arXiv:2305.01669v2 [math.NA] 5 May 2023

无矩阵牛顿法

UWE NAUMANN*

摘要。 本文提出了一种用于求解 n 非线性方程组的牛顿法改进方法。该新型矩阵自由方法依赖于根据微分链式法则，将残差的可逆雅可比矩阵分解为给定的可逆稀疏局部雅可比矩阵。该方法的提出背景是带宽为 $2m+1$ 、适用于 $m \ll n$ 的局部雅可比矩阵。可以观察到计算成本降低了 $\mathcal{O}(\frac{n}{m})$。本文还给出了针对三对角情形的运行时间测量结果，表明计算成本降低了 $\mathcal{O}(n)$。

泛化产生了组合 Matrix-Free Newton Step 问题。我们证明了其 NP 完全性，并提出了用于构建近似解方法的算法组件。受伴随算法微分的启发，这种新方法也面临着后者所面临的若干挑战，包括 DAG Reversal 问题。此外，稀疏线性代数中的组合问题也带来了进一步的挑战，例如 Bandwidth 或 Directed Elimination Ordering。

关键词。 牛顿法、矩阵自由、算法微分、伴随量

MSC 分类号。 49M15、47A05、68N99

1. 引言。 我们重新研究牛顿法 [18, 25] ，用于计算 $x \in \mathbb{R}^n$ 具有可逆稠密雅可比矩阵的可微多元向量函数 F 的根 $\mathbb{R}^n \to \mathbb{R}^n$ ： $y = F(x)$

$$F'(x) \equiv \left(\frac{dy_j}{dx_i}\right)_{i,j=1,\ldots,n} \in \mathbb{R}^{n\times n} ,$$

并且其中 $x = (x_i)$ 和 $y = (y_i)$ 对于 $i = 1, \ldots, n.$ 满足 $F(x) \approx 0$ 的近似解按如下方式迭代计算:

$$\begin{aligned} \Delta x &:= -F'(x)^{-1} \cdot F(x) \\ x &:= x + \Delta x \end{aligned} \tag{1.1}$$

对于给定的起始点 $x \in \mathbb{R}^n.$ ，通常将经过 $p \geq 0$ 次迭代后的收敛定义为残差的范数 $F(x)$ 降至某个给定阈值 $0 < \epsilon \ll 1.$ 以下。例如，关于牛顿法的进一步信息，请参见 [3, 13] 。

我们使用 $=$表示数学等式，使用 $\equiv$ 表示“定义为”，并使用 $:=$ 按照命令式编程表示赋值。近似等式表示为 $\approx$。我们区分偏导数（∂）和全导数（d）。乘法用点号表示。为使记号更加紧凑，点号也可以省略。

残差 F 通常由使用某种高级编程语言编写的可微程序给出。方程 (1.1) 表明，牛顿法的两个主要组成部分是计算雅可比矩阵 $F'(x)$，以及求解线性牛顿方程组 $F'(x) \cdot \Delta x = -F(x)$。对$F$进行算法微分（AD） [10] ，可在机器精度下得到 $F'(x)$，其计算成本为 $\mathcal{O}(n) \cdot \text{Cost}(F)$。随后直接求解线性牛顿方程组，得到牛顿步 $\Delta x \in \mathbb{R}^n$，其计算成本为 $\mathcal{O}(n^3)$。总体计算成本可能由这两部分中的任一部分主导，具体取决于比值 $\text{Cost}(F)\,\mathcal{O}(n^2)$ ·

利用 F 的特殊结构可以显著降低计算成本。因此，本文提出了一种无矩阵牛顿法

*计算工程软件与工具，亚琛工业大学，德国亚琛，naumann@stce.rwth-aachen.de

其动机源于在相关场景中获得的节省。我们介绍该方法背后的基本思想，并进一步讨论其泛化。在第2节中，我们回顾自动微分的基本原理，并推导出关于牛顿步计算的结论。第3节通过实践中具有重要意义的带状局部雅可比矩阵这一特殊情形，说明了降低计算成本的潜力。第4节讨论进一步的泛化与形式化。第5节给出结论。

2. 伴随自动微分的启示。一阶自动微分有两种基本 l
形式。切向自动微分可得

$$\dot{y} = \dot{F}(x, \dot{x}) \equiv F'(x) \cdot \dot{x} \tag{2.1}$$

因此，在 $\mathcal{O}(n)$处以机器精度得到（稠密）雅可比矩阵， $\cdot$ Cost(F) 的计算方法是令$\dot{x}$ 遍历 $\mathbb{R}^n$. 的笛卡尔基。伴随自动微分计算

$$\bar{x} = \bar{F}(x, \bar{y}) \equiv F'(x)^T \cdot \bar{y} \ , \tag{2.2}$$

从而在 $\mathcal{O}(n)$处以机器精度得到相同的雅可比矩阵， $\cdot$ Cost(F) 的计算方法是令 $\bar{y}$ 遍历 $\mathbb{R}^n$. 的笛卡尔基。切向自动微分和伴随自动微分都是无矩阵方法，因为在计算方程 (2.1) 或 (2.2) 时，无须显式构造雅可比矩阵。两种方法都可用于累积 F'. 它们的计算成本之比为

$$\mathcal{R} \equiv \frac{\mathrm{Cost}(\bar{F})}{\mathrm{Cost}(\dot{F})} \ ,$$

其中，通常 $\mathcal{R} \geq 1$. 因此，在牛顿法中计算残差的雅可比矩阵时，切向自动微分通常是首选方法。残差雅可比矩阵的稀疏性可能会改变这一情况 [6]。

二阶切向量和伴随自动微分自然随之而来；参见 [10]。凸优化目标函数所需的牛顿法海森矩阵 $f : \mathbb{R}^n \to \mathbb{R}$ 可通过二阶伴随版本 f 在 $\mathcal{O}(n) \cdot$ Cost(f) 处计算，
. 其
f. 。相应的矩阵自由牛顿-克雷洛夫方法 [14]也可以推导出来。这些方法基于这样一个观察：克雷洛夫子空间方法（例如共轭梯度法 [11]）所需的海森矩阵-向量积，可以通过二阶伴随量f 计算，而无需预先累积海森矩阵。本文提出的方法则有所不同。它利用 F 的特殊结构和局部稀疏性，而非全局传播切向量或伴随量。

2.1. 术语。 自动微分要求给定程序 F 可微。本文仅对 [10]中的记法作了少许修改。

Definition 2.1（可微程序）。一个 可微程序

$$F : \mathbb{R}^n \to \mathbb{R}^m : \ y := F(x)$$

分解为一段单赋值代码

$$v_j := \varphi(v_i)_{i \in P_j} \quad for \ j = 1, \ldots, p + m \ , \tag{2.3}$$

其中 $v_{i-n} = x_i$ 对于 $i = 1$、.....、 n， 以及 $y_k = v_{p+k}$ 对于 $k = 1$、.....、 m，且具有可微的 基本函数 φ_j、 $j = 1$、.....、 $p+m$，并具有元素偏导数

$$\partial_{j,i} \equiv \frac{\partial \varphi_j}{\partial v_i} \quad for \ i \in P_j \ .$$

φ_j （ v_j的直接前驱）的参数集合记为 P_j.

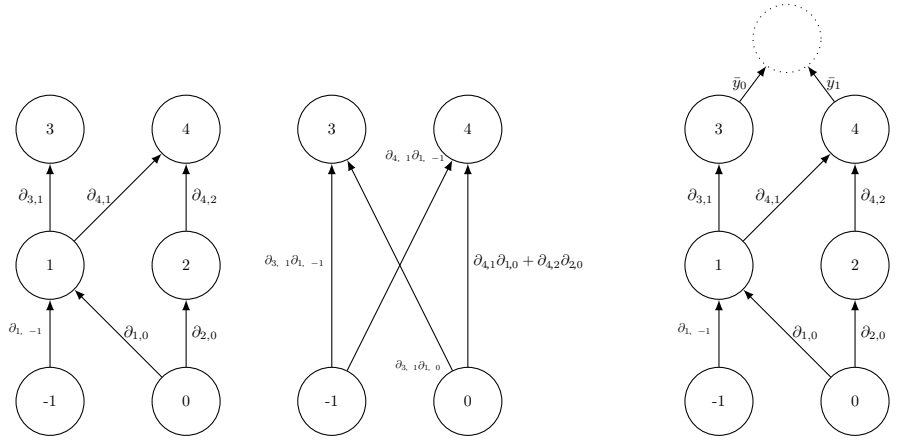


FIG. 1。 *Sample DAGs*

DEFINITION 2.2（有向无环图）。方程 *(2.3)* 导出一个有向无环图（DAG） $G=(V,E)$ $V=(X,Z,Y)$ $X\cap Z\cap Y=\varnothing$,[1] $E\subseteq V\times V$，使得 $X=\{1-n,\ldots,0\}$, $Z=\{1,\ldots,p\}$, $Y=p+1,\ldots,p+m$ 以及$(i,j)\in E\Leftrightarrow i\in P_j$.所有边$(i,j)$ 均标记为 $\partial_{j,i}$.

连接顶点 i和 j的路径记作（i，.....， j）。

DEFINITION 2.3（分层有向无环图和局部雅可比矩阵）。矩阵链乘积的有向无环图

$$F'_q\cdot\ldots\cdot F'_1=F'\in\mathbb{R}^{m\times n}$$

由局部雅可比矩阵 F'_i、 $i=1$、.....、 q组成的分层有向无环图。

图 1(a) 中的有向无环图是分层的。

DEFINITION 2.4（均匀分层有向无环图）。矩阵链乘积的分层有向无环图

$$F'_q\cdot\ldots\cdot F'_1=F'\in\mathbb{R}^{n\times n}$$

称为均匀分层有向无环图，如果 $F'_i\in\mathbb{R}^{n\times n}$ 对于 $i=1,\ \ldots,\ q$.

图 1(a) 中的有向无环图是均匀分层的。

DEFINITION 2.5（可逆有向无环图）。我们将雅可比矩阵可逆的可微程序的有向无环图 $F'\in\mathbb{R}^{n\times n}$ 称为可逆有向无环图。相应的牛顿步定义为$(F')^{-1}\cdot y$，给定 $y\in\mathbb{R}^n$.

伴随自动微分可以通过在合适的编程语言（如 C++ [9, 12, 21]）中使用运算符和函数重载来实现。将给定的 F 作为可微程序的实现置于重载算术中运行，从而在计算函数值的同时，记录相应的有向无环图（也称为磁带）。例如，对于q=2 ，可微C++程序

对于(整数i =0; i <q ; i++) {

[1]最简单的“空操作”程序 x=x 将x 区分为输入（$\in X$）和输出（$\in Y$），而 $Z=\varnothing$ 在此情形下。相应的更大规模场景也自然地随之而来。

```
x[i%2]+=x[(i+1)%2]; x[(i+1)%2]=sin(x[(i+1)%2]);}
```

得到图 1(a) 中的有向无环图，其中，$\partial_{1,-1}$ =x[1]，$\partial_{3,1}$ =cos(x[0])，等等。雅可比矩阵

$$F' = \begin{pmatrix} \partial_{3,1}\partial_{1,-1} & \partial_{3,1}\partial_{1,0} \\ \partial_{4,1}\partial_{1,-1} & \partial_{4,1}\partial_{1,0} + \partial_{4,2}\partial_{2,0} \end{pmatrix} = \begin{pmatrix} \partial_{3,1} & 0 \\ \partial_{4,1} & \partial_{4,2} \end{pmatrix} \begin{pmatrix} \partial_{1,-1} & \partial_{1,0} \\ 0 & \partial_{2,0} \end{pmatrix}$$

可表示为图 1(b) 中的二部有向无环图。计算伴随量的两种方案都可行。

$$\left(\begin{pmatrix} \partial_{3,1} & 0 \\ \partial_{4,1} & \partial_{4,2} \end{pmatrix} \begin{pmatrix} \partial_{1,-1} & \partial_{1,0} \\ 0 & \partial_{2,0} \end{pmatrix} \right)^T \begin{pmatrix} \bar{y}_0 \\ \bar{y}_1 \end{pmatrix} = \begin{pmatrix} \partial_{1,-1} & 0 \\ \partial_{1,0} & \partial_{2,0} \end{pmatrix} \left(\begin{pmatrix} \partial_{3,1} & \partial_{4,1} \\ 0 & \partial_{4,2} \end{pmatrix} \begin{pmatrix} \bar{y}_0 \\ \bar{y}_1 \end{pmatrix} \right)$$

（其中 $\bar{y} = (0\ 1)^T$ 和 $\bar{y} = (1\ 0)^T$ 分别给出雅可比矩阵的各行）得到相同的结果。注意，计算左侧表达式需要 12融合乘加[2] ($\mathcal{O}(qn^3)$); 而右侧表达式只需要 8融合乘加 ($\mathcal{O}(qn^2)$)。微分链式法则的结合律（等价地，矩阵乘法的结合律）[3] 所产生的相同效果，也被反向传播应用于人工神经网络的训练过程中 [7]。伴随量可以表示为图 1(c) 中的有向无环图。

关于自动微分（AD）的诸多方面及其实现的讨论超出了本文的讨论范围。感兴趣的读者可参阅 [10]，以获取有关该主题的更多信息。此外，自动微分社区的网络门户 www.autodiff.org 提供了全面的参考文献目录，以及指向各种自动微分研究和工具开发项目的链接。

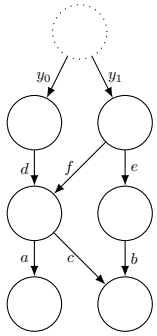


FIG. 2. *Inverse DAG*

2.2. 经验教训。牛顿法如何受益于伴随自动微分中的经验？为回答这一问题，假设给定的 F 实现作为一个可微程序，由基本函数 $F_i : \mathbb{R}^n \to \mathbb{R}^n : x^i = F_i(x^{i-1})$ 组成，其中$i = 1, \ldots, q$ as

$$y = x^q = F(x^0) = (F_q \circ \ldots \circ F_1)(x^0) \tag{2.4}$$

其中 $x^0 = x$. 对方程 (2.4) 应用微分的链式法则可得

$$F'(x^0) = F'_q(x^{q-1}) \cdot \ldots \cdot F'_1(x^0) \tag{2.5}$$

意味着

$$\begin{aligned} \Delta x &= -F'(x)^{-1} \cdot y \\ &= -(F'_q(x^{q-1}) \cdot \ldots \cdot F'_1(x^0))^{-1} \cdot y \\ &= -F'_1(x^0)^{-1} \cdot \ldots \cdot F'_q(x^{q-1})^{-1} \cdot y\ . \end{aligned} \tag{2.6}$$

[2]融合乘加浮点运算[3]假设采用无限精度算术

$F'(x)$需要是可逆的。所有 F_i'的可逆性立即可得。其（不失一般性）LU 因式分解 [4] 得到

$$F_i \equiv F_i(x^{i-1}) = L_i(x^{i-1}) \cdot U_i(x^{i-1}) \equiv L_i \cdot U_i \tag{2.7}$$

具有下三角 L_i 和上三角 U_i. 原位分解的计算成本为 $L_i - I_n + U_i$ $\mathcal{O}$（n^3）。 矩阵 $I_n \in \mathbb{R}^{n\times n}$ 表示 $\mathbb{R}^n$. 中的单位矩阵。由方程（2.6）可得

$$\begin{aligned} \Delta x &= -U_1^{-1} \cdot L_1^{-1} \cdot \ldots \cdot U_q^{-1} \cdot L_q^{-1} \cdot y \\ &= -U_1^{-1} \cdot \left(L_1^{-1} \cdot \ldots \cdot \left(U_q^{-1} \cdot \left(L_q^{-1} \cdot y\right)\right)\ldots\right) \end{aligned} \tag{2.8}$$

从而得到 $2q$ 个线性系统，需要尽可能高效地求解，以降低相较于标准牛顿法的计算成本。新方法之所以是矩阵自由的，是因为 F' 不会被显式累积。图2展示了与伴随自动微分的类比，其中反向的边表示向量与局部雅可比矩阵逆的乘积。对于图1中的例子，标准的“先累积，再分解”方法得到的牛顿步为

$$\underbrace{\underbrace{\underbrace{\left(\begin{pmatrix}\partial_{3,1} & 0\\ \partial_{4,1} & \partial_{4,2}\end{pmatrix}\begin{pmatrix}\partial_{1,-1} & \partial_{1,0}\\ 0 & \partial_{2,0}\end{pmatrix}\right)^{-1}}_{\mathcal{O}(qn^3)\ \hat{=}\ 5}\begin{pmatrix}y_0\\ y_1\end{pmatrix}}_{\mathcal{O}(n^3)\ \hat{=}\ 3+3+1=7}}_{\Sigma=12}$$

代价是 12融合乘加。 另一种“先分解，再累积”方法

$$\underbrace{\begin{pmatrix}\partial_{1,-1} & \partial_{1,0}\\ 0 & \partial_{2,0}\end{pmatrix}^{-1}\underbrace{\left(\begin{pmatrix}\partial_{3,1} & 0\\ \partial_{4,1} & \partial_{4,2}\end{pmatrix}^{-1}\begin{pmatrix}y_0\\ y_1\end{pmatrix}\right)}_{O(n^2)\ \hat{=}\ 3}}_{O(qn^2)\ \hat{=}\ 6}$$

完成相同的任务只需 6融合乘加。 诚然，这个例子较为极端，因为所有元素雅可比矩阵本身已经是三角矩阵。因此，与向量相乘时，其逆矩阵的作用可以通过简单的前代或回代非常高效地计算出来。更实际的情况下，需要先将各个因子转换为三角形式。理想情况下，利用元素雅可比矩阵的稀疏性，可以将相应的额外工作量控制在较低水平。然而，稀疏线性代数直接法中通常面临的挑战 [4] 仍需解决。下一节将讨论一个具有代表性的特例。该例展示了新方法的潜力，并为第4节中进一步推广该方法提供了动机，以使其能够适用于更广泛的实际问题。

3. 带状元素雅可比矩阵。设所有 F_i' 的（最大）带宽为$2m+1$ （m 条非对角线），其中 $m \ll n.$。按照方程（2.7）进行原位 LU 因式分解后，得到的 $L_i - I_n + U_i$ 具有相同的带宽，计算成本为$\mathcal{O}$（m^2n）。所得的 $2q$ 三角线性系统可分别通过简单代入法高效求解，计算成本为 $\mathcal{O}$（mn）。因此，无矩阵牛顿法的总计算成本可估算为 $\mathcal{O}$（qm^2n）。

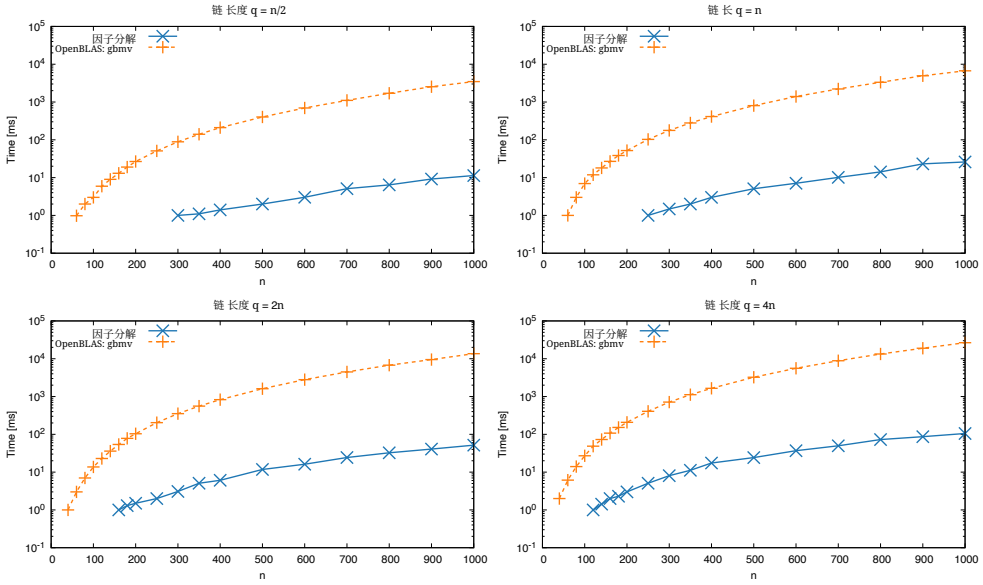


FIG. 3. *Tridiagonal Elemental Jacobians*

要使 F' 变为稠密矩阵，我们需要$q \geq \underline{q} \equiv (n-m-1)/m.$。对于 $q \gg \underline{q}$，计算 F' 的成本可估算为 $\mathcal{O}(qmn^2)$。需要计算一个长度为 q 的矩阵链乘积，其中一个因子的带宽为 $2m+1$，另一个因子在 $q \geq \underline{q}$. 时变为稠密矩阵。对于 $m \ll n$，矩阵自由牛顿法的优越性显而易见，因为 $\mathcal{O}(\underline{q}m^2n) < \mathcal{O}(qmn^2)$。

在大多数现实场景中，F' 的累积很可能会主导总体计算成本，从而使矩阵自由牛顿法的计算成本较标准牛顿法的计算成本降低 $\mathcal{O}\binom{n}{m}$。然而，设牛顿步的计算成本由线性牛顿方程组的求解主导。矩阵自由牛顿法的优越性要求$\mathcal{O}(\overline{q}m^2n) < \mathcal{O}(n^3)$，只有在假设 $n = \mathcal{O}(m^2)$ 时，才会违反这一点。

令所有 F'_i 均为三对角矩阵。例如，在采用中心有限差分进行空间离散的一维扩散隐式欧拉积分中，就会出现这种情形。使用托马斯算法进行原地 LU 因式分解 [23]，可在计算成本为 $\mathcal{O}(n)$ 的情况下得到三对角 $L_i - I_n - U_i$。方程（2.8）中的 $2q$ 线性系统分别也可按相同数量级的成本求解。矩阵自由牛顿法的总计算成本为 $\mathcal{O}(qn) < \max(\mathcal{O}(qn^2), \mathcal{O}(n^3))$。$\mathcal{O}\binom{n}{m} = \mathcal{O}(n)$ 的加速效果如图3所示，其中展示了针对单个牛顿步的多项实验结果以及$q/n = 0.5$，1、2、4、$\overline{0} < n \leq 10^3$. “先因式分解，再累积” 方法实现了新的矩阵自由牛顿法。它在图3的四个子图中均对应底部曲线。“先累积，再分解”方法使用 OpenBLAS 的 gbmv 方法（www.openblas.net）将带状矩阵与稠密向量相乘。其计算成本约为新方法的 n. 倍。我们寻找用于带状矩阵乘法的专用方法，但最终未能成功。我们预计，这样的算法对于

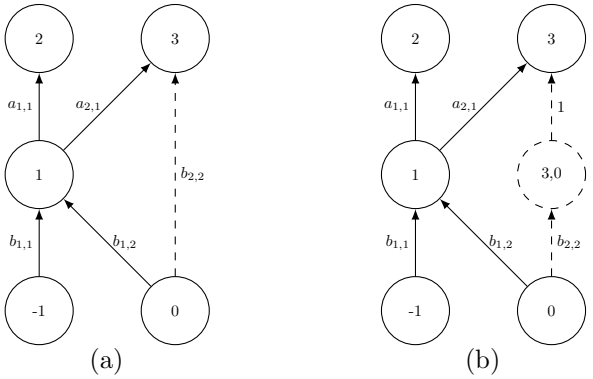


FIG. 4。 *Edge Splitting*

$q \ll n$. 不过，我们的基本参考实现并未优于 gbmv。由于更常见的情形是 $q \gg n$.，因此我们没有投入太多精力对其进行优化，因为它不会成为我们的首选方法

4. 泛化。 由伴随自动微分记录的有向无环图通常既不是分层的，也不是均匀的。根据定义 2.3，“分层”一词指将有向无环图分解为一系列表示元素雅可比矩阵的二分子有向无环图 $F_i'.$。若有向无环图是均匀分层的，则所有层都必须包含相同数量（n）的顶点，定义 2.4 对此作了正式表述。

下面我们将介绍如何使可逆有向无环图成为均匀分层图的思路。我们将对相应方法进行形式化，并分别证明其数值正确性。所有证明都基于以下微分链式法则的表述。

LEMMA 4.1（有向无环图上的微分链式法则）。设 G 为可微程序 $F\colon\ \mathbb{R}^n \to \mathbb{R}^m$ 的有向无环图，如定义 *2.2* 所示，其雅可比矩阵为

$$F' \equiv \left(\frac{dy_k}{dx_i}\right)_{i=1,\ldots,n}^{k=1,\ldots,m} = \left(\frac{dv_{p+k}}{dv_{i-n}}\right)_{i=1,\ldots,n}^{k=1,\ldots,m} \in \mathbb{R}^m \times \mathbb{R}^n \ .$$

Then

$$\frac{dv_{p+k}}{dv_{i-n}} = \sum_{(i-n,\ldots,p+k)} \prod_{(s,t)\in(i-n,\ldots,p+k)} \partial_{t,s} \tag{4.1}$$

对于 $i = 1, \ldots, n$ 和 $k = 1, \ldots, m$.

证明。参见 [1]。 ■

每种方法的直观说明之后，接着给出其形式化描述以及数值正确性的证明。

4.1. 边分裂。 考虑对第2节中的示例程序稍作修改，得到

```
对于(整数i=0;i<q;i++){x[i%2]*=x[(i+1)%
2];如果(!(i%2))x[(i+1)%2]=sin(x[(i+1)%2]);
}
```

对于q=2，记录图 4(a) 中的有向无环图。标记有局部偏导数$b_{2,2}$ (=x [0] for i=1) 的边跨越了两层，因此该有向无环图并非分层结构。我们的目标是将其转换为一个分层三部有向无环图，以表示局部雅可比乘积$A \cdot B$, 其中 $A \equiv (a_{j,i})$ 以及 $B \equiv (b_{j,i})$. 将 $\partial_{j',i'}$ 替换为 $a_{j,i}$ 或 $b_{j,i}$。这种细微的符号记法调整旨在使接下来的示例更易于理解。

通过边分裂，在连接层 i 和 j 中顶点的所有边上插入 $l-1=j-i-1$ 虚拟顶点，从而使有向无环图分层（此处，$l=2-0=2$ 会产生一个额外顶点，标记为 3，0，以表明它是边 (0，3) 的分裂顶点）。生成的新边中，一条保留原标签，其余边则按照图 4(b) 中的方式标记为 1（$l-1=1$ 虚拟顶点）。根据微分链式法则，雅可比矩阵 F' 显然在边分裂下保持不变。例如，

$$F' = \begin{pmatrix} a_{1,1} & 0 \\ a_{2,1} & 1 \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} \\ 0 & b_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,1} & a_{1,1}b_{1,2} \\ a_{2,1}b_{1,1} & a_{2,1}b_{1,2} + b_{2,2} \end{pmatrix} .$$

F' 的可逆性意味着所得均匀分层有向无环图中的所有元素雅可比矩阵均可逆。

Definition 4.2（边分裂）。设 $G=(V,E)$ 是定义 *2.2* 中的有向无环图。带标签 λ 的边 (i,j) 通过将其替换为两条新边 (i,k) 和 (k,j)，$k \notin V$并分别将它们标记为 $\partial_{k,\ i}=\lambda$ 和 $\partial_{j,\ k}=1$（或反之）而进行分裂。
新边的两种标记选择所带来的后果仍在研究之中。在接下来的讨论中，两种方案均视为等价。

Lemma 4.3. 设有向无环图 G，其在微分链式法则下诱导出一个可逆雅可比矩阵 $F' \in \mathbb{R}^{n \times n}$，如引理*4.1*所示。

1. 雅可比矩阵在定义 *4.2* 中所述的边分裂下保持不变。2. 如果反复进行边分裂后得到定义 *2.4* 中所述的均匀分层有向无环图，那么每一层的局部雅可比矩阵都是可逆的。

证明。
1. 方程 (4.1) 在边分裂下保持不变。2. 设均匀分层有向无环图的深度为 $q.$，则方程 (4.1) 得出 $F' = F'_q \cdot \ldots \cdot F'_1$，因此

$$(F')^{-1} = (F'_q \cdot \ldots \cdot F'_1)^{-1} = (F'_1)^{-1} \cdot \ldots \cdot (F'_q)^{-1}$$

这意味着所有 $F'_j, j=1,\ldots,q.$ 都是可逆的。
■

显然，一旦有向无环图变为分层有向无环图，边分裂就会终止。

4.2. 预累积。通过边分裂使有向无环图分层，结果相当直接。遗憾的是，由此得到的分层有向无环图很少是均匀分层的。此外，雅可比矩阵的可逆性意味着所有层的大小都至少等于 n. 我们提出先对局部子雅可比矩阵进行预累积，再将所得的二分子有向无环图分解为分层三部子有向无环图，以此作为使这类有向无环图均匀分层的方法。

图 5(a) 展示了一个非均匀分层三部有向无环图 G，分别具有 $n=m=3$个源、汇和 $p=4$ 个中间顶点。二分子有向无环图的连通分量诱导出 G 的纯分层三部子有向无环图，正式地而言，

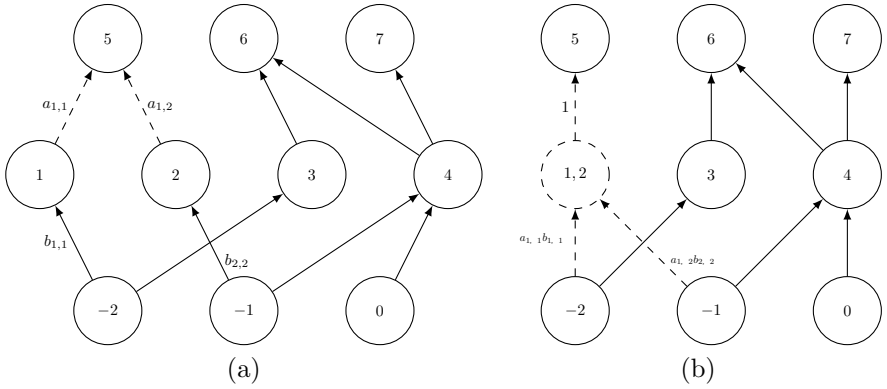


FIG. 5。 *Preaccumulation*

如定义 4.4 所述。例如，由三个顶点 1,2,5 张成的二分分量诱导出 G 的纯分层三部子有向无环图，其中包括 -2 和 -1.（类似地，由 -2,1, 3 张成的二分分量给出 G 的纯分层三部子有向无环图，其中包括 5 和 6.）。对 G 应用引理 4.1 中的链式法则，可得到一个局部雅可比矩阵 $A \in \mathbb{R}^{m\times n}$ （此处，$A \in \mathbb{R}^{1\times 2}$ 是 v_5 关于 v_{-2} 的梯度，而 v_{-1}）。

替换替换为 $\hat{A}$ 的二部有向无环图，违反了 G 分层的要求。将 A 分解为 $A = I_m \cdot A$ （或 $A = A \cdot I_n$），通过替换为修改后的纯分层三部子有向无环图 G，解决了这一违反问题，如图5(b)所示。 G 中的中间顶点数量等于 $\tilde{p} = \min(n, m)$，这意味着 G 中的中间顶点数量减少；如果 $\tilde{p} < p$. 图5得到 $1 = \tilde{p} < p = 2$，则生成的有向无环图将是均匀分层的。

最终，预累积总会生成均匀分层有向无环图，并保留原有向无环图中大小为 n 的所有中间层。最坏情况下，我们会得到 F'. 的二部有向无环图。我们的目标是像上文所示那样有选择地应用预累积。所得的均匀分层有向无环图旨在尽可能保留所有局部雅可比矩阵的稀疏性。

将二部图（类似地，有向无环图） $G = (V, E)$分解为连通分量，过程相当直接。每个顶点恰好属于一个连通分量。从尚未分配的顶点开始反复搜索（例如深度优先搜索），即可将图分解为连通分量，计算成本为$\mathcal{O}(|V| + |E|)$。

以下定义将上述内容形式化，随后证明预累积与二部有向无环图拆分相结合时在数值上的正确性。

DEFINITION 4.4（纯分层三部子有向无环图）。如定义*2.3*所示，分层三部子有向无环图 $G = (V, E)$ $V = (X, Z, Y)$，若满足 $P(Z) = X$ 且$S(Z) = Y$.，则称其为纯。

DEFINITION 4.5（预累积）。令 G 为定义 *4.4* 中表示局部雅可比乘积 $F' = F'_2 \cdot F'_1$. 的纯分层三部子有向无环图。预累积将 G 替换为表示 F'.的二部子有向无环图

LEMMA 4.6。设有向无环图 G 按照引理*4.1*中的微分链式法则诱导出可逆雅可比矩阵 $F' \in \mathbb{R}^{n\times n}$ 。

1. 雅可比矩阵在定义*4.5*所述的预累积下保持不变。

2. 预累积会终止。

证明。
1. 方程 (4.1) 在预累积下保持不变。2. 连接源与汇的所有路径的累积和在预累积下单调递减。因此，完全预累积得到表示 F'. 的二部有向无环图

■

Definition 4.7（二部有向无环图拆分）。一个表示局部雅可比矩阵 $F'_j \in \mathbb{R}^{m\times n}$ 的二部有向无环图 G ，通过将其替换为一个分层三部有向无环图来进行拆分，该图表示矩阵乘积 $F' \cdot I_n$（若满足 $n \leq m$），否则表示 $I_m \cdot F'$。
若我们的唯一目标是数值正确性，则可以省略条件 $n \leq m$ 。然而，为了实现最大稀疏性，我们尽可能减少新生成的边数。这两种选择带来的进一步影响仍在研究中。

Lemma 4.8。设有向无环图 G 在微分链式法则下诱导出可逆雅可比矩阵 $F' \in \mathbb{R}^{n\times n}$ ，如引理*4.1*所述。

1. 雅可比矩阵在二部有向无环图拆分下保持不变，如定义 *4.7* 所示。2. 如果重复进行二部有向无环图拆分，得到如定义 *2.4* 所示的均匀分层有向无环图，那么每一层的局部雅可比矩阵都是可逆的。

证明。
1. 方程 (4.1) 在二部有向无环图拆分下保持不变。 2. 设该均匀分层有向无环图的深度为 q. 方程 (4.1) 得出 $F' = F'_q \cdot \ldots \cdot F'_1$，因此 $(F')^{-1} = (F'_1)^{-1} \cdot \ldots \cdot (F'_q)^{-1}$，这表明所有 $F'_j,\ j = 1, \ldots, q.$ 都是可逆的。

■

4.3. 组合学。 接下来将以判定版本表述的组合 Matrix-Free Newton Step 问题，事实证明在计算上是难以处理的。

Problem 1 (Matrix-Free Newton Step).实例: 一个可逆有向无环图和一个整数 $K \geq 0$.问题: 能否用至多 K 次浮点运算[4]计算相应的牛顿步？

Theorem 4.9. Matrix-Free Newton Step是 *NP* 完全的。

证明。 Adjoint Computation 问题的 NP 完全性证明见 [16] ，它将 Ensemble Computation [5] 规约为由对角雅可比矩阵上的矩阵链乘积表示的均匀分层有向无环图 $\in \mathbb{R}^{n\times n}$ ，如下所示

$$F' = F'_q \cdot \ldots \cdot F'_1 \in \mathbb{R}^{n\times n} .$$

相应的牛顿步等于

$$(F')^{-1} \cdot y = (F'_q \cdot \ldots \cdot F'_1)^{-1} \cdot y = (F'_1)^{-1} \cdot \ldots \cdot (F'_q)^{-1} \cdot y .$$

事实证明，“先累积，再分解”方法更优，因为 F' 仍是对角矩阵。对其求逆会使任意给定的 Adjoint Computation实例的浮点运算次数增加一个固定的 n （倒数运算）偏移量。 Matrix-Free的一种解

[4]浮点运算

Newton Step 因此也能求解 Adjoint Computation，这意味着前者具有 NP困难性。此外，只需统计至多 $(q+1)n$ 次浮点运算，即可轻松高效地验证所提出的解。 ■

证明 Matrix-Free Newton Step problem 的 NP 完全性完全依赖于元素偏导数之间的代数依赖关系（尤其是相等关系）。底层有向无环图的结构最终证明十分简单。另一方面，使有向无环图均匀分层的方法，其动机纯粹来自对有向无环图结构的考量。所有元素偏导数都被假定为彼此独立。在这一阶段，定理 4.9 的证明中所用的归约对求解 Matrix-Free Newton Step 实例近似解的算法没有任何影响。事实上，针对 [16] 的相关证明也适用类似的说法。预计未来的研究将弥补这一空白。

5. 结论。 新的矩阵自由牛顿法有望显著降低求解非线性方程组的计算成本。这一主张得到了针对具有三对角元素雅可比矩阵的问题所进行的运行时间测量的支持。对于具有不规则稀疏模式的元素雅可比矩阵，以及计算开销大的残差，仍有进一步的挑战需要应对。

无矩阵牛顿法依赖于有向无环图反转。对于规模极大的问题，有向无环图的大小可能超出可用内存资源。组合 DAG Reversal 问题要求对可用存储空间进行分配，以使总体计算成本最小化。众所周知，该问题是 NP 完全的[17] ，相关的 Call Tree Reversal 问题 [15]也是如此。检查点方法[8, 22, 24] 为多种特殊情形提供了解决方案。

具有不规则稀疏模式的元素雅可比矩阵因式分解，会在稀疏线性代数中产生若干著名的组合问题，例如 Bandwidth [19]或 Directed Elimination Ordering [20]。为了使我们的方法优于其他最先进的方案，必须尽量减少填充。在这一十分活跃的研究领域，过去和正在进行的工作已积累了丰富的成果，可供我们借鉴。

另一类潜在目标是采用带状层的替代模型，用于处理通过机器学习（ML）获得的计算成本高昂的非线性残差。尽管训练此类模型可能颇具挑战，但采用新方法求逆其雅可比矩阵必将受益。目前的研究旨在探索将该方法扩展至机器学习模型的可能性，这些模型的层具有其他合适的稀疏模式，例如具有三角形元素雅可比矩阵的层。

致谢。 图3中报告的实验由 Gero Kauerauf 在亚琛工业大学攻读理学硕士期间作为其硕士论文项目的一部分完成。

参考文献

[1] W. Baur and V. Strassen, *The complexity of partial derivatives*, 理论计算机科学，22（1983），第317–330页。

[2] C. Bischof, M. Bücker, P. Hovland, U. Naumann, and J. Utke, 编者, *Advances in Automatic Differentiation* 64 ，计算科学与工程讲义（LNCSE）第{v号}，施普林格，2008年。

[3] P. Deuflhard, *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms* 35 ，计算数学卷，施普林格国际出版社，2004年。

[4] I. S. Duff, A. M. Erisman, and J. K. Reid, *Direct Methods for Sparse Matrices*, 克拉伦登出版社，1986年。

[5] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences)* ，W. H. Freeman，初版

版，1979年。[6] A. Gebremedhin, F. Manne, and A. Pothen, *What color is your Jacobian? Graph coloring for computing derivatives* 47 2005 629 705 , SIAM Review, （），第–页。[7] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT出版社，美国马萨诸塞州剑桥，2016年。http://www.deeplearningbook.org. [8] A. Griewank, *Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation* 1 1992 35 54 , Optimization Methods and Software, （），第–页。[9] A. Griewank, D. Juedes, and J. Utke, *Algorithm 755: ADOL-C: A package for the auto-matic differentiation of algorithms written in C/C++*, ACM Transactions on Mathematical Software，22（1996），第131–167页。
[10] A. Griewank and A. Walther, *Evaluating Derivatives. Principles and Techniques of Algo-* , , 应用数学其他著 *rithmic Differentiation Second Edition* 105
作丛书OT号，SIAM，2008年。[11] M. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards，42（1952）。[12] R. Hogan, *Fast reverse-mode automatic differentiation using expression templates in C++*, ACM Transactions on Mathematical Software，40（2014），第26:1–26:24页。[13] C. T. Kelley, *Solving Nonlinear Equations with Newton's Methods*, SIAM，2003年。[14] D. Knoll and D. Keyes, *Jacobian-free Newton–Krylov methods: A survey of 348 approaches and applications* 193 2004 357 397 , J. Comp. Phys., （），第–页。[15] U. Naumann, *Call tree reversal is NP-complete*，载于 [2], 施普林格，2008年，第13–22页。[16] U. Naumann, *Optimal Jacobian accumulation is NP-complete*, Mathematical Programming，112（2008），第427–441页。[17] U. Naumann, *DAG reversal is NP-complete*, Journal of Discrete Algorithms，7（2009），第402–410页。[18] I. Newton, *Philosophiae naturalis principia mathematica*, Colonia Allobrogum：由Cl. et Ant. Philibert出版，1760年。[19] C. H. Papadimitriou, *The NP-completeness of the bandwidth minimization problem*, Computing，16（1976），第263–270页。[20] D. J. Rose and R. E. Tarjan, *Algorithmic aspects of vertex elimination on directed graphs*, SIAM J. Appl. Math.，34（1978），第176–197页。[21] M. Sagebaum, T. Albring, and N. Gauger, *High-performance derivative computations using CoDiPack* 45 2019 , ACM Transactions on Mathematical Software, （）。[22] P. Stumm and A. Walther, *Multistage approaches for optimal offline checkpointing*, SIAM Journal on Scientific Computing，31（2009），第1946–1967页。[23] L. H. Thomas, *Elliptic problems in linear differential equations over a network*，技术报告，Watson科学计算实验室报告，哥伦比亚大学，纽约，1949年。[24] Q. Wang, P. Moin, and G. Iaccarino, *Minimal repetition dynamic checkpointing algo-rithm for unsteady adjoint calculation*, SIAM Journal on Scientific Computing，31（2009），第2549–2567页。[25] D. T. Whiteside, *The Mathematical Papers of Isaac Newton (7 volumes)*, 剑桥大学出版社，1967年至1976年。