

融入几何元素的并行计算课程教学实践 ——以卡茨马兹型算法为例

吴念慈

中南民族大学数学与统计学学院, 湖北 武汉

收稿日期: 2024年5月6日; 录用日期: 2024年6月7日; 发布日期: 2024年6月14日

摘要

并行计算作为数据科学与大数据技术专业的一门重要课程, 以介绍各种数值并行算法原理为主要内容。本文以卡茨马兹型算法为例, 采用案例教学法, 将几何元素融入到并行计算课程教学实践中, 旨在帮助学生能够更直观地理解卡茨马兹算法及其变种之间的联系, 建立更牢固的知识结构。

关键词

并行计算, 卡茨马兹算法, 几何元素, 案例教学

Teaching Practice of Parallel Computing Course with Incorporated Geometric Elements

—A Case Study of the Kaczmarz Algorithms

Nianci Wu

School of Mathematics and Statistics, South-Central Minzu University, Wuhan Hubei

Received: May 6th, 2024; accepted: Jun. 7th, 2024; published: Jun. 14th, 2024

Abstract

Parallel computing, as an important course of data science and big data technology, mainly introduces the principles of various numerical parallel algorithms. Taking Kaczmarz algorithm as an example, this paper uses case teaching method incorporated geometric elements into the teaching practice of parallel computing course, aiming to help students understand the relationship be-

tween Kaczmarz algorithm and its variants intuitively, and build a more solid knowledge structure.

Keywords

Parallel Computing, Kaczmarz Algorithm, Geometric Elements, Case Teaching

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

数值并行计算为解决实际问题提供了有效方案,可进一步促进科学技术的发展和社会经济的进步[1]。随着计算机硬件技术的发展,单个处理器的性能已经达到了瓶颈,而并行计算技术能够充分利用多个处理器或计算节点的并行计算能力,提高计算效率。同时,数值并行计算也是计算科学和工程领域的重要分支之一,对推动学科发展具有重要意义。此外,数值并行计算还可以促进不同学科之间的交叉融合,推动学科的跨界发展[2]-[5]。

线性方程组 $Ax=b$ ($A \in R^{m \times n}$, $x \in R^n$, $b \in R^m$) 的求解是并行计算课程的一个重要部分(见教材[6]第五章)。当 m 和 n 很大时,卡茨马兹算法(Kaczmarz Algorithm)因其迭代格式简单、几何意义明显、理论分析简洁,常被用来求解此类问题[7]。若用 a_i^T 表示系数矩阵 A 的第 i 行, b_i 表示右端项 b 的第 i 个分量,其迭代格式在 1 个处理器上的代数表达式为

$$x^{(k)} = x^{(k-1)} + \frac{b_{i_k} - a_{i_k}^T x^{(k-1)}}{\|a_{i_k}\|^2} a_{i_k}, \quad (1)$$

其中迭代步数 $k=1,2,\dots$, 行指标 $i_k=1,2,\dots,m$ 。

考虑一般的形式,选取 $m(m \geq 2)$ 个处理器,可得多投影迭代格式

$$x^{(k)} = \frac{1}{m} \sum_{i=1}^m \left(x^{(k-1)} + \frac{b_i - a_i^T x^{(k-1)}}{\|a_i\|^2} a_i \right).$$

经过简单计算可知,

$$x^{(k)} = x^{(k-1)} + \frac{1}{m} \sum_{i=1}^m \left(\frac{b_i - a_i^T x^{(k-1)}}{\|a_i\|^2} a_i \right) = x^{(k-1)} + \frac{1}{m} A^T M (b - Ax^{(k-1)}),$$

其中对角矩阵 $M = \text{diag}(1/\|a_i\|^2)$, 此方法称为奇米诺算法(Cimmino Algorithm)。因而,奇米诺算法是卡茨马兹算法的一种并行形式。

几何解释。若按照顺序 $1,2,\dots,m,1,2,\dots,m,\dots$ 依次选取行指标,则迭代向量 $x^{(k)}$ 是 $x^{(k-1)}$ 到超平面 H_{i_k} 上的正交投影。迭代过程如图 1 所示(以 $m=6$),其中实心圆表示迭代向量,超平面 $H_i = \{x | a_i^T x = b_i\}$, $i=1,2,\dots,6$,超平面的交点表示方程组的解。

几何解释能够与实际情境相联,学生通过视觉感知理解概念,这种直观性可以帮助学生对图形化的抽象概念有更好的理解[8]。

2. 卡茨马兹算法及案例分析

卡茨马兹算法及其变种具有明显的几何意义，本节将结合具体案例，将此类算法蕴含的几何元素融入到教学实践中。

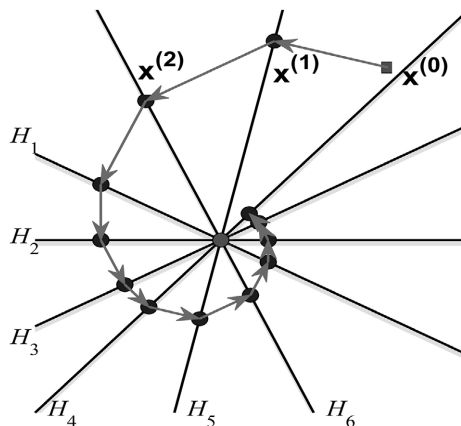


Figure 1. Demonstration of the iterative process for the Kaczmarz algorithm

图 1. 卡茨马兹算法的迭代过程演示

2.1. 教学案例

考虑用卡茨马兹算法求解下列线性方程组

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \Leftrightarrow Ax = b.$$

每个方程看作是一个约束条件，而求解线性方程组就是要找到一个满足所有约束条件的点 $x^* = [1 \ 1]^T$ 。

2.2. 案例分析

第一步：当 $k=1$ 时，取行指标 $i=1$ ，根据迭代公式(1)计算

$$x^{(1)} = x^{(0)} + \frac{b_1 - a_1^T x^{(0)}}{\|a_1\|_2^2} a_1 = \begin{bmatrix} 7/5 & 4/5 \end{bmatrix}^T,$$

其中初值向量 $x^{(0)} = [2 \ 2]^T$ 。记 $X_0(2,2)$ 和 $X_1\left(\frac{7}{5}, \frac{4}{5}\right)$ ，易验证 $x_1^{(1)} + 2x_2^{(1)} = 3$ ，说明 X_1 在第一个超平面 H_1 上；同时， $\overline{X_0 X_1}$ 与第一个方程的法向量 a_1 共线，因而 X_1 是 X_0 在 H_1 上的投影。此时误差为 $\|x^{(1)} - x^*\|_2 \approx 0.447 < \|x^{(0)} - x^*\|_2 \approx 1.414$ 。

第二步：当 $k=2$ 时，取行指标 $i=2$ ，根据迭代公式(1)计算

$$x^{(2)} = x^{(1)} + \frac{b_2 - a_2^T x^{(1)}}{\|a_2\|_2^2} a_2 = \begin{bmatrix} 71/50 & 43/50 \end{bmatrix}^T.$$

记 $X_2\left(\frac{71}{50}, \frac{43}{50}\right)$ ，易验证 $x_1^{(2)} + 3x_2^{(2)} = 4$ ，说明 X_2 在第二个超平面 H_2 上；同时， $\overline{X_1 X_2}$ 与第二个方程的法

向量 a_2 共线, 因而 X_2 是 X_1 在 H_2 上的投影。此时误差为 $\|x^{(2)} - x^*\|_2 \approx 0.443 < 0.447$ 。

第三步: 当 $k=3$ 时, 取行指标 $i=3$, 根据迭代公式(1)计算

$$x^{(3)} = x^{(2)} + \frac{b_3 - a_3^T x^{(2)}}{\|a_3\|_2^2} a_3 = \begin{bmatrix} 881 & 496 \\ 650 & 650 \end{bmatrix}^T。$$

记 $X_3 \left(\frac{881}{650}, \frac{496}{650} \right)$, 易验证 $2x_1^{(3)} + 3x_2^{(3)} = 5$, 说明 X_3 在第三个超平面 H_3 上; 同时, $\overline{X_2 X_3}$ 与第三个方程的法向量 a_3 共线, 因而 X_3 是 X_2 在 H_3 上的投影。此时误差为 $\|x^{(3)} - x^*\|_2 \approx 0.427 < 0.443$ 。

引导学生发现规律: 当 k 增大时, $x^{(k)}$ 将不断趋近 x^* 。上述计算利用具体的几何形状或者图形来演示抽象的理论和概念, 不仅将卡茨马兹算法的迭代过程可视化, 也说明了卡茨马兹算法的收敛性。

案例教学是一种实践性很强的教学方法, 在向量序列 $x^{(k)}$ 的计算中, 学生掌握了卡茨马兹算法的基本过程, 提升了学生实际应用能力, 也有助于学生将理论知识转化为实际操作的能力。教学过程中使用的案例要求具有情节性或故事性, 如卡茨马兹算法具有几何背景, 在案例讲述中, 学生更容易投入到学习中, 积极参与讨论和思考。在线性方程组的求解中, 若按照不同方式选取行指标, 要求学生对卡茨马兹算法的收敛性进行评估, 这种过程培养了学生的批判性思维, 提升了学生的分析能力。本文采用的案例涉及到了代数和几何等多个学科领域的知识, 通过对交叉学科的学习和思考, 有助于学生理解各学科知识的联系。在共同分析案例、讨论解决方案、协调分工等教学实践中, 学生的团队合作和沟通能力也得到了培养。

2.3. 案例延申

案例延申是对问题的深入探索, 通过教师与学生的讨论, 学生可以从不同的角度思考和理解新概念, 有助于培养学生创新思维 and 创新能力。按照“选择适当的案例” - “确定延申的方向” - “引导学生分析和讨论” - “探索相关知识” - “总结和反思”等基本步骤进行案例延申, 得到如下所述的松弛卡茨马兹算法。

若在卡茨马兹迭代格式中引入松弛参数, 即

$$x^{(k)} = x^{(k-1)} + \boxed{\lambda} \frac{b_i - a_i^T x^{(k-1)}}{\|a_i\|^2} a_i。 \quad (2)$$

当 $\lambda=1$ 时, 即为卡茨马兹迭算法, 计算过程为正交投影; 当 $\lambda=2$ 时, 计算过程变为反射运算, 如图 2 所示(以 $m=2$ 为例), 超平面的交点表示方程组的解。

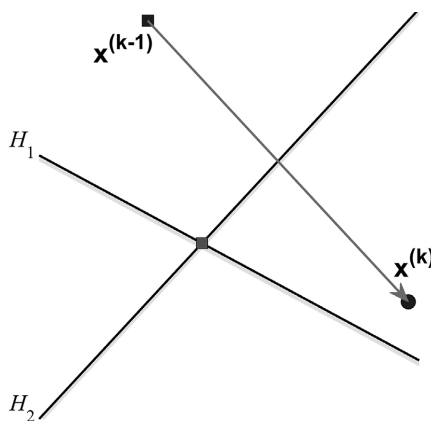


Figure 2. Demonstration of the iterative process for the reflection algorithm

图 2. 反射运算的迭代过程演示

考虑第 2.1 节中的教学案例, 当 $k=1$ 时, 取初值 $x^{(0)} = [2 \ 2]^T$, 以及行指标 $i=1$, 计算

$$x^{(1)} = x^{(0)} + 2 \frac{b_1 - a_1^T x^{(0)}}{\|a_1\|_2^2} a_1 = \left[\frac{4}{5} \quad -\frac{2}{5} \right]^T,$$

由此可知, $\frac{1}{2}(x^{(0)} + x^{(1)}) = \left[\frac{7}{5} \quad \frac{4}{5} \right]^T$ 。易验证 $\frac{7}{5} + \frac{4}{5} \times 2 = 3$, 说明 $X_0(2,2)$ 和 $X_1\left(\frac{4}{5}, -\frac{2}{5}\right)$ 的中点在第一个方程上, 又因为 $\overline{X_0 X_1}$ 与第一个方程的法向量 a_1 共线, 因而 $x^{(1)}$ 是 $x^{(0)}$ 关于 H_1 的对称点, 即说明上述过程是一个反射运算。

以上步骤可以根据具体教学目标和案例特点进行灵活调整和组合, 以实现最佳的教学效果。

3. 基于卡茨马兹算法的延拓式教学

延拓式教学旨在学生在已有知识基础上的延伸和拓展, 促进学生对新知识的深入理解 and 应用, 本节将继续引导学生对卡茨马兹算法做进一步的学习。

1) **知识评估**。通过课堂或小组讨论、问答等方式评估学生对卡茨马兹算法的理解程度。

2) **情境引入**。引导学生思考如何对卡茨马兹算法并行化, 这有助于激发学生的学习兴趣和提高自己的学习动机。

3) **总结规律**。通过解释、示范、演示等方式对卡茨马兹算法进行总结, 帮助学生理解 and 应用新知识。

几何解释。首先将 $x^{(k-1)}$ 依次投影到第 $i_1, \dots, i_p$ 个超平面 $H_{i_1}, \dots, H_{i_p}$, 分别得到 $x^{(k-1,1)}, \dots, x^{(k-1,p)}$; 然后取其平均值, 得到下一个迭代向量 $x^{(k)}$ 。迭代过程如图 3 所示(以 $m=2, t=2$ 为例), 超平面的交点表示方程组的解。

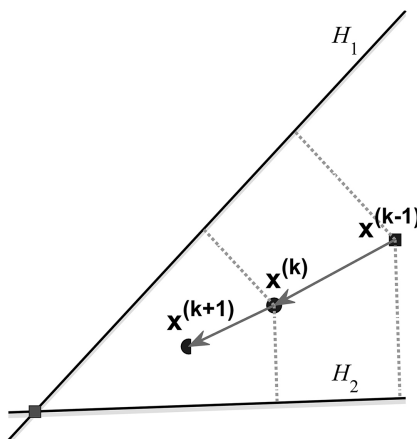


Figure 3. Demonstration of the iterative process for the Average Block Kaczmarz Algorithm

图 3. 平均块卡茨马兹算法的迭代过程演示

4) **循序善诱**。通过理论分析与实践等方式, 鼓励学生积极参与到学习中, 探索和发现新知识, 培养学生的自主学习能力。

若处理器个数为 p , 选取系数矩阵 A 的第 $i_1, i_2, \dots, i_p$ 行, 可得迭代格式

$$x^{(k)} = \frac{1}{p} \sum_{t=1}^p \left(x^{(k-1)} + \frac{b_{i_t} - a_{i_t}^T x^{(k-1)}}{\|a_{i_t}\|^2} a_{i_t} \right), \quad (3)$$

或等价地表示为 $x^{(k)} = (x^{(k-1,1)} + x^{(k-1,2)} + \cdots + x^{(k-1,p)})/p$ ，其中

$$x^{(k-1,t)} = x^{(k-1)} + \frac{b_t - a_t^T x^{(k-1)}}{\|a_t\|^2} a_t, \quad t = 1, 2, \dots, p。$$

此方法称为平均块卡茨马兹算法(Average Block Kaczmarz Algorithm) [9]。

5) 组织讨论。以小组活动或实践任务的形式组织学生参与讨论，帮助学生共同探讨和应用延拓内容，以第 2.1 节中的教学案例为例。

第一步(1): 当 $k=1$ 时，取行指标 $i_1=1$ ， $i_2=2$ ，根据迭代公式(3)计算

$$x^{(0,1)} = x^{(0)} + \frac{b_1 - a_1^T x^{(0)}}{\|a_1\|^2} a_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix} \text{ 和 } x^{(0,2)} = x^{(0)} + \frac{b_2 - a_2^T x^{(0)}}{\|a_2\|^2} a_2 = \begin{bmatrix} 8 \\ 5 \end{bmatrix}^T,$$

其中初值向量 $x^{(0)} = [2 \quad 2]^T$ 。

第一步(2): 计算

$$x^{(1)} = \frac{1}{2} (x^{(0,1)} + x^{(0,2)}) = \begin{bmatrix} 15 \\ 10 \end{bmatrix}^T,$$

以及误差 $\|x^{(1)} - x^*\|_2 \approx 0.539 < \|x^{(0)} - x^*\|_2 \approx 1.414$ 。

第二步(1): 当 $k=2$ 时，取行指标 $i_1=3$ ， $i_2=1$ ，根据迭代公式(3)计算

$$x^{(1,1)} = x^{(1)} + \frac{b_3 - a_3^T x^{(1)}}{\|a_3\|^2} a_3 = \begin{bmatrix} 187 \\ 130 \end{bmatrix}^T \text{ 和 } x^{(1,2)} = x^{(1)} + \frac{b_1 - a_1^T x^{(1)}}{\|a_1\|^2} a_1 = \begin{bmatrix} 37 \\ 25 \end{bmatrix}^T。$$

第二步(2): 计算

$$x^{(2)} = \frac{1}{2} (x^{(1,1)} + x^{(1,2)}) = \begin{bmatrix} 1897 \\ 1300 \end{bmatrix}^T,$$

以及误差 $\|x^{(2)} - x^*\|_2 \approx 0.531 < 0.539$ 。

第三步: 当 $k=3$ 时，取行指标 $i_1=2$ ， $i_2=3$ ，根据迭代公式(3)计算 $x^{(2,1)}$ ， $x^{(2,2)}$ 和 $x^{(3)}$ ，计算过程与上述类似，不再赘述。

6) 反馈评价。及时提供反馈和评价，帮助学生了解自己的学习进展并指导学生进一步学习。例如，若取 $p=m$ ，即可得到引言中的奇米诺算法，因而说明平均块卡茨马兹算法也是卡茨马兹算法的一种并行形式。

7) 总结反思。在学习过程结束时，帮助学生总结所学内容，并鼓励学生将新知识应用到实际问题中，例如，行指标的选取顺序对算法误差大小的影响；算法收敛速率的理论分析；若对系数矩阵的列做运算应该如何设计算法。巩固学习成果、梳理思路、反思与启发可进一步深化学生对知识的理解和记忆，并提高学生的学习成果转化能力。

4. 小结

几何解释在不同学科的案例教学中应用广泛，例如数学、物理、工程、生物等领域，几何都能提供一种通用的语言和工具，帮助学生跨学科地理解和应用所学知识。几何解释往往能够激发学生的好奇心和求知欲，因为几何化教学可以呈现出一些有趣的形态或者规律，促使学生更积极地参与到案例教学中，并通过自己的思考和探索来深化对知识的理解。本文以卡茨马兹型算法为例，将几何元素融入到并行计

算课程教学实践中,旨在帮助学生能够更直观地感受到卡茨马兹算法及其变种算法之间的联系,能够建立更牢固的知识结构,并容易地将所学知识与实际问题联系起来。

融入延拓式教学,深挖卡茨马兹型算法与逐次超松弛迭代法(参考文献[6]第五章)之间的联系是本课程的一个研究方向,这有助于理解卡茨马兹型算法收敛速度、收敛条件、有效性等方面的性质;结合机器学习和优化方法,可以进一步提高卡茨马兹型算法的性能,如利用机器学习技术来预测迭代过程中需要更新的行指标,从而加速收敛过程。

在教学实践中,及时更新教学资源,引入最新的研究成果,保持教学内容的前沿性是本课程的一个改进内容。结合教学大纲,紧贴重点和难点,提取和整合具体的问题个案,如 Google 的网页排序模型计算大规模矩阵的特征值和特征向量(参考文献[6]第六章)、拟合高维数据求解大规模线性方程组(参考文献[6]第五章)等,建立一个具有一定关联性、典型性和真实性的案例库是本课程的另一个改进内容。

基金项目

中南民族大学校级教研项目“《并行计算》课程教学案例库的建设路径与实践”(项目编号:JYX23017)。

参考文献

- [1] 石钟慈. 第三种科学方法——计算机时代的科学计算[M]. 北京: 清华大学出版社, 2000.
- [2] 谭立湘, 李斌. 人工智能时代“并行计算”课程的改革与探索研究[J]. 工业和信息化教育, 2023(5): 26-29.
- [3] 梁毅, 楼佳明, 方娟, 赵昱, 王秀娟. 产学协同育人背景下的并行计算课程改革与实践[J]. 计算机教育, 2023(4): 38-42.
- [4] 李文正, 贾克斌, 沈多加. 中外合作办学模式下并行计算课程教学体系的研究[J]. 高教学刊, 2021, 7(31): 22-24.
- [5] 范培勤, 韩梅. 《并行计算》课程教学方法探讨[J]. 教育现代化, 2019, 6(62): 236-237.
- [6] 李晓梅, 吴建平. 数值并行算法[M]. 北京: 科学出版社, 2014.
- [7] Kaczmarz, S. (1937) Angenaherte auflosung von systemen linearer gleichungen. *Bulletin International de l'Academie Polonaise des Sciences A*, **35**, 355-357.
- [8] 雍龙泉, 刘三阳. 线性方程组迭代法的几何解释[J]. 大学数学, 2023, 39(2): 92-99.
- [9] Necoara, I. (2019) Faster Randomized Block Kaczmarz Algorithms. *SIAM Journal on Matrix Analysis and Applications*, **40**, 1425-1452. <https://doi.org/10.1137/19M1251643>