^{(4)}$ 是严格下三角块阵, 而 $U_k^{(4)}$ 是严格上三角块阵.

比较(60)的严格上三角部分有

$$\hat{\xi}_{k+1}(\hat{W}_P^{2k})^T \hat{v}_{k+1} e_{2k}^T = \hat{H}_P^{2k} J_P^{2k} C_k^T - J_P^{2k} C_k^T \hat{H}_P^{2k} + U_k^{(4)} - U_k^{(2)}, \quad (65)$$

$$\hat{\xi}_{k+1}(\hat{S}_P^{2k})^T J_P^{2k} \hat{v}_{k+1} e_{2k}^T = (\hat{H}_P^{2k})^T C_k^T + C_k^T \hat{H}_P^{2k} + J_P^{2k} (U_k^{(4)} - U_k^{(2)}). \quad (66)$$

(66)左乘 y_{2j}^T 右乘 y_{2j-1}^T 及左乘 y_{2j-1}^T 右乘 y_{2j}^T 得

$$\hat{\xi}_{k+1} y_{2j}^T (\hat{S}_P^{2k})^T J_P^{2k} \hat{v}_{k+1} (e_{2k}^T y_{2j-1}) = y_{2j}^T J_P^{2k} (U_k^{(4)} - U_k^{(2)}) y_{2j-1}, \quad (67)$$

$$\hat{\xi}_{k+1} y_{2j-1}^T (\hat{S}_P^{2k})^T J_P^{2k} \hat{v}_{k+1} (e_{2k}^T y_{2j}) = y_{2j-1}^T J_P^{2k} (U_k^{(4)} - U_k^{(2)}) y_{2j}, \quad (68)$$

于是有(56), (57)式成立.

定理 3 推出的(56)和(57)与 Bai 在非对称 Lanczos 算法中推出的方程是类似的. 因此关于辛 Lanczos 算法中 J-正交性的失去问题有相似的结果. 注意到 $\hat{\xi}_{k+1}$ 与特征向量的最后一个分量 $e_{2k}^T y_{2j}$ 及 $e_{2k}^T y_{2j-1}$ 是 $\{\lambda_j, z_j (J x_j)^T\}, \{-\lambda_j, x_j (J z_j)^T\}$ 向后误差标准中至关重要的量. 因此如果这两个方程中的 $|\Psi_1|$ 和 $|\Psi_2|$ 有界, 且不为零. 这两个方程反映出了辛 Lanczos 过程收敛 ($\hat{\xi}_{k+1}, e_{2k}^T y_{2j}$ 及 $e_{2k}^T y_{2j-1}$) 和 J-正交性失去的倒数关系.

例 4.1 我们对例 3.1 中的矩阵 H 继续进行数值实验. 第一次实验的初始向量 v_1 是随机产生的. 表 4.1 以 $z_1^T J_P \hat{v}$ 的形式给出了 J-正交性的损失和以剩余 $\hat{\xi}_{k+1} (e_{2k}^T y_2)$ 的形式给出了 Ritz 对的收敛程度. 正如我们所预料的, J-正交性的失去伴随着一个 Ritz 值 λ_1 以小剩余的形式收敛到 H_P 最大特征值.

当辛 Lanczos 算法在 $k=12$ 停止时, 计算出的最大 Ritz 值 λ_1 的相对精度为

$$\frac{|200 - \lambda_1|}{200} = 2.8421e - 015.$$

表 4.1: J-正交性的损失和 Ritz 值的收敛, 初始向量随机

k	$z_1^T J_P \hat{v}_{k+1}$	$\hat{\xi}_{k+1} (e_{2k}^T y_2)$
1	1.235990476633475e-016	7.910070568576705e+001
2	2.786399583287746e-017	2.978701659339106e+001
3	2.533238376012381e-016	-4.747053263321223e+000
4	1.340860270580314e-014	-9.243074736396117e-002
5	8.335203373464341e-013	1.487498631858310e-003
6	-4.597315842439093e-011	-2.696907247419584e-005
7	-2.598028558815933e-009	-4.772285230066042e-007
8	-1.434447726410595e-007	-8.643419434706768e-009
9	-1.049546975609927e-005	-1.181322384306966e-010
10	8.406588488809229e-004	1.474859046195846e-012
11	-5.291673224817185e-002	-2.342232829530912e-014
12	-1.295390693357952e-001	1.231138513574290e-015

当取初始化为 $v_1 = [1, 1, 10^{-11}, \dots, 10^{-11}]^T$ 时, 表 4.2 给出了相应的结果, 如例 3.1 一样 J-正交性会很快失去, 与之相伴的是 Ritz 值 λ_1 很快以小剩余的形式收敛到最大特征值.

表 4.2: J-正交性的损失和 Ritz 值的收敛, 特殊的初始向量

k	$z_j^T J_P \hat{v}_{k+1}$	$\hat{\xi}_{k+1}(e_{2k}^T y_2)$
1	$-2.447422744535689e-006$	$9.687061715810921e-009$
2	$-6.656099209538462e-005$	$3.561896303622030e-010$
3	$4.648969985606251e-004$	$-5.099696329380495e-011$
4	$-2.351185556743768e-002$	$1.010816906609733e-012$
5	$-6.310754262924400e-001$	$3.618595069906540e-014$

当 Lanczos 算法进行到 $k=7$ 时, 计算出的 Hamilton J-三对角矩阵有两个特征值逼近 200 (一个的相对精度为 $1.421085471520200e-16$, 另一个为的相对精度 $5.009455419582309e-6$), 一个特征值逼近 100. 因此正如标准的非对称 Lanczos 算法一样, J-正交性的失去会导致鬼特征值的出现. 文^[3]中的例 3 也说明了此现象. 因此辛 Lanczos 算法也需某种形式的重新 J-正交化^[3].

下面对 ϕ_1 和 ϕ_2 的绝对值进行估计. 为了简化分析细节, 假设 $\Delta_k=0$, 即 $\hat{w}_j^T J_P \hat{v}_j = -1$ (由引理 2 知, 这种假设是合乎情理的). 在此假设下, $U_k^{(2)}=0$, 而且 $|\phi_l| \leq \|U_k^{(4)}\|_F \|y_{2j}\|_2 \|y_{2j-1}\|_2, l=1, 2$. 接下来对 $\|U_k^{(4)}\|_F$ 进行估计. $U_k^{(4)}$ 是 $F_k^T \hat{S}_P^{2k} - (\hat{W}_P^{2k})^T E_k$ 的严格下三角部分. 因此

$$\begin{aligned} \|U_k^{(4)}\|_F &\leq \|F_k^T\|_F \|\hat{S}^{2k}\|_F + \|\hat{W}^{2k}\|_F \|E_k\|_F \\ &\leq 2u \|\hat{S}^{2k}\|_F \{(m+4) \|H_P\|_F + 5 \|\hat{H}_P^{2k}\|_F\} + O(u^2). \end{aligned} \quad (69)$$

由以上分析得到如下推论.

推论 4 假设定理 3 中的 $\Delta_k=0$, 则 $|\phi| \leq u \text{Cond}(\lambda_j) \{(m+4) \|H_P\|_F + 5 \|\hat{H}_P^{2k}\|_F\} + O(u^2)$, 其中 $\phi \in \{\phi_1, \phi_2\}$, $\text{Cond}(\lambda_j) = \text{Cond}(-\lambda_j) = \|\hat{S}^{2k}\|_F^2 \|y_{2j}\|_2 \|y_{2j-1}\|_2$ 是 Ritz 值 λ_j 和 $-\lambda_j$ 的条件数.

注 推论 4 说明了 J-正交性的失去与 Ritz 值的条件数的关系. 通过累加 $\sum_{j=1}^k (\|\hat{v}_j\|_2^2 + \|\hat{w}_j\|_2^2)$ (大约需要 $4nk$ 次运算), 可以计算出 Ritz 值的条件数. 不幸的是 Lanczos 过程中, 有可能发生近似中断, 在这种情况下, $\|\hat{v}_j\|_2$ 和 $\|\hat{w}_j\|_2$ 无法界定.

例 4.2 对例 4.1 中的矩阵继续进行试验. 表 4.3 给出了在随机的初始向量的情况下, ϕ_2 和推论 4 中 ϕ 的界的比较.

从试验结果可以看出, 我们推论中的误差界不象文[8]中的相应的误差界那样悲观, 原因可能是我们没有文[8]中的 $\|K^{2k, 2k}\|_F$, $\|\hat{N}^{2k, 2k}\|_F$ 等项.

当取初始向量为 $v_1 = [1, 1, 10^{-11}, \dots, 10^{-11}]^T$ 时, 结果是类似的.

表 4.3: ψ_2 和推论 4 中 ψ 的界的比较, 随机的初始向量

k	ψ_2	$ \psi $ 的界
1	1.005625767992502e-017	6.931048994062058e-012
2	-1.418942785406774e-018	9.799538350339220e-012
3	-8.577847306819934e-019	1.643258321470685e-011
4	6.092347337387886e-019	1.889820768387689e-011
5	-3.793074950943500e-019	1.635948681678490e-010
6	3.868623128104302e-019	1.820455112883205e-010
7	-1.364446447182767e-019	6.839412365404314e-009
8	-6.055288015532179e-018	7.794517917124763e-009
9	1.083347707796175e-019	6.823871388297669e-008
10	3.017459522995292e-019	8.068158035253201e-008
11	-1.395796963719727e-019	1.454226433859354e-007
12	7.567460010253422e-018	1.185803195819155e-006
13	-1.720023209773639e-019	1.676992089876161e-006
14	-1.707310261325535e-019	2.136144922015552e-006
15	-1.187125825079363e-018	3.408271132642330e-005

致谢 本文第一作者感谢贾仲孝导师、熊西文教授的精心指导和热情鼓励。

参考文献:

- [1] BAI Z. *Error analysis of the Lanczos algorithm for nonsymmetric eigenvalue problem* [J]. Math. Comp., 1994, **62**: 209-226.
- [2] BENNER P, FABBENDER H. *An implicitly restarted symplectic Lanczos method for the symplectic eigenvalue problem* [J]. SIAM. J. Matrix Anal. Appl., 2001, **22**: 103-125.
- [3] BENNER P, FABBENDER H. *An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem* [J]. Linear Algebra Appl., 1997, **263**: 75-111.
- [4] BENNER P, FABBENDER H. *A restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem* [R]. Tech. Rep. SPC 95-28, Fak f. Mathematik, TU Chemnitz-Zwickau, 09107 Chemnitz, FRG, 1995.
- [5] BENNER P, MEHRMANN V, XU H. *A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils* [J]. Numer. Math., 1998, **78**: 329-358.
- [6] BUNSE-GERSTNER A, MEHRMANN V. *A symplectic QR-like Algorithm for the solution of the real algebraic Riccati equation* [J]. IEEE Trans. Automat. Contr., AC-31, 1986, 1104-1113.
- [7] BUNSE-GERSTNER A, MEHRMANN V, WATKINS D. *An SR algorithm for Hamiltonian matrices, based on Gaussian elimination* [J]. Methods Oper. Res., 1989, **58**: 339-358.
- [8] FABBENDER H. *Error analysis of the symplectic Lanczos method for symplectic eigenvalue problem* [J]. BIT, 2000, **40**: 471-496.
- [9] FABBENDER H. *Symplectic methods for symplectic eigenproblems* [R]. Habilitationsschrift, Fachbere-

ich 3-Mathematik und Informatik, Universität Bremen, Bremen, Germany, 1998.

- [10] FREUND R, MEHRMANN V. *A symplectic Look-ahead Lanczos algorithm for the Hamiltonian eigenvalue problem* [J]. Manuscript.
- [11] GOLUB G H, LOAN C V. *Matrix Computations* [M]. Third Edition, The Johns Hopkins University Press, Baltimore, 1996.
- [12] KAHAN W, PARLETT B, JIANG E. *Residual bounds on approximate eigensystems of nonnormal matrices* [J]. SIAM J. Numer. Anal. , 1982, **19**: 470—484.
- [13] LANCZOS C. *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators* [J]. J. Res. Natur. Bur. Stand. , 1950, **45**: 225—280.
- [14] MEHRMANN V, WATKINS D. *Structure preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils* [R]. Tech. Report SFB393/00—02, Fak. F. Mathematik, TU Chemnitz, Chemnitz, FRG, 2000.
- [15] OLSON J, JENSEN H, Jorgensen. *Solution of large matrix equations which occur in response theory* [J]. J. Comput. Phys. , 1994, **74**: 265—282.
- [16] PAIGE C. *Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix* [J]. J. Inst. Math. Appl. , 1976, **18**: 341—349.
- [17] PARLETT B N. *The Symmetric Eigenvalue Problem* [M]. Prentice-Hall, Englewood Cliffs, New Jersey, 1980.
- [18] WILKINSON J H. *The Algebraic Eigenvalue Problem* [M]. Oxford, Clarendon, 1965.

Error Analysis of Symplectic Lanczos Method for Hamiltonian Eigenvalue Problem

YAN Qing-you^{1,2}, WEI Xiao-peng³

(1. Dept. of Economics and Statistics, Shandong Finance Institute, Ji'nan 250014, China;

2. School of Business Administration, North China Electric University, Beijing 102206, China;

3. Center of Advanced Design Technology, Dalian University, Liaoning 116622, China)

Abstract: A rounding error analysis of the symplectic Lanczos method is given for the Hamiltonian eigenvalue problem. It is applicable when no break down occurs and shows that the restriction of preserving the Hamiltonian structure does not destroy the characteristic feature of nonsymmetric Lanczos processes. An analog of Paige's theory on the relationship between the loss of orthogonality among the Lanczos vectors and the convergence of Ritz values in the symmetric Lanczos algorithm is discussed. All analysis follows the lines of Bai's analysis of the nonsymmetric Lanczos algorithm and the lines of H. FaBbender's analysis of the symplectic Lanczos algorithm for the symplectic eigenvalue problem. As is expected, it follows that (under certain assumptions) the computed J-orthogonal Lanczos vectors lose J-orthogonality when some Ritz values begin to converge.

Key words: Symplectic Lanczos method; Hamiltonian matrix; eigenvalues; error analysis; Ritz values; Ritz vectors.