2021 年 11 月 计 算 数 学 第 43 卷第 4 期
Nov., 2021 MATHEMATICA NUMERICA SINICA Vol.43, No.4

doi: 10.12286/jssx.j2020-0713

非线性特征值问题平移对称幂法的收敛率估计*1)

唐耀宗 杨庆之
(喀什大学数学与统计学院, 喀什 844000;
南开大学数学科学学院, 天津 300071)

摘 要

平移对称幂法 (SS-HOPM) 在求解源自玻色 - 爱因斯坦凝聚态的非线性特征值问题时, 不仅具有较高的计算效率, 而且具有点列收敛性, 但其收敛率尚未得到有效估计. 本文通过将多项式 Kurdyka-Łojasiewicz(K-Ł) 指数界的相关结果应用到所涉及优化问题的 Lagrange 函数上, 得到了平移对称幂法的次线性收敛率估计, 从理论上解释了平移对称幂法的计算效率.

关键词: 非线性特征值问题; 玻色 - 爱因斯坦凝聚态; 平移对称幂法; 收敛率估计.

MR (2010) 主题分类: 15A18, 15A69, 49D37, 49M99, 65K05, 90C26.

1. 引 言

本文研究一类球面约束下的非线性特征值问题

$$\begin{cases} \mathcal{A}\mathbf{x}^3 + \mathbf{B}\mathbf{x} = \lambda\mathbf{x} \\ \|\mathbf{x}\|_2 = 1, \end{cases} \tag{1.1}$$

其中 $\mathcal{A} \in \mathbb{R}^{n\times n\times n\times n}$ 是 4 阶对称张量, $\mathbf{B} \in \mathbb{R}^{n\times n}$ 是半正定矩阵, $\mathbf{x} \in \mathbb{R}^n$ 是向量. 这类非线性特征值问题源自玻色 - 爱因斯坦凝聚态 (简称 BEC)[1−5], 称其为 BEC 类非线性特征值问题. 目前在量子力学领域, BEC 问题, 特别是 BEC 基态 (能量最低时的状态) 是一个非常重要和活跃的研究课题 [6−10]. BEC 基态及其推广可由以下非凸最小化问题来描述 [11]

$$\begin{aligned} \min \quad & \frac{1}{2}\mathcal{A}\mathbf{x}^4 + \mathbf{x}^\top \mathbf{B}\mathbf{x} \\ \text{s.t.} \quad & \|\mathbf{x}\|_2 = 1. \end{aligned} \tag{1.2}$$

容易验证其 Karush-Kuhn-Tucker(KKT) 系统即为非线性特征值问题 (1.1).

BEC 基态解对 BEC 问题的理论研究以及实验设计具有重要意义. 有关它的存在性和唯一性研究进展, 可以参考 [1, 12] 及其中的参考文献. BEC 基态问题是 NP 难的 [13], 通常采取两类数值方法进行求解. 第一类数值方法是在球面约束条件下, 利用非线性特征值格式设计的不同类型的数值方法 [14−21]. 第二类数值方法是求解球面约束下的极小化优化问题格式的方法, 即使用不同的梯度技术来处理极小化以及投影技术来处理球形约束的各类方法 [22−28].

* 2020 年 7 月 24 日收到.

1) 基金项目: 国家自然科学基金项目 (12071234, 11671217); 新疆维吾尔自治区自然科学基金面上项目 (2018D01A01) 资助.

此外, 半定规划 (SDP) 方法 [13,29] 以及交替方向乘子法 (ADMM)[29] 等也用于求解 BEC 的基态解. 文献 [11] 中对更一般的 NP 难的 BEC 类非线性特征值问题采用了平移对称幂法 (SS-HOPM) 进行求解, 表现了较好的数值计算效率, 并且能够保证点列收敛性 [30].

收敛率对于算法研究而言也是一个十分重要的研究内容. 不过目前, 关于 SS-HOPM 的收敛率尚未有相关研究, 即不能从理论上 SS-HOPM 的收敛快慢进行解释. 对于非凸优化问题而言, 其迭代算法的收敛率可以借助于 K-Ł指数进行估计. 不过一般情况下, K-Ł指数不易得到. 本文在文献 [11, 30] 的基础上进一步对 SS-HOPM 的收敛率进行估计. 具体地, 利用文献 [31] 给出的多项式 K-Ł指数的一个界, 结合非凸优化问题 (1.2) 的 Lagrange 函数及相关定理, 获得了 BEC 类非线性特征值问题的 SS-HOPM 的次线性收敛率估计.

文章结构如下. 第二节简单介绍了 BEC 类非线性特征值问题的 SS-HOPM 算法及相关概念. 第三节对 SS-HOPM 算法的收敛率进行了估计. 第四节进行了总结.

2. BEC 类非线性特征值问题的 SS-HOPM 及相关概念

这篇文章中, 用 Σ 表示 $\mathbb{R}^n$ 上的单位球面, 即

$$\Sigma = \{\mathbf{x} \in \mathbb{R}^n: \ ||\mathbf{x}|| = 1\}.$$

文章采用文献 [32] 中的张量、矩阵以及向量的表示方法. 即向量 (一阶张量) 用黑体小写字母表示, 例如 $\mathbf{a}$. 矩阵 (二阶张量) 用黑体大写字母表示, 例如 $\mathbf{A}$. 高阶张量 (三阶或更高阶张量) 由 Euler 手写体表示, 例如 $\mathcal{A}$. 标量用小写字母表示, 例如 a.

对称张量在指标任意排列下其元素都是不变的. 下面给出对称张量及相关概念的定义.

定义 2.1. (对称张量)[33] 如果张量 $\mathcal{A} \in \mathbb{R}^{\overbrace{n\times\cdots\times n}^{m}}$ 对所有 $i_1\cdots i_m \in \{1,\cdots,n\}$ 以及 $p\in\Pi_m$, 均有

$$a_{i_{p(1)}\cdots i_{p(m)}} = a_{i_1\cdots i_m}$$

成立, 则称张量 $\mathcal{A}$ 是对称的, 其中 Π_m 表示了 $(1,\cdots,m)$ 所有置换的集合.

定义 2.2. (对称张量 - 向量乘积)[33] 设 $\mathcal{A} \in \mathbb{R}^{\overbrace{n\times\cdots\times n}^{m}}$ 是对称张量, 且 $\mathbf{x}\in\mathbb{R}^n$. 那么张量 $\mathcal{A}$ 与向量 $\mathbf{x}$ 的 $m-r$ 次乘积可以表示为

$$\mathcal{A}\mathbf{x}^{m-r} \in \mathbb{R}^{\overbrace{n\times\cdots\times n}^{r}}, \quad 0\le r\le m-1,$$

其具体表达式为

$$(\mathcal{A}\mathbf{x}^{m-r})_{i_1\cdots i_r} \equiv \sum_{i_{r+1},\cdots,i_m} a_{i_1\cdots i_m} x_{r+1,\cdots,i_m} \quad 对所有 \ \ i_1,\cdots,i_r\in\{1,\cdots,n\}.$$

文献 [11] 构建了非齐次约束优化问题

$$\begin{aligned} \min \quad & \frac{1}{2}\mathcal{A}\mathbf{x}^4 + \mathbf{x}^\top\mathbf{B}\mathbf{x} \\ \text{s.t.} \quad & ||\mathbf{x}||_2 = 1. \end{aligned}$$

的稳定点与 BEC 类非线性特征值问题

$$\begin{cases}\mathcal{A}\mathbf{x}^3+\mathbf{B}\mathbf{x}=\lambda\mathbf{x}\\ ||\mathbf{x}||_2=1,\end{cases}$$

的特征对之间的联系, 并将由 Kolda 和 Mayor[34] 提出的用于齐次优化问题求解的 SS-HOPM 推广到了上述非齐次约束优化问题, 以求解 BEC 类非线性特征值问题.

从文献 [30] 可知, 对于凸的情形, 平移项 $\alpha>0$ 必须要大于某个参数 β 才能确保 SS-HOPM 算法收敛, 其中 β 定义如下

$$\beta\equiv\max_{\mathbf{x}\in\Sigma}\rho(3\mathcal{A}\mathbf{x}^2+\mathbf{B}).$$

其中 $\rho(\cdot)$ 表示矩阵的谱半径. SS-HOPM 步骤如下:

算法 1 平移对称幂法 (SS-HOPM)[11]

Require:

1. 给定 4 阶对称张量 $\mathcal{A}\in\mathbb{R}^{n\times n\times n\times n}$ 和半正定矩阵 $\mathbf{B}\in\mathbb{R}^{n\times n}$.
2. 平移项 α;

Ensure:

特征值 λ 及相应的特征向量 $\mathbf{x}$.

```
取初值 x_0 ∈ R^n, 使其满足 ||x_0||_2 = 1. 令 λ_0 = A x_0^4 + x_0^T B x_0
for k = 0, 1, ··· , do
  if α ≥ 0 then
    x̂_{k+1} ← A x_k^3 + B x_k + α x_k
  else
    x̂_{k+1} ← −(A x_k^3 + B x_k + α x_k)
  end if
  x_{k+1} ← x̂_{k+1} / ||x̂_{k+1}||_2
  λ_{k+1} = A x_{k+1}^4 + x_{k+1}^T B x_{k+1}
end for
return λ = λ_{k+1}, x = x_{k+1}.
```

3. 收敛率估计

文献 [11] 给出了非线性特征值问题在实特征向量个数有限假设下的SS-HOPM 算法的收敛性, 而 [30] 在去除这个假设条件后, 借助于 K-Ł不等式等工具, 给出了 SS-HOPM 算法的点列收敛性. 这篇文章在此基础上对 BEC 类非线性特征值问题的 SS-HOPM 算法的收敛率进行分析.

3.1. 几个引理

首先给出几个收敛率估计时将会用到的引理.

引理 3.1. [30] 设 $\alpha > \beta$, $\{\lambda_k, \mathbf{x}_k\}$ 由 SS-HOPM 算法生成, 那么对足够大的 k 和 $\gamma > 0$, 有

$$\lambda_{k+1} - \lambda_k \geq \gamma \|\mathbf{x}_{k+1} - \mathbf{x}_k\|^2.$$

定理 3.2. (点列收敛性)[30] 设 $\alpha > \beta$, $\{\lambda_k, \mathbf{x}_k\}$ 由 SS-HOPM 算法生成, λ_* 为 $\{\lambda_k\}$ 的极限, $\mathbf{y}_k \equiv (\lambda_k + \alpha)^{\frac{1}{4}} \mathbf{x}_k$. 那么 $\{\mathbf{y}_k\}$ 收敛到 $\mathbf{y}_*$. 因此 $\{\mathbf{x}_k\}$ 收敛到 $\mathbf{x}_* = (\lambda_* + \alpha)^{-\frac{1}{4}} \mathbf{y}_*$.

定理 3.3. [35,36] 设 $f: V \to \mathbb{R}$ 为有限维实向量空间 V 上的实解析函数, 且对于充分大的 k, 存在 $\kappa > 0$ 满足

$$f(\mathbf{y}_k) - f(\mathbf{y}_{k+1}) \geq \kappa \|\nabla f(\mathbf{y}_k)\|^2,$$

并假定

$$[f(\mathbf{y}_{k+1}) = f(\mathbf{y}_k)] \Longrightarrow \mathbf{y}_{k+1} = \mathbf{y}_k.$$

则 $\nabla f(\mathbf{y}_*) = 0$, 且 $\{\mathbf{y}_k\} \to \mathbf{y}_*$, 而且收敛率可以进行如下估计:

$$\|y_* - y_k\| \lesssim \begin{cases} q^k & \text{如果 } \theta = \dfrac{1}{2}(\text{ 对某些} 0 < q < 1) \\ k^{-\frac{\theta}{1-2\theta}} & \text{如果 } 0 < \theta < \dfrac{1}{2} \end{cases}$$

采用 K-Ł指数估计非凸优化问题算法的收敛率一般都采用上面的估计框架. 不过通常在 $\mathbf{y}_*$ 附近, 不知道 K-Ł梯度不等式中的 θ. 因此在一般情况下, 不能从定理 3.3 得到收敛率的显式估计. 当解析函数 f 为多项式时, 在文献 [29] 中给出了 K-Ł梯度不等式中 θ 的界, 即下面的引理.

引理 3.4. [31] 设 f 是 $\mathbb{R}^n$ 上的实多项式, 阶数为 $d \in \mathbb{N}$. 假设 $f(0) = 0$, 并且 $\nabla f(0) = 0$, 那么存在常数 c, ϵ, 使得对于所有的 $\|\mathbf{x}\| \leq \epsilon$, 有

$$\|\nabla f(\mathbf{x})\| \geq c|f(\mathbf{x})|^\tau, \quad \tau = 1 - \frac{1}{d(3d-3)^{n-1}}.$$

本文即是针对非凸优化问题 (1.2) 构造满足引理 3.4 的多项式, 然后利用此结果及相关引理来估计 SS-HOPM 的收敛率.

3.2. 收敛率估计

对于非凸优化问题

$$\begin{aligned} \min \quad & \frac{1}{2}\mathcal{A}\mathbf{x}^4 + \mathbf{x}^\top \mathbf{B}\mathbf{x} \\ \text{s.t.} \quad & ||\mathbf{x}||_2 = 1, \end{aligned}$$

其 Lagrange 函数为

$$\mathcal{L}(\mathbf{x}, \mu) = \frac{1}{2}\mathcal{A}\mathbf{x}^4 + \mathbf{x}^\top \mathbf{B}\mathbf{x} + \mu(\mathbf{x}^\top \mathbf{x} - 1),$$

其中 μ 为 Lagrange 乘子. 从定理 3.2 可知, 由 SS-HOPM 算法生成的序列 $\{\mathbf{x}_k\}$ 收敛到 $\mathbf{x}_*$, 而且 $(\mathcal{A}\mathbf{x}_*^4 + \mathbf{x}_*^\top \mathbf{B}\mathbf{x}_*, \mathbf{x}_*)$ 为上述非凸优化问题的 KKT 点, 即

$$\mathcal{A}\mathbf{x}_*^3 + \mathbf{B}\mathbf{x}_* - (\mathcal{A}\mathbf{x}_*^4 + \mathbf{x}_*^\top \mathbf{B}\mathbf{x}_*)\mathbf{x}_* = 0.$$

记

$$\begin{aligned} \mu &= -(\mathcal{A}\mathbf{x}_*^4 + \mathbf{x}_*^\top \mathbf{B}\mathbf{x}_*) \\ \hat{\mathcal{L}}(\mathbf{x}, \mu) &= \mathcal{L}(\mathbf{x}, \mu) - \mathcal{L}(\mathbf{x}_*, \mu_*), \end{aligned}$$

则

$$\hat{\mathcal{L}}(\mathbf{x}_*, \mu_*) = 0.$$

$\hat{\mathcal{L}}$ 的梯度记为 $\nabla\hat{\mathcal{L}}$, 而且

$$\nabla\hat{\mathcal{L}} = \begin{pmatrix} 2\mathcal{A}\mathbf{x}^3 + 2\mathbf{B}\mathbf{x} + 2\mu\mathbf{x} \\ \mathbf{x}^\top\mathbf{x} - 1 \end{pmatrix}$$

则

$$\nabla\hat{\mathcal{L}}(\mathbf{x}_*, \mu_*) = 0.$$

由此可知, $\mathbb{R}^{n+1}$ 上的多项式 $\hat{\mathcal{L}}$ 是满足引理 3.4 的条件的, 那么就可以以此为基础给出 SS-HOPM 的收敛率估计, 即 $\{\lambda_k, \mathbf{x}_k\}$ 的收敛率. 即

定理 3.5. 设 $\alpha > \beta$, $\{\lambda_k, \mathbf{x}_k\}$ 由 SS-HOPM 算法生成, 那么下列陈述成立.

(1) $\{\lambda_k\}$ 以至少 $O(k^{\frac{1}{1-2\tau}})$ 的收敛速度收敛到 λ_*, 即存在 $M_1 > 0$ 和 k_1 使得对所有 $k > k_1$, 有

$$0 \leq \lambda_* - \lambda_k \leq M_1 k^{\frac{1}{1-2\tau}},$$

其中 $\tau = 1 - (4 \cdot 9^{n-1})^{-1}$.

(2) $\{\mathbf{x}_k\}$ 以至少 $O(k^{\frac{1-\tau}{1-2\tau}})$ 的收敛速度收敛到 $\mathbf{x}_*$, 即存在 $M_2 > 0$ 和 k_2 使得对所有 $k > k_2$, 有

$$\|\mathbf{x}_k - \mathbf{x}_*\| \leq M_2 k^{\frac{1-\tau}{1-2\tau}},$$

其中 $\tau = 1 - (4 \cdot 9^{n-1})^{-1}$.

证明. (1) 因为 $\hat{\mathcal{L}}(\mathbf{x}, \mu) = \mathcal{L}(\mathbf{x}, \mu) - \mathcal{L}(\mathbf{x}_*, \mu_*)$ 为 $\mathbb{R}^{n+1}$ 上的多项式, 而且满足 $\hat{\mathcal{L}}(\mathbf{x}_*, \mu_*) = 0$ 和 $\nabla\hat{\mathcal{L}}(\mathbf{x}_*, \mu_*) = 0$, 那么由引理 3.4 可知, 存在 c, ϵ, 使得当 $\|(\mathbf{x}, \mu) - (\mathbf{x}_*, \mu_*)\| < \epsilon$ 时, 有

$$\|\nabla\hat{\mathcal{L}}(\mathbf{x}, \mu)\| \geq c|\hat{\mathcal{L}}(\mathbf{x}, \mu)|^\tau,$$

其中 $\tau = 1 - (4 \cdot 9^{n-1})^{-1}$. 因为 $(\mathbf{x}_k, \mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k)$ 收敛到 $(\mathbf{x}_*, \mathcal{A}\mathbf{x}_*^4 + \mathbf{x}_*^\top\mathbf{B}\mathbf{x}_*)$, 所以存在 N_1 使得对所有 $k \geq N_1$, 有

$$\|(\mathbf{x}_k, \mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k) - (\mathbf{x}_*, \mathcal{A}\mathbf{x}_*^4 + \mathbf{x}_*^\top\mathbf{B}\mathbf{x}_*)\| < \epsilon.$$

在上述不等式中用 $(\mathbf{x}_k, -(\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k))$ 替换 $(\mathbf{x}, \mu)$, 再结合引理 3.1 可以得到

$$\begin{aligned}
& 2\left\|\mathcal{A}\mathbf{x}_k^3 + \mathbf{B}\mathbf{x}_k - (\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k)\mathbf{x}_k\right\| \\
\geq\ & c\left|\hat{\mathcal{L}}\left(\mathbf{x}_k, -(\mathcal{A}x_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k)\right)\right|^\tau \\
=\ & c\left|\frac{1}{2}\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k - \frac{1}{2}\mathcal{A}\mathbf{x}_*^4 - \mathbf{x}_*^\top\mathbf{B}\mathbf{x}_*\right|^\tau \\
=\ & c\left|\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k - \mathcal{A}\mathbf{x}_*^4 - \mathbf{x}_*^\top\mathbf{B}\mathbf{x}_* + \frac{1}{2}(\mathcal{A}\mathbf{x}_*^4 - \mathcal{A}\mathbf{x}_k^4)\right|^\tau \\
=\ & c\left|\lambda_k - \lambda_* + \frac{1}{2}(\mathcal{A}\mathbf{x}_*^4 - \mathcal{A}\mathbf{x}_k^4)\right|^\tau \\
=\ & c\left|\lambda_k - \lambda_* + \frac{1}{2}(\lambda_* - \lambda_k) + \frac{1}{2}(\mathbf{x}_k - \mathbf{x}_*)^\top\mathbf{B}(\mathbf{x}_k - \mathbf{x}_*)\right|^\tau \\
\geq\ & c\left|\frac{1}{4}(\lambda_k - \lambda_*)\right|^\tau.
\end{aligned} \tag{3.1}$$

由于在凸的情况下有

$$\hat{\mathbf{x}}_{k+1} = \mathcal{A}\mathbf{x}_k^3 + \mathbf{B}\mathbf{x}_k + \alpha\mathbf{x}_k, \quad \mathbf{x}_{k+1} = \frac{\hat{\mathbf{x}}_{k+1}}{\|\hat{\mathbf{x}}_{k+1}\|}.$$

那么

$$\begin{aligned}\mathcal{A}\mathbf{x}_k^3 + \mathbf{B}\mathbf{x}_k - \lambda_k\mathbf{x}_k &= \|\hat{\mathbf{x}}_{k+1}\|\mathbf{x}_{k+1} - \alpha\mathbf{x}_k - \lambda_k\mathbf{x}_k \\ &= \|\hat{\mathbf{x}}_{k+1}\|\left(\mathbf{x}_{k+1} - \frac{\alpha+\lambda_k}{\|\hat{\mathbf{x}}_{k+1}\|}\mathbf{x}_k\right).\end{aligned}$$

因为

$$\lambda_k + \alpha = \|\hat{\mathbf{x}}_{k+1}\|\mathbf{x}_{k+1}^\top\mathbf{x}_k,$$

所以

$$\left(\mathbf{x}_{k+1} - \frac{\lambda_k+\alpha}{\|\hat{\mathbf{x}}_{k+1}\|}\mathbf{x}_k\right)^\top \mathbf{x}_k = \mathbf{x}_{k+1}^\top\mathbf{x}_k - \frac{\alpha+\lambda_k}{\|\hat{\mathbf{x}}_{k+1}\|} = 0.$$

进一步有

$$\left\|\mathbf{x}_{k+1} - \frac{\lambda_k+\alpha}{\|\hat{\mathbf{x}}_{k+1}\|}\mathbf{x}_k\right\| \le \|\mathbf{x}_{k+1} - \mathbf{x}_k\|.$$

那么

$$\left\|\mathcal{A}\mathbf{x}_k^3 + \mathbf{B}\mathbf{x}_k - (\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k)\mathbf{x}_k\right\| \le \|\hat{\mathbf{x}}_{k+1}\|\,\|\mathbf{x}_{k+1} - \mathbf{x}_k\|.$$

因为 $\|\hat{\mathbf{x}}_{k+1}\| \to \lambda_* + \alpha > 0$, 那么存在 N_2, 使得对所有 $k \ge N_2$, 满足

$$\left\|\mathcal{A}\mathbf{x}_k^3 + \mathbf{B}\mathbf{x}_k - (\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top\mathbf{B}\mathbf{x}_k)\mathbf{x}_k\right\| \le 2(\lambda_* + \alpha)\|\mathbf{x}_{k+1} - \mathbf{x}_k\|. \tag{3.2}$$

结合不等式 (3.1), (3.2) 以及引理 3.1 可知, 对于充分大的 k, 有

$$\lambda_{k+1} - \lambda_k \ge \gamma\left(\frac{1}{16}\right)^{1+\tau}\left(\frac{c}{\alpha+\lambda_*}\right)^2(\lambda_* - \lambda_k)^{2\tau}.$$

记 $\beta_k = \lambda_* - \lambda_k$ 和 $\pi = \gamma\left(\frac{1}{16}\right)^{1+\tau}\left(\frac{c}{\alpha+\lambda_*}\right)^2$, 那么

$$\beta_k \ge \beta_{k+1} + \pi\beta_k^{2\tau}.$$

进一步

$$\begin{aligned}\pi &\le (\beta_k - \beta_{k+1})/(\beta_k^{2\tau}) \\ &\le \int_{\beta_{k+1}}^{\beta_k} 1/\sigma^{2\tau}d\sigma \\ &= \frac{1}{1-2\tau}(\beta_k^{1-2\tau} - \beta_{k+1}^{1-2\tau}).\end{aligned}$$

对上述不等式从 k_0 到 k 求和, 可得

$$(k-k_0)\pi = \frac{1}{1-2\tau}(\beta_{k_0}^{1-2\tau} - \beta_k^{1-2\tau}).$$

则

$$\beta_k^{1-2\tau} = \beta_{k_0}^{1-2\tau} + (2\tau-1)(k-k_0)\pi \ge (2\tau-1)(k-k_0)\pi.$$

因为 $1-2\tau<0$, 所以

$$\beta_k \le \big((2\tau-1)\pi\big)^{\frac{1}{1-2\tau}}(k-k_0)^{\frac{1}{1-2\tau}} = \big((2\tau-1)\pi\big)^{\frac{1}{1-2\tau}}\left(\frac{k-k_0}{k}\right)^{\frac{1}{1-2\tau}}k^{\frac{1}{1-2\tau}}.$$

因此, 对充分大的 k, 存在 $M_1>0$ 满足

$$0 \le \lambda_* - \lambda_k \le M_1 k^{\frac{1}{1-2\tau}}.$$

(2) 从不等式 (3.1) 和 (3.2) 可知, 对充分大的 k, 有

$$\left(\frac{1}{2}\mathcal{A}\mathbf{x}_k^4 + \mathbf{x}_k^\top \mathbf{B}\mathbf{x}_k - \frac{1}{2}\mathcal{A}\mathbf{x}_*^4 - \mathbf{x}_*^\top \mathbf{B}\mathbf{x}_*\right)^\tau \le \frac{4(\lambda_*+\alpha)}{c}\left\|\mathbf{x}_{k+1}-\mathbf{x}_k\right\|.$$

记 $s_k = \|\mathbf{x}_{k+1}-\mathbf{x}_k\|$, 那么

$$\beta_k^\tau \le \frac{4(\lambda_*+\alpha)}{c} s_k.$$

由于

$$\begin{aligned}(\lambda_{k+1}-\lambda_k)/\beta_k^\tau &= (\beta_k - \beta_{k+1})/(\beta_k^\tau)\\ &\le \int_{\beta_{k+1}}^{\beta_k} 1/\sigma^\tau d\sigma\\ &= \frac{1}{1-\tau}(\beta_k^{1-\tau} - \beta_{k+1}^{1-\tau}),\end{aligned}$$

根据引理 3.1 以及上面的不等式, 对充分大的 k, 有

$$\frac{1}{1-\tau}(\beta_k^{1-\tau}-\beta_{k+1}^{1-\tau}) \ge \gamma s_k^2/\beta_k^\tau \ge \frac{\gamma c}{4(\lambda_*+\alpha)} s_k,$$

即存在 $\omega>0$ 使得 $s_k \le \omega(\beta_k^{1-\tau}-\beta_{k+1}^{1-\tau})$ 成立. 将上述不等式从 k_1 到 k_2 求和, 可得

$$\sum_{i=k_1}^{k_2} s_i \le \omega(\beta_{k_1}^{1-\tau} - \beta_{k_2}^{1-\tau}) \le \omega\beta_{k_1}^{1-\tau}.$$

由此表明 $\sum_{i=k_1}^{k_2} s_i < +\infty$. 记 $\delta_k = \sum_{i=k_1}^{k_2} s_i$, 那么对于某 $\omega_1>0$, 有

$$\delta_k \le \omega\beta_k^{1-\tau} \le \omega_1 s_k^{\frac{1-\tau}{\tau}},$$

即存在 $\chi>0$ 满足

$$\delta_k^{\frac{\tau}{1-\tau}} \le \chi s_k \le \chi(\delta_k - \delta_{k-1}).$$

进一步使用 (1) 中一样的方式, 可知存在 $M_2>0$, 对于充分大的 k, 有

$$\delta_k \le M_2 k^{\frac{1-\tau}{1-2\tau}}.$$

那么

$$\|\mathbf{x}_k - \mathbf{x}_*\| \le \delta_k \le M_2 k^{\frac{1-\tau}{1-2\tau}}.$$

□

4. 结论及展望

非凸优化问题采用迭代法求解时, 其收敛率可借助于 K-Ł指数直接估计. 不过 K-Ł指数通常难于求得. 本文针对此问题, 通过构造非凸优化问题相应的 Lagrange 函数, 以其 K-Ł指数的界为基础, 获得 BEC 类非线性特征值问题求解的 SS-HOPM 的次线性收敛率估计.

从文献 [11] 中的数值结果可知, SS-HOPM 在求解中小规模的 BEC 问题的基态解时, 能以很好的效率收敛. 随着规模 n 的增大, 其值不断递增逼近某一固定值, 即随着划分变细, 离

散问题的结果与真实解之间的误差越来越小. 此结果与文献 [8] 中的误差分析结果一致. 不过或许因为问题中的稀疏结构未能充分利用到算法设计中, 或者平移项的选择不够好等原因, SS-HOPM 算法不能如文献 [7] 中一样进行大规模的高效计算. 故而结合问题结构设计更高效的算法和对 SS-HOPM 中的平移项进行自适应改进以提高其效率等都是值得进一步研究的问题.

此外, Hu 和 Li 在 [37] 中给出了张量秩一逼近的高阶幂法 (HOPM) 的线性收敛率估计. 对比文献 [11] 中的数值算例, 本文所得到的收敛率估计还是比较保守的, 实际结果表明应该有更好的收敛效率. 那么 Hu 和 Li 的收敛率估计方法是否适用于 BEC 类非线性特征值问题的 SS-HOPM 算法? 能否得到更好的收敛率估计? 我们将在今后的工作中进一步研究这个问题.

致谢 感谢两名审稿人提出的宝贵建议; 感谢罗刚博士在本文撰写过程中给予的启发和帮助.

参 考 文 献

[1] Bao W Z, Cai Y Y. Mathematical theory and numerical methods for Bose-Einstein condensation[J]. Kinetic and Related Models, 2013, 6(1): 1–135.

[2] Bao W Z, Chern I L, Zhang Y Z. Efficient numerical methods for computing ground states of spin-1 Bose-Einstein condensates based on their characterizations[J]. Journal of Computational Physics, 2013, 253: 189–208.

[3] Antoine X, Duboscq R. Modeling and computation of Bose-Einstein condensates: stationary states, nucleation, dynamics, stochasticity[C]. Lecture Notes in Mathematics, 2015, 49–145.

[4] Gross E P. Structure of a quantized vortex in Boson systems[J]. Nuovo Cimento, 1961, 20(3): 454.

[5] Pitaevskii L P. Vortex lines in an imperfect Bose gas[J]. Soviet Physics JETP, 1961, 13(2): 451–454.

[6] Zhang N, Xu F, Xie H H. An efficient multigrid for ground state solution of Bose-Einstein condensates[J]. International Journal of Numerical Analysis and Modeling, 2019, 16(5): 789–803.

[7] Jia S H, et al. A full multigrid method for nonlinear eigenvalue problems[J]. Science China Mathematics, 2016, 59(10): 2037-2048.

[8] 谢和虎. 非线性特征值的多重网格算法 [J]. 中国科学: 数学, 2015, 45(8): 1193–1204.

[9] Wu X M, Wen Z W, Bao W Z. A regularized Newton method for computing ground states of Bose-Einstein condensates[J]. Journal of Scientific Computing, 2017, 73(1): 303–329.

[10] Cancès E, Chakir R, Maday Y. Numerical analysis of nonlinear eigenvalue problems, Journal of Scientific Computing[J]. 2010, 45(1–3): 90–117.

[11] Tang Y Z, Yang Q Z, Huang P F. SS-HOPM for BEC-like nonlinear eigenvalue problems[J]. 高等学校计算数学学报, 2020, 42(2): 163–192.

[12] Lieb E H, Seiringer R, Yngvason J. Bosons in a trap: A rigorous derivation of the Gross-Pitaevskii energy functional, Physical Review[J]. 2001, A61: 043602.

[13] Hu J, et al. A note on semi-definite programming relaxations for polynomial optimization over a single sphere[J]. Science China Mathematics, 2016, 59(8): 1543–1560.

[14] 谢和虎. 子空间扩展算法及应用 [J]. 数值计算与计算机应用, 2020, 41(3): 169–191.

[15] Adhikari S K. Numerical solution of the two-dimensional Gross-Pitaevskii equation for trapped interacting atoms[J]. Physics Letters A, 2000, 265(1): 91–96.

[16] Edwards M, Burnett K. Numerical solution of the nonlinear Schrödinger equation for small samples of trapped neutral atoms[J]. Physical Review A, 1995, 51(2): 1382–1386.

[17] Dodd R J. Approximate solutions of the nonlinear Schrödinger equation for ground and excited states of Bose-Einstein condensates[J]. Journal of Research of the National Institute of Standards and Technology, 1996, 101(2): 545–552.

[18] Schneider B I , Fede D L. Numerical approach to the ground and excited states of a Bose-Einstein condensed gas confined in a completely anisotropic trap[J]. Physical Review, 1999, 59(3): 2232–2242.

[19] Bao W Z, Tang W J. Ground-state solution of Bose-Einstein condensate by directly minimizing the energy functional[J]. Journal of Computational Physics, 2003, 187(1): 230–254.

[20] Chang S L, Chien C S, Jeng B W. Computing wave functions of nonlinear Schrödinger equations: A time-independent approach[J]. Journal of Computational Physics, 2007, 226(1): 104–130.

[21] Chang S M, Lin W W, Shieh S F. Gauss-Seidel type methods for energy states of a multi-component Bose-Einstein condensate[J]. Journal of Computational Physics, 2005, 202(1): 367–390.

[22] Aftalion A, Danaila I. Three-dimensional vortex configurations in a rotating Bose Einstein condensate[J]. Physical Review, 2003, 68(2): 92–94.

[23] Cerimele M M, et al. Numerical solution of the Gross-Pitaevskii equation using an explicit finite-difference scheme: An application to trapped Bose-Einstein condensates[J]. Physical Review, 2009, 62(1): 1382–1389.

[24] Chiofalo M, Succi S, Tosi M. Ground state of trapped interacting Bose-Einstein condensates by an explicit imaginary-time algorithm[J]. Physical Review E, 2000, 62(5): 7438–7444.

[25] Ruprecht P A, et al. Time-dependent solution of the nonlinear Schrödinger equation for Bose-condensed trapped neutral atoms[J]. Physical Review A, 1995, 51(6): 4704–4711.

[26] Garcia-Ripoll J J, Perez-Garcia V M. Optimizing Schrödinger functionals using sobolev gradients: Applications to quantum mechanics and nonlinear optics[J]. SIAM Journal on Scientific Computing, 2001, 23(4): 1316–1334.

[27] Bao W Z, Wang H Q. A mass and magnetization conservative and energy-diminishing numerical method for computing ground state of spin-1 Bose-Einstein condensates[J]. SIAM Journal on Numerical Analysis, 2007, 45(5): 2177–2200.

[28] Danaila I, Kazemi P. A new Sobolev gradient method for direct minimization of the Gross-Pitaevskii energy with rotation[J]. SIAM Journal on Scientific Computing, 2010, 32(5): 2447–2467.

[29] 杨庆之, 黄鹏斐, 刘亚君. 解一类非线性特征值的数值算例 [J]. 数值计算与计算机应用, 2019, 40(2): 130–142.

[30] Tang Y Z, Yang Q Z, Luo G. Convergence analysis on SS-HOPM for BEC-like nonlinear eigenvalue problems[J]. Journal of Computational Mathematics, 2021, 39 (4): 621–632.

[31] Didier D, Krzysztof K. Explicit bounds for the Łojasiewicz exponent in the gradient inequality for polynomials[J]. Annal Polonici Mathematici, 2005, 87(1): 51–61.

[32] Kolda T G, Bader B W, Kenny J P. Higher-order web link analysis using multilinear algebra[C]. Proceedings of the 5th IEEE International Conference on Data Mining, Houston, TX., 2005, 242–249.

[33] Pierre Comon, et al. Symmetric tensors and symmetric tensor rank, SCCM Technical Report 06-02, Stanford University, 2006.

[34] Kolda T G, Mayo J R. Shifted power method for computing tensor eigenpairs[J]. SIAM Journal on Matrix Analysis and Applications, 2011, 32(4): 1095–1124.

[35] Absil P A, Manhony R, Andrews B. Convergence of the iterates of descent methods for analytic cost functions[J]. SIAM Journal on Optimization, 2005, 16(2): 531–547.

[36] Uschmajew A. A new convergence proof for the higher-order power method and generalizations[J]. Pacific Journal of Optimization, 2015, 11(2): 309–321.

[37] Hu S L, Li G Y. Convergence rate analysis for the higher order power method in best rank one approximations of tensors[J]. Numerische Mathematik, 2018, 140(4): 993–1031.

CONVERGENCE RATE ESTIMATION ON SS-HOPM FOR NONLINEAR EIGENVALUE PROBLEMS

Tang Yaozong Yang Qingzhi

(*School of Mathematics and Statistics, Kashi University, Kashi 844000, China;*
School of Mathematical Sciences, Nankai University, Tianjin 300071, China)

Abstract

In solving the nonlinear eigenvalue problems originated from Bose-Einstein Condensation, the shifted symmetric higher-order power method (SS-HOPM for short) not only has high computational efficiency, but also has point-wise convergence. However, the convergence rate of SS-HOPM has not been given. We apply the bound of the Kurdyka-Łojasiewicz (K-Ł) exponent of polynomial to the Lagrange function of the optimization problem involved in this paper, then we obtain sublinear convergence rate of the SS-HOPM, which can explain the calculation efficiency of the algorithm theoretically.

Keywords: nonlinear eigenvalues; Bose-Einstein Condensation; SS-HOPM; convergence rate estimation.

2010 Mathematics Subject Classification: 15A18, 15A69, 49D37, 49M99, 65K05, 90C26.