无标题稿件编号（将由编辑插入）

一种基于 MINRES 的线搜索算法：用于非凸优化及非正曲率检测

Hanfeng Zeng · Yang Liu · Wenqing Ouyang · AndreMilzarek

arXiv:2601.01575v1 [math.OC] 4 Jan 2026

收稿日期：日期 / 接受日期：日期

摘要 我们提出了一种基于 MINRES 的牛顿型算法，用于求解无约束非凸优化问题。该方法采用最小残差方法（MINRES）——一种用于求解不定对称线性方程组的知名求解器——计算能够利用二阶信息和非正曲率（NPC）信息的下降方向。在标准假设下，我们推导出了全面的渐近收敛性。特别地，在 Kurdyka-Łojasiewicz 不等式和温和的 NPC 可检测条件下，我们证明了该算法能够避开严格鞍点并收敛到二阶临界点。这主要通过沿 NPC 方向整合适当的正则化技术和前向线搜索机制来实现。此外，当满足局部 Polyak-Łojasiewicz 条件时，我们还证明了算法可快速局部超线性收敛到可能非孤立的极小值。在 CUTEst 测试集和深度自编码器问题上的数值实验展示了所提方法的有效性。

关键词 非凸优化 · MINRES · 非正曲率方向 · Kurdyka-Łojasiewicz 性质 · 前向线搜索

数学主题分类（2020年） 90C30 · 90C06 · 65K05

Hanfeng Zeng & Andre Milzarek 数据科学学院，香港中文大学（深圳）（CUHK-Shenzhen），广东，518172，中华人民共和国。Wenqing Ouyang 工业工程与运筹学系，哥伦比亚大学，美国纽约。（wo2205@columbia.edu）Yang Liu 牛津大学数学研究所，英国。（yang.liu@maths.ox.ac.uk）通讯作者：Andre Milzarek（andremilzarek@cuhk.edu.cn）

1 引言

本文考虑如下无约束优化问题：

$$\min_{x\in\mathbb{R}^n} \; f(x), \tag{1}$$

其中，$f:\ \mathbb{R}^n \to \mathbb{R}$ 是二阶连续可微的，但不一定是凸的。非凸性是许多现代应用中普遍且重要的特征，例如机器学习和 [8], 深度学习、统计学和回归 [19], 以及非凸 M 估计量 [48]。众所周知，非凸问题会表现出几种不同类型的临界点。其中，鞍点，即海森矩阵包含非正曲率（NPC）信息的临界点，通常是不理想的，而局部极小值则是优化所要达到的目标。配备适当内部线性系统求解器的截断牛顿法，是求解问题（1）的一种经典方法。在这里，嵌入式线性系统求解器能够以有效的方式处理并利用底层问题的二阶信息。因此，这类方法不仅能够识别临界点，通常还具备逃离鞍点的能力。本文研究一种基于线搜索的截断牛顿法，其中下降方向 d 通过求解以下系统的近似解得到

$$Bd = -g, \quad g = \nabla f(x), \quad B \approx \nabla^2 f(x). \tag{2}$$

用于计算此类 d 的常见内层求解器包括共轭梯度法（CG） [28], 、共轭残差法（CR） [73], 和最小残差法（MINRES 或 MR）[62], 等。

最近，Liu 和 Roosta [46] 证明了MINRES 能够有效检测非正曲率信息并以较低的计算成本生成非正曲率方向。受这一优良性质的启发，我们提出了一种二阶类算法，使用MINRES 求解（2）的正则化版本：

$$\bar{B}d = -g, \quad \text{where} \quad \bar{B} := B + \zeta I, \tag{3}$$

其中 ζ 为正则化参数。我们证明了所提出的方法具有多种良好的渐近性质，包括在 Kurdyka-Łojasiewicz（KL）不等式条件下的迭代收敛保证、对严格鞍点的规避，以及超线性收敛至非孤立局部极小值。

相关工作。 截断牛顿法是以相对较低的成本利用二阶信息的基础方法。经典地，共轭梯度法（CG） [28] 一直被用作此类算法的内层求解器，以获得牛顿方程的不精确且成本更低的解，见[17,68]。在 [16], 中，Dahito 和 Orban 使用 CR 作为 CG 的替代方法，并展示了相当的性能。此外， [45,67]中的作者考虑了基于 MINRES 的变体；MINRES 是 CR 的一种更一般形式，其中 B 可以是不定的，但其分析仅限于不变凸问题。在 [47], 中，Liu 和 Roosta 讨论了此类 Newton-MR 方法针对一般

非凸问题的非渐近复杂度界。最近，在 [42], 中，Lim 和 Roosta 将 [47] 中推导出的整体设定和复杂度界扩展到了具有不精确海森信息的 Newton 型 MR 方法。据我们所知，对于一般非凸设定下基于截断MINRES的牛顿法，目前文献中似乎尚无渐近收敛性结果。

正则化是使牛顿型方法实现全局化并增强稳健性的另一种重要策略。如式（3）所示，该策略通过引入正则化参数 ζ_k，并将牛顿更新修改为 $x_{k+1} = x_k - (\nabla^2 f(x_k) + \zeta_k I)^{-1} \nabla f(x_k)$ 来实现。经典结果表明，当满足条件 $\zeta_k \leq C\|\nabla f(x_k)\|, C > 0$, [38] 时，正则化牛顿方法仍能保持二次收敛。在开创性工作 [57], 中，Nesterov 和 Polyak 提出了三次正则化方法，其中 $\zeta_k = C\|x_{k+1} - x_k\|$，并建立了全局最坏情况复杂度保证。近期，Mishchenko[54] 提出了一种结合三次正则化与 Levenberg-Marquardt（LM）方法 [52] 特点的正则化方案。在该方法中，ζ_k 取为 $\zeta_k = \sqrt{C\|\nabla f(x_k)\|}$, 在每次迭代仅需进行一次矩阵求逆的同时，实现了三次正则化方法的快速全局收敛速度。截断与正则化技术常结合使用，以增强理论保证并提高实际效率，参见例如 [22,40]。

零阶和一阶方法的鞍点规避结果通常依赖稳定流形定理，并保证这些方法在渐近意义下，几乎总能通过适当的步长规避严格鞍点，[37,63,71,60,74]。对于二阶方法，逃逸机制通常利用能够表明海森矩阵 $B = \nabla^2 f(x)$具有非正曲率的方向 v，即 $\langle v, Bv\rangle \leq 0$。理论上，此类方向有助于规避鞍点，并提升数值表现，[24,27,55,25,14]。检测非正曲率相当直接，因为大多数内层求解器已经会将项 $\langle v, Bv\rangle$ 作为副产品计算出来（对于某些 $v \in \mathbb{R}^n$）。实践中，方向 v 通常选为与 B的最小特征值对应的特征向量，即$\langle v, Bv\rangle \leq \lambda_{\min}(B)\|v\|^2 < 0$, [27,14]，或通过特征对相关方法获得，[64]。计算与特征对相关的方向可能代价高昂，例如，应用 Lanczos 过程需要存储额外的矩阵或重新执行某些递推， [13,27]。相比之下，在非渐近情形下，Royer 等人 [68] 提出了阻尼 Newton-CG 方法：当梯度范数足够小时，该方法采用截断 CG 和最小特征值预言机（MEO），并使用随机生成的向量。通过这种方式，可以以无特征对方式构造 NPC方向，并建立规避鞍点的概率性结果。文献 [15]还提出了用于信赖域方法的类似 MEO 方法。然而，要确保 v 是有效的下降方向并带来充分下降，在算法设计中还需格外谨慎。值得注意的是，近期工作 [46] 表明，MINRES 返回的残差向量 r 能够自然地保证这种充分下降；这说明 MINRES 或许是利用非凸优化问题中非正曲率的更适用工具。虽然 Liu 和 Roosta 在 [47], 中为 Newton-MR 提供了非渐近的二阶复杂度保证，但据我们所知，尚无确保规避严格鞍点的渐近结果——

方法	Ref.	分析类型	鞍点规避	EP-Free	非孤立极小值	
					收敛率	收敛率 ω
信赖域+牛顿共轭梯度 (a)	[13]	渐近	✓	✗	✗	–
信赖域牛顿共轭梯度	[15]	非渐近	✓	✓	✗	–
信赖域牛顿共轭梯度	[65]	渐近	(✓)	✓	✓(b)	<1 (✗)
牛顿+NC (a)	[27,14]	(非)渐近	✓	✗	✗	–
NCN	[64]	非渐近	✓	✗	✗	–
Newton-CG	[68]	非渐近	✓	✓	✗	–
Newton-MR	[47]	非渐近	✓	✓	✗	–
共轭残差牛顿法	[57]	（非）渐近	✓	– (c)	✓	$=\frac{1}{3}$ $(=\frac{1}{3})$
共轭残差牛顿法	[76]	渐近	✓	– (c)	✓	$=1$ $(=1)$
ARC	[66]	渐近	(✓)	– (c)	✓(b)	$=1$ (✗)
本文		渐近	✓	✓	✓	<1 $(\leq\frac{1}{3})$

表 1 可避免鞍点和／或收敛至满足（局部）Łojasiewicz 不等式的非孤立极小值的二阶方法的经典及近期研究综述，该不等式的指数为 $\theta=\frac{1}{2}$。此处，“无 EP”表示不需要（隐式）计算特征对的方法。参数 $\omega\in(0,1]$（位于最后一列）刻画了 $\{\|\nabla f(x_k)\|\}$ 的整体超线性收敛速率，即 $\|\nabla f(x_{k+1})\|=\mathcal{O}(\|\nabla f(x_k)\|^{1+\omega})$，$k\to\infty$。括号中所示的速率，是各方法能够确保避免（严格）鞍点的最佳速率。(a) [13,27,14] 中的算法是利用特征值检验并生成适当方向来发挥负曲率（NC）作用的（信赖域）Newton型方法。(b) [66,65] 中的结果要求 $x_k\to\bar{x}$，其中 $\bar{x}$ 是一个（非孤立）局部极小值。(c) 在 [57,76]，中，假设三次正则化子问题可被精确求解；在 [66]，中，则考虑此类模型的非精确解。

目前似乎仍没有针对基于 MINRES 的方法的相关结果。值得注意的是，无论方法的阶数如何，都可以采用概率技术来避开鞍点；一阶方法参见 [23,21,31,33,32]，三阶变体参见 [2]。然而，在二阶方法的局部渐近分析中，这类思路似乎并不适用。

非孤立极小值常见于过参数化优化或具有连续对称性的问题中 [51,57,43]。三次正则化方法 [57] 及其自适应变体（ARC） [10,11] 在广义条件下可实现局部超线性收敛，而这些条件不要求海森矩阵正定——例如，[57,66] 采用 Polyak-Łojasiewicz（PL）条件，[77] 采用 Kurdyka-Łojasiewicz（KL）性质，以及 [76]采用误差界（EB）条件。最近，[65], Rebjock 和 Boumal 首次证明，在（局部）PL 条件下，以共轭梯度法作为内层求解器的信赖域方法具有超线性收敛性。具体而言，他们对共轭梯度法在非孤立极小值附近的表现进行了新颖的探讨。正如 [65], 所指出的，目前尚无针对基于 MINRES 的方法的类似局部收敛保证。此外，[65] 中的结果具有局部性质，并假设迭代点最终会进入极小值的一个足够小的邻域。在此背景下，一个值得探讨的问题是，能否推导出全面的全局到局部收敛性质，

从而确保迭代点能够到达这样的局部邻域。

贡献。 我们的思路是应用 MINRES 求解(2)，并遵循截断方法的一般范式，利用非正曲率信息。总之，我们的贡献包括：

– 我们提出了一种基于 MINRES 的牛顿型算法，该算法结合了（前向和后向）线搜索全局化、内置的非正曲率检测以及自适应正则化。在标准假设下，我们证明所生成的迭代点序列 $\{x_k\}$ 的聚点对应于该问题的驻点。此外，在众所周知的 Kurdyka-Łojasiewicz (KL) 不等式成立时，我们还建立了更强的收敛性结论，包括迭代点收敛 $x_k \to \bar{x}$。我们的分析在采用更现代且更弱的假设的同时，重新得到了截断牛顿法的经典结果，[17,16], 据我们所知，这是首批为基于 MINRES 的截断算法提供深入渐近收敛性保证的研究之一。

– 当矩阵 B_k 在（2）中取为精确海森矩阵时，我们证明，所提出的算法框架无需额外开销或任何特征值相关计算，即可避免严格鞍点。我们进一步分析了在 Łojasiewicz 性质下，所提方法在非孤立极小值附近的局部收敛行为，并证明 x_k）$\|\}$ 以任意接近二次的速率超线性收敛。这些理想的保证部分得益于我们的算法机制：该机制能够自适应地控制并选择容差、曲率检验和正则化参数。

– 最后，我们展示了所提出的基于 MINRES 的算法的效率，并报告其在深度自编码器问题和 CUTEst 测试集上的数值表现。

在表 1中，我们进一步概述并比较了相关工作和二阶方法。

符号说明。 我们用 **0** 表示零向量或零矩阵。用 I_t 表示维数为 $t \times t$ 的单位矩阵，$e_j^{(t)}$ 表示 I_t的第 j列。对于$t \geq 1$，利用 b 和 A 生成的次数为 t 的 Krylov 子空间定义为 $\mathcal{K}_t(A,b) = \text{span}\{b, Ab, \ldots, A^{t-1}b\}$。我们用下标 k表示主算法的迭代次数。上标 t 专用于表示子问题求解器为求得 d_k而进行的迭代次数，例如 $d_k = p_k^{(t)}$ 或 $d_k = r_k^{(t-1)}$。我们常使用缩写 $d_k = p_t$、$d_k = r_{t-1}$或 $d_t = d_k^{(t)}$（当对应关系 $p_t \equiv p_k^{(t)}$、$r_{t-1} \equiv r_k^{(t-1)}$ 明确时）。第 $(t-1)$ 次内迭代的残差向量记为 $r_{t-1} = b - Ap_{t-1}$。此外，令 $g_k \equiv \nabla f(x_k) \in \mathbb{R}^n$和 $B_k \approx \nabla^2 f(x_k) \in \mathbb{R}^{n\times n}$。对于对称矩阵 $A \in \mathbb{S}_n \subset \mathbb{R}^{n\times n}$，我们记作 $A \succeq (\succ) 0$ A 为半正定（正定）。向量的欧几里得范数和矩阵的谱范数记为 $\|\cdot\|$。

组织结构。 在第 2 节中，我们回顾 MINRES 及其主要性质。在

第 3 节中，我们介绍一种基于 MINRES 的线搜索算法，并对其进行详细的收敛分析。在第 4 节中，我们展示所提出算法良好的数值性能。

2 MINRES 回顾

本节简要回顾 MINRES，并列出若干在我们的分析中会用到的有用性质。感兴趣的读者可参阅 [46]，了解更多细节。我们考虑一个一般的对称（可能是不定或奇异）矩阵 $A \in \mathbb{S}_n \subset \mathbb{R}^{n\times n}$ 和一个向量 $b \in \mathbb{R}^n$。此外，为简单起见，我们令 $p_0 = \mathbf{0}$，这意味着 $\mathcal{K}_t(A,\ r_0) = \mathcal{K}_t(A,\ b)$。

2.1 MINRES：回顾与基本组成部分

我们首先给出非正曲率（NPC）方向的定义。

定义 2.1。 满足 $v^\top Av \le 0$ 的向量 $v \in \mathbb{R}^n\backslash\{\mathbf{0}\}$ 称为（针对 A的）非正曲率方向。

A 与 b 之间的相互作用对 MINRES 的收敛性至关重要。这种联系完全由 b 关于. A的阶描述，见 [69]。

定义 2.2。 b 关于. A 的阶是满足以下条件的正整数 $g \in \mathbb{N}$

$$\dim(\mathcal{K}_t(A,b)) = t \ \ if\ t \le g \quad and \quad \dim(\mathcal{K}_t(A,b)) = g \ \ if\ t > g.$$

MINRES 求解该问题的以下最小二乘形式（2）：

$$p_t = \mathrm{argmin}_{p\in\mathcal{K}_t(A,b)} \|b - Ap\|. \tag{4}$$

MINRES 由三个主要步骤组成，如算法 1所示：Lanczos 过程、QR分解和迭代更新。

Lanczos 过程。 令 $v_1 = b/\|b\|$，经过 t 次迭代后，Lanczos向量构成一个正交矩阵$V_{t+1} = \begin{bmatrix} v_1 & v_2 & \dots & v_{t+1}\end{bmatrix} \in \mathbb{R}^{n\times(t+1)}$，其列张成 $\mathcal{K}_{t+1}$（A，b），并满足

$$AV_t = V_{t+1}\tilde{T}_t = V_{t+1}\begin{bmatrix} T_t \\ \beta_{t+1}(e_t^{(t)})^\top\end{bmatrix} = V_{t+1}\begin{bmatrix} V_t^\top AV_t \\ \beta_{t+1}(e_t^{(t)})^\top\end{bmatrix}, \tag{5}$$

其中 $\tilde{T}_t \in \mathbb{R}^{(t+1)\times t}$ 是一个上Hessenberg矩阵，而 $\beta_{t+1} \in \mathbb{R}$ 是一个标量。设 $p_t = V_tq_t$ 给定某个 $q_t \in \mathbb{R}^t$。于是我们可以将（4）改写为

$$\min_{q_t\in\mathbb{R}^t} \|\beta_1 e_1^{(t+1)} - \tilde{T}_tq_t\|, \quad \beta_1 = \|b\|. \tag{6}$$

QR分解。 最小二乘问题 (6) 通过对 $\tilde{}$ 进行 QR 分解 T_t。由此可得

$$Q_t\tilde{T}_t = \tilde{R}_t \equiv \begin{bmatrix} R_t \\ \mathbf{0}^\top\end{bmatrix}, \quad \min_{q_t\in\mathbb{R}^t} \|\beta_1 e_1^{(t+1)} - \tilde{T}_tq_t\| = \min_{q_t\in\mathbb{R}^t} \left\|\begin{bmatrix} t_t \\ \phi_t\end{bmatrix} - \begin{bmatrix} R_t \\ \mathbf{0}^\top\end{bmatrix} q_t\right\|,$$

算法 1MINRES **内置** 非正曲率 **检测，用于** (2)

1: **Input:** the matrix A, the vector b, the tolerance parameter θ,
2: $\beta_1 = \|b\|$, $r_0 = b$, $v_1 = b/\beta_1$, $v_0 = p_0 = d_0 = d_{-1} = \mathbf{0}$, $c_0 = -1$, $s_0 = \delta_1^{(1)} = \epsilon_1 = 0$, $\phi_0 = \beta_1$, $t = 1$;
3: **while** true **do**
4: $z_t = Av_t$, $\alpha_t = v_t^\top z_t$, $z_t = z_t - \beta_t v_{t-1}$, $z_t = z_t - \alpha_t v_t$, $\beta_{t+1} = \|z_t\|$,
5: $\begin{bmatrix} \delta_t^{(2)} & \epsilon_{t+1} \\ \gamma_t^{(1)} & \delta_{t+1}^{(1)} \end{bmatrix} = \begin{bmatrix} c_{t-1} & s_{t-1} \\ s_{t-1} & -c_{t-1} \end{bmatrix} \begin{bmatrix} \delta_t^{(1)} & 0 \\ \alpha_t & \beta_{t+1} \end{bmatrix}$;
6: **if** $c_{t-1}\gamma_t^{(1)} \geq 0$ **then**
7: flag = NPC, **return** $d = \beta_1 \frac{r_{t-1}}{\|r_{t-1}\|}$ as an NPC direction;
8: **end if**
9: $\gamma_t^{(2)} = \sqrt{(\gamma_t^{(1)})^2 + \beta_{t+1}^2}$;
10: $c_t = \gamma_t^{(1)}/\gamma_t^{(2)}$[1], $s_t = \beta_{t+1}/\gamma_t^{(2)}$, $\tau_t = c_t\phi_{t-1}$, $\phi_t = s_t\phi_{t-1}$;
11: $d_t = (v_t - \delta_t^{(2)} d_{t-1} - \epsilon_t d_{t-2})/\gamma_t^{(2)}$, $p_t = p_{t-1} + \tau_t d_t$;
12: **if** $\phi_t \leq \theta\beta_1$ **then**
13: flag = SOL, **return** $d = p_t$ as a solution to (2);
14: **end if**
15: $v_{t+1} = z_t/\beta_{t+1}$[2], $r_t = s_t^2 r_{t-1} - \phi_t c_t v_{t+1}$;
16: $t \leftarrow t + 1$;
17: **end while**
18: **Output:** NPC or SOL direction d, r_t, flag;

其中 $Q_t \in \mathbb{R}^{(t+1)\times(t+1)}$ 是正交的， $R_t \in \mathbb{R}^{t\times t}$ 是上三角的，并且有 $Q_t\beta_1 e_1^{(t+1)} = [t_t; \phi_t]^\top$，其中 $t_t \in \mathbb{R}^t$、 $\phi_t \in \mathbb{R}$。因此， q_t 满足 $R_t q_t = t_t$，从而得到 $\phi_t = \|r_t\|$。显然，我们还有 $\phi_0 = \beta_1 = \|b\|$。

迭代更新。设 $t < g$，并令 D_t 由下三角方程组 $R_t^\top D_t^\top = V_t^\top$给出。现在，令 $V_t = [V_{t-1} \mid v_t]$，并利用 R_t是上三角的这一事实，可得递推关系 $D_t = [D_{t-1} \mid d_t]$，其中某个向量$d_t \in \mathbb{R}^n$。因此，利用 $R_t q_t = t_t$，我们可以通过以下方式更新迭代点 p_{t-1}

$$p_t = V_t q_t = D_t R_t q_t = D_t t_t = \begin{bmatrix} D_{t-1} \mid d_t \end{bmatrix} \begin{bmatrix} t_{t-1} \\ \tau_t \end{bmatrix} = p_{t-1} + \tau_t d_t.$$

这里， τ_t 是一个可由前几次迭代计算得到的标量。根据 $R_t^\top D_t^\top = V_t^\top$，可以利用 $\{v_i\}_{i=1}^{t-1}$ 和 $\{d_i\}_{i=1}^{t-1}$ 递推更新方向 d_t。

2.2 MINRES 的基本性质

接下来，我们总结 MINRES 的三项核心性质： 非正曲率检测、用于证明 A 正定性的证书，以及其他有用的结果。我们将省略证明，并请感兴趣的读者参阅 [46,70]。定理 2.3 基本上表明，在 MINRES 中无需额外开销即可检测 非正曲率方向。

[1] 这里，我们有 $\gamma_t^{(2)} \neq 0$。注意，由第 9 行可知， $\gamma_t^{(2)} = 0$ 意味着 $\gamma_t^{(1)} = 0$。因此，将检测到非正曲率，算法于第 7 行终止。[2] 这里，我们有 $\beta_{t+1} \neq 0$。注意，如果 $\beta_{t+1} = 0$，那么 $s_t = 0$，并且 $\phi_t = 0$ 于第 10 行。因此，算法会更早地在第 13 行终止。

定理 2.3（非正曲率检测）。令 g 为 A 关于 b 的阶。假设残差向量 $\{r_t\}$ 由 MINRES 生成。则有

$$r_{t-1}^\top A r_{t-1} = -\phi_{t-1}^2 c_{t-1}\gamma_t^{(1)}, \quad 1 \le t \le g, \tag{7a}$$

$$r_t^\top b = \|r_t\|^2, \quad 1 \le t \le g. \tag{7b}$$

这里，ϕ_{t-1}，c_{t-1} 和 $\gamma_t^{(1)}$ 是在MINRES中计算得到的参数，参见算法1。由式（7a），若 $\phi_{t-1}^2 c_{t-1}\gamma_t^{(1)} \ge 0$，则有 $r_{t-1}^\top A r_{t-1} \le 0$ 且r_{t-1} 可被判定为一个非正曲率方向。在经典截断方法中，例如 Newton-CG、 [17,68], 和 Newton-CR, [16], 这类方向也不难获得，但不一定是充分下降方向。相比之下，式（7b）表明，MINRES 中的残差向量可以作为一个非正曲率方向，为 $b = -g_k$带来充分下降。

我们注意到，矩阵 $T_t = V_t^\top A V_t \in \mathbb{R}^{t\times t}$ 在式（5）中与 A 密切相关，并在MINRES的分析中发挥重要作用。

定理 2.4（正定性证书）。给定 $A \in \mathbb{S}_n$ 和 $b \in \mathbb{R}^n$，并回顾 g 和 T_t（参见定义 2.2 和 (5)），有：

(i) 只要 $T_t \succ 0$，且 MINRES 中未检测到 非正曲率，就有。(ii)矩阵 i 的第 T_t 个特征值不大于 i的第 A个特征值。(iii) MINRES 在恰好经过 (4) 后返回一个解 g 次迭代（在精确算术下）。

矩阵 T_t 在采用精确算术时已知会收敛至 A；参见 [70,第 6.6 节。2]。定理 2.4 的第 (iii) 部分基本可由g 的定义得出，这证明了 MINRES 应用于低阶系统时的效率。

只要尚未检测到非正曲率，迭代点 $\{p_t\}$ 就具有以下性质；在我们框架的收敛分析中将频繁用到该性质；参见. [46, 定理 3.8 和 3.11]。

定理 2.5 (MINRES的性质)。设 g 是 b 关于 A 的阶。只要尚未检测到非正曲率，迭代点 $\{p_t\}$, $1 \le t < g$ 就满足：

(i) $p_t^\top b > p_t^\top A p_t$.

(ii) $p_1 = \frac{b^\top A b}{\|Ab\|^2} b$ *and* $p_t^\top b \ge p_1^\top b = \frac{b^\top A b}{\|Ab\|^2}\|b\|^2$.

3 基于 MINRES 的线搜索算法

我们的算法框架受到基于线搜索的截断牛顿方法的启发。与以往使用 CR 或共轭梯度法生成截断牛顿方向的工作不同， [16,17], 我们旨在利用 MINRES 的优良性质来实现这一目的。具体而言，在第 k 次迭代中，我们应用 MINRES 计算方向 d_k，其求解

$$\min_{d\in\mathbb{R}^n} \|B_k d + g_k\|.$$

完整的算法框架和全局收敛结果将在第 3.1 节至第 3.3节中讨论。在第 3.4 节和第 3.5节中，我们将验证所提出的 MINRES方法能够有效避开严格鞍点，并给出收敛到潜在非孤立极小值的局部收敛结果。

算法 2MINRES**基于线搜索算法**

1：输入：初始点 $x_0 \in \mathbb{R}^n$，参数 $\rho, \sigma \in (0,1)$，$\alpha \in (0,1]$, A $s > 0$，容差 $\{\theta_k\}$，参数序列 $\{a_k\}$, $\{\zeta_k\}$。2：对于 $k = 0, 1, \ldots$，执行 3：计算 $g_k = \nabla f(x_k)$，选择 $B_k \in \mathbb{S}_n$，并设置 $\bar{B}_k = B_k + \zeta_k I$；4：如果 $\|g_k\| = 0$，则 5：返回 $x^* = x_k$；6：结束 if 7：$(d_k, r_k, \text{flag}_k) = \text{MINRES}(\bar{B}_k, -g_k, \theta_k)$，通过算法 1；8：如果 $\text{flag}_k =$ 为 SOL 且 $d_k^\top \bar{B}_k d_k < \min\{$A $a_k \|g_k\|^\alpha\} \|d_k\|^2$，则 9：$d_k = -g_k$，$\text{flag}_k =$为梯度下降；10：结束 if 11：使用算法 3 计算满足 (9) 或 (10) 的步长 λ_k；12：设置 $x_{k+1} = x_k + \lambda_k d_k$, $k \leftarrow k + 1$；13：结束 for

算法 3 线搜索方案

1：输入：当前迭代点 $x_k \in \mathbb{R}^n$、方向 d_k、参数 $\rho, \sigma \in (0,1)$、$s > 0$；2：设置 $\lambda = s$；3：如果 flag = 为 SOL 或梯度下降，则 4：当 (9) 对 λ 不成立时，执行 $\lambda = \rho\lambda$；结束循环 5：否则，如果 flag = 为非正曲率，则 6：如果 (10) 对 $\lambda = s$ 不成立，则 7：当 (10) 对 λ 不成立时，执行 $\lambda = \rho\lambda$；结束循环 8：否则 9：当 (10) 对 λ 成立时，执行 $\lambda = \lambda/\rho$；结束循环 10：$\lambda = \rho\lambda$；11：结束条件判断 12：结束条件判断 13：返回 $\lambda_k = \lambda$；

3.1 算法框架

完整框架如算法2所示，其中调用了算法1 和算法3。该框架具有两个主要特征。首先，与 [16,17], 类似，我们使用三种方向：解方向 p_t、缩放后的非正曲率方向$\frac{\|g_k\|}{\|r_{t-1}\|} r_{t-1}$以及负梯度方向 $-g_k$。给定参数 θ_k、$\zeta_k \geq 0$后，我们首先使用算法1，

$$(d_k, r_k, \text{flag}_k) = \text{MINRES}(\bar{B}_k, -g_k, \theta_k), \quad \bar{B}_k := B_k + \zeta_k I,$$

计算正则化牛顿方程组的非精确解 d_k。根据（7a），如果$c_{t-1}\gamma_t^{(1)} \geq 0$，则检测到非正曲率，并且 $\frac{\|g_k\|}{\|r_{t-1}\|} r_{t-1}$ 是一个~~非正~~曲率方向，其满足 $r_{t-1}^\top B_k r_{t-1} \leq -\zeta_k \|r_{t-1}\|^2$，其中 $-\zeta_k$ 可以解释为 B_k的伪负特征值。正则化参数 ζ_k 在避开严格鞍点和实现局部收敛方面发挥着重要作用。如果 $d_k = p_t$，我们将执行以下附加曲率检验

$$p_t^\top \bar{B}_k p_t \geq \min\{\text{A}, a_k \|g_k\|^\alpha\} \|p_t\|^2, \quad \text{A}, \alpha \in (0,1],\ a_k \geq \bar{a} > 0, \tag{8}$$

在算法 2中。恰当选择 $\{a_k\}$ 使我们能够应用基于库尔迪卡–罗亚谢维奇–Łojasiewicz 不等式的分析技术，并获得统一的全局收敛性和局部收敛性

结果。若（ $d_k = -g_k$ 8）检验未通过，则设 8）。检验（8）旨在避免较小的正曲率。由于 $r_t = -g_k - \bar{B}_k p_t$，可得 $p_t^\top \bar{B}_k p_t = -p_t^\top (g_k + r_t)$ 对所有 t 均成立。因此，如果返回 8，便可高效地检验（8） r_t。

其次，我们根据 d_k 的最终取值采用两种不同的线搜索策略。如果 $d_k = p_t$ 或 $d_k = -g_k$，我们通过回溯搜索 选择步长 $\lambda_k = s\rho^j$ ，以满足 Armijo 条件

$$f(x_k + \lambda_k d_k) - f(x_k) \leq \sigma \lambda_k d_k^\top g_k, \tag{9}$$

其中 $s > 0$， ρ， $\sigma \in (0, 1)$ 且 j 是满足（9）的最小非负整数。在 非正曲率情况下， $d_k = \frac{\|g_k\|}{\|r_{t-1}\|} r_{t-1}$，我们检查以下条件：

$$f(x_k + \lambda_k d_k) - f(x_k) \leq \sigma \lambda_k g_k^\top d_k + \frac{\sigma}{2} \lambda_k^2 d_k^\top B_k d_k. \tag{10}$$

当（10）对 $\lambda_k = s$成立时，我们将 $\lambda_k = s\rho^{-j}$ 按因子 $1/\rho > 1$， $j = 1, 2,$ ..... 逐步增大，使得（10）仍然成立（即，此处 j 是满足线搜索条件的最大整数）。这是一种 前向线搜索。如果（10）对 $\lambda_k = s$不成立，则改用回溯搜索。我们的主要动机是迫使该步远离当前位置（因为当前迭代点可能位于严格鞍点的区域内）。由于 $d_k^\top B_k d_k \leq -\zeta_k \|d_k\|^2 \leq 0$， （10）可被视为（9）的更严格版本。前向线搜索和修改后的条件（10）将在鞍点规避中发挥重要作用，参见第 3.4节。类似策略也曾用于 [27,47]。

3.2 全局收敛分析

在介绍主定理之前，我们先陈述一些假设和准备性引理。我们的分析基于以下关键假设：

(A.1) 函数 $f\colon \mathbb{R}^n \to \mathbb{R}$ 二阶连续可微。**(B.1)** 存在 $M > 0$ ，使得对所有 $k \in \mathbb{N}$，都有 $\|B_k\| \leq M$。**(B.2)** 容差参数 $\{\theta_k\}$ 满足 $\theta_k = \min\{B$ $b_k \|g_k\|^\beta\}$ for 某个 B $\beta > 0$, $b_k \geq b > 0$， 以及所有 k。**(B.3)** 我们令 $\zeta_k = \min\{Z$ $z_k \|g_k\|^\zeta\}$ ，其中 Z > 0, $\zeta \in (0, 1]$, $z_k \geq \bar{z} > 0$。

我们还将使用以下 (B.1) 的另一种表述：

(B.1)′ 对所有 $k \in \mathbb{N}$，均有 $B_k = \nabla^2 f(x^k)$。

在开始分析之前，我们先对算法2 和3 的良定性作一些说明。由于算法3 采用前向线搜索机制，该方法可能在有限次迭代后于某个循环中停止，此时存在一个方向，使得 f 趋近于 $-\infty$。下文将证明，当 f 假设有下界时，可以避免这种情况。

引理 3.1。设 $f\colon \mathbb{R}^n \to \mathbb{R}$ 连续可微，并给定 x、$d \in \mathbb{R}^n$、$B \in \mathbb{S}_n$ 和 $\sigma \in (0, 1)$，其中 $\|d\| \neq 0$、 $\nabla f(x)^\top d < 0$ 以及 $d^\top B d \leq 0$。

(i) 存在 $\lambda_\ell = \lambda_\ell(x, d, B) > 0$，使得

$$f(x+\lambda d) - f(x) \le \sigma\lambda\nabla f(x)^\top d + \frac{\sigma}{2}\lambda^2 d^\top B d \quad \forall\ \lambda \in [0, \bar{\lambda}_\ell]. \tag{11}$$

(ii) 此外，假设 f 有下界。则存在 $\lambda_u = \lambda_u(x, d, B) > 0$ 使得（11）对所有 $\lambda > \lambda_u$ 均不成立。

证明。第 (i) 部分可直接由

$$\lim_{\lambda\to 0} \frac{f(x+\lambda d) - f(x) - \sigma\lambda\nabla f(x)^\top d - \frac{\sigma}{2}\lambda^2 d^\top B d}{\lambda} = (1-\sigma)\nabla f(x)^\top d < 0;$$

另见 [27, 引理 2.1]。第 (ii) 部分显然。 □

在引理 3.1中的假设下，算法 2 和 3 均定义良好，即，算法 3 中的线搜索过程将在有限步内终止。下文中，每当我们称算法 2 生成了一列迭代点 $\{x_k\}$ 时，均默认算法 3中的前向线搜索会在有限次迭代内停止；也就是说，我们不会显式假设 f有下界。虽然可以为 λ_k （在算法 3的第 9 步中）引入额外的自适应上界，但为使表述简洁，本文不采用这种做法。

在 (A.1)–(B.1)条件下，使用 Powell-Wolfe 线搜索的 Newton-CG（及 Newton-CR）方法生成满足 d_k 的方向 $-d_k^\top g_k \ge C_k^1\|g_k\|^2$ 以及 $\|d_k\| \le C_k^2\|g_k\|$，其中 $C_k^1,\ C_k^2 > 0$，见 [16,17,72]。接下来，当 $-d_k^\top g_k$、$\|d_k\|$以及 $\|g_k\|$ 分别为 $d_k = p_t$ 由 MINRES得到，且 λ_k 通过回溯搜索选取时，我们将建立它们之间的类似关系。

引理 3.2 （SOL 方向的性质）。设 $d_k = p_t$ 由算法 2 生成（即在算法 2的第 11 步中，有 flag_k =SOL），并设

$$s_k := x_{k+1} - x_k = \lambda_k d_k, \quad C_k := (\|\bar{B}_k\| + \|\bar{B}_k\|^2)^{-1}.$$

则有：

(i) $-d_k^\top g_k > C_k \min\{A\ a_k\|g_k\|^\alpha\}\|g_k\|^2$。(ii) $\|d_k\| = \|p_t\| \le \max\{A^{-1}\|g_k\|,\ a_k^{-1}\|g_k\|^{1-\alpha}\}$。(iii) 如果 $\alpha < 1$，则 $f(x_k) - f(x_{k+1}) \ge C_k\sigma\|g_k\| \min\{A^2\|s_k\|,\ (\frac{a_k}{s^\alpha})^{\frac{2}{1-\alpha}}\|s_k\|^{\frac{1+\alpha}{1-\alpha}}\}$。

证明。首先，根据定理 2.5 （ii），我们有 $p_1 = -\frac{g_k^\top \bar{B}_k g_k}{\|\bar{B}_k g_k\|^2} g_k$ 且

$$-p_t^\top g_k \ge -p_1^\top g_k = \frac{g_k^\top \bar{B}_k g_k}{\|\bar{B}_k g_k\|^2}\|g_k\|^2, \quad t \ge 1. \tag{12}$$

当 $d_k = p_1$，即 $\phi_1 \le \theta_k\beta_1$ 在算法 1中，此时不等式（12）的右端可根据曲率检验（8）界定如下：

$$\frac{g_k^\top \bar{B}_k g_k}{\|\bar{B}_k g_k\|^2}\|g_k\|^2 \ge \frac{\min\{A, a_k\|g_k\|^\alpha\}}{\|\bar{B}_k\|^2}\|g_k\|^2.$$

现在考虑 $d_k = p_t$ 对于 $t \geq 2$的情形，即 $\phi_1 > \theta_k\beta_1$ 在算法 1中。利用 $V_1 = [v_1] \in \mathbb{R}^{n\times 1}$， $p_1 = V_1 q_1 = -\frac{g_k^\top \bar{B}_k g_k}{\|\bar{B}_k g_k\|^2} g_k$可得

$$\frac{g_k^\top \bar{B}_k g_k}{\|\bar{B}_k g_k\|^2} = \frac{p_1^\top \bar{B}_k p_1}{\|\bar{B}_k p_1\|^2} = \frac{q_1^\top V_1^\top \bar{B}_k V_1 q_1}{\|\bar{B}_k V_1 q_1\|^2}. \tag{13}$$

根据构造方式和 Krylov 子空间的定义，我们有 $B_k V_t q_t \in \mathcal{K}_{t+1}(\bar{B}_k, -g_k)$ 且 $V_{t+1}V_{t+1}^\top$ 是投影到 $\mathcal{K}_{t+1}(\bar{B}_k, -g_k)$上的正交投影算子。因此，有 $V_{t+1}V_{t+1}^\top \bar{B}_k V_t q_t = \bar{B}_k V_t q_t$ $\mathcal{K}_t\ \bar{B}_k\ -g_k\ \subset$。此外，$\mathcal{K}_{t+1}\ \bar{B}_k\ -g_k$ $V_t q_t = V_{t+1}\tilde{q}_t$ $\tilde{q}_t = [q_t\ 0] \in \mathbb{R}^{t+1}$ 由于 (,)(,)，我们有，其中； 。此外，根据 Lanczos 过程，有 $T_t = V_t^\top \bar{B}_k V_t$ 。现在，当 13) 中使用这些性质 $t = 1$；由此得到

$$\frac{q_1^\top V_1^\top \bar{B}_k V_1 q_1}{\|\bar{B}_k V_1 q_1\|^2} = \frac{\tilde{q}_1^\top V_2^\top \bar{B}_k V_2 \tilde{q}_1}{\|V_2^\top \bar{B}_k V_2 \tilde{q}_1\|^2} = \frac{\tilde{q}_1^\top T_2 \tilde{q}_1}{\|T_2\tilde{q}_1\|^2}.$$

由于 $d_k = p_t$， $t \geq 2$，可得 $r_0^\top \bar{B}_k r_0 > 0$ 且 $r_1^\top \bar{B}_k r_1 > 0$，即 非正曲率在前两次迭代中均未被检测到。根据定理 2.4 (i)，于是有 $T_2 \succ 0$。设 $T_2 = Q\Sigma Q^\top$ 是 T_2 的特征值分解，其中 Σ =diag(λ_1， λ_2) 满足 λ_1， $\lambda_2 > 0$ 且 $Q \in \mathbb{R}^{2\times 2}$ 是一个正交矩阵。因此，令 $u = [u_1;\ u_2] = Q^\top \tilde{q}_1$，可得

$$\frac{\tilde{q}_1^\top T_2 \tilde{q}_1}{\|T_2\tilde{q}_1\|^2} = \frac{\sum_{i=1}^2 u_i^2\lambda_i}{\sum_{i=1}^2 u_i^2\lambda_i^2} \geq \frac{\sum_{i=1}^2 u_i^2\lambda_i}{\max\{\lambda_1,\lambda_2\}\sum_{i=1}^2 u_i^2\lambda_i} = \frac{1}{\|T_2\|} \geq \frac{1}{\|\bar{B}_k\|}.$$

最后一个不等式由定理 2.4 (ii) 得出。利用 A $\in (0, 1]$, ，这就完成了第 (i) 部分的证明。接下来讨论第 (ii) 部分。由 (7b)，对于 $t \geq 1$

$$\begin{aligned}\|\bar{B}_k p_t\|^2 &= \|-g_k - r_t\|^2 = \|g_k\|^2 + 2g_k^\top r_t + \|r_t\|^2\\ &= \|g_k\|^2 - \|r_t\|^2 \leq \|g_k\|^2.\end{aligned}$$

这意味着 min{A $a_k\|g_k\|^\alpha\}\|p_t\|^2 \leq p_t^\top \bar{B}_k p_t \leq \|p_t\|\|\bar{B}_k p_t\| \leq \|p_t\|\|g_k\|$，即有 $\|d_k\| = \|p_t\| \leq$ max{A$^{-1}\|g_k\|$, $a_k^{-1}\|g_k\|^{1-\alpha}$}，由此证明了第 (ii) 部分。若 A $\leq a_k\|g_k\|^\alpha$，则可得 $\|d_k\| \leq$ A$^{-1}\|g_k\|$ ，并由 (9) 和第 (i) 部分可知，(f(x_k) $- f(x_{k+1})$)/($\sigma\|g_k\|$) $\geq C_k\lambda_k$A$\|g_k\| \geq C_k$A$^2\|s_k\|$。否则，若 A $\geq a_k\|g_k\|^\alpha$，则可推得 $\|d_k\| \leq a_k^{-1}\|g_k\|^{1-\alpha}$ ，且由于 $\lambda_k \leq s$, $\alpha \in$ (0, 1)，有

$$\begin{aligned}(f(x_k) - f(x_{k+1}))/(\sigma\|g_k\|) &\geq C_k\lambda_k a_k\|g_k\|^{1+\alpha} \geq C_k\lambda_k a_k^{\frac{2}{1-\alpha}}\|d_k\|^{\frac{1+\alpha}{1-\alpha}}\\ &= C_k\lambda_k^{-\frac{2\alpha}{1-\alpha}} a_k^{\frac{2}{1-\alpha}}\|s_k\|^{\frac{1+\alpha}{1-\alpha}} \geq C_k(a_k/s^\alpha)^{\frac{2}{1-\alpha}}\|s_k\|^{\frac{1+\alpha}{1-\alpha}}.\end{aligned}$$

结合这两个估计，部分 (iii) 的证明至此完成。 □

注 3.3。 引理 *3.2 (i)* 的证明讨论了以下两种情况： $t = 1$ 和 $t \geq 2$，并分别加以处理，这看起来可能不太自然。然而，如果 $t = 1$ 和 B_k 是不定的，我们可能有 $g_k^\top \bar{B}_k g_k/\|\bar{B}_k g_k\|^2 < 1/\|\bar{B}_k\|$，尽管 $g_k^\top \bar{B}_k g_k > 0$，即 非正曲率未能在第 $t = 1$步检测到。例如，考虑 $B_k = \mathrm{diag}(1, -1)$,

$-g_k=(a,b)^\top$，以及 $|a|>|b|$。于是可得 $g_k^\top \bar{B}_k g_k = a^2-b^2>0$，并且我们可以轻易计算出 $T_1 = v_1^\top \bar{B}_k v_1 = g_k^\top \bar{B}_k g_k/\|g_k\|^2 > 0$ 和 $g_k^\top \bar{B}_k g_k/\|\bar{B}_k g_k\|^2 = (a^2-b^2)/(a^2+b^2) < 1 = 1/\|\bar{B}_k\|$。

若将 d_k 设为 $\frac{\|g_k\|}{\|r_{t-1}\|} r_{t-1}$ 或 $-g_k$，我们可以推导出类似的界。

引理 3.4 (非正曲率-方向的性质)。设 $d_k = \frac{\|g_k\|}{\|r_{t-1}\|} r_{t-1}$ 由算法 2 (即在第 *11* 步中，我们有 标志$_k$ =非正曲率)。有：

(i) $-d_k^\top g_k > \theta_k \|g_k\|^2$。(ii) $\|d_k\| = \|g_k\|$。(iii) 若满足(B.2)–(B.3)，则有 $f(x_k) - f(x_{k+1}) \geq \sigma \|g_k\| \|s_k\|$ 最小{B $^{s_Z}{}_2$, $s^{-\beta} b_k \|s_k\|^\beta$, $s^{\frac{1-\zeta}{2}} z_k \|s_k\|^\zeta$} 其中 $s_k = \lambda_k d_k = \lambda_k \frac{\|g_k\|}{\|r_{t-1}\|} r_{t-1}$。

证明。在标记$_k$ =非正曲率的情况下，不精确性判据不成立，即有 $\|r_{t-1}\| > \theta_k \|g_k\|$。根据定理 7b中的（2.3），可得 $-d_k^\top g_k = \|r_{t-1}\| \|g_k\| > \theta_k \|g_k\|^2$。第 (ii) 部分由 d_k的定义可知。进一步利用线搜索条件（10），可得

$$\begin{aligned} \frac{f(x_k) - f(x_{k+1})}{\sigma \|g_k\|} &\geq -\lambda_k \frac{g_k^\top d_k}{\|g_k\|} - \frac{\lambda_k^2}{2} \frac{d_k^\top B_k d_k}{\|g_k\|} \\ &\geq \lambda_k \|r_{t-1}\| + \frac{\lambda_k^2}{2} \frac{\zeta_k \|d_k\|^2}{\|g_k\|} > \theta_k \|s_k\| + \frac{\lambda_k}{2} \zeta_k \|s_k\|. \end{aligned} \tag{14}$$

如果 B $\leq b_k \|g_k\|^\beta$，这意味着 $f(x_k) - f(x_{k+1}) \geq$ B$\sigma \|g_k\| \|s_k\|$。否则，若 B $\geq b_k \|g_k\|^\beta$，我们分两种情况讨论。首先，若 $\lambda_k \leq s$，则有

$$(f(x_k) - f(x_{k+1}))/(\sigma \|g_k\|) \geq b_k \|g_k\|^\beta \|s_k\| = b_k \|d_k\|^\beta \|s_k\| \geq s^{-\beta} b_k \|s_k\|^{1+\beta}.$$

若 $\lambda_k > s$，那么，由于 $\zeta \leq 1$ 在(B.3)中，我们可以推得

$$2(f(x_k) - f(x_{k+1}))/(\sigma \|g_k\|) \geq \lambda_k \zeta_k \|s_k\| \geq \min\{sZ, s^{1-\zeta} z_k \|s_k\|^\zeta\} \|s_k\|.$$

结合前面的估计，证明至此完成。 □

引理 3.5 （GD 方向的性质）。设 $d_k = -g_k$ 由算法2 生成（即在算法2的第 *11* 步中，我们有标志$_k$ =GD）。则：

(i) $-d_k^\top g_k = \|g_k\|^2$。(ii) $\|d_k\| = \|g_k\|$。(iii) $f(x_k) - f(x_{k+1}) \geq \sigma \|g_k\| \|s_k\|$ （其中 $s_k = \lambda_k d_k = -\lambda_k g_k$）。

引理3.5 中列出的性质可由 $d_k = -g_k$立即得出。

在截断牛顿法的经典结果中，在满足 (A.1)、(B.1) 且水平集$\mathcal{L}_f(x_0) := \{x : f(x) \leq f(x_0)\}$ 有界的条件下，可以证明 [17,16]，全局收敛。这里，我们采用不同的策略。首先，在满足 (A.1)、(B.1)–(B.3) 或 (A.1)、(B.1)′、(B.3) 的条件下，我们验证 $\{x_k\}$ 的每个聚点都是驻点。其次，在第 3.3 节中，我们证明，如果众所周知的 Kurdyka-Łojasiewicz 性质在聚点 x^*处成立，则整个序列 $x_k \to x^*$ 收敛。这种几何方法使我们无需额外的有界性假设。现在，我们陈述第一个全局收敛结果，其灵感来自 [27, 定理 3.1]。

定理 3.6 (全局收敛)。 设条件 (A.1)、 (B.1)–(B.3)或 (A.1)、 (B.1)′、 (B.3) 成立，并设 $\{x_k\}$ 由算法 2生成。那么， $\{x_k\}$ 的每个聚点都是 (1)的驻点。

证明。 我们首先假设条件 (A.1)、 (B.1)–(B.3) 均得到满足。

设 x^* 是 $\{x_k\}$ 的一个聚点，并设 $\{x_{k_\ell}\}$ 是收敛到 x^*的一个子序列。根据线搜索条件（9）和（10）， $\{f(x_k)\}$ 是非递增的，并收敛到某个极限 $\xi \in \mathbb{R} \cup \{-\infty\}$。根据$f$的连续性，我们可以推得 $f(x_{k_\ell}) \to f(x^*)$，由极限唯一性可知，这意味着$\xi = f(x^*) \in \mathbb{R}$。因此，定义 $\mathcal{N}_{\rm npc} := \{k : \mathsf{flag}_k = \text{非正曲率}\}$，我们有

$$\begin{aligned} \infty > f(x_0) - f(x^*) &= \lim_{i\to\infty} \sum_{k=0}^{i-1} f(x_k) - f(x_{k+1}) \\ &\geq -\sigma \sum_{k=0}^{\infty} \lambda_k d_k^\top g_k - \frac{\sigma}{2} \sum_{k\in\mathcal{N}_{\rm npc}} \lambda_k^2 d_k^\top B_k d_k \\ &\geq -\sigma \sum_{k\notin\mathcal{N}_{\rm npc}} \lambda_k d_k^\top g_k + \frac{\sigma}{2} \sum_{k\in\mathcal{N}_{\rm npc}} \lambda_k^2 \zeta_k \|d_k\|^2. \end{aligned} \tag{15}$$

如果 $\|\nabla f(x^*)\| \neq 0$，则存在 $\epsilon > 0$ 和 $L \in \mathbb{N}$ ，使得 $\|g_{k_\ell}\| > \epsilon$ 对于$\ell \geq L$成立。此外，根据 (B.1)、 $a_k \geq \bar{a}$、 $\zeta_k \leq \mathsf{Z}$以及引理3.2 和 3.5中的（i）–（ii）部分，存在 $\bar{\epsilon} > 0$ ，使得对于所有 $k_\ell \notin \mathcal{N}_{\rm npc}$ 和 $\ell \geq L$，我们有

$$-\frac{d_{k_\ell}^\top g_{k_\ell}}{\|d_{k_\ell}\|} \geq \frac{\min\left\{\frac{\min\{\mathsf{A}, a_{k_\ell}\|g_{k_\ell}\|^\alpha\}}{\|B_{k_\ell}\|+\|B_{k_\ell}\|^2}, 1\right\}\|g_{k_\ell}\|^2}{\max\{\mathsf{A}^{-1}\|g_{k_\ell}\|, a_{k_\ell}^{-1}\|g_{k_\ell}\|^{1-\alpha}, \|g_{k_\ell}\|\}} > \bar{\epsilon}. \tag{16}$$

结合（15）、（16并注意到 $\zeta_{k_\ell} \geq$ 最小值$\{\mathsf{Z}\ \bar{z}\epsilon^\zeta\} =: \bar{\zeta}$（对所有 $\ell \geq L$），这表明 $\lambda_{k_\ell}\|d_{k_\ell}\| \to 0$ 为 $\ell \to \infty$。

我们首先考虑使用 Armijo 条件（9）确定步长 λ_k 对所有 k 充分大时的情形，即假设 $|\mathcal{N}_{\rm npc}| < \infty$。由于步长 $\rho^{-1}\lambda_{k_\ell}$ 不满足（9），因此有

$$f(x_{k_\ell} + \rho^{-1}\lambda_{k_\ell} d_{k_\ell}) - f(x_{k_\ell}) > \sigma\rho^{-1}\lambda_{k_\ell} g_{k_\ell}^\top d_{k_\ell}. \tag{17}$$

此外，根据中值定理，存在 $\tau_{k_\ell} \in [0,1]$ 使得

$$f(x_{k_\ell} + \rho^{-1}\lambda_{k_\ell} d_{k_\ell}) - f(x_{k_\ell}) = \rho^{-1}\lambda_{k_\ell} \nabla f(x_{k_\ell} + \tau_{k_\ell}\rho^{-1}\lambda_{k_\ell} d_{k_\ell})^\top d_{k_\ell}. \tag{18}$$

令 $\hat{}d_{k_\ell} = \tau_{k_\ell}\rho^{-1}\lambda_{k_\ell} d_{k_\ell}$ 并结合（17）和（18），可得

$$\frac{\sigma d_{k_\ell}^\top g_{k_\ell}}{\|d_{k_\ell}\|} < \frac{\nabla f(x_{k_\ell} + \hat{d}_{k_\ell})^\top d_{k_\ell}}{\|d_{k_\ell}\|} \leq \|\nabla f(x_{k_\ell} + \hat{d}_{k_\ell}) - \nabla f(x_{k_\ell})\| + \frac{d_{k_\ell}^\top g_{k_\ell}}{\|d_{k_\ell}\|}.$$

由于 $\{x_{k_\ell}\}$ 收敛于 x^*，且 ∇f 在紧集上一致连续，因此存在 $r > 0$ 和 $L' > L$ 使得

$$\|\nabla f(x_{k_\ell} + d) - \nabla f(x_{k_\ell})\| \leq (1-\sigma)\bar{\epsilon}, \quad \forall\, d \in B_r(0), \quad \forall\, \ell > L'.$$

^由于 $\lambda_{k_\ell}\|d_{k_\ell}\| \to 0$，可得 $\|d_{k_\ell}\| \to 0$ 因为 $\ell \to \infty$ 并结合（16），得到矛盾 $(1-\sigma)\,\bar{\epsilon} > -(1-\sigma)\, d_{k_\ell}^\top g_{k_\ell}/\|d_{k_\ell}\| > (1-\sigma)\,\bar{\epsilon}$对所有 ℓ 充分大的情形均成立。

接下来，我们考虑采用另一种线搜索条件（10）的情况。我们只需讨论子序列 $\{k_\ell\}$ 中有无穷多个元素满足应用（10）的情形，即 $|\{k_\ell\}\cap\mathcal{N}_{\rm npc}|=\infty$。与前面一样，对于所有这样的 k_ℓ，由于步长 $\rho^{-1}\lambda_{k_\ell}$ 不满足（10），因此有

$$f(x_{k_\ell}+\rho^{-1}\lambda_{k_\ell}d_{k_\ell})-f(x_{k_\ell})>\sigma(\rho^{-1}\lambda_{k_\ell}g_{k_\ell}^\top d_{k_\ell}+\rho^{-2}\lambda_{k_\ell}^2 d_{k_\ell}^\top B_{k_\ell}d_{k_\ell}/2). \quad (19)$$

此外，由中值定理，再次存在 $\tau_{k_\ell}\in[0,1]$ 使得

$$f(x_{k_\ell}+\rho^{-1}\lambda_{k_\ell}d_{k_\ell})-f(x_{k_\ell})=\rho^{-1}\lambda_{k_\ell}g_{k_\ell}^\top d_{k_\ell}+\rho^{-2}\lambda_{k_\ell}^2 d_{k_\ell}^\top \hat{B}_{k_\ell}d_{k_\ell}/2, \quad (20)$$

其中 $\hat{B}_{k_\ell}=\nabla^2 f(x_{k_\ell}+\hat{d}_{k_\ell})$ 且 $\hat{d}_{k_\ell}=\tau_{k_\ell}\rho^{-1}\lambda_{k_\ell}d_{k_\ell}$。因此，结合 (19)、(20)、引理 3.4 (i)，$\|g_{k_\ell}\|\geq\epsilon$，$\theta_{k_\ell}\geq\min\{\bar{b}\epsilon^\beta\}=:\bar{\theta}$，以及 $d_{k_\ell}^\top B_{k_\ell}d_{k_\ell}\leq-\zeta_{k_\ell}\|d_{k_\ell}\|^2\leq 0$，可得

$$\begin{aligned}&\rho^{-1}\lambda_{k_\ell}d_{k_\ell}^\top(\hat{B}_{k_\ell}-B_{k_\ell})d_{k_\ell}/2>(\sigma-1)(g_{k_\ell}^\top d_{k_\ell}+\rho^{-1}\lambda_{k_\ell}d_{k_\ell}^\top B_{k_\ell}d_{k_\ell}/2)\\&\geq(\sigma-1)g_{k_\ell}^\top d_{k_\ell}>(1-\sigma)\theta_{k_\ell}\|d_{k_\ell}\|\|g_{k_\ell}\|>(1-\sigma)\bar{\theta}\epsilon\|d_{k_\ell}\|,\quad \ell\geq L.\end{aligned} \quad (21)$$

将 (21) 除以 $\|d_{k_\ell}\|$ 并利用 $\lambda_{k_\ell}\|d_{k_\ell}\|$，$\|d_{k_\ell}\|\to 0$，$\ell\to\infty$ 以及的有界性 B_{k_ℓ} 和 B_{k_ℓ}，我们得到矛盾。这证明了 $\nabla f(x^*)=0$。

最后，假设 (A.1)、(B.1)′、(B.3) 均成立。在这种情况下，我们之前的步骤仍然适用。特别是，由于 15）和（16）中的界仍然有效，因为 $\|B_{k_\ell}\|=\|\nabla^2 f(x_{k_\ell})\|$ 对所有 $\ell\geq L$都有界。因此，当 $\ell\to\infty$ 时，我们仍有 $\lambda_{k_\ell}\|d_{k_\ell}\|\to 0$。由于 $|\mathcal{N}_{\rm npc}|<\infty$ 情形下的推导不依赖于 (B.1) 和(B.2)，我们只需在使用（10）时证明会导致矛盾。结合（19）和（20），并应用 $d_{k_\ell}^\top B_{k_\ell}d_{k_\ell}\leq-\zeta_{k_\ell}\|d_{k_\ell}\|^2\leq-\bar{\zeta}\|d_{k_\ell}\|^2$，可得

$$\begin{aligned}&\rho^{-1}\lambda_{k_\ell}d_{k_\ell}^\top(\hat{B}_{k_\ell}-B_{k_\ell})d_{k_\ell}/2>(\sigma-1)(g_{k_\ell}^\top d_{k_\ell}+\rho^{-1}\lambda_{k_\ell}d_{k_\ell}^\top B_{k_\ell}d_{k_\ell}/2)\\&\geq(\sigma-1)\rho^{-1}\lambda_{k_\ell}d_{k_\ell}^\top B_{k_\ell}d_{k_\ell}/2>(1-\sigma)\rho^{-1}\lambda_{k_\ell}\bar{\zeta}\|d_{k_\ell}\|^2/2.\end{aligned} \quad (22)$$

将（22）除以 $\rho^{-1}\lambda_{k_\ell}\|d_{k_\ell}\|^2$，利用 $\|\hat{d}_{k_\ell}\|\to 0$，$\ell\to\infty$ 以及 $\nabla^2 f$ 的连续性，我们再次得到矛盾。□

3.3 Kurdyka-Łojasiewicz 不等式下的收敛性

现在，我们利用著名的 Kurdyka-Łojasiewicz（KL）不等式，为算法 2 提供更多收敛性结果。KL 性质刻画了函数 f 在临界点附近的增长行为，并可用于为迭代点建立更强的收敛保证。

用 $\mathfrak{L}_\eta$ 表示所有连续且凹的去奇异函数 $\varrho:[0,\eta)\to\mathbb{R}_+$ 所构成的类，使得

$$\varrho\in C^1((0,\eta)),\quad \varrho(0)=0,\quad \varrho'(x)>0,\quad \forall\, x\in(0,\eta).$$

下面给出关于 f 的 KL 性质定义，参见 [1,3]。

定义 3.7。 如果存在 $\eta \in (0, \infty]$、 $\bar{x}$ 的一个邻域 V 以及一个函数 $\varrho \in \mathfrak{L}_\eta$，使得对于所有 $x \in V \cap \{x \in \mathbb{R}^n : \ 0 < f(x) - f(\bar{x}) < \eta\}$，*KL*不等式成立，即我们称 f 在 $\bar{x}$ 处具有 *Kurdyka-Łojasiewicz* 性质，

$$\varrho'(f(x) - f(\bar{x}))\|\nabla f(x)\| \geq 1. \tag{23}$$

如果映射 ϱ 还满足 $\varrho(t) = ct^{1-\theta}$，其中某个 $c > 0$ 且 $\theta \in [0, 1)$，那么我们称 f 在 $\bar{x}$ 处具有 *Łojasiewicz* 性质，指数为 θ。

基于 KL 的分析技术已取得巨大成功，并在实践中具有广泛的适用性，[1,3,4,7]。定义 3.7 中引入的 KL 性质和 Łojasiewicz 性质适用于丰富且普遍存在的温和函数、次解析函数和半代数函数类别，参见例如 [49,50,36,6]。

设 $\{x_k\}$ 由算法 2 生成，并定义聚点集 $\mathcal{A} := \{x^* : \liminf_{k\to\infty} \|x_k - x^*\| = 0\}$。我们考虑以下关键假设。

（A.2） $\nabla^2 f$ 在 V_H 的某个邻域内局部Lipschitz连续，其中 $\mathcal{A}$。（C.1）参数序列 $\{a_k\}$ 和 $\{z_k\}$ 满足 $\sum_{k=1}^{\infty} a_k^{-1/\alpha} < \infty$和 $\sum_{k=1}^{\infty} z_k^{-1/\zeta} < \infty$。（C.2） f 满足 KL 性质（23），且某个聚点 $x^* \in \mathcal{A}$。（C.3）函数 f 在某个聚点 $x^* \in \mathcal{A}$ 处满足 Łojasiewicz 性质，其指数为 $\theta \in [0, 1)$。

(C.1)规定了参数 $\{a_k\}$ 和 $\{z_k\}$的增长行为。一个简单的示例性选择是 $a_k = (k\log(k)^2)^\alpha$ 和 $z_k = (k\log(k)^2)^\zeta$。基于(C.2)，我们证明了由算法 2生成的整个迭代点序列 $\{x_k\}$收敛到临界点 x^*。我们的证明依赖于引理3.2、3.4和3.5(iii)中的下降估计，并将[1] 中的原始分析以及 [30] 中近期采用的数学归纳法应用于我们的基于MINRES的框架。

基于假设(A.2)，我们现在首先推导当迭代点接近某个聚点时步长的下界。

引理 3.8。设 $\{x_k\}$ 由算法 *2*生成，且*(A.1)–(A.2)*成立。设 $r > 0$ 和 $x^* \in \mathcal{A}$ 满足 $B_{2r}(x^*) \subset V_H$，并令 $L_g \equiv L_g(x^*)$和 $L_H \equiv L_H(x^*)$分别表示 ∇f 和 $\nabla^2 f$ 在 $B_{2r}(x^*)$上的*Lipschitz* 常数。进一步假设 $x_k,\ x_{k+1} \in B_{2r}(x^*)$。则有：

(i) $\lambda_k \geq \min\{s, \frac{2\rho(1-\sigma)}{L_g}\min\{\mathsf{A}, a_k\|g_k\|^\alpha\}\}$ *if* $\text{flag}_k \neq \text{NPC}$.

(ii) $\lambda_k \geq \frac{3\rho(\sigma-1)d_k^\top B_k d_k}{L_H\|d_k\|^3}$ *if* $\text{flag}_k = \text{NPC}$ *and* $B_k = \nabla^2 f(x_k)$.

证明。 根据下降引理，可得

$$f(x_{k+1}) - f(x_k) - \sigma\lambda_k d_k^\top g_k \leq (1-\sigma)\lambda_k d_k^\top g_k + \frac{L_g}{2}\lambda_k^2\|d_k\|^2. \tag{24}$$

因此，(9) 对所有 $\lambda_k \in (0, \mu_k^1]$ 其中 $\mu_k^1 := -2(1-\sigma)\frac{d_k^\top g_k}{L_g\|d_k\|^2}$ 且我们有 $\lambda_k \geq \rho\mu_k^1$。具体而言，对于 $d_k = p_t$，应用定理 2.5 (i) 和 (8)，可得

$$\lambda_k \geq \frac{2\rho(1-\sigma)p_t^\top \bar{B}_k p_t}{L_g\|p_t\|^2} \geq \frac{2\rho(1-\sigma)}{L_g}\min\{\mathsf{A}, a_k\|g_k\|^\alpha\}.$$

对于 $d_k = -g_k$，利用 $-d_k^\top g_k = \|d_k\|^2$，我们得到 $\lambda_k \geq 2\rho(1-\sigma)/L_g$。注意到 $\mathsf{A} \leq 1$，这就证明了 (i)。接下来，应用 $\nabla^2 f$ (在$B_{2r}(x^*)$上) 的 Lipschitz 连续性以及 $g_k^\top d_k \leq 0$，可得

$$\begin{aligned} & f(x_{k+1}) - f(x_k) - \sigma(\lambda_k g_k^\top d_k + \lambda_k^2 d_k^\top B_k d_k/2) \\ & \quad \leq (1-\sigma)\lambda_k(g_k^\top d_k + \lambda_k d_k^\top B_k d_k/2) + L_H\lambda_k^3\|d_k\|^3/6 \\ & \quad \leq 0.5\lambda_k^2((1-\sigma)d_k^\top B_k d_k + L_H\lambda_k\|d_k\|^3/3). \end{aligned}$$

因此，(10) 对所有 $\lambda_k \in (0, \mu_k^2]$ 其中 $\mu_k^2 := 3(\sigma-1)d_k^\top B_k d_k/(L_H\|d_k\|^3)$并且，根据 非正曲率方向的线搜索机制（参见算法 3），可以推得 $\lambda_k \geq \rho\mu_k^2$。证明完毕。□

下文中，基于引理 3.8中的下界，我们给出算法 2的一个主要收敛性结果。

定理 3.9 （KL 性质下的收敛性，I）。设 (A.1)–(A.2)、(B.1)′、(B.3)以及 (C.1) 均成立，并设 $\{x_k\}$ 由算法2生成。假设 $x^* \in \mathcal{A}$ 是 $\{x_k\}$ 的一个聚点，且假设(C.2)在该点成立。那么，算法要么在有限次迭代后于某个驻点处停止，要么整个序列 $\{x_k\}$ 收敛到 x^*，且 $\|g_k\| \to 0$。

证明。不妨假设算法不会在有限步后停止。那么，根据定理 3.6，x^* 是 f的一个驻点。由于 $f(x_k) \searrow f(x^*)$，$k \to \infty$ （如（15）所示），存在 $k_\eta \in \mathbb{N}$，使得对所有 $k \geq k_\eta$，均有 $0 < f(x_k) - f(x^*) \leq \eta$。

我们首先关注 $\alpha < 1$ 的情形，并回顾 $\mathcal{N}_{\rm npc} := \{k : \text{标记}_k = \text{非正曲率}\}$。由（14），可得

$$f(x_k) - f(x_{k+1}) \geq \sigma\|g_k\|\|s_k\| \cdot \lambda_k\zeta_k/2$$

若 标记$_k$ =非正曲率。（注意，我们不假设条件 (B.2) 成立，因此引理 3.4 （iii）不适用）。结合引理 3.2 （iii）、引理3.5 （iii）以及（14），可得

$$f(x_k) - f(x_{k+1}) \geq \tau_k\sigma\|g_k\|\|s_k\|, \quad \tau_k := \begin{cases} \min\{T_{1,k}, T_{2,k}\} & \text{if } k \notin \mathcal{N}_{\rm npc}, \\ \min\{T_{3,k}, T_{4,k}\} & \text{if } k \in \mathcal{N}_{\rm npc}, \end{cases} \tag{25}$$

其中 $T_{1,\,k} := \min\{C_k\mathsf{A}^2,\ 1\}$, $T_{2,\,k} := C_k(a_k/s^\alpha)^{\frac{2}{1-\alpha}}\|s_k\|^{\frac{2\alpha}{1-\alpha}}$, $T_{3,\,k} := \lambda_k\mathsf{Z}/2$以及 $T_{4,\,k} := \lambda_k^{1-\zeta}z_k\|s_k\|^\zeta/2$。（$T_{4,\,k}$ 的表达式源自 $\|g_k\|^\zeta = \|d_k\|^\zeta = \lambda_k^{-\zeta}\|s_k\|^\zeta$，参见引理 3.4 （ii））。证明的主要步骤之一是确保 $\{\lambda_k\}$ （在 $T_{3,\,k}$ 和 $T_{4,\,k}$中）有下界。

令 $\{x_{k_\ell}\}$ 为收敛到 x^* 的子序列。此外，设给定 $r > 0$，使得 $B_{2r}(x^*) \subset V_H$，并令 L_g、L_H 表示 Lipschitz 常数

关于 ∇f和 $\nabla^2 f$在 $B_{2r}(x^*)$。利用 ∇f的连续性， $\zeta_k =$最小{Z $z_k\|g_k\|^\zeta\}$，$\zeta \le 1$，以及引理 3.4 (ii)，存在 $\delta > 0$ 使得对所有 $x_k \in B_\delta(x^*)$和 $k \in \mathcal{N}_{\rm npc}$，均有 $\|g_k\| \le$最小$\{r/s, 1\}$ 且

$$\frac{3\rho(\sigma-1)d_k^\top B_k d_k}{L_H\|d_k\|^3} \ge \frac{3\rho(1-\sigma)\zeta_k}{L_H\|d_k\|} \ge \frac{3\rho(1-\sigma)}{L_H}\min\left\{\frac{\mathsf{Z}}{\|g_k\|}, \frac{\bar{z}}{\|g_k\|^{1-\zeta}}\right\} \ge \bar{\tau}, \tag{26}$$

其中 $\bar{\tau} := 3\rho(1-\sigma)\min\{\mathsf{Z}\ \bar{z}\}/L_H$。定义

$$\sigma_{1,k} := \sum\nolimits_{i=k}^{\infty} a_i^{-\frac{1}{\alpha}}, \quad \sigma_{2,k} := \sum\nolimits_{i=k}^{\infty} z_i^{-\frac{1}{\zeta}}, \quad \varrho_k := \varrho(f(x_k) - f(x^*)). \tag{27}$$

接下来，存在 $k_{\ell_0} \ge k_\eta$，使得 $\nu \le \min\{r, \delta\}$ 且 $B_\nu(x^*) \subset V$，其中

$$\nu := \|x_{k_{\ell_0}} - x^*\| + \frac{\tilde{\varrho}}{\tilde{T}\sigma} + (s\tilde{\sigma}_1)^{1-p_1}\left(\frac{\tilde{\varrho}}{\tilde{C}\sigma}\right)^{p_1} + \tilde{\sigma}_2^{1-p_2}\left(\frac{\tilde{\varrho}}{\tilde{T}_2\sigma}\right)^{p_2},$$

$\tilde{\varrho} := \varrho_{k_{\ell_0}}$, $\tilde{\sigma}_i := \sigma_{i,k_{\ell_0}}$, $i = 1, 2$, $p_1 := \frac{1-\alpha}{1+\alpha}$, $p_2 := \frac{1}{1+\zeta}$, $\tilde{C} := [(L_g + \mathsf{Z})(L_g + \mathsf{Z} + 1)]^{-1}$, $\tilde{T} = \min\{\tilde{C}\mathsf{A}^2, 1, \bar{\lambda}\mathsf{Z}/2\}/2$, $\tilde{T}_2 := \bar{\lambda}^{1-\zeta}/2$, $\bar{\lambda} := \min\{s, \bar{\tau}\}$，且集合$V$ 定义见 (C.2)。该指标 k_{ℓ_0} 的存在性由 $x_{k_\ell} \to x^*$,$\varrho_{k_\ell} \to 0$ 和 $\sigma_{i,k_\ell} \to 0$, $i = 1,2$ 得出，因为 $\ell \to \infty$。现在我们通过归纳法证明以下三个命题对所有 $k \ge k_{\ell_0}$均成立:

$$x_k \in B_\nu(x^*), \quad \lambda_k \ge \bar{\lambda} \text{ if } k \in \mathcal{N}_{\rm npc}, \quad \sum\nolimits_{i=k_{\ell_0}}^{k} \|s_i\| \le \nu - \|x_{k_{\ell_0}} - x^*\|. \tag{28}$$

根据 ν 的定义，有 $x_{k_{\ell_0}} \in B_\nu(x^*)$。假设我们有 $\lambda_{k_{\ell_0}} < \bar{\lambda}$和 $k_{\ell_0} \in \mathcal{N}_{\rm npc}$。那么，由于 $x_{k_{\ell_0}} \in B_\nu(x^*) \subset B_\delta(x^*)$、$\|d_{k_{\ell_0}}\| = \|g_{k_{\ell_0}}\|$、$\bar{}$和 $\lambda \le s$，可得 $\|x_{k_{\ell_0}+1} - x_{k_{\ell_0}}\| = \lambda_{k_{\ell_0}}\|d_{k_{\ell_0}}\| \le s\|g_{k_{\ell_0}}\| \le r$。利用 $x_{k_{\ell_0}} \in B_\nu(x^*) \subset B_r(x^*)$，这意味着 $x_{k_{\ell_0}+1} \in B_{2r}(x^*)$。因此，适用引理 3.8 的第 (ii) 部分，并调用 $x_{k_{\ell_0}} \in B_\delta(x^*)$ 和 (26)，可得

$$\lambda_{k_{\ell_0}} \ge \frac{3\rho(\sigma-1)d_{k_{\ell_0}}^\top B_{k_{\ell_0}} d_{k_{\ell_0}}}{L_H\|d_{k_{\ell_0}}\|^3} \ge \min\{s, \bar{\tau}\} = \bar{\lambda}.$$

这就产生了矛盾，因此我们必须有 $\lambda_{k_{\ell_0}} \ge \lambda$ 和 $k_{\ell_0} \in \mathcal{N}_{\rm npc}$，或者有 $k_{\ell_0} \notin \mathcal{N}_{\rm npc}$。注意到 $B_\nu(x^*) \subset B_{2r}(x^*)$、$\max_{x\in B_{2r}(x^*)}\|\nabla^2 f(x)\| \le L_g$，以及 $\zeta_k \le \mathsf{Z}$，根据 C_k 和 C的定义，我们有 $C_{k_{\ell_0}} \ge C$。此外，由 (23)、(25)、$x_{k_{\ell_0}} \in V$以及 ϱ的凹性和非负性，可知

$$\begin{aligned}\tau_{k_{\ell_0}}\sigma\|s_{k_{\ell_0}}\| &\le \frac{f(x_{k_{\ell_0}}) - f(x_{k_{\ell_0}+1})}{\|g_{k_{\ell_0}}\|} \\ &\le \varrho'(f(x_{k_{\ell_0}}) - f(x^*))(f(x_{k_{\ell_0}}) - f(x_{k_{\ell_0}+1})) \le \varrho_{k_{\ell_0}} - \varrho_{k_{\ell_0}+1} \le \tilde{\varrho}.\end{aligned} \tag{29}$$

最后，我们定义 $\mathcal{K}_1 := \{k \ge k_{\ell_0}, k \notin \mathcal{N}_{\rm npc} : T_{1,k} \le T_{2,k}\}$, $\mathcal{K}_2 := \{k \ge k_{\ell_0}, k \notin \mathcal{N}_{\rm npc} : T_{1,k} > T_{2,k}\}$, $\mathcal{K}_3 := \{k \ge k_{\ell_0}, k \in \mathcal{N}_{\rm npc} : T_{3,k} \le T_{4,k}\}$, 以及

$\mathcal{K}_4 := \{k \geq k_{\ell_0}, k \in \mathcal{N}_{\text{npc}} : T_{3,k} > T_{4,k}\}$。调用 (25)、(29) 和 $C_{k_{\ell_0}} \geq C$，条件 $T_{2,k_{\ell_0}}\|s_{k_{\ell_0}}\| \leq \tilde{\varrho}/\sigma$ 意味着

$$\|s_{k_{\ell_0}}\| \leq s^{\frac{2\alpha}{1+\alpha}} a_{k_{\ell_0}}^{-\frac{2}{1+\alpha}} (\tilde{\varrho}/(\tilde{C}\sigma))^{\frac{1-\alpha}{1+\alpha}} \leq (s\tilde{\sigma}_1)^{\frac{2\alpha}{1+\alpha}} (\tilde{\varrho}/(\tilde{C}\sigma))^{\frac{1-\alpha}{1+\alpha}}.$$

类似地，在情形 $k_{\ell_0} \notin \mathcal{K}_2$ 下，我们可以得到如下结果

$$\begin{cases} T_{1,k_{\ell_0}}\|s_{k_{\ell_0}}\| \leq \frac{\tilde{\varrho}}{\sigma} \implies \|s_{k_{\ell_0}}\| \leq \frac{\tilde{\varrho}}{2\tilde{T}\sigma} & \text{if } k_{\ell_0} \in \mathcal{K}_1, \\ T_{3,k_{\ell_0}}\|s_{k_{\ell_0}}\| \leq \frac{\tilde{\varrho}}{\sigma} \implies \|s_{k_{\ell_0}}\| \leq \frac{2\tilde{\varrho}}{\mathsf{Z}\bar{\lambda}\sigma} \leq \frac{\tilde{\varrho}}{2\tilde{T}\sigma} & \text{if } k_{\ell_0} \in \mathcal{K}_3, \\ T_{4,k_{\ell_0}}\|s_{k_{\ell_0}}\| \leq \frac{\tilde{\varrho}}{\sigma} \implies \|s_{k_{\ell_0}}\| \leq \tilde{\sigma}_2^{\frac{\zeta}{1+\zeta}} (\frac{\tilde{\varrho}}{\tilde{T}_2\sigma})^{\frac{1}{1+\zeta}} & \text{if } k_{\ell_0} \in \mathcal{K}_4, \end{cases}$$

其中我们使用了 $\lambda_{k_{\ell_0}} \geq \lambda$ 若 $k_{\ell_0} \in \mathcal{K}_3 \cup \mathcal{K}_4$。因此，(28) 对 $k = k_{\ell_0}$成立。

假设对于某个 $k > k_{\ell_0}$，且对于所有 $k_{\ell_0} \leq i \leq k$，（28）均成立。根据三角不等式以及 ν的定义，我们得到

$$\|x_{k+1} - x^*\| \leq \|x_{k_{\ell_0}} - x^*\| + \sum_{i=k_{\ell_0}}^{k} \|s_i\| \leq \nu.$$

这证明了 $x_{k+1} \in B_\nu(x^*)$。条件 $\lambda_{k+1} \geq \lambda$ 若 $k+1 \in \mathcal{N}_{\text{npc}}$ 可再次通过反证法验证，且与基例 $k = k_{\ell_0}$相同。对迭代点 x_i、$i = k_{\ell_0}$、……、$k+1$ 应用估计式 ($x_i \in B_\nu(x^*)$（这对于所有 $i = k_{\ell_0}$，……，$k+1$均成立）是可行的），我们首先注意到

$$\sum_{i=k_{\ell_0}}^{k+1} \tau_i \|s_i\| \leq \sum_{i=k_{\ell_0}}^{k+1} \frac{f(x_i) - f(x_{i+1})}{\sigma\|g_i\|} \leq \sum_{i=k_{\ell_0}}^{k+1} \frac{\varrho_i - \varrho_{i+1}}{\sigma} \leq \frac{\tilde{\varrho}}{\sigma}. \quad (30)$$

根据逆霍尔德不等式，并回顾 $p_1 = \frac{1-\alpha}{1+\alpha} < 1$，我们还可以得到

$$\begin{aligned} \sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} \tau_i\|s_i\| &\geq \tilde{C}s^{-\frac{2\alpha}{1-\alpha}} \sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} a_i^{\frac{2}{1-\alpha}} \|s_i\|^{\frac{1+\alpha}{1-\alpha}} \\ &\geq \tilde{C}s^{-\frac{2\alpha}{1-\alpha}} \left(\sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} (a_i^{\frac{2}{1-\alpha}})^{\frac{-1}{\frac{1}{p_1}-1}}\right)^{1-\frac{1}{p_1}} \left(\sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} \|s_i\|\right)^{\frac{1}{p_1}} \\ &= \tilde{C}s^{1-\frac{1}{p_1}} \left(\sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} a_i^{-\frac{1}{\alpha}}\right)^{1-\frac{1}{p_1}} \left(\sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} \|s_i\|\right)^{\frac{1}{p_1}} \\ &\geq \tilde{C}(s\tilde{\sigma}_1)^{1-\frac{1}{p_1}} \left(\sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} \|s_i\|\right)^{\frac{1}{p_1}}. \end{aligned}$$

这里，估计式 $C_i \geq C$ 成立，这是因为对所有$i = k_{\ell_0}, \ldots, k+1$，$x_i \in B_\nu(x^*) \subset B_{2\tau}(x^*)$。结合式（30），我们可以推得 $\sum_{i=k_{\ell_0}, i\in\mathcal{K}_2}^{k+1} \|s_i\| \leq (s\tilde{\sigma}_1)^{1-p_1}(\tilde{\varrho}/(\tilde{C}\sigma))^{p_1}$。仿照这些步骤并运用归纳法假设，我们还可以得到

$$\sum_{i=k_{\ell_0}, i\in\mathcal{K}_1}^{k+1} \|s_i\| \leq \frac{\tilde{\varrho}}{2\tilde{T}\sigma}, \quad \sum_{i=k_{\ell_0}, i\in\mathcal{K}_3}^{k+1} \|s_i\| \leq \frac{2\tilde{\varrho}}{\sigma\bar{\lambda}\mathsf{Z}} \leq \frac{\tilde{\varrho}}{2\tilde{T}\sigma},$$

且 $\sum_{i=k_{\ell_0}, i\in\mathcal{K}_4}^{k+1} \|s_i\| \leq \tilde{\sigma}_2^{1-p_2}(\tilde{\varrho}/(\tilde{T}_2\sigma))^{p_2}$。这表明 (28) 对 $k+1$ 也成立，从而完成归纳法。由于 (28) 对所有 $k \geq k_{\ell_0}$都成立，取

$k \to \infty$，我们可以推断 $\{x_k\}$ 是一个柯西序列，因此可得$x_k \to x^*$ 且 $\lim_{k\to\infty} \|g_k\| = \|\nabla f(x^*)\| = 0$。

最后，考虑特殊情况 $\alpha = 1$。在这种情况下，如果 A，引理3.2 (iii) 不适用。然而，如果 flag$_k$ =SOL 且 A $\geq a_k\|g_k\|$，那么引理 3.2 (ii) 意味着 $1 \geq a_k\|d_k\|$。因此，重新定义指标集 $\mathcal{K}_2 := \{k \geq k_{\ell_0} : \text{flag}_k = \text{SOL},\ \text{A} \geq a_k\|g_k\|\}$ 和 $\mathcal{K}_1 := \{k \geq k_{\ell_0} :\ k \notin \mathcal{N}_{\text{npc}},\ k \notin \mathcal{K}_2\}$，可得 $\bigcup_{i=1}^{4} \mathcal{K}_i \supset \{k \geq k_{\ell_0}\}$ 和 $\sum_{i=k_{\ell_0},\ i\in\mathcal{K}_2}^{\infty} \|s_i\| \leq \sum_{i=k_{\ell_0},\ i\in\mathcal{K}_2}^{\infty} \lambda_i/a_i \leq s\tilde{\sigma}_1$。因此，注意到 $p_1 = 0$，在情形 $\alpha < 1$ 下，前面的步骤仍然全部有效，并且使用相同的 ν[3] 定义。证毕。 □

定理 3.9 中推导出的结果在以下情况下仍然成立：将条件 (A.2) 和 (B.1)′ 替换为(B.1)–(B.2)，即矩阵 $\{B_k\}$保持有界，并采用容差参数 $\{\theta_k\}$ 的特定形式。现陈述定理 3.9的这一变体。

定理 3.10 （KL 性质下的收敛性，II）。设 (A.1)、(B.1)–(B.3)以及 (C.1) 成立，并设 $\{x_k\}$ 由算法 2生成。进一步假设 $\sum_{k=1}^{\infty} b_k^{-1/\beta} < \infty$，并设 $x^* \in \mathcal{A}$ 是$\{x_k\}$ 的一个聚点，且在该点满足假设 (C.2) 。则算法要么在有限次迭代后于某个驻点处停止，要么整个序列 $\{x_k\}$收敛于 x^*，且 $\text{Lim}_{k\to\infty} \|g_k\| = 0$。

证明。 定理 3.10 的证明与定理 3.9 的推导类似。特别是，引理3.4 (iii) 适用，且不再需要引理 3.8 中关于步长的界 $\{\lambda_k\}$。若 $\alpha < 1$ 并结合引理 3.2、3.4 和 3.5 (iii)，关键估计式（25）现在可写为

$$f(x_k) - f(x_{k+1}) \geq \sigma\|g_k\|\|s_k\| \min\{\hat{T}, \hat{T}_{2,k}, \hat{T}_{3,k}, \hat{T}_{4,k}\},$$

其中 $\hat{T} := \min\{\bar{C}\text{A}^2, \text{B}^2, 2, 1\}$, $\hat{T}_{2,k} := \bar{C}(a_k/s^\alpha)^{\frac{2}{1-\alpha}}\|s_k\|^{\frac{2\alpha}{1-\alpha}}$, $\hat{T}_{3,k} := \frac{b_k}{s^\beta}\|s_k\|^\beta$, 以及 $\hat{T}_{4,k} := \frac{s^{1-\zeta}}{2} z_k\|s_k\|^\zeta$。这里，$\bar{C} > 0$ 是 C_k的一个（统一）下界，即对于所有 $k \in \mathbb{N}$，我们都有 $C_k = (\|\bar{B}_k\| + \|\bar{B}_k\|^2)^{-1} \geq \bar{C}$。（由于（B.1）和 $\zeta_k \leq 2$，存在这样的 $\bar{C}$。）其余步骤与定理 3.9 的证明基本相同。此处不再赘述。 □

最后，我们给出一个关于 $\{\|g_k\|\}$收敛速率的技术性结果，该结果基于 Łojasiewicz 条件(C.3)。引理3.11 将在下一节中用于确立算法2的鞍点规避性质。

引理 3.11 (收敛速率)。设满足(A.1)–(A.2)、(B.1)′、(B.3)和(C.1)，且 $\{x_k\}$ 由算法2生成。假设 $x^* \in \mathcal{A}$是 $\{x_k\}$ 的一个聚点，且性质(C.3) 在该点以指数$\theta \in [0, 1)$成立。令 $\omega = \frac{(1+\text{最大值}\{2\alpha,\zeta\})\ \theta)}{1-\theta}$，则对所有 $p \in (0, \frac{1}{\omega} - 1)$，有 $\liminf_{k\to\infty} k^{1+p}\|g_k\| = 0$。

引理3.11 的证明放在A节。

[3] 在标志位为k = SOL 且 A $< a_k\|g_k\|$的情况下，我们有 $f(x_k) - f(x_{k+1}) \geq C_k\text{A}^2\sigma\|g_k\|\|s_k\|$，如引理 3.2 的证明所示。因此，利用引理 3.5 (iii)，可确保对所有 $k \in \mathcal{K}_1$，(25) 在 $\tau_k = T_{1,k}$ 下成立。

3.4 鞍点规避

本节的目标是证明，当我们选择使用海森矩阵信息时，算法2 能够有效避开严格鞍点， $B_k = \nabla^2 f(x_k)$。更具体地说，我们旨在证明，算法2生成的迭代点只能收敛到满足 $x^* \quad \nabla f(x^*) \ = 0$和 $\nabla^2 f(x^*) \succeq 0$的二阶临界点。我们考虑以下附加假设。

（D.1）定义 $\mathcal{N} := \{k : \lambda_{\min}(\bar{B}_k) \le 0\}$ 并回顾 $\mathcal{N}_{\rm npc} \equiv \{k \in \mathcal{N} : \text{flag}_k = \text{非正曲率}\}$，我们假设 $|\mathcal{N}| = \infty \implies |\mathcal{N}_{\rm npc}| = \infty$。（D.2）我们有 $\zeta_k = \min\{\text{Z}\ z_k\|g_k\|^\zeta\}$，其中 Z $> 0, \zeta \in (0, \frac{1}{3}], z_k \ge \bar{z} > 0$，且 $z_k = \mathcal{O}((k\log(k)^2)^\zeta), \ k \to \infty$，以及 $\sum_{k=1}^{\infty} a_k^{-1/\alpha} < \infty$。

在 （D.2） 下，前几节中使用的假设 （B.3） 和 （C.1） 自然得到满足。现在我们给出主要的规避结果。

定理 *3.12*（鞍点规避）。设（*A.1*）–（*A.2*）、（*B.1*）′以及（*D.1*）–（*D.2*）成立，并设 $\{x_k\}$ 由算法 *2* 生成。假设该算法经过有限次迭代后不会停止。令 $x^* \in \mathcal{A}$ 为 $\{x_k\}$ 的一个聚点，且该聚点满足性质（*C.3*）。那么，$\{x_k\}$ 收敛于 x^*，并且有 $\nabla f(x^*) = 0$ 以及 $\nabla^2 f(x^*) \succeq 0$。

证明。 根据定理 3.9，我们有 $x_k \to x^*$ 且 $\|g_k\| \to 0$，不失一般性，假设 $\{x_k\} \subset B_{2r}(x^*) \subset V_H$， $r > 0$。令 L_H 进一步表示 $\nabla^2 f$ 在 $B_{2r}(x^*)$上的 Lipschitz 常数。根据引理 3.11 和 (D.2)，我们有

$$0 \le \text{下极限}_{k\to\infty}\ \zeta_k \le \text{下极限}_{k\to\infty}\ z_k\|g_k\|^\zeta \le \text{上极限}_{k\to\infty} \frac{M\log(\quad k\ ^{2\zeta})}{k^{p\zeta}}\ \text{下极限}_{k\to\infty}(k^{1+p}\|g_k\|)^\zeta = 0,$$

对于某个 $M > 0$和 $p \in (0, \frac{1}{\omega-1})$。现在假设 $\nabla^2 f(x^*)$ 是不定的。这意味着下极限 $_{k\to\infty}\lambda_{\min}(\bar{B}_k) =$下极限$_{k\to\infty}\lambda_{\min}(\nabla^2 f(x_k) + \zeta_k I) < 0$，根据 (D.1)，我们可以推出 $|\mathcal{N}| = \infty$ 和 $|\mathcal{N}_{\rm npc}| = \infty$。不失一般性，假设 $\lim_{\mathcal{N}_{\rm npc}\ni k\to\infty}\zeta_k = 0$。利用 (D.2) 和引理 3.4 (ii)，可得

$$\frac{3\rho(\sigma-1)d_k^\top B_k d_k}{L_H\|d_k\|^3} \ge \frac{3\rho(1-\sigma)\zeta_k}{L_H\|d_k\|} = \frac{3\rho(1-\sigma)}{L_H}\frac{z_k}{\|g_k\|^{1-\zeta}} \to \infty$$

时 $\mathcal{N}_{\rm npc} \ni k \to \infty$。因此，根据引理 3.8 (ii)，存在 $k \in \mathbb{N}$ 使得$\lambda_k \ge \frac{3\rho(\sigma-1\ d_k^\top B_k d_k}{L_H\|d_k\|^3} > s)$对所有 $\mathcal{N}_{\rm npc} \ni k \ge \bar{k}$成立。令 $\mathcal{K} := \{k \in \mathcal{N}_{\rm npc} : \quad k \ge \bar{k}\}$并利用 (D.2)、（10）和（30）（必要时增大 k），可得

$$\begin{aligned}
\infty > \varrho_{\bar{k}} &> \sum_{k\in\mathcal{K}} \frac{f(x_k) - f(x_{k+1})}{\sigma\|g_k\|} \ge \sum_{k\in\mathcal{K}} -\frac{\lambda_k d_k^\top g_k + \frac{1}{2}\lambda_k^2 d_k^\top B_k d_k}{\|g_k\|} \\
&\ge \sum_{k\in\mathcal{K}} -\frac{\lambda_k^2 d_k^\top B_k d_k}{2\|g_k\|} \ge \sum_{k\in\mathcal{K}} \frac{9\rho^2(\sigma-1)^2(-d_k^\top B_k d_k)^3}{2L_H^2\|d_k\|^6\|g_k\|} \\
&\ge \sum_{k\in\mathcal{K}} \frac{4\rho^2(\sigma-1)^2}{L_H^2}\frac{\zeta_k^3}{\|g_k\|} = \frac{4\rho^2(\sigma-1)^2}{L_H^2}\sum_{k\in\mathcal{K}}\frac{z_k^3}{\|g_k\|^{1-3\zeta}} = \infty,
\end{aligned}$$

其中 $\varrho_{\bar{k}} = c(f(x_{\bar{k}}) - f(x^*))^{1-\theta}$，参见 (C.3)。这就产生了矛盾。 □

备注 3.13。我们的分析表明，严格鞍点规避不需要直接使用与 $\lambda_{min}(B_k)$相关的非正曲率方向；相反，伪负特征值 $-\zeta_k$ 就足够了。这种无需特征对的机制降低了计算成本，并能提升实际性能。我们推导中的关键洞见是：如果 x_k 趋近于一个严格鞍点 x^*，条件 (10) 以及前向线搜索迫使步长增大其阶为 $\sim 1/\|g_k\|^{1-\zeta}$，因为 $\|g_k\| \to 0$。结合 (10)，可推得这些项 $(f(x^k) - f(x^{k+1}))/\|g_k\|$, $k \in \mathcal{N}_{\rm npc}$ 发散且不可求和。然而，这与定理 *3.9* 中给出的基于*KL*的保证相矛盾。我们的论证受到 *[27, 14, 47]*, 中分析的启发，但我们进一步利用 (C.2)–(C.3) 所提供的几何性质。

¯注 *3.14*。条件 *(D.1)* 意味着，如果矩阵 B_k 在无限多次迭代中包含 *NPC* 信息，那么 *MINRES* 就能检测到此类 *NPC* 信息——至少在一个无限子序列上。我们注意到，在 θ_k 中令 $b_k = (k\log(k)^2)^\beta$，并仿照定理 *3.12* 证明中关于 ζ_k 的推导，可得 $\liminf_{k\to\infty} \theta_k = 0$，因此 *MINRES* 子问题的求解精度会不断提高。在这种情形下，在合理的假设下可以保证条件 *(D.1)* 成立。例如，若 g_k 在由 B_k 的非正特征向量张成的特征子空间上的正交投影¯非平凡，那么根据 *MINRES* 的*[61,* 式 *(4)*，在最后一次内层迭代时必有 $\lambda_{\min}(T_k) = \lambda_{\min}(B_k) \leq 0$，从而保证检测到 *NPC*。此外，取 $\theta_k = 0$, $k \in \mathbb{N}$时，*(D.1)* 以概率 *1* 成立。在这种情况下，如果 B_k不是正定的，那么与正特征值对应的特征向量所张成空间的维数严格小于 n。因此，由于 $\theta_k = 0$，且以概率 *1*，*MINRES* 将返回一个负曲率方向，而不会提前停止。类似的概率论证也见于例如 *[27,*第 *4]*节。注意，如果线性系统不相容，*MINRES* 会将非零残差向量 r_{g-1} 作为零曲率方向输出，参见 *[46,* 引理 *3.4 (iii)]*, ，其中 g 是该系统的阶。

3.5 快速局部收敛到非孤立局部极小值

如前面各小节所示，基于 Kurdyka-Łojasiewicz 不等式的分析技术使我们能够确保迭代点 $x_k \to x^*$收敛。现在，在假设（更强的）Łojasiewicz 不等式 (C.3) 在 x^*处成立的条件下，我们将讨论更精细的局部收敛性质；也就是说，存在 V 的一个邻域 x^*， $\theta \in (0, 1)$，以及 η， $c > 0$ 使得

$$[f(x) - f(x^*)]^{2\theta} \leq (2\mu)^{-1}\|\nabla f(x)\|^2, \tag{31}$$

对所有 $x \in V \cap \{x : 0 < f(x) - f(x^*) < \eta\}$，其中 $\mu := 1/(2c^2(1-\theta)^2)$。此外，我们关注 x^* 是目标函数 f 的（可能非孤立的）局部极小值的情形。为此，我们定义

$$\mathcal{S} := \{x \in \mathbb{R}^n : x \text{ is a local minimum and } f(x) = f(x^*)\}.$$

Łojasiewicz 不等式（31）已知可推出二次增长条件

$$f(x) - f(x^*) \geq [(1-\theta)\sqrt{2\mu}]^{\frac{1}{1-\theta}} \operatorname{dist}(x, \mathcal{S})^{\frac{1}{1-\theta}} \quad \forall\, x \in V', \tag{32}$$

其中 V' 是 $x^* \in \mathcal{S}$的一个合适邻域，参见 [66, 命题 2。2]。结合（31）和（32），局部极小值 $x^* \in \mathcal{S}$的某个邻域内必须满足以下误差界条件：

$$(1-\theta)(2\mu)^{\frac{1}{2\theta}} \operatorname{dist}(x, \mathcal{S}) \leq \|\nabla f(x)\|^{\frac{1-\theta}{\theta}}. \tag{33}$$

关于条件（31）、（32）和（33）的更多讨论，请参见 [20,5,66] 。基于这些特定条件，我们现在证明，可以实现快速局部收敛，其收敛速率与 [65]中的结果相匹配。我们的证明策略基于对正则化参数 $\{\zeta_k\}$的适当选择。特别是，容差 $\{\theta_k\}$ 以及 MINRES 的固有性质所起的作用并不那么关键——这与 [65] 中的整体推导形成对比。在以下引理中，我们首先证明，在 x^*的一个局部区域内，线性系统（3）是一致的，并且 MINRES 总会返回 SOL-方向。这也使我们能够加强引理 3.2中的一些先前结果。

引理 3.15. 设 *(A.1)–(A.2)*、*(B.1)′*、*(B.3)* 和 *(C.1)* 均成立，且$\{x_k\}$ 由算法 *2* 生成。假设 $x^* \in \mathcal{A} \cap \mathcal{S}$是 $\{x_k\}$ 的一个聚点，且在该点性质 *(C.3)* 以指数 $\theta \in (0, 1)$ 成立。进一步假设，对于所有 k，$\zeta \leq \frac{1-\theta}{\theta}$，$\zeta_k \geq 2\min\{A\ a_k\|g_k\|^\alpha\}$，且 $\theta_k = \mathcal{O}(\zeta_k)$，$k \to \infty$。那么存在 $k \in \mathbb{N}$，使得：

（i）由MINRES 返回的所有方向 d_k, $k \geq k$均为SOL方向。（ii）对于所有 $k \geq \bar{k}$，均有 $-d_k^\top g_k \geq \|g_k\|^2/M_1$, $\|d_k\| \leq M_2 \max\{\|g_k\|, \|g_k\|^{\frac{1-\theta}{\theta}}\}$，且 $\bar{B}_k \succeq \frac{1}{2}\zeta_k I$，其中常数 $M_1, M_2 > 0$。

证明。我们注意到，存在 x^* 的一个邻域，使得欧几里得投影 $\mathcal{P}_\mathcal{S}$ 到 $\mathcal{S}$ 上的投影是良定义的（即集合 $\mathcal{P}_\mathcal{S}(x)$ 对于所有 x充分接近 x^*的 [66, 引理 1。4]）均非空。此外，由于 (C.3) 和 $x^* \in \mathcal{S}$，误差界条件（33）需要在 x^*的一个邻域内成立。因此，根据矩阵特征值的局部 Lipschitz 连续性，我们可以推得

$$\lambda_i(\nabla^2 f(x)) - \lambda_i(\nabla^2 f(\mathcal{P}_\mathcal{S}(x))) = \mathcal{O}(\operatorname{dist}(x, \mathcal{S})) = \mathcal{O}(\|\nabla f(x)\|^{\frac{1-\theta}{\theta}}) \tag{34}$$

对于邻域内的所有 $i = 1$、.....、n 和 x。特别地，根据 $\mathcal{S}$ 的定义，存在 $M > 0$，使得对于所有接近 x^* 的 x，min$_{i=1}$, ..., $_n\,\lambda_i$（$\nabla^2 f$（x））$\geq -M\|\nabla f$（x）$\|^{\frac{1-\theta}{\theta}}$。由于 $z_k \to \infty$（这由（C.1）推出）、$\zeta \leq \frac{1-\theta}{\theta}$、（34）、以及 $x_k \to x^*$（根据定理 3.9），存在 k 和 $M_1 > 0$，使得

$$B_k + \frac{1}{2}\zeta_k I \succeq 0, \quad \|\bar{B}_k\| \leq \|B_k - \nabla^2 f(x^*)\| + \|\nabla^2 f(x^*)\| + \zeta_k \leq M_1$$

对所有 $k \geq k$ 均成立。因此，B_k 是正定的，算法最终将不再返回 NPC方向。此外，我们有 $d_k^\top \bar{B}_k d_k \geq \frac{1}{2}\zeta_k\|d_k\|^2$，并且由于 $\zeta_k \geq 2$min{A $a_k\|g_k\|^\alpha$}，曲率检验将通过。因此，不会

返回梯度下降方向。由于 B_k 现已确保为正定，我们可以利用 (12) 得到

$$-d_k^\top g_k \geq \frac{g_k^\top \bar{B}_k g_k}{g_k^\top \bar{B}_k \bar{B}_k g_k}\|g_k\|^2 \geq \frac{g_k^\top \bar{B}_k g_k}{\|\bar{B}_k\|\|\bar{B}_k^{1/2} g_k\|}\|g_k\|^2 \geq \frac{1}{M_1}\|g_k\|^2.$$

接下来，令 $y_k := \bar{B}_k^{-1} g_k = (B_k + \zeta_k I)^{-1} g_k$ 以及 $\hat{x}_k := \mathcal{P}_{\mathcal{S}}(x_k)$。那么，根据泰勒展开和 (A.2)，可得

$$\begin{aligned} y_k &= \bar{B}_k^{-1}[\nabla f(\hat{x}_k) + \nabla^2 f(\hat{x}_k)(x_k - \hat{x}_k) + \mathcal{O}(\|\hat{x}_k - x_k\|^2)] \\ &= \bar{B}_k^{-1}[B_k(x_k - \hat{x}_k) + \mathcal{O}(\|\hat{x}_k - x_k\|^2)] \\ &= [x_k - \hat{x}_k] + \bar{B}_k^{-1}[\zeta_k(\hat{x}_k - x_k) + \mathcal{O}(\|\hat{x}_k - x_k\|^2)] \end{aligned}$$

因此，利用 $\|\hat{x}_k - x_k\| = \mathrm{dist}(x_k, \mathcal{S})$ 和 $\bar{B}_k \succeq \frac{1}{2}\zeta_k I$，可得 $\|y_k\| = \mathcal{O}(\mathrm{dist}(x_k, \mathcal{S})) + \mathcal{O}(\zeta_k^{-1}\,\mathrm{dist}(x_k, \mathcal{S})^2)$。应用 $\bar{B}_k d_k + r_k = -g_k$、误差界（33）、$\theta_k = \mathcal{O}(\zeta_k)$ 和 $\zeta \leq \frac{1-\theta}{\theta}$，存在某个 $M_2 > 0$ 使得

$$\begin{aligned} \|d_k\| &= \|\bar{B}_k^{-1}(r_k + g_k)\| \leq \|\bar{B}_k^{-1}\|\|r_k\| + \|y_k\| \\ &\leq 2\zeta_k^{-1}\theta_k\|g_k\| + \mathcal{O}(\max\{1, \zeta_k^{-1}\mathcal{O}(\mathrm{dist}(x_k, \mathcal{S}))\}\mathcal{O}(\mathrm{dist}(x_k, \mathcal{S}))) \\ &\leq 2\zeta_k^{-1}\theta_k\|g_k\| + \mathcal{O}(\max\{\|g_k\|^{\frac{1-\theta}{\theta}}, \zeta_k^{-1}\|g_k\|^{\frac{2-2\theta}{\theta}}\}) \\ &\leq M_2 \max\{\|g_k\|, \|g_k\|^{\frac{1-\theta}{\theta}}\} \end{aligned}$$

（如有必要，增大 k）。证明至此完成。 □

现在，我们准备给出 $\{\|g_k\|\}$ 的局部超线性收敛性。

定理3.16。 设(A.1)–(A.2)、(B.1)′、(B.2)–(B.3)和(C.1) 均成立，并且令 $\{x_k\}$ 由算法2生成。假设该算法不会在有限次迭代后停止，且 $x^* \in \mathcal{A} \cap \mathcal{S}$ 是 $\{x_k\}$ 的一个聚点，在此处性质(C.3) 以指数 $\theta = \frac{1}{2}$ 成立。此外，假设 $b_{k+1}/b_k \to 1$，$z_{k+1}/z_k \to 1$，$\zeta_k \geq 2\min\{\mathsf{A}\ a_k\|g_k\|^\alpha\}$，$\theta_k = \mathcal{O}(\zeta_k)$，且$\sigma \in (0, \bar{1}2)$。那么，序列 $\{\|\nabla f(x_k)\|\}$ 以阶数 $1 + \omega$ 超线性收敛至零，其中 $\omega \in (0, \min\{\beta, \zeta\})$。

证明。根据定理 3.9，我们有 $x_k \to x^*$ 和 $\|g_k\| \to 0$，并且不妨设 $\{x_k\} \subset B_{2r}(x^*) \subset V_H$，$r > 0$。再次令 L_g, L_H 表示 ∇f 和 $\nabla^2 f$ 在 $B_{2r}(x^*)$ 上的 Lipschitz 常数。根据我们对参数的假设，引理 3.15 适用（取 $\theta = \frac{1}{2}$），即存在 $\bar{k} \in \mathbb{N}$，使得对所有 $k \geq \bar{k}$，都有 $d_k = p_t$。由定理 2.5 (i)、$\bar{\sigma} \in (0, \bar{1}2)$ 以及引理 3.15 (ii)，可得

$$\begin{aligned} &f(x_k + d_k) - f(x_k) - \sigma g_k^\top d_k \leq (1 - \sigma) g_k^\top d_k + \frac{1}{2} d_k^\top B_k d_k + \frac{L_H}{6}\|d_k\|^3 \\ &\leq (\sigma - 1) d_k^\top \bar{B}_k d_k + \frac{1}{2} d_k^\top B_k d_k + \frac{L_H}{6}\|d_k\|^3 \\ &= \left(\sigma - \frac{1}{2}\right) d_k^\top \bar{B}_k d_k - \frac{1}{2}\zeta_k\|d_k\|^2 + \frac{L_H}{6}\|d_k\|^3 \end{aligned}$$

$$\leq \Big(\Big(\frac{\sigma}{2}-\frac{3}{4}\Big)\zeta_k+\frac{L_H}{6}\|d_k\|\Big)\|d_k\|^2$$
$$\leq \Big(\Big(\frac{\sigma}{2}-\frac{3}{4}\Big)\min\{\mathsf{Z},z_k\|g_k\|^{\zeta}\}+\frac{L_H M_2}{6}\|g_k\|\Big)\|d_k\|^2 \tag{35}$$

对所有 $k\geq k$。此外，由于 $\|g_k\|\to 0$ 为 $k\to\infty$ 且 $\zeta\leq 1$，存在 $K\geq k$，使得对所有 $k\geq K$，(35) 右端均非正。因此，步长 $\lambda_k=1$ 将在 Armijo 线搜索中对所有$k\geq K$被接受。于是，利用引理 3.15 (ii) 和 (9)，可得

$$\frac{f(x_k)-f(x_{k+1})}{\|g_k\|}\geq \sigma\lambda_k\frac{-d_k^{\top}g_k}{\|g_k\|}\geq\frac{\sigma}{M_1}\|g_k\|\quad\forall\, k\geq K. \tag{36}$$

按照我们先前采用的策略（参见定理 3.9 的证明中的（29）和（30）），对 36 式求和，其中 $k\geq j\geq K$，并结合 31 式及 $\theta=\frac{1}{2}$，可得

$$\Gamma_j:=\sum_{k=j}^{\infty}\|g_k\|\leq\frac{M_1 c}{\sigma}\sqrt{f(x_j)-f(x^*)}\leq\frac{M_1 c}{\sqrt{2\mu\sigma}}\|g_j\|=C_\theta(\Gamma_j-\Gamma_{j+1}),$$

其中 $C_\theta=M_1c/(\sqrt{2\mu\sigma})$。若 $C_\theta\leq 1$，则得到 $\Gamma_{j+1}\leq 0$，并且有限步收敛。否则，可推得 $\Gamma_{j+1}\leq(1-1/C_\theta)\Gamma_j$，这证明了 $\{\Gamma_j\}$ 的 q-线性收敛性以及 $\{\|g_j\|\}$ 的 r-线性收敛性。接下来，我们将使用以下技术性辅助结果（证明见第 B 节）。

引理 3.17。设 $\{u_k\}\subset\mathbb{R}_+$ 是一个 r-线性收敛到0 的序列，并给定 $\{v_k\}\subset\mathbb{R}_+$，满足 $v_{k+1}/v_k\to 1$。则有 $u_k v_k\to 0$。

因此，应用引理3.17，我们可以推知 $b_k^{\gamma_1}\|g_k\|^{\gamma_2}\to 0$ 且 $z_k^{\gamma_1}\|g_k\|^{\gamma_2}\to 0$，$k\to\infty$，对于任意固定的 γ_1，$\gamma_2>0$。因此，根据(A.2) 和引理3.15，并利用 $\|B_kd_k+g_k\|=\|r_k\|\leq\theta_k\|g_k\|$，我们有

$$\begin{aligned}\|g_{k+1}\|&=\|g_{k+1}-g_k-B_kd_k+B_kd_k+g_k\|\\&\leq\|\nabla f(x_{k+1})-\nabla f(x_k)-\nabla^2 f(x_k)d_k\|+\zeta_k\|d_k\|+\|r_k\|\\&\leq\frac{L_H}{2}\|d_k\|^2+\min\{\mathsf{Z},z_k\|g_k\|^{\zeta}\}M_2\|g_k\|+\min\{\mathsf{B},b_k\|g_k\|^{\beta}\}\|g_k\|\\&\leq\frac{L_H M_2^2}{2}\|g_k\|^2+M_2z_k\|g_k\|^{1+\zeta}+b_k\|g_k\|^{1+\beta}\end{aligned}$$

如 $k\to\infty$。至此，证明完成。 □

定理3.16 中对参数的不同假设可通过以下简单且典型的选择得到满足：

$$\alpha=\beta=\zeta,\quad 2a_k=z_k=(k\log(k)^2)^{\zeta},\quad b_k\leq z_k,\quad 2\mathsf{A}=\mathsf{Z},\quad \mathsf{B}\leq\mathsf{Z},$$

以及 $b_{k+1}/b_k\to 1$。在这种情况下，令 $\alpha=\beta=\zeta=1$，收敛速度可以任意接近二次收敛。

注 3.18。 参数 ζ_k 在我们的推导中起着至关重要的作用。在全局收敛分析中，它使我们能够沿非正曲率方向获得下降估计。此外， ζ_k 作为矩阵 B_k 的伪特征值，有助于以无特征对方式避开严格鞍点。通过适当选择 ζ_k 以及曲率检验 (8)中的参数，我们可以进一步确保 B_k 正定。特别地，对于满足 *Łojasiewicz* 不等式且具有 $\theta = \frac{1}{2}$ 的非孤立极小值附近，线性系统 (3)是一致的。这为局部分析带来了许多有益性质。

4 数值实验

本节报告所提出的线搜索方法在不同问题上的数值表现。我们首先在第 4.1 节列出实现细节。随后，我们在第 1 节中对算法 3.16（如定理 4.2 中所讨论）的局部超线性收敛性质进行数值验证和说明。该测试使用 Matlab 在配备 2 GHz 四核 Intel Core i5 处理器和 16 GB 内存的 MacBook Pro 上完成。我们还在第 4.3 节和 4.4 节中，针对大规模自编码器问题以及 CUTEst 测试集合中的一系列问题进行了额外测试。这些实验使用 Python 在配备 NVIDIA GeForce RTX 3090 和 24 GB 内存的服务器上运行。

4.1 实现细节

我们的数值比较将基于以下算法：

– **Newton-CR** [16], **Newton-CG** [17]: 采用 CR 或共轭梯度法作为线性系统求解器的经典牛顿法；步长使用强Wolfe-Powell线搜索条件 [56] 生成。– **Armijo-L-BFGS** [39,34]: 我们测试采用 Armijo 线搜索和谨慎更新方案的 L-BFGS 方法；具体而言，仅当 $y_k^\top s_k \geq 10^{-18}\|s_k\|^2$ 时才使用曲率对 $\{y_k, s_k\}$ 。–**Wolfe-L-BFGS** [44]: 我们实现带有强Wolfe-Powell线搜索条件的经典 Liu-Nocedal L-BFGS方法。[4]

我们实现了2 算法，采用 L-BFGS 近似和完整海森矩阵，分别称为**L-BFGS-MR** 和 **Newton-MR**。**L-BFGS-MR**采用略作调整的曲率检验（见（37）），而 **Newton-MR** 采用原始检验（8）。L-BFGS 矩阵通过以下紧凑形式构造 [9]:

$$S_k = \left[s_{j_0}, \ldots, s_{j_{\tilde{m}-1}}\right], \quad Y_k = \left[y_{j_0}, \ldots, y_{j_{\tilde{m}-1}}\right], \quad \tilde{m} = \min\{m, k+1\}$$
$$B_k^{\mathrm{BFGS}} = \gamma_k I - \left[\gamma_k S_k \; Y_k\right] \begin{bmatrix} \gamma_k S_k^\top S_k & L_k \\ L_k^\top & -D_k \end{bmatrix}^{-1} \left[\gamma_k S_k \; Y_k\right]^\top.$$

[4] 请注意，**Newton-CR**、**Newton-CG**和**Wolfe-L-BFGS** 中的强Wolfe-Powell线搜索条件基于Moré-Thuente线搜索 [56], , 这与 [47, 第 4] 节中使用的带缩放步骤的线搜索 [58, 算法 35] 不同。

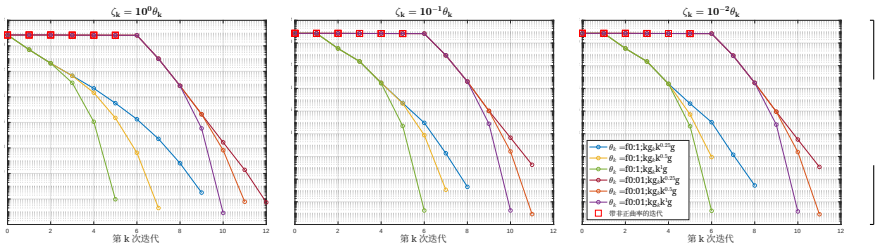


图1 的收敛性 $\{\|g_k\|\}$ 在玩具函数（38）上，不同 θ_k 和 ζ_k取值的结果。检测到 NPC 信息的迭代以红色方块标记。

其中， L_k 是 $S_k^\top Y_k$的严格下三角部分， D_k 是 $S_k^\top Y_k$的对角部分， R_k是 $S_k^\top Y_k$的非严格上三角部分。我们选择 $m = 10$ 和 $\gamma_k = y_k^\top y_k / y_k^\top s_k$。与**Armijo-L-BFGS**不同，当 $|y_k^\top s_k| \geq 10^{-18}\|s_k\|^2$时，我们使用配对 $\{y_k, s_k\}$ 来更新 B_k^{BFGS}。因此，L-BFGS 矩阵通常可以是不定的，这使我们能够探索可能出现的非正曲率。

所有算法均使用从 $[0,1]^n$ 上的均匀分布生成的相同初始点。我们将所有子问题求解器的最大迭代次数设为 1000。当 $\|g_k\| \leq 10^{-10}$、步长小于 10^{-18} 或预言机调用次数达到 10^5 时，终止算法。在算法 3 中，我们设定 $s = 1$、 $\rho = 0.5$、 $\sigma = 10^{-4}$ ，并在 Wolfe-Powell 线搜索中设定 $\sigma_{\text{Wolfe}} = 0.9$ 。在（8）、（37）和（B.3）中，我们设定 A $= 10^{-12}/2$、Z $= 10^{-12}$、$a_k =$ （$k$log（k）^{2}）$^\alpha/2$ 以及 $z_k =$ （$k$log（k）^{2}）$^\zeta$ ，其中 $\alpha = \zeta = 1$ 和 $\bar{M} = 10^8$。对于所有截断牛顿法，我们选择 $\theta_k = \min\{0.1, \sqrt{\|g_k\|}\}$ ；对于 L-BFGS-MR，则选择 $\theta_k = \min\{0.1,\log(k)\sqrt{k\|g_k\|}\}$ 。

注4.1。 如果我们将 (8) 替换为以下精化曲率检验

$$\begin{cases} p_t^\top \bar{B}_k p_t \geq \min\{\mathsf{A}, a_k\|g_k\|^\alpha\}\max\{\|p_t\|^2, \|g_k\|^2\}, & \text{if } d_k = p_t, \\ |d_k^\top B_k d_k| < \bar{M}\|d_k\|^2, & \text{if } d_k = r_{t-1}, \end{cases} \tag{37}$$

$\bar{M} > 0$，那么定理 *3.6* 和 *3.10* 无需 (B.1)即可成立。具体而言，根据定理*3.2(i)* 和 $C_k = 1$ *2.5 (i)* 以及不等式 （37）中的第一个不等式，引理 *3.2*(i) 得以满足。此外，利用不等式 （37）中的第二个不等式，我们可以将 （21）重写为

$$(\sigma - 1)\theta_{k_\ell}\|g_{k_\ell}\|^2 < \rho^{-1}\lambda_{k_\ell} d_{k_\ell}^\top (\hat{B}_{k_\ell} - B_{k_\ell}) d_{k_\ell}/2 \leq \rho^{-1}\lambda_{k_\ell}(\|\hat{B}_{k_\ell}\| + \bar{M})\|d_{k_\ell}\|^2/2,$$

这同样会导致矛盾，因为 $\ell \to \infty$。因此，定理 *3.6* 和 *3.10*仍然成立。当近似矩阵 $\{B_k\}$ 被选为不定且可能无界的（拟牛顿）矩阵时，可以使用检验 (37)

4.2 局部超线性收敛性检验

我们首先使用测试函数 f ，在 [65, 附录 A] 中对**Newton-MR** (如定理3.16中所讨论的)的局部超线性收敛性进行实证验证：

$$f(x, y) := \frac{1}{2}\|y - \sin(x)\|^2, \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^n. \tag{38}$$

数据集	n	d	编码器网络架构	参考值
CIFAR10	50，000	1，664，232	$3,072-256-128-64-32-16-8$	[35]
STL10	5，000	8,365,504	27，$648-150-100-50-25-6$	[12]

表 2 CIFAR10和 STL10 数据集的自编码器模型。

数据集： CIFAR10	Iter	Fun	$\|g_k\|$	预言机	时间（秒）	比例
L-BFGS-MR	303.7	$8.629\cdot10^{-2}$	$7.94\cdot10^{-11}$	687.8	163.2	100%
Armijo-L-BFGS	323.7	$8.784\cdot10^{-2}$	$4.16\cdot10^{-10}$	789.7	160.2	0%
Wolfe-L-BFGS	286.4	$8.759\cdot10^{-2}$	$4.06\cdot10^{-10}$	1250.0	498.7	0%
Newton-MR	35.7	$8.287\cdot10^{-2}$	$9.97\cdot10^{-12}$	601.0	333.4	100%
Newton-CR	49.8	$8.306\cdot10^{-2}$	$9.17\cdot10^{-12}$	799.3	424.5	100%
Newton-CG	45.6	$8.632\cdot10^{-2}$	$9.71\cdot10^{-12}$	844.8	447.9	100%

数据集： STL10	Iter	Fun	$\|g_k\|$	预言机	时间（秒）	比例
L-BFGS-MR	150.1	$1.265\cdot10^{-1}$	$7.81\cdot10^{-11}$	370.0	101.4	100%
Armijo-L-BFGS	212.9	$1.345\cdot10^{-1}$	$3.14\cdot10^{-10}$	545.3	79.6	0%
Wolfe-L-BFGS	156.4	$1.341\cdot10^{-1}$	$3.12\cdot10^{-10}$	949.6	239.6	0%
Newton-MR	25.2	$1.247\cdot10^{-1}$	$1.38\cdot10^{-11}$	493.3	168.5	100%
Newton-CR	42.3	$1.292\cdot10^{-1}$	$1.10\cdot10^{-10}$	775.0	230.5	97%
Newton-CG	33.1	$1.254\cdot10^{-1}$	$7.14\cdot10^{-12}$	675.1	211.7	100%

表3 自编码器问题的数值结果。记录的数据是在算法终止时采集的，终止条件为“收敛性”（$\|g_k\|<10^{-10}$）或“步长停滞”（$\alpha_k<10^{-18}$）。具体而言，“迭代次数”、“函数”和“预言机调用”分别表示（外层）迭代总次数、函数值和预言机调用次数。一次预言机调用指的是在复杂度上等同于一次函数评估的操作。时间以秒为单位。结果为100次独立运行的平均值。最后，“比例”表示100次运行中算法以收敛性为终止条件的百分比。

集合 $\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$ 中的点： $y=\sin(x)\}$，是该问题的全局解。此外，[65, 附录 A], 中证明了 f 满足 PL 条件，因此 (C.3) 在 $\theta=\frac{1}{2}$ 下成立。我们选取 $n=200$，非精确参数 $\theta_k=\min\{\mathrm{B}\ \|g_k\|^\beta\}$，其中 $\mathrm{B}\in\{0.01,0.1\}$，$\beta\in\{1,\frac{1}{2},\frac{1}{4}\}$，并且 $\zeta_k=\zeta\theta_k$，其中 $\zeta\in\{1,10^{-1},10^{-2}\}$。结果如图1所示；对于 β 和 ζ 的不同取值，$\{\|g_k\|\}$ 的超线性收敛和近二次收敛清晰可见，与定理 3.16 中证明的保证一致。此外，较小的 θ_k 和 ζ_k 似乎会导致检测到 NPC 信息所需的迭代次数增加。

4.3 自编码器问题

我们考虑以下非凸深度自编码器问题

$$f(\mathbf{x}_\mathcal{E},\mathbf{x}_\mathcal{D})=\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{a}_i-\mathcal{D}(\mathcal{E}(\mathbf{a}_i;\mathbf{x}_\mathcal{E});\mathbf{x}_\mathcal{D})\|^2+\lambda\psi(\mathbf{x}_\mathcal{E},\mathbf{x}_\mathcal{D}),\tag{39}$$

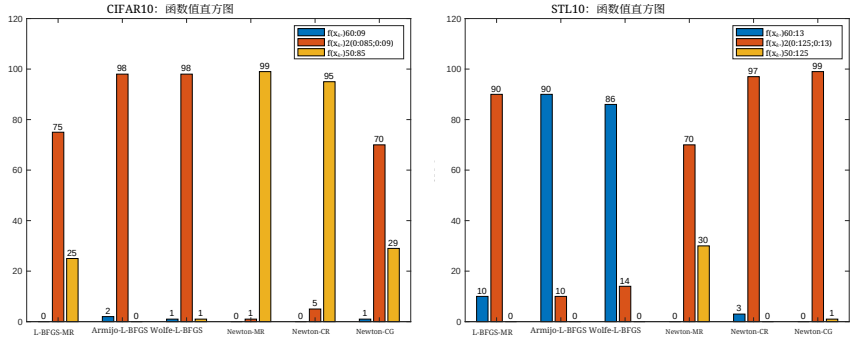


图 2 不同算法应用于自编码器问题时返回的最终函数值直方图。报告的函数值 $f(x_{k^*})$ 是 100 次独立运行的平均值。

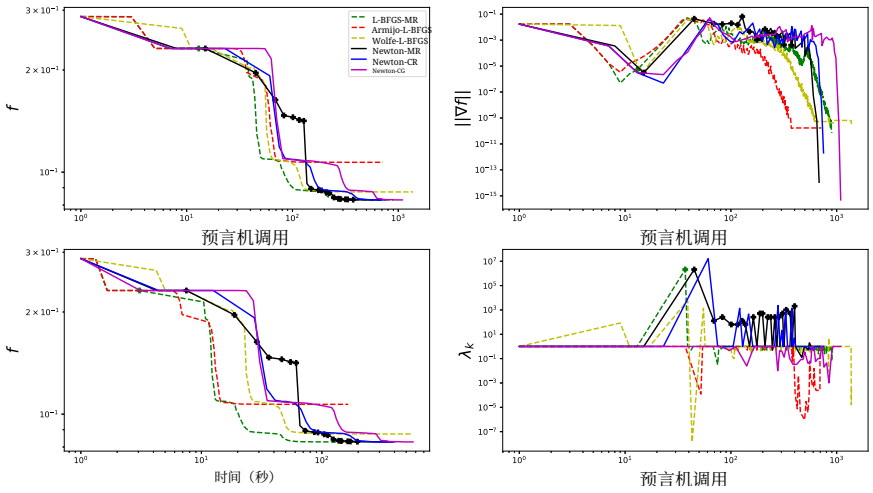


图 3 不同算法在深度自编码器问题（39）上使用 CIFAR10 数据集的性能。报告了函数值、梯度范数和步长相对于预言机调用次数及 CPU 时间的变化。标记表示检测到负曲率的迭代次数。

其中，$\mathcal{E}: \mathbb{R}^p \to \mathbb{R}^q$ 和 $\mathcal{D}: \mathbb{R}^\ell \to \mathbb{R}^p$ 分别是编码器和解码器映射。这里，$x \in \mathbb{R}^d$ 通过 $x = (\mathbf{x}_\mathcal{E}^\top, \mathbf{x}_\mathcal{D}^\top)^\top$进行分解；更多详情请参见 [29,53,75] 。我们采用与 [47]中相同的神经网络结构和一般设置。具体而言，正则化函数 ψ 选择为 $\psi(x) = \sum_{i=1}^d x_i^2/(1+x_i^2)$，并设置 $\lambda = 10^{-3}$。初始点取自均值为零、标准差为10^{-8}的正态分布。我们在数据集 CIFAR10 和 STL10[5]上进行实验；更多信息请参见表 2。

[5] STL10 数据集包含十个类别的彩色图像。我们将偶数类别重新标记为“0”，奇数类别重新标记为“1”。

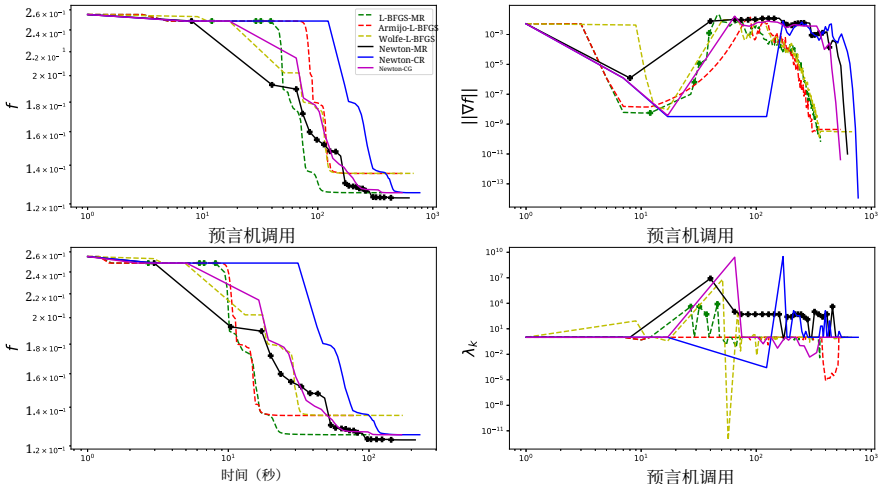
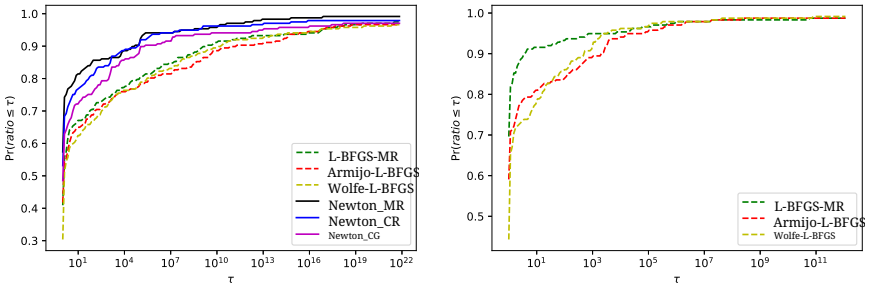


图 4 不同算法在深度自编码器问题（39）上的性能，采用 STL10 数据集。报告了函数值、梯度范数和步长随预言机调用次数及 CPU 时间的变化情况。标记表示检测到负曲率的迭代。

100 次独立运行的平均结果列于表 3。在三种截断牛顿法中，**Newton-MR** 在迭代次数、目标函数值、预言机调用次数和计算时间方面似乎表现最佳。令人意外的是，在拟牛顿方法中，只有 **L-BFGS-MR** 满足了收敛准则，而 **Armijo-L-BFGS**和 **Wolfe-L-BFGS** 都因步长过小而提前停止。与经典 L-BFGS 方法相比，**L-BFGS-MR**通过结合负曲率检测和截断，能够更好地应对非凸景观，并取得较低的函数值，尽管其计算成本略高于 **Armijo-L-BFGS**。图 2也对此作了说明，其中展示了不同方法所得最优函数值的直方图/分布。在图 3 和 4中，我们给出了更详细的收敛图，比较函数值、梯度值和步长相对于预言机调用次数及 CPU 时间的变化。这些图补充了表 3 和图 2 中的观察结果，并强调了在非凸问题中利用 NPC 信息的优势，无论是对牛顿方法还是拟牛顿方法而言。

4.4 CUTEst 测试问题

最后，我们报告了不同算法在 CUTEst 环境中的 237 个测试问题上的性能。我们使用性能剖面进行比较 [26,18]，并遵循 [47] 中的设置。我们注意到，通常并不清楚这些测试问题是否为凸问题。在本次测试中，对于 **L-BFGS-MR**，我们设置 $\theta_k = \min\{10^{-10}, \log(k)\sqrt{k}\|g_k\|\}$，以进一步提升其性能，并在较早的迭代中利用更多非正曲率信息。



(a) Newton型和 L-BFGS型 (b) L-BFGS型

图 5 以 $\{f(x_k)\}$ 衡量的性能剖面。对于 x 轴上给定的 τ， y 轴上的对应值表示：在所有运行中，给定求解器的性能处于最佳性能的 τ 倍范围内的比例。

我们在图 5 中总结了结果。在图 5(a) 中，在所有测试算法中，**Newton-MR** 表现最佳，略胜于另外两种截断牛顿法。图 5(b) 更详细地比较了三种 L-BFGS 类算法。 **L-BFGS-MR** 在函数值方面明显优于 **Armijo-L-BFGS** 和**Wolfe-L-BFGS** $\{f(x_k)\}$。这些结果进一步展现了非正曲率检测以及使用MINRES 作为内层线性系统求解器的优势。

5 结论与讨论

本文提出了一种基于 MINRES 的线搜索算法框架，用于求解无约束非凸优化问题。该方法采用著名的对称线性系统求解器 MINRES 来生成下降方向，并通过自适应机制控制正则化、截断和接受参数。当采用一般的海森矩阵近似$B_k \approx \nabla^2 f(x_k)$ 时，我们在 Kurdyka-Łojasiewicz 性质下给出了全面的收敛性保证。此外，当 B_k 为真实海森矩阵时，我们建立了更强的局部性质，包括避免严格鞍点以及超线性收敛到非孤立极小值。最后，在自编码器问题和 CUTEst问题集上的数值实验凸显了在非凸问题中利用非正曲率信息的优势。

由于 CR 与 MINRES 等价[41], ，我们用于避免严格鞍点的无特征对方法在将 MINRES 替换为 CR 时仍然适用。将我们的策略扩展到使用共轭梯度法的截断牛顿法，是未来研究的一个有趣方向。此外，与 [65], 不同，后者在海森矩阵仅为半正定且满足 PL 条件时，需要对共轭梯度法进行复杂分析；本工作采用自适应正则化来确保线性系统正定，从而在局部分析中避免了这类讨论。分析.....也将很有价值

.....的性能MINRES 步（不进行正则化） [65] ，并建立具有可比性的理论保证。

声明与陈述

经费资助和/或利益冲突/竞争性利益。Andre Milzarek 获得了深圳市科技计划项目（编号. RCYX20221008093033010）以及深圳市2023年稳定科技计划和深圳市多模态认知计算重点实验室的部分资助。

引理3.11的证明

证明。首先注意到，定理3.9 证明中的所有步骤均适用——从而确保$x_k \to x^*$。引理3.11 的推导是定理3.9证明的一个特例。利用 $x_k \to x^*$ 和(C.3)，存在 $\bar{k} \in \mathbb{N}$ 和 $M > 0$ 使得 $\|\bar{B}_k\| \le \|B_k\| + \mathsf{Z} \le M$ 并且

$$\varrho_k = \varrho(f(x_k) - f(x^*)) = c\Big[\frac{c(1-\theta)}{\varrho'(f(x_k) - f(x^*))}\Big]^{\frac{1-\theta}{\theta}} \le c(c(1-\theta))^{\frac{1-\theta}{\theta}} \|g_k\|^{\frac{1-\theta}{\theta}} \tag{40}$$

对所有 $k \ge \bar{k}$。这里，我们应用了 $\varrho(t) = ct^{1-\theta}$ 和 $\varrho(t) = c[c(1-\theta)/\varrho'(t)]^{(1-\theta)/\theta}$。当 $d_k = p_t$时，结合引理3.2 和3.8 (i)，可得

$$\begin{aligned}\frac{f(x_k) - f(x_{k+1})}{\|g_k\|} &\ge \frac{\sigma\lambda_k\|g_k\| \min\{\mathsf{A}, a_k\|g_k\|^\alpha\}}{\|\bar{B}_k\| + \|\bar{B}_k\|^2} \\ &\ge \sigma\|g_k\| \min\Big\{s, \frac{2\rho(1-\sigma)}{L_g}\min\{\mathsf{A}, a_k\|g_k\|^\alpha\}\Big\}\frac{\min\{\mathsf{A}, a_k\|g_k\|^\alpha\}}{M + M^2} \\ &\ge \bar{D}\|g_k\| \min\{\mathsf{A}^2, a_k^2\|g_k\|^{2\alpha}\},\end{aligned}$$

其中 $\bar{D} := \sigma$ 最小值$\{s, 2\rho(1-\sigma)/L_g\}/(M + M^2)$ 且 L_g 表示∇f 在 $\mathrm{conv}(\{x_k\})$上的 Lipschitz 常数。当 $d_k = \frac{\|g_k\|}{\|r_{t-1}\|}$ ~~r_{t-1}~~时，由（14），可得

$$[f(x_k) - f(x_{k+1})]/\|g_k\| \ge 0.5\sigma\lambda_k^2\zeta_k\|g_k\| \ge \bar{E}\|g_k\| \min\{\mathsf{Z}, z_k\|g_k\|^\zeta\}.$$

其中 $\bar{E} := \sigma\bar{\lambda}^2/2$ 且 $\bar{\lambda}$是在定理 3.9的证明中引入的。最后，若 $d_k = -g_k$，结合引理 3 5 24 $f\ x_k\ -f\ x_{k+1}\ /\|g_k\| \ge \sigma$ $\{s, 2\rho\ 1-$。(i) 和 ( )，可得 (() ())min():。$\sigma\ /L_g\}\|g_k\| = \bar{F}\|g_k\|$ 因此，我们可以推断

$$[f(x_k) - f(x_{k+1})]/\|g_k\| \ge \bar{G}\|g_k\| \min\{1, a_k^2\|g_k\|^{2\alpha}, z_k\|g_k\|^\zeta\}, \tag{41}$$

对所有 $k \ge \bar{k}$，其中 $\bar{G} :=$最小化$\{\bar{D}\mathsf{A}^2,\ \bar{E}$最小化$\{1, \mathsf{Z}\},\ \bar{F}\}$。接下来，我们介绍 $\mathcal{I}_1 = \{k \ge \bar{k} :\ 1 <$最小化 $\{a_k^2\|g_k\|^{2\alpha}, z_k\|g_k\|^\zeta\}\}$ $a_k^2\|g_k\|^{2\alpha} <$ $\{1, z_k\|g_k\|^\zeta\}\}$ $k \notin$， $\mathcal{I}_2 = \{k \ge \bar{k} :$ 最小 $\mathcal{I}_1 \cup \mathcal{I}_2\}$ $\Gamma_k = \sum_{i=k}^\infty \|g_i\|$ $\Gamma_{k\,\mathcal{I}_j} = \sum_{i=k\ i\in\mathcal{I}_j}^\infty \|g_i\|$ $j = 1\ 2\ 3$ 29 30 化， $\mathcal{I}_3 = \{k \ge \bar{k} :$ 、， 以及、. . ，、，。仿照（ ）–（ ）并利用（41）和反向Hölder不等式，我们有

$$\begin{aligned}\varrho_k/\bar{G} &\ge \sum_{i=k}^\infty (f(x_i) - f(x_{i+1}))/\|g_i\| \\ &\ge \sum\nolimits_{i\in\mathcal{I}_1}\|g_i\| + \sum\nolimits_{i\in\mathcal{I}_2} a_i^2\|g_i\|^{1+2\alpha} + \sum\nolimits_{i\in\mathcal{I}_3} z_i\|g_i\|^{1+\zeta} \\ &\ge \Gamma_{k,\mathcal{I}_1} + \Big(\sum\nolimits_{i\ge k, i\in\mathcal{I}_2} a_i^{-2/(2\alpha)}\Big)^{-2\alpha}\Gamma_{k,\mathcal{I}_2}^{1+2\alpha} + \Big(\sum\nolimits_{i\ge k, i\in\mathcal{I}_3} z_i^{-1/\zeta}\Big)^{-\zeta}\Gamma_{k,\mathcal{I}_3}^{1+\zeta} \\ &= \Gamma_{k,\mathcal{I}_1} + \sigma_{1,k}^{-2\alpha}\Gamma_{k,\mathcal{I}_2}^{1+2\alpha} + \sigma_{2,k}^{-\zeta}\Gamma_{k,\mathcal{I}_3}^{1+\zeta}.\end{aligned}$$

这里，$\sigma_{1,k}$ 和 $\sigma_{2,k}$ 定义于（27）中。此外，由于 $\varrho_k \to 0$ 当 $k \to \infty$时，存在某个 $k' \geq \bar{k}$ 和某个常数 $\bar{G}' > 0$ 使得对所有 $k \geq k'$，都有

$$\Gamma_k \leq \frac{\varrho_k}{\bar{G}} + \left(\frac{\varrho_k \sigma_{1,k}^{2\alpha}}{\bar{G}}\right)^{\frac{1}{1+2\alpha}} + \left(\frac{\varrho_k \sigma_{2,k}^{\zeta}}{\bar{G}}\right)^{\frac{1}{1+\zeta}} \leq \left(\frac{\varrho_k}{\bar{G}'}\right)^{\frac{1}{1+\max\{2\alpha,\zeta\}}}. \tag{42}$$

设定 $\omega = \frac{(1+\max\{2\alpha,\zeta\}\ \theta)}{1-\theta}$，$C_\theta = [(\bar{G}'/c)^{\theta/(1-\theta)}/(c(1-\theta))]^{-\frac{1}{\omega}}$ 并结合（40）、（42），可得

$$\Gamma_k^{-\omega}(\Gamma_k - \Gamma_{k+1}) \geq C_\theta^{-\omega}.$$

该递推使我们能够推导出 $\|g_k\|$ 的所述收敛速率。我们省略详细计算，参见 [59, 定理 5.5，第 (ii) 部分]。（[59] 处理 $\max\{2\alpha,\zeta\} = 0$ 的特殊情形。） □

B 引理 3.17 的证明

证明。根据假设，存在某个序列 $\{w_k\}$，使得 $u_k \leq w_k$ 和 $w_k \to 0$ q-线性。因此，存在某个 $k_1 \in \mathbb{N}$ 和 $\mu \in (0, 1)$，使得 $w_{k+1}/w_k \leq \mu$ 对于 $k \geq k_1$ 成立。然后，由于 $v_{k+1}/v_k \to 1$，对于任意 $q \in (\mu, 1)$，存在 $k_2 \in \mathbb{N}$，使得 $v_{k+1}/v_k \leq q/\mu$ 对于 $k \geq k_2$ 成立。因此，对于所有 $k \geq \max\{k_1, k_2\}$，我们有 $\frac{w_{k+1}v_{k+1}}{w_k v_k} \leq q < 1$，从而可以推出 $\lim_{k\to\infty} w_k v_k = 0$。最后，~~根据 $u_k \leq$~~ w_k，可得 $\lim_{k\to\infty} u_k v_k = 0$。 □

参考值

1. Absil, P.A., Mahony, R., Andrews, B.：解析代价函数下降方法的迭代点收敛性。《SIAM 优化杂志》**16**(2), 531–547 (2005) 2. Anandkumar, A., Ge, R.：非凸优化中逃离高阶鞍点的有效方法。载于：学习理论会议，页码 81–102。机器学习研究论文集 (2016) 3. Attouch, H., Bolte, J.：关于涉及解析特征的非光滑函数的近端算法收敛性。《数学规划》**116**(1-2, Ser. B), 5–16 (2009) 4. Attouch, H., Bolte, J., Svaiter, B.F.：半代数问题与温和问题中下降方法的收敛性：近端算法、前向-后向分裂及正则化高斯-赛德尔方法。《数学规划》**137**(1-2, Ser. A), 91–129 (2013) 5. Bai, S., Li, M., Lu, C., Zhu, D., Deng, S.：弱凸函数与近似凸函数的三类误差界等价性。《优化理论与应用杂志》**194**(1), 220– 245 (2022) 6. Bolte, J., Daniilidis, A., Lewis, A.：非光滑次解析函数的Łojasiewicz不等式及其在次梯度动力系统中的应用。《SIAM 优化杂志》**17**(4), 1205–1223 (2006) 7. Bolte, J., Sabach, S., Teboulle, M.：用于非凸非光滑问题的近端交替线性化最小化。《数学规划》**146**(1-2, Ser. A), 459–494 (2014) 8. Bottou, L., Curtis, F.E., Nocedal, J.：大规模机器学习的优化方法。《SIAM 综述》**60**(2), 223–311 (2018) 9. Byrd, R.H., Nocedal, J., Schnabel, R.B.：拟牛顿矩阵的表示及其在有限内存方法中的应用。《数学规划》**63**(1), 129–156 (1994) 10. Cartis, C., Gould, N.I., Toint, P.L.：无约束优化的自适应三次正则化方法。第一部分：动机、收敛性与数值结果。《数学规划》**127**(2), 245–295 (2011) 11. Cartis, C., Gould, N.I., Toint, P.L.：无约束优化的自适应三次正则化方法。第二部分：最坏情况下的函数与导数评估复杂度。《数学规划》**130**(2), 295–319 (2011) 12. Coates, A., Ng, A., Lee, H.：无监督特征学习中单层网络的分析。载于：第14届国际人工智能与统计学会议论文集，页码 215–223 (2011)

13. Conn, A.R., Gould, N.I., Toint, P.L.：信赖域方法。SIAM (2000) 14. Curtis, F.E., Robinson, D.P.：在确定性与随机优化中利用负曲率。《数学规划》**176**, 69–94 (2019) 15. Curtis, F.E., Robinson, D.P., Royer, C.W., Wright, S.J.：具有强二阶复杂度保证的非凸优化信赖域牛顿共轭梯度法。《SIAM 优化杂志》**31**(1), 518–544 (2021) 16. Dahito, M.A., Orban, D.：线搜索与信赖域方法中的共轭残量法。《SIAM 优化杂志》**29**(3), 1988–2025 (2019) 17. Dembo, R.S., Steihaug, T.：大规模无约束优化的截断牛顿算法。《数学规划》**26**(2), 190–212 (1983) 18. Dolan, E.D., Moré, J.J.：使用性能剖面对优化软件进行基准测试。《数学规划》**91**, 201–213 (2002) 19. Draper, N.R., Smith, H.：应用回归分析，第326卷。约翰·威立父子出版公司 (1998) 20. Drusvyatskiy, D., Ioffe, A.D., Lewis, A.S.：利用类泰勒模型进行非光滑优化：误差界、收敛性与终止准则。《数学规划》**185**(1-2), 357–383 (2021) 21. Du, S.S., Jin, C., Lee, J.D., Jordan, M.I., Singh, A., Poczos, B.：梯度下降可能需要指数级时间才能逃离鞍点。《神经信息处理系统进展》**30** (2017) 22. Fasano, G., Lampariello, F., Sciandrone, M.：用于大规模非线性最小二乘问题的截断非单调高斯–牛顿法。《计算优化与应用》**34**, 343 358 2006– ( ) 23. Ge, R., Huang, F., Jin, C., Yuan, Y.：逃离鞍点——用于张量分解的在线随机梯度。载于：第28届学习理论会议论文集，页码 797–842 (2015) 24. Goldfarb, D.：利用负曲率方向进行最小化的曲线搜索步长算法。《数学规划》**18**(1), 31–40 (1980) 25. Goldfarb, D., Mu, C., Wright, J., Zhou, C.：在求解非线性规划时利用负曲率。《计算优化与应用》**68**, 479–502 (2017) 26. Gould, N., Scott, J.：关于用于软件基准测试的性能剖面的说明。《ACM 数学软件汇刊》**43**(2), 1–5 (2016) 27. Gould, N.I., Lucidi, S., Roma, M., Toint, P.L.：在无约束优化的线搜索方法中利用负曲率方向。《优化方法与软件》**14**(1-2), 75–98 (2000) 28. Hestenes, M.R., Stiefel, E.：求解问题的共轭梯度法。《美国国家标准局研究杂志》**49**(6), 409 (1952) 29. Hinton, G.E., Salakhutdinov, R.R.：利用神经网络降低数据维度。《科学》**313**(5786), 504–507 (2006) 30. Jia, X., Kanzow, C., Mehlitz, P.：在不具备全局利普希茨假设的情况下，满足Kurdyka–Łojasiewicz性质时近端梯度法的收敛分析。《SIAM 优化杂志》**33**(4), 3038–3056 (2023) 31. Jin, C., Ge, R., Netrapalli, P., Kakade, S.M., Jordan, M.I.：如何高效逃离鞍点。载于：国际机器学习会议，页码 1724–1732。机器学习研究论文集 (2017) 32. Jin, C., Netrapalli, P., Ge, R., Kakade, S.M., Jordan, M.I.：机器学习中的非凸优化：梯度、随机性与鞍点。《ACM 期刊》**68**(2), 1–29 (2021) 33. Jin, C., Netrapalli, P., Jordan, M.I.：加速梯度下降比梯度下降更快地逃离鞍点。载于：学习理论会议，页码 1042–1085。机器学习研究论文集 (2018) 34. Kanzow, C., Steck, D.：面向大规模非凸最小化的有限内存拟牛顿法正则化。《数学规划计算》**15**(3), 417–444 (2023) 35. Krizhevsky, A.：从微小图像中学习多层特征。硕士论文，特伦特大学 (2009) 36. Kurdyka, K.：关于o-极小结构中可定义函数的梯度。载于：《傅里叶研究所年刊》，第48卷，页码 769–783 (1998)

37. Lee, J.D., Panageas, I., Piliouras, G., Simchowitz, M., Jordan, M.I., Recht, B.：一阶方法几乎总能避开严格鞍点。《数学规划》**176**，311–337（2019年） 38. Li, D.H., Fukushima, M.：修正 BFGS 方法及其在非凸最小化中的全局收敛性。《计算与应用数学杂志》**129**(1-2)，15–35（2001年） 39. Li, D.H., Fukushima, M.：关于 BFGS 方法在非凸无约束优化问题中的全局收敛性。SIAM 优化杂志 **11**(4)，1054–1064（2001年） 40. Li, Y.J., Li, D.H.：用于凸最小化的截断正则化牛顿法。《计算优化与应用》**43**(1)，119–131（2009） 41. Lim, A., Liu, Y., Roosta, F.：不相容系统下的共轭方向法。arXiv 预印本 arXiv:2401.11714（2024年） 42. Lim, A., Roosta, F.：不精确海森信息下非凸 Newton-MR 的复杂度保证。《IMA 数值分析杂志》页码 1–28（2025） 43. Liu, C., Zhu, L., Belkin, M.：过参数化非线性系统和神经网络中的损失景观与优化。《应用计算调和分析》**59**，85–116（2022年） 44. Liu, D.C., Nocedal, J.：关于大规模优化的有限内存 BFGS 方法。《数学规划》**45**(1)，503–528（1989） 45. Liu, Y., Roosta, F.：不精确海森信息下 Newton-MR 的收敛性。SIAM 优化杂志 **31**(1)，59–90（2021年） 46. Liu, Y., Roosta, F.：MINRES：从负曲率检测到单调性性质。SIAM 优化杂志 **32**(4)，2636–2661（2022年） 47. Liu, Y., Roosta, F.：具有复杂度保证的非凸光滑无约束优化 Newton-MR 算法。arXiv 预印本 arXiv:2208.07095（2022年） 48. Loh, P.L., Wainwright, M.J.：具有非凸性的正则化 M 估计量：局部最优解的统计与算法理论。神经信息处理系统进展 **26** （2013） 49. Łojasiewicz, S.：实解析子集的一个拓扑性质。偏微分方程 **117**，87–89（1963） 50. Łojasiewicz, S.：关于半解析与次解析几何。《傅里叶研究所年刊》（格勒诺布尔）**43**，1575–1595（1993） 51. Luo, Z.Q., Tseng, P.：可行下降法的误差界与收敛分析：一种一般方法。《运筹学年刊》**46**(1)，157–178（1993） 52. Marquardt, D.W.：非线性参数最小二乘估计的一种算法。工业与应用数学学会杂志 **11**(2)，431–441（1963） 53. Martens, J.：通过无海森优化进行深度学习。载于：第 27 届国际机器学习会议论文集，第 27 卷，第 735–742 页（2010） 54. Mishchenko, K.：具有全局收敛性的正则化牛顿法。SIAM 优化杂志 **33**(3)，1440–1462（2023） 55. Moré, J.J., Sorensen, D.C.：关于在修正牛顿法中使用负曲率方向。《数学规划》**16**(1)，1–20（1979） 56. Moré, J.J., Thuente, D.J.：具有保证充分下降的线搜索算法。《ACM 数学软件汇刊》**20**(3)，286–307（1994） 57. Nesterov, Y., Polyak, B.T.：牛顿法的三次正则化及其全局性能。《数学规划》**108**(1)，177–205（2006） 58. Nocedal, J., Wright, S.J.：数值优化，第二版。运筹学与金融工程 Springer 丛书。Springer，纽约（2006） 59. Ouyang, W., Milzarek, A.：一种基于信赖域型正规映射的半光滑牛顿法，用于非光滑非凸复合优化。arXiv 预印本 arXiv:2106.09340（2021年） 60. ONeill, M., Wright, S.J.：加速梯度法在非凸函数临界点附近的行为。《数学规划》 **176**，403–427（2019年） 61. Paige, C.C.：超大型稀疏矩阵特征值和特征向量的计算。博士学位论文，伦敦大学（1971） 62. Paige, C.C., Saunders, M.A.：稀疏不定线性方程系统的求解。SIAM 数值分析杂志 **12**(4)，617–629（1975） 63. Panageas, I., Piliouras, G., Wang, X.：一阶方法几乎总能避开鞍点：趋零步长情形。神经信息处理系统进展 **32**（2019）

64. Paternain, S., Mokhtari, A., Ribeiro, A.：一种用于非凸优化且能快速避开鞍点的基于牛顿的方法。SIAM 优化杂志 **29**(1)，343–368（2019年） 65. Rebjock, Q., Boumal, N.：通过分析不定矩阵上的共轭梯度法，实现非孤立极小值下信赖域的快速收敛。《数学规划》页码 1–42（2024年） 66. Rebjock, Q., Boumal, N.：向非孤立极小值的快速收敛：二阶连续可微函数的四个等价条件。《数学规划》页码 1–49（2024年） 67. Roosta, F., Liu, Y., Xu, P., Mahoney, M.W.：Newton-MR：带有最小残差子问题求解器的非精确牛顿法。EURO 计算优化杂志 **10**，100035（2022年） 68. Royer, C.W., ONeill, M., Wright, S.J.：具有复杂度保证的光滑无约束优化牛顿-共轭梯度算法。《数学规划》**180**(1)，451–488（2020年） 69. Saad, Y.：稀疏线性系统的迭代方法，第 82 卷。SIAM（2003年） 70. Saad, Y.：大型特征值问题的数值方法：修订版。SIAM（2011年） 71. Schaeffer, H., McCalla, S.G.：扩展梯度下降的步长限制以避开严格鞍点。SIAM 数学与数据科学杂志 **2**(4)，1181–1197（2020年） 72. Steihaug, T.：大规模非线性问题的拟牛顿法。耶鲁大学（1981年） 73. Stiefel, E.：求解线性方程组的最佳策略松弛法。《赫尔维数学评论》**29**(1)，157–179（1955年） 74. Vlatakis-Gkaragkounis, E.V., Flokas, L., Piliouras, G.：利用零阶方法高效避开鞍点：无需梯度。神经信息处理系统进展 **32**（2019） 75. Xu, P., Roosta, F., Mahoney, M.W.：非凸机器学习的二阶优化：一项实证研究。载于：2020 年 SIAM 国际数据挖掘会议论文集，第 199–207 页。SIAM（2020年） 76. Yue, M.C., Zhou, Z., Man-Cho So, A.：局部误差界条件下三次正则化方法的二次收敛性。SIAM 优化杂志 **29**(1)，904–932（2019年） 77. Zhou, Y., Wang, Z., Liang, Y.：KL 性质下非凸优化三次正则化的收敛性。神经信息处理系统进展 **31**（2018年）