

来自 Quanta Magazine

作者：Steve Nadis

机器之心编译

从你网购的包裹如何以最快速度送达，到航空公司如何规划数千架飞机的航线以节省燃料，背后都有一个近 80 岁「高龄」的数学方法在默默工作。它被誉为优化领域的基石，高效又令人信赖。然而，一个奇怪的事实是：几十年来，没有人能从理论上完美解释它为何如此高效。现在，这个谜题的最后一块拼图，终于被找到了。

1939 年，当时还是加州大学伯克利分校一年级研究生的乔治·丹齐格（George Dantzig）在一次统计学课上迟到了。他从黑板上抄下了两个问题，以为是家庭作业。他后来回忆说，他发现这次的作业「比平时难得多」，并为自己多花了好几天才完成而向教授道歉。

几周后，他的教授告诉他，他成功解决了统计学领域两个尚待解决的著名问题。

丹齐格的这项成果为他的博士论文奠定了基础，并在几十年后成为了电影《心灵捕手》的灵感来源。



George B. Dantzig 图：算法爱好者

乔治·丹齐格（George Dantzig，1914—2005），美国著名数学家，1947 年提出了单纯形法，被称为线性规划之父。

丹齐格在 1946 年，也就是二战刚结束后不久，获得了博士学位，并很快成为了新成立的美国空军的数学顾问。

与所有现代战争一样，第二次世界大战的胜负取决于对有限资源的审慎分配。但与以往的战争不同，这场冲突的规模是全球性的，其胜利在很大程度上是依靠纯粹的工业实力取得的。当时，美国能够生产比其敌人更多的坦克、航空母舰和轰炸机。

了解到这一点后，军方对优化问题产生了浓厚兴趣，即：**在可能涉及成百上千个变量的情况下，如何战略性地分配有限的资源？**

美国空军交给丹齐格的任务是，找出解决这类优化问题的新方法。作为回应，他发明了**单纯形法（simplex method）**，这个算法借鉴了他在近十年前解决黑板问题时所发展出的一些数学技巧。

近 80 年后，当需要在复杂的约束条件下做出后勤或供应链决策时，单纯形法仍然是使用最广泛的工具之一。它高效且行之有效。法国国家科学研究中心（CNRS）的 Sophie Huiberts 说道：「它一直运行得很快，没人见过它变慢过。」

与此同时，一个奇特的性质长期以来为丹齐格的方法蒙上了一层阴影。

1972 年，数学家们证明：**该算法完成任务所需的时间可能会随着约束条件的数量呈指数级增长。**

因此，无论该方法在实践中运行得多快，理论分析始终给出最坏情况下的场景，暗示它可能需要指数级更长的时间。对于单纯形法，「我们研究算法的传统工具并不奏效，」Huiberts 说。

但在即将于 12 月在计算机科学基础（Foundations of Computer Science）会议上发表的一篇新论文中，Huiberts 和慕尼黑工业大学的博士生 Eleon Bach 似乎已经克服了这个问题。

Optimal Smoothed Analysis of the Simplex Method

Eleon Bach^{*1} and Sophie Huiberts^{†2}

¹EPFL

²CNRS, Université Clermont Auvergne Clermont Auvergne INP, Mines Saint-Etienne, LIMOS

April 5, 2025

公众号：算法爱好者

- 论文标题：Optimal Smoothed Analysis of the Simplex Method
- 论文地址：<https://arxiv.org/abs/2504.04197>

他们不仅使该算法变得更快，还从理论上解释了这一现象：**长期以来令人担忧的指数级运行时间在实践中并未出现。**

这项成果建立在 Daniel Spielman 和滕尚华（Shang-Hua Teng）2001 年里程碑式成果《Smoothed Analysis of Algorithms: Why the Simplex Algorithm Usually Takes Polynomial Time》（arXiv:cs/0111050）之上。滕尚华表示，这项成果「卓越且优美」。

「这是一项非常了不起的技术工作，它巧妙地结合了以往研究中发展的许多思想，同时增加了一些非常好的新技术构想，」波恩大学的数学家 László Végh 说道，他没有参与这项研究。

从几个角度来看，**单纯形算法的整个执行过程，就是寻找一条路径（比方说，从底部顶点到最高点），并且这条路径包含的步数最少，或者说，经过的边最少。**总步数则与算法的运行时间（或「复杂度」）相关，目标是在最少的步数内解决问题。在这张图中找出从底部到顶部的最短路线，就相当于找出了这类算法所能采取的最有效形式。

如果不是因为以下两个问题，找到一条快速直接的路径可能会很容易：

- 第一，原始形状可能比我们的例子复杂得多，有更多的面、顶点和边。

- 第二，没有地图可以引导你。你无法看到你试图导航的整个物体。相反，你从一个顶点开始，选择要先沿着哪条边走。在下一个顶点你做同样的事情，依此类推，并不确定你选择的边会把你带到哪里。

如果你运气特别差，你可能会在每个顶点都走错方向，最终陷入迷宫。「你可能会走上从 A 点到 B 点的最长路径。」Bach 说，「因为在每个拐角处，都有人告诉你该走向错误的方向。」正是这种情况可能导致需要指数级时间才能完成的最坏情况。

2001 年，Spielman 和滕尚华证明：引入一点点随机性可以帮助避免这样的结果。



滕尚华（左）和 Daniel Spielman 在他们 2001 年的里程碑式成果中运用了随机性。

回到前面的例子（注意这里极大地简化了 Spielman 和滕尚华的论证），让我们假设你到达的每个拐角处都有六个选择。如果总是有人告诉你走最差的方向，你就会陷入困境。然而，如果你依靠掷骰子来决定，你将有六分之五的机会避开最差的选择，从而可能避免灾难。在过程中注入随机性和不确定性的元素是合理的，因为在现实世界中，测量永远不是精确的。

当然，Spielman 和滕尚华引入随机性的方式完全不同，但他们证明：运行时间绝不会比约束数量的某个固定次幂更差（例如 n^2 ）——这就是所谓的多项式时间 (polynomial time)。这比指数时间（例如 2^n 的形式）要好得多。

最优几何

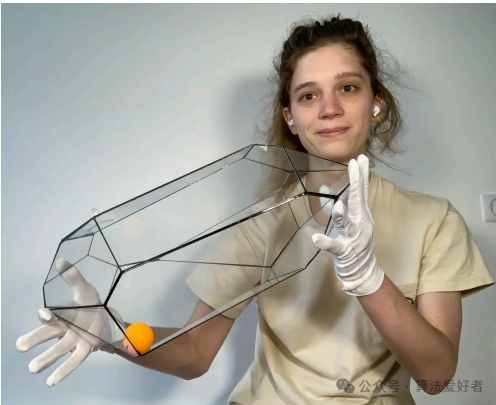
单纯形法旨在解决这样一类问题：假设一家家具公司生产衣柜、床和椅子。巧合的是，每个衣柜的利润是椅子的三倍，而每张床的利润是椅子的两倍。如果我们想把这写成一个表达式，用 a、b 和 c 分别代表生产的家具数量，那么总利润将与 $3a+2b+c$ 成正比。

为了实现利润最大化，公司应该各种物品各生产多少件呢？

答案取决于它面临的约束条件。比方说，该公司每月最多能生产 50 件产品，因此 $a+b+c \leq 50$ 。衣柜的制造难度更大——假设产量不能超过 20 个，所以 $a \leq 20$ ；椅子需要特殊的木材，而且供应有限，所以 c 必须小于 24。

单纯形法可将这类情况（通常涉及更多变量）转化为一个几何问题。

想象一下在一个三维图表中绘制我们对 a、b 和 c 的约束条件。如果 $a \leq 20$ ，我们可以想象在三维图上有垂直于 a 轴的平面，在 $a=20$ 的位置将其切开。我们会规定，我们的解必须位于该平面上或其下方的某个位置。同样，我们可以为其他约束条件创建边界。这些边界组合起来可以将空间分割成一个复杂的三维形状，称为多面体 (polyhedron)。



Sophie Huiberts 手持一个优化问题的几何模型。

然而，这并没有完全解决问题。

Huiberts 指出，他们的多项式时间值仍然相当高，其中包含一个 30 次幂的项，对于指数来说，这是一个相当大的数字。

因此，在过去二十年里，研究人员一直在逐步降低这个数值。

在 Bach 和 Huiberts 的新论文中，他们在算法中应用了更多的随机性，以证明运行时间可以保证远低于先前确立的水平。他们还证明：基于 Spielman 和滕尚华所建立方法的算法，其运行速度无法超越他们获得的值。Huiberts 说，这告诉我们，「我们完全理解了单纯形法的这个模型。」

1.1 Our results

Our main contribution is a substantially improved shadow bound presented below. An overview about our and previous upper and lower bounds are summarized in Table 1.

	Expected number of vertices	Model
Borgwardt	$\Theta(d^{3/2}\sqrt{\log n})$	Average-case, Gaussian
Spielman, Teng'04	$O(\sigma^{-6}d^3n + d^6n\log^3 n)$	$D(n, d, \sigma)$
Deshpande, Spielman'05	$O(\sigma^{-2}dn^2\log n + d^2n^2\log^3 n)$	$D(n, d, \sigma)$
Vershynin'09	$O(\sigma^{-4}d^3 + d^6\log^3 n)$	$D(n, d, \sigma)$
Dadush, Huijberts'18	$O(\sigma^{-2}d^2\sqrt{\log n} + d^6\log^{1.5} n)$	$D(n, d, \sigma)$
Huijberts, Lee, Zhang'23	$O(\sigma^{-3/2}d^{3/4}\log^{7/4} n + d^{19/4}\log^{13/4} n)$	$D(n, d, \sigma)$
This paper	$O(\sigma^{-1/2}d^{1/4}\log(n)^{7/4} + d^3\log(n)^2)$	$R(n, d, \sigma)$
Huijberts, Lee, Zhang'23	$\Omega(\min(\frac{1}{\sqrt{cd\sqrt{\log n}}}, 2^d))$	$D(4d - 13, d, \sigma)$
This paper	$\Omega(\frac{\sqrt{d}}{\sqrt{\sigma\sqrt{\ln(4/\sigma)}}})$	$R(\lfloor (4/\sigma)^d \rfloor, d, \sigma)$

Table 1: Bounds of expected number of pivots in previous literature. $\sigma = \sqrt{d}/\sqrt{\ln(4/\sigma)}$ is a technical factor, factors are simplified. The lower bound of [Bor87] holds in the smoothed models as well.

研究主要结果

尽管如此，Huijberts 并不认为这是终点。

从 2001 年 Spielman 和滕尚华开始的方法展示了如何将运行时间从指数级降低到多项式级。下一个合理的目标是发明一种能够与约束数量成线性关系扩展的方法。「这是所有这项研究的北极星 (North Star)」她说。但这需要一种全新的策略。「我们短期内还不大可能实现它。」

虽然 Bach 和 Huijberts 的努力对他们领域的同行具有理论意义，但这项工作尚未产生任何直接的实际应用。



Eleon Bach 是这项新成果的作者之一。

尽管如此，它可能会平息一些人对于依赖当今基于单纯形法的软件可能产生的担忧。爱丁堡大学数学讲师、线性规划软件设计师 Julian Hall 表示：「虽然实践实验表明这些问题总是在多项式时间内解决」，但 Bach 和 Huijberts 的研究为这一直觉提供了更强的数学支持。「因此，现在更容易说服那些担心指数级复杂度的人了。」

参考链接: <https://www.quantamagazine.org/researchers-discover-the-optimal-way-to-optimize-20251013/>

- EOF -

推荐阅读 — 点击标题可跳转

1、

OpenAI 山姆官宣 ChatGPT 解锁成人内容。网友：涩涩是第一生产力

2、

蚂蚁凌晨 1 点发面试邮件，要求 6 小时内响应。网友：提前筛选加班牛马？

3、

芯片战升级！美国拟立法要求 AMD 英伟达“先内后外”，中国企业拿货更难了？

4、

突发！美国不仅对华加税，而且对“所有关键软件”出口管制