

多边形边的中点迭代与极限图形研究

PB23010356 张竞一

2025 年 11 月 9 日

1 问题导入

在数学中，我们经常能够发现对称美，尤其是在分形图和黄金比例的迭代图中。这种美不仅体现在自然界的许多结构中，也体现在数学的抽象构造中。曾发现过一个现象：如果对一个正 n 边形，取每条边的中点并依次连接这些中点，得到的仍然是一个正 n 边形。通过不断重复这一过程，我们会得到一个美丽的迭代图。

于是自然地想到，如果我们将这个过程推广到更一般的情况，不仅限于正 n 边形，而是考虑任意的 n 边形，甚至允许边交叉，情况会变的如何？如果我们从一个任意形状的 n 边形开始，取每条边的中点并连接这些中点，会得到什么样的图形呢？

首先，我们可以观察到，中点连接的过程本质上是一种平均化操作，它会在一定程度上“平滑”原始图形的不规则性。然而，这种对称性可能不再是正多边形那样的完美对称，而是某种更复杂的对称结构。

其次，如果原始的 n 边形有边交叉，那么取中点并连接后得到的图形可能会出现更多的交叉和重叠。这是因为中点连接的过程会将原始图形的交叉结构在一定程度上保留下来。这可能会导致得到的图形变得更加复杂，甚至可能出现一些意想不到的结构。

通过这些思考，我们可以看到，数学中的对称美不仅仅局限于正多边形这样的简单结构，而是可以扩展到更一般的图形和更复杂的迭代过程中。

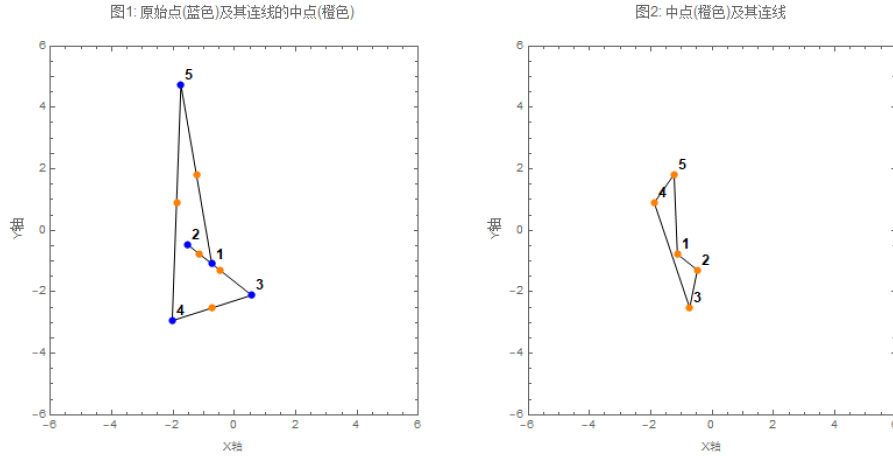
2 介绍

假设 x 和 y 是 n 维单位向量，其分量和为 0。令 $\mathcal{P}(x, y)$ 表示通过连接 $(x_1, y_1), \dots, (x_n, y_n), (x_1, y_1)$ 形成的多边形。如果从 $\mathcal{P}(x, y)$ 连接到其边中点形成新多边形，则得到“平均”多边形 $\hat{\mathcal{P}}$ ， $\hat{\mathcal{P}}$ 可视为 \mathcal{P} 的“平均版本”。图 1 展示了随机五边形的平均过程。

可以看出其结构似乎变得更加简单了，我们大胆猜测如果重复此平均过程，多边形可能会解开并在收敛于有序结构。我们的目标是找出此现象并且用数学语言精确表述。

2.1 中点连接操作的矩阵表述

因为我们需要对每个边执行相同的操作，中点连接操作可以自然地表述为向量运算。在标量层面，单个中点计算涉及两个坐标的平均值，如：

图 1: 五边形 \mathcal{P} (左) 和 “平均五边形” $\tilde{\mathcal{P}}$ (右)

$$\tilde{x}_1 = \frac{x_1 + x_2}{2}, \quad \tilde{y}_1 = \frac{y_1 + y_2}{2}.$$

在向量层面，整个操作可以描述为 x 向量与其上移向量的平均， y 向量也类似。我们用上面的五边形来举例：

$$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \\ \tilde{x}_4 \\ \tilde{x}_5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_4 \\ x_4 + x_5 \\ x_5 + x_1 \end{bmatrix}, \quad \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \\ \tilde{y}_4 \\ \tilde{y}_5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} y_1 + y_2 \\ y_2 + y_3 \\ y_3 + y_4 \\ y_4 + y_5 \\ y_5 + y_1 \end{bmatrix}.$$

这些变换是线性的，可用矩阵形式表示：

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \\ \tilde{x}_4 \\ \tilde{x}_5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x} \equiv M_5 \mathbf{x},$$

$$\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \\ \tilde{y}_4 \\ \tilde{y}_5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{y} \equiv M_5 \mathbf{y}.$$

其中矩阵 M_5 是 5 维情况下的平均矩阵。一般地，从 \mathcal{P} 到其平均 $\tilde{\mathcal{P}}$ 的变换为：

$$\tilde{\mathbf{x}} = M_n \mathbf{x}, \quad \tilde{\mathbf{y}} = M_n \mathbf{y},$$

其中 $n \times n$ 平均矩阵定义为：

$$(2.1) \quad M_n = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

向量 $\mathbf{x}, \mathbf{y}, \bar{\mathbf{x}}, \bar{\mathbf{y}}$ 称为顶点向量。 $\mathcal{P}(x, y)$ 用于表示由顶点向量 \mathbf{x} 和 \mathbf{y} 定义的多边形。在经过 k 次平均步骤后，我们得到多边形 $\mathcal{P}(M_n^k \mathbf{x}, M_n^k \mathbf{y})$ 。

我们之所以使用多边形平均过程的矩阵表述，是因为它使我们能够更简洁地描述重复平均过程，并且将问题重心转化为理解向量 $M_n^k \mathbf{x}, M_n^k \mathbf{y}$ 的变化，将问题简化为数学计算。

2.2 初次尝试迭代

用算法 1 描述这个迭代过程：

给出 n 维单位向量 $x^{(0)}$ 和 $y^{(0)}$ ，其分量和为零。

定义 $\mathcal{P}_0 = \mathcal{P}(x^{(0)}, y^{(0)})$ 。

用 $\mathcal{P}_{k-1} = \mathcal{P}(x^{(k-1)}, y^{(k-1)})$ 计算 $\mathcal{P}_k = \mathcal{P}(x^{(k)}, y^{(k)}) (k = 1, 2, \dots)$

其中 $x^{(k)} = M_n x^{(k-1)}, y^{(k)} = M_n y^{(k-1)}$

得到 \mathcal{P}_k 。

实验表明，多边形序列收敛到一个点，图 2 展示了一个典型的 $n=15$ 示例，显示了分别迭代 0, 5, 20, 50 次的“平均多边形”。

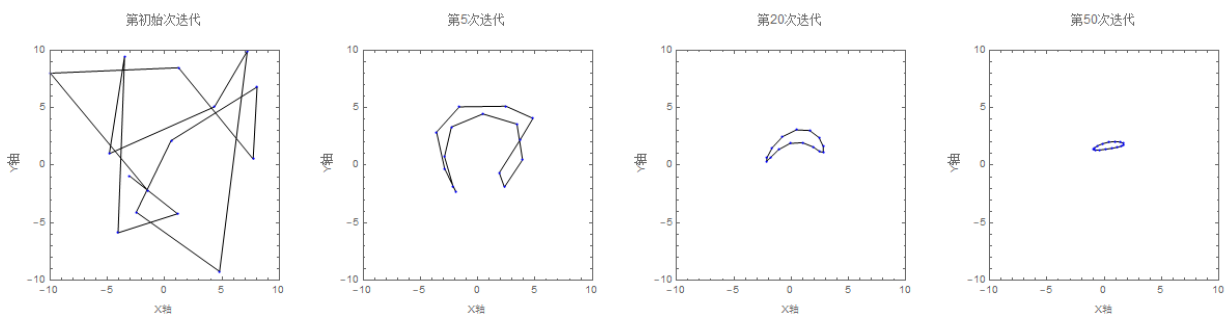


图 2: 多次迭代图形

下面我们严格解释算法 1 的极限行为：首先解释为什么极限为 \mathcal{P}_0 的重心，其中 \mathcal{P}_k 的重心 (\bar{x}, \bar{y}) 定义为：

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{e^T \mathbf{x}}{n}, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{e^T \mathbf{y}}{n},$$

其中 e 是 n 维全 1 向量。

$$e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

注意到 $e^T M_n = e^T$, 有:

$$\frac{e^T x^{(k)}}{n} = \frac{e^T M_n x^{(k-1)}}{n} = \frac{e^T x^{(k-1)}}{n}$$

$$\frac{e^T y^{(k)}}{n} = \frac{e^T M_n y^{(k-1)}}{n} = \frac{e^T y^{(k-1)}}{n}$$

这表明算法 1 生成的多边形序列具有相同的重心。如果序列 \mathcal{P}_k 收敛到一个点, 该点一定是 \mathcal{P}_0 的重心。

2.3 改进迭代过程

图 2 表明, 在算法 1 中多边形在从视野中消失之前会解开交叉边。为了更仔细地研究这一点, 我们引入归一化方法, 以防止 \mathcal{P}_k 坍缩为一个点:

1. 我们假设 $\mathcal{P}(x^{(0)}, y^{(0)})$ 的重心为 $(0, 0)$ 。
2. 每次更新后, 我们对顶点向量进行缩放, 使其模长为 1 (即多边形的分量平方和为 1)。

这些调整是为了多边形序列 $\{\mathcal{P}_k\}$ 保持合理的大小以免收敛至一个点, 并以原点为中心。用算法 2 描述这个迭代过程:

给出模长为 1 的 n 维向量 $x^{(0)}$ 和 $y^{(0)}$, 其分量和为零。

定义 $\mathcal{P}_0 = \mathcal{P}(x^{(0)}, y^{(0)})$ 。

使用 $\mathcal{P}_{k-1} = \mathcal{P}(x^{(k-1)}, y^{(k-1)})$ 计算 $\mathcal{P}_k = \mathcal{P}(x^{(k)}, y^{(k)})$ ($k = 1, 2, \dots$)

$\xi = M_n x^{(k-1)}, x^{(k)} = \xi / \|\xi\|_2$

$\eta = M_n y^{(k-1)}, y^{(k)} = \eta / \|\eta\|_2$

得到 \mathcal{P}_k 。

根据上一节的说明, 我们知道每个 \mathcal{P}_k 的重心都是 $(0, 0)$, 于是只需要缩放而无需再进行中心的调整, 这样的缩放对整体的图形没有影响。

算法 2 相当于幂方法的应用, 将其应用于 $n \times n$ 矩阵 A 和模长为 1 的起始向量 $w^{(0)}$ 时, 幂方法涉及重复的矩阵-向量乘法和缩放:

$$w^{(k)} = Aw^{(k-1)} / \|Aw^{(k-1)}\|_2, \quad k = 1, 2, \dots$$

需要注意的是缩放不一定在每次迭代时都进行, 理论上可以在终止时进行 (因为形状一直不变), 这样可以减少迭代的时间复杂度。而由于本实验中未涉及很高边数与很高迭代次数的情况并且要观察在迭代过程中的图形情况, 使用的是每一步都进行归一化的方法。

下面我们简要介绍其工作原理，假设 A 有一组完整的 n 个独立特征向量 z_1, \dots, z_n 和特征值 $\lambda_1, \dots, \lambda_n$ ，满足 $Az_i = \lambda_i z_i$ 。如果

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

且

$$w^{(0)} = \gamma_1 z_1 + \gamma_2 z_2 + \dots + \gamma_n z_n,$$

那么 $w^{(k)}$ 是以下向量方向上的单位向量：

$$(2.2) \quad A^k w^{(0)} = \lambda_1^k \left(\gamma_1 z_1 + \sum_{i=2}^n \gamma_i \left(\frac{\lambda_i}{\lambda_1} \right)^k z_i \right).$$

可以看出只要 $\gamma_1 \neq 0$ ，该向量在 z_1 方向上的分量相比之下会越来越多。特征向量 z_1 被称为**主特征向量**，因为它与**主特征值** λ_1 相关联。

回到我们的问题：矩阵 M_n 的主特征值是 1，对应的主特征向量是 e 。理由如下：注意到 $Me = e$ ，因此 1 是具有特征向量 e 的特征值。 M_n 的任何特征值的大小都不会超过其 1-范数（这是由于范数有次乘性），由矩阵 1-范数的定义

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

可以得到 $\|M_n\|_1 = 1$ 。因此，1 是主特征值。

这些观察阐明了算法 1 中的多边形为何会收敛到一个点：两个顶点向量序列在 e 方向上的分量都越来越多，这相当于说 $x^{(k)}$ 和 $y^{(k)}$ 的分量越来越趋于一致。由于迭代保持重心不变，极限值分别是初始顶点向量 $x^{(0)}$ 和 $y^{(0)}$ 的重心。

2.4 从实验到猜想

如果我们用算法 2 进行实验并显示图像 \mathcal{P}_k ，会发现以下现象：

猜想 1: 无论初始 \mathcal{P}_0 多么随机， \mathcal{P}_k 的边最终都会不相交，并且在极限情况下，顶点都会围绕一个与坐标轴成 45 度倾斜的椭圆排列。

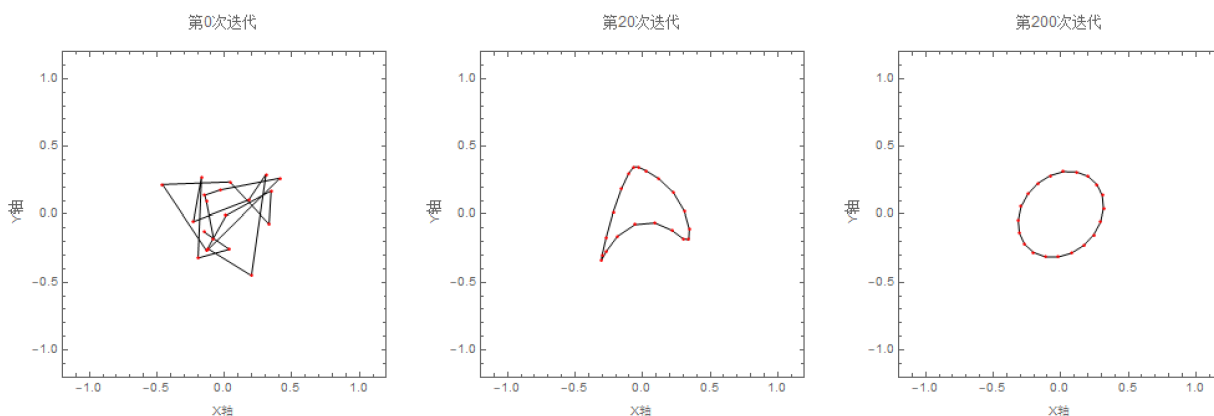


图 3: $n=20$ 时初始图形在算法 2 下不同次数迭代

图 3 描绘了一个典型的 $n = 20$ 的例子。我们的目标是解释从交叉到有序的转变。由于其与幂方法联系密切，解释可以围绕平均矩阵 M_n 的特征系统性质展开。

算法 2 中的顶点向量迭代 $x^{(k)}$ 和 $y^{(k)}$ 重心为零, 即 $\frac{1}{n}\sum_{i=1}^n x_i^{(k)} = 0$ 与 $\frac{1}{n}\sum_{i=1}^n y_i^{(k)} = 0$, 这可以写为 $e^T x^{(k)} = 0$ 与 $e^T y^{(k)} = 0$, 因此与 M_n 的主特征向量 e 正交。于是在式 (2.2) 的符号中, $x^{(k)}$ 和 $y^{(k)}$ 的 γ_1 项都不存在。因此, 多边形在所有方向上均匀收缩, 从而不会收敛到单个点。相反, 这些向量收敛到一个二维不变子空间, 我们在下一章会提到。

3 子空间 \mathcal{D}_2

要分析算法 2 中顶点向量 $x^{(k)}$ 和 $y^{(k)}$ 的行为就需要对平均矩阵 M_n 的特征系统进行透彻分析。

3.1 上移矩阵

定义 $n \times n$ 上移矩阵 S_n 为:

$$S_n = \begin{pmatrix} e_n & e_1 & e_2 & \cdots & e_{n-1} \end{pmatrix}$$

其中 e_k 是 $n \times n$ 单位矩阵的第 k 列。例如:

$$S_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

从 (2.1) 可知, 平均矩阵 M_n 可表示为:

$$M_n = \frac{1}{2}(I_n + S_n)$$

S_n 的特征向量是已知的, 为:

$$v_j = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \omega_j \\ \omega_j^2 \\ \vdots \\ \omega_j^{n-1} \end{pmatrix}, \quad j = 0, 1, \dots, n-1$$

$$\omega_j = \cos\left(\frac{2\pi j}{n}\right) + i \cdot \sin\left(\frac{2\pi j}{n}\right), \quad 0, 1, \dots, n-1.$$

其中 ω_j 是 n 次单位根。

显然 v_j 模长为 1 且满足 $S_n v_j = \omega_j v_j$ 。此外, v_0, \dots, v_{n-1} 两两正交。简要证明这一点: 我们观察到

$$v_i^H S_n v_j = v_i^H (S_n v_j) = \omega_j v_i^H v_j.$$

另一方面，由于 $S_n^H = S_n^T = S_n^{-1}$ ，我们还有：

$$v_i^H S_n v_j = (v_i^H S_n) v_j = (S_n^{-1} v_i)^H v_j = (v_i / \omega_i)^H v_j = \omega_i v_i^H v_j.$$

因此，当 $i \neq j$ 时， $v_i^H v_j = 0$ 。上标“H”表示 Hermite 共轭转置。

3.2 M_n 的特征值和特征向量

平均矩阵 $M_n = (I_n + S_n)/2$ 与 S_n 具有相同的特征向量。其特征值 $\lambda_1, \dots, \lambda_n$ 由下式给出：

$$\lambda_1 = (1 + \omega_0)/2 = 1,$$

$$\lambda_2 = (1 + \omega_1)/2 = (1 + \cos(2\pi/n) + i \cdot \sin(2\pi/n))/2,$$

$$\lambda_3 = (1 + \omega_{n-1})/2 = \bar{\lambda}_2,$$

$$\lambda_4 = (1 + \omega_2)/2 = (1 + \cos(4\pi/n) + i \cdot \sin(4\pi/n))/2,$$

$$\lambda_5 = (1 + \omega_{n-2})/2 = \bar{\lambda}_4,$$

⋮

$$\lambda_n = (1 + \omega_m)/2,$$

其中 $m = \lfloor \frac{n}{2} \rfloor$ 。我们选择按这种方式排序特征值，因为这样可以将共轭特征值对组合在一起。如图 4 所示， λ_i 位于复平面上的一个圆上，该圆以 $(0.5, 0)$ 为圆心，直径为 1。

特征值 λ_1 到 λ_7 ：

$$\lambda_1 = 1.00000$$

$$\lambda_2 = 0.811745 + 0.390916 I$$

$$\lambda_3 = 0.811745 - 0.390916 I$$

$$\lambda_4 = 0.388740 + 0.487464 I$$

$$\lambda_5 = 0.388740 - 0.487464 I$$

$$\lambda_6 = 0.049516 + 0.216942 I$$

$$\lambda_7 = 0.049516 - 0.216942 I$$

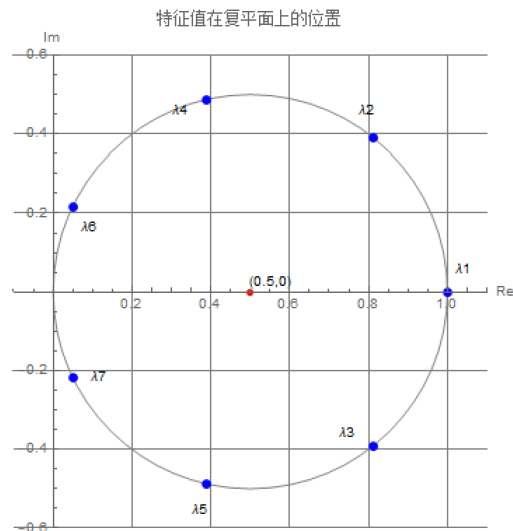


图 4: $n=7$ 时 M_7 的特征值

此外，

$$(3.1) \quad |\lambda_1| > |\lambda_2| = |\lambda_3| > \cdots > |\lambda_{n-1}| \geq |\lambda_n|.$$

令 z_1, \dots, z_n 是特征向量 v_0, \dots, v_{n-1} 的重新排序, 使得

$$M_n z_k = \lambda_k z_k, \quad k = 1, 2, \dots, n.$$

3.3 阻尼因子

设 $w \in \mathbf{C}^n$ 为一个单位向量, 其分量之和为 0。由于它具有重心 $\mathbf{0}$, 因此有特征向量展开式:

$$w = \gamma_2 z_2 + \cdots + \gamma_n z_n,$$

其中 $|\gamma_2|^2 + \cdots + |\gamma_n|^2 = 1$ 。这是因为 $\{z_1, \dots, z_n\}$ 是一个正交基, 且 z_1 是 e 的倍数, 因此具有重心零的向量 w 满足 $e^T w = 0$ 。

由此可得:

$$(3.2) \quad M_n^k w = |\lambda_2|^k \left(\gamma_2 \left(\frac{\lambda_2}{|\lambda_2|} \right)^k z_2 + \gamma_3 \left(\frac{\lambda_3}{|\lambda_2|} \right)^k z_3 + \sum_{j=4}^n \gamma_j \left(\frac{\lambda_j}{|\lambda_2|} \right)^k z_j \right).$$

从假设 (3.1) 可以看出, 向量 $M_n^k w$ 在 z_2 和 z_3 方向上的分量不断增加, 而在 z_4, \dots, z_n 方向上的分量明显衰减。(这里的衰减指的是相对衰减, 非绝对衰减) 衰减的速率取决于商:

$$(3.3) \quad \rho_n \equiv \left| \frac{\lambda_4}{\lambda_2} \right| = \max \left\{ \left| \frac{\lambda_4}{\lambda_2} \right|, \dots, \left| \frac{\lambda_n}{\lambda_2} \right| \right\}.$$

由于:

$$\left| \frac{\lambda_4}{\lambda_2} \right|^2 = \frac{|1 + \cos(4\pi/n) + i \sin(4\pi/n)|^2}{|1 + \cos(2\pi/n) + i \sin(2\pi/n)|^2} = \frac{1 + \cos(4\pi/n)}{1 + \cos(2\pi/n)} = \frac{\cos^2(2\pi/n)}{\cos^2(\pi/n)},$$

因此可得:

$$\rho_n = \left| \frac{\lambda_4}{\lambda_2} \right| = \frac{\cos^2(\pi/n) - \sin^2(\pi/n)}{\cos(\pi/n)} = \cos(\pi/n) (1 - \tan^2(\pi/n)).$$

表 3.1 阻尼因子 ρ_n 的数值随 n 的变化

n	ρ_n
5	0.38197
10	0.85065
20	0.96291
50	0.99408
100	0.99852
500	0.99994

表 3.1 给出了 ρ_n 的一些取值。通过操作近似：

$$\cos(\pi/n) = 1 - \frac{1}{2} \left(\frac{\pi}{n}\right)^2 + O\left(\frac{1}{n^4}\right)$$

以及

$$\cos(2\pi/n) = 1 - \frac{1}{2} \left(\frac{2\pi}{n}\right)^2 + O\left(\frac{1}{n^4}\right)$$

可以证明：

$$\rho_n = 1 - \frac{3}{2} \left(\frac{\pi}{n}\right)^2 + O\left(\frac{1}{n^4}\right).$$

因此，随着 n 的增大，阻尼过程变得越来越慢，这一点在表 3.1 中可以看出。

3.4 实复数运算

式 (3.2) 左侧的向量 $M_n^k w$ 是实数，但右侧的线性组合涉及复数特征向量。实矩阵 M_n 的特征值和特征向量成共轭对出现，复数部分在求和中相互抵消。特别是无阻尼向量

$$\tilde{w}^{(k)} = \gamma_2 \left(\frac{\lambda_2}{|\lambda_2|}\right)^k z_2 + \gamma_3 \left(\frac{\lambda_3}{|\lambda_3|}\right)^k z_3$$

在 (3.2) 中是实数，因为 $\gamma_3 = \bar{\gamma}_2$ ， $\lambda_3 = \bar{\lambda}_2$ 且 $z_3 = \bar{z}_2$ 。此外，它属于由下式定义的实子空间 \mathcal{D}_2 ：

$$(3.4) \quad \mathcal{D}_2 = \text{span}\{\Re(z_2), \Im(z_2)\} = \text{span}\{\Re(z_3), \Im(z_3)\}.$$

这个子空间是矩阵 M_n 的不变子空间。为了证明这一点，我们比较矩阵方程 $M_n z_2 = \lambda_2 z_2$ 的实部和虚部：

$$M_n(\Re(z_2)) + i \cdot M_n(\Im(z_2)) = (\cos(2\pi/n) + i \cdot \sin(2\pi/n))(\Re(z_2) + i \cdot \Im(z_2)).$$

显然， $M_n \cdot \Re(z_2) \in \mathcal{D}_2$ 且 $M_n \cdot \Im(z_2) \in \mathcal{D}_2$ ，证毕。

3.5 收敛到 \mathcal{D}_2

定理 3.1 假设

$$w^{(0)} = \sum_{j=2}^n \gamma_j z_j$$

是一个实的模长为 1 的向量，且

$$\mu = \sqrt{|\gamma_2|^2 + |\gamma_3|^2} > 0.$$

如果 P_n 是正交投影到 $\text{span}\{z_2, z_3\}^\perp$ 上的算子, 且 $w^{(k)}$ 是模长为 1 的向量, 方向为 $M_n^k w^{(0)}$, 则

$$\|P_n w^{(k)}\|_2 \leq \rho_n^k \frac{\sqrt{1-\mu^2}}{\mu}.$$

证明: 由于特征向量集合 $\{z_2, \dots, z_n\}$ 是正交的, 从 (3.2) 可得

$$w^{(k)} = \left(\sum_{j=2}^n \gamma_j \lambda_j^k z_j \right) / \sqrt{\sum_{j=2}^n |\gamma_j|^2 |\lambda_j|^{2k}}.$$

投影 $P_n w^{(k)}$ 的形式与分子中没有 z_2 和 z_3 项的形式相同。因此,

$$\|P_n w^{(k)}\|_2^2 = \frac{\sum_{j=4}^n |\gamma_j|^2 |\lambda_j|^{2k}}{\sum_{j=2}^n |\gamma_j|^2 |\lambda_j|^{2k}} \leq \frac{|\lambda_4|^{2k} \sum_{j=4}^n |\gamma_j|^2}{|\lambda_2|^{2k} (|\gamma_2|^2 + |\gamma_3|^2)} = \rho_n^{2k} \frac{1-\mu^2}{\mu^2},$$

其中我们使用了 (3.3), $|\lambda_2| = |\lambda_3|$, 且 $|\gamma_2|^2 + \dots + |\gamma_n|^2 = 1$ 。□

$\|P_n w^{(k)}\|_2$ 可以看作是从 $w^{(k)}$ 到 \mathcal{D}_2 的距离。当 $n=3$ 时, 子空间 \mathcal{D}_2 是一个平面, $w^{(k)}$ 位于平面外, 而 $\|P_n w^{(k)}\|_2$ 是垂线的长度。

我们提到算法 2 中的顶点向量重心为零。然而, 由于舍入误差不断累积, 因此有时需要将顶点向量重新居中, 即:

$$x^{(k)} - \frac{(e^T x^{(k)})}{n} e \rightarrow x^{(k)}$$

观察到当 k 趋于无穷时, \mathcal{P}_k 中的顶点成对共线。如果初始顶点向量仅与 $\Re(z_2)$ 或 $\Im(z_2)$ 中的一个正交, 则多边形收敛到一条 45 度线。

如果仅有一个初始顶点向量与 \mathcal{D}_2 正交, 则多边形收敛到一条包含两个椭圆状区域的曲线。最后, 如果一个顶点向量最初与 $\Re(z_2)$ 正交, 另一个顶点向量与 $\Im(z_2)$ 正交, 则顶点收敛到一个圆。

在本文的其余部分, 我们将假设没有顶点向量与子空间 \mathcal{D}_2 正交。

3.6 子空间 \mathcal{D}_2 的实正交基

特征向量 z_2 (及其共轭 z_3) 的实部和虚部具有高度的结构化特征。如果

$$(3.5) \quad \tau = \begin{bmatrix} 0 \\ 2\pi/n \\ 4\pi/n \\ \vdots \\ 2(n-1)\pi/n \end{bmatrix},$$

那么用向量表示, $z_2 = (\cos(\tau) + i \cdot \sin(\tau)) / \sqrt{n}$ 。利用三角恒等式, 很容易证明:

$$\cos(\tau)^T \cos(\tau) = \sum_{j=1}^n \cos(\tau_j)^2 = \sum_{j=1}^n \frac{1 + \cos(2\tau_j)}{2} = \frac{n}{2},$$

$$\sin(\tau)^T \sin(\tau) = \sum_{j=1}^n \sin(\tau_j)^2 = \sum_{j=1}^n \frac{1 - \cos(2\tau_j)}{2} = \frac{n}{2},$$

以及

$$\sin(\tau)^T \cos(\tau) = \sum_{j=1}^n \sin(\tau_j) \cos(\tau_j) = \sum_{j=1}^n \frac{\sin(2\tau_j)}{2} = 0.$$

从 (3.4) 可得, 向量

$$(3.6) \quad c = \sqrt{\frac{2}{n}} \begin{bmatrix} \cos(\tau_1) \\ \cos(\tau_2) \\ \vdots \\ \cos(\tau_n) \end{bmatrix}, \quad s = \sqrt{\frac{2}{n}} \begin{bmatrix} \sin(\tau_1) \\ \sin(\tau_2) \\ \vdots \\ \sin(\tau_n) \end{bmatrix}$$

构成了子空间 \mathcal{D}_2 的实正交基。

4 顶点向量

因为顶点向量与 z_1 (\mathbf{e} 的倍数) 正交, 并且我们假设在算法 2 中它们不与 \mathcal{D}_2 正交, 因此初始向量 $x^{(0)}$ 和 $y^{(0)}$ 可以表示为正交向量集合 $\{c, s, z_4, \dots, z_n\}$ 的线性组合:

$$(4.1) \quad x^{(0)} = \alpha_1 c + \alpha_2 s + \text{span}\{z_4, \dots, z_n\} \text{ 中的向量,}$$

$$y^{(0)} = \beta_1 c + \beta_2 s + \text{span}\{z_4, \dots, z_n\} \text{ 中的向量.}$$

根据定理 3.1, 对于充分大的 k , 有:

$$(4.2) \quad x^{(k)} = u^{(k)} + O(\rho_n^k),$$

$$y^{(k)} = v^{(k)} + O(\rho_n^k),$$

其中 $u^{(k)}$ 和 $v^{(k)}$ 是单位向量:

$$(4.3) \quad u^{(k)} = \frac{\alpha_1 M_n^k c + \alpha_2 M_n^k s}{\|\alpha_1 M_n^k c + \alpha_2 M_n^k s\|_2},$$

$$v^{(k)} = \frac{\beta_1 M_n^k c + \beta_2 M_n^k s}{\|\beta_1 M_n^k c + \beta_2 M_n^k s\|_2}.$$

注意到这些向量在子空间 \mathcal{D}_2 中, 因为极限行为 $\{\mathcal{P}(u^{(k)}, v^{(k)})\}$ 与多边形序列 $\{\mathcal{P}(x^{(k)}, y^{(k)})\}$ 的极限行为一致。我们的计划是研究多边形序列 $\{\mathcal{P}(u^{(k)}, v^{(k)})\}$ 。

换句话说，与 \mathcal{P}_k 相关联的顶点向量是与 \mathcal{P}_{k-2} 相关联的顶点向量的上移。表达如下：

$$\mathcal{P}_{2k} = \mathcal{P}(S_n^k u^{(0)}, S_n^k v^{(0)}) = \mathcal{P}(u^{(0)}, v^{(0)}),$$

$$\mathcal{P}_{2k+1} = \mathcal{P}(S_n^k u^{(1)}, S_n^k v^{(1)}) = \mathcal{P}(u^{(1)}, v^{(1)}).$$

我们会在下一章对其进行证明。

4.2 顶点向量 c 和 s 的平均

我们需要向量 $M_n^k c$ 和 $M_n^k s$ ，因为它们显然来自算法 3 中的第 n 维向量 $c(\tau + \Delta)$ 和 $s(\tau + \Delta)$ ，其中

$$c(\tau + \Delta) = \sqrt{\frac{2}{n}} \begin{bmatrix} \cos(\tau_1 + \Delta) \\ \vdots \\ \cos(\tau_n + \Delta) \end{bmatrix},$$

和

$$s(\tau + \Delta) = \sqrt{\frac{2}{n}} \begin{bmatrix} \sin(\tau_1 + \Delta) \\ \vdots \\ \sin(\tau_n + \Delta) \end{bmatrix},$$

其中 $\tau \in \mathbb{R}^n$ 且 $\Delta \in \mathbb{R}$ 由 (3.5) 给出。记号 $\tau + \Delta$ 表示向量，其每个分量都是向量 τ 对应分量与标量 Δ 的和。

注意到，若 $\Delta = 0$ ，则 $c(\tau + \Delta) = c$ 且 $s(\tau + \Delta) = s$ ，其中 c 和 s 由 (3.6) 定义。此外， $c(\tau + \Delta)$ 和 $s(\tau + \Delta)$ 模长都是 1，并且彼此正交，于是二者可以视为子空间 \mathcal{D}_2 中对标准正交基 c 和 s 进行旋转角度 Δ 得到的。因此，有

$$c(\tau + \Delta) = \cos(\Delta)c - \sin(\Delta)s,$$

$$s(\tau + \Delta) = \sin(\Delta)c + \cos(\Delta)s,$$

定义 $\tau_{n+1} = \tau_1 = 0$ 。如果 $\mu_i = \tau_i + \Delta + \pi/n$ 对 $i = 1, n$ 成立，则利用三角恒等式

$$\cos(\alpha + \beta) + \cos(\alpha - \beta) = 2 \cos(\alpha) \cos(\beta),$$

$$\sin(\alpha + \beta) + \sin(\alpha - \beta) = 2 \sin(\alpha) \cos(\beta),$$

我们有

$$(M_n c(\tau + \Delta))_i = \frac{1}{2} \sqrt{\frac{2}{n}} (\cos(\tau_i + \Delta) + \cos(\tau_{i+1} + \Delta))$$

$$\begin{aligned}
&= \frac{1}{2} \sqrt{\frac{2}{n}} (\cos(\mu_i - \pi/n) + \cos(\mu_i + \pi/n)) \\
&= \sqrt{\frac{2}{n}} \cos\left(\frac{\pi}{n}\right) \cos(\mu_i),
\end{aligned}$$

以及

$$\begin{aligned}
(M_n s(\tau + \Delta))_i &= \frac{1}{2} \sqrt{\frac{2}{n}} (\sin(\tau_i + \Delta) + \sin(\tau_{i+1} + \Delta)) \\
&= \frac{1}{2} \sqrt{\frac{2}{n}} (\sin(\mu_i - \pi/n) + \sin(\mu_i + \pi/n)) \\
&= \sqrt{\frac{2}{n}} \cos\left(\frac{\pi}{n}\right) \sin(\mu_i).
\end{aligned}$$

因此,

$$\begin{aligned}
M_n c(\tau + \Delta) &= \cos\left(\frac{\pi}{n}\right) c\left(\tau + \Delta + \frac{\pi}{n}\right), \\
M_n s(\tau + \Delta) &= \cos\left(\frac{\pi}{n}\right) s\left(\tau + \Delta + \frac{\pi}{n}\right),
\end{aligned}$$

因此通过归纳法, 我们有

$$\begin{aligned}
M_n^k c &= \cos^k\left(\frac{\pi}{n}\right) c\left(\tau + \frac{k\pi}{n}\right), \\
M_n^k s &= \cos^k\left(\frac{\pi}{n}\right) s\left(\tau + \frac{k\pi}{n}\right).
\end{aligned}$$

从算法 3 我们看到 $u^{(k)}$ 和 $v^{(k)}$ 是方向为

$$\begin{aligned}
M_n^k u^{(0)} &= \cos^k\left(\frac{\pi}{n}\right) \left(\cos(\theta_u) c\left(\tau + \frac{k\pi}{n}\right) + \sin(\theta_u) s\left(\tau + \frac{k\pi}{n}\right) \right), \\
M_n^k v^{(0)} &= \cos^k\left(\frac{\pi}{n}\right) \left(\cos(\theta_v) c\left(\tau + \frac{k\pi}{n}\right) + \sin(\theta_v) s\left(\tau + \frac{k\pi}{n}\right) \right).
\end{aligned}$$

的单位向量。

由于 $\{c, s\}$ 是一个正交集, 我们得出结论:

$$\begin{aligned}
u^{(k)} &= \cos(\theta_u) c\left(\tau + \frac{k\pi}{n}\right) + \sin(\theta_u) s\left(\tau + \frac{k\pi}{n}\right), \\
v^{(k)} &= \cos(\theta_v) c\left(\tau + \frac{k\pi}{n}\right) + \sin(\theta_v) s\left(\tau + \frac{k\pi}{n}\right).
\end{aligned}$$

根据 τ 、 c 和 s 的定义, 可以得出:

$$(4.4) \quad u^{(k+2)} = S_n u^{(k)}, \quad v^{(k+2)} = S_n v^{(k)}.$$

5 极限椭圆

我们证明对于所有 i 和 k , 算法 3 中的点 $(u_i^{(k)}, v_i^{(k)})$ 都位于同一个椭圆 \mathcal{E} 上, 并且该椭圆 \mathcal{E} 具有 45 度倾角。我们称这个极限椭圆为 \mathcal{D}_2 椭圆, 因为它依赖于顶点向量在不变子空间中的投影。

5.1 倾斜椭圆

当 t 遍历所有实数值时, 点集 $(u(t), v(t))$ 由下式给出:

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

这定义了一个以 $(0,0)$ 为中心、倾角为 ϕ 、半轴长分别为 $|\sigma_1|$ 和 $|\sigma_2|$ 的椭圆。它是通过对

椭圆

$$\left(\frac{u}{\sigma_1}\right)^2 + \left(\frac{v}{\sigma_2}\right)^2 = 1$$

逆时针旋转 ϕ 弧度得到的, 如图 7 所示。

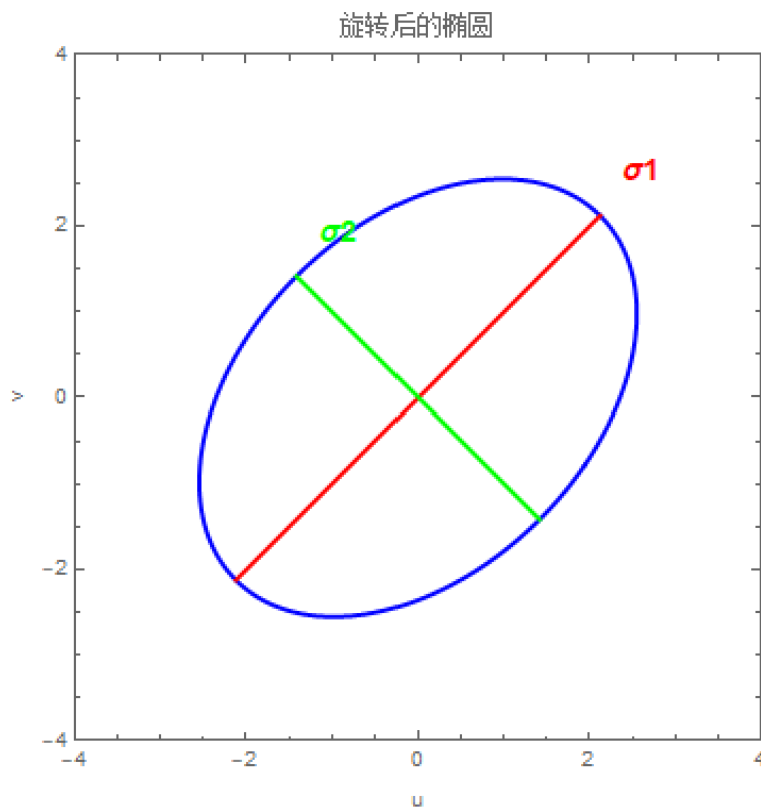


图 6: 倾斜椭圆

由下式定义的点集 $(u(t), v(t))$:

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \equiv A \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

也是一个以 $(0,0)$ 为中心的椭圆。其半轴和倾角由矩阵 A 的奇异值分解 (SVD) 指定。实数 $n \times n$ 矩阵 A 的 SVD 保证我们可以找到正交矩阵 U 和 V , 使得 $A = U \text{diag}(\sigma_1, \dots, \sigma_n) V^T$ 。通常, 我们假设奇异值 σ_i 按非负顺序排列。然而, 在本应用中, 我们不需要这种归一化。特别地, 对于 2×2 的情况, 我们有

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix}^T.$$

因此, 椭圆 (4.2) 的半轴为 $|\sigma_1|$ 和 $|\sigma_2|$, 倾角为 ϕ 。注意, 指定椭圆倾角的 U 矩阵就是上面的旋转矩阵。

5.2 \mathcal{D}_2 椭圆

如果 $t_i = \tau_i + k\pi/n$, 对于 $i = 1, 2, \dots, n$, 并且顶点向量 $u^{(k)}$ 和 $v^{(k)}$ 由算法 3 生成, 则有

$$\begin{bmatrix} u_i^{(k)} \\ v_i^{(k)} \end{bmatrix} = \sqrt{\frac{2}{n}} \begin{bmatrix} \cos(\theta_u) & \sin(\theta_u) \\ \cos(\theta_v) & \sin(\theta_v) \end{bmatrix} \begin{bmatrix} \cos(t_i) \\ \sin(t_i) \end{bmatrix}.$$

这表明多边形 $\mathcal{P}(u^{(k)}, v^{(k)})$ 的顶点位于由点集 $(u(t), v(t))$ 定义的椭圆 \mathcal{E} 上, 其中

$$(5.1) \quad \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \sqrt{\frac{2}{n}} \begin{bmatrix} \cos(\theta_u) & \sin(\theta_u) \\ \cos(\theta_v) & \sin(\theta_v) \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \quad -\infty < t < \infty.$$

为了指定椭圆 \mathcal{E} 的倾角和半轴, 我们需要对 2×2 的变换矩阵进行 SVD。

定理 5.1 若 θ_u 和 θ_v 是实数, 且

$$A = \begin{bmatrix} \cos(\theta_u) & \sin(\theta_u) \\ \cos(\theta_v) & \sin(\theta_v) \end{bmatrix},$$

则有 $U^T A V = \Sigma$, 其中

$$U = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix}, \quad V = \begin{bmatrix} \cos(a) & -\sin(a) \\ \sin(a) & \cos(a) \end{bmatrix},$$

以及

$$\Sigma = \begin{bmatrix} \sqrt{2} \cos(b) & 0 \\ 0 & \sqrt{2} \sin(b) \end{bmatrix}.$$

其中

$$a = \frac{\theta_v + \theta_u}{2}, \quad b = \frac{\theta_v - \theta_u}{2}.$$

证明: 使用三角恒等式, 我们有

$$U^T A = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos(\theta_v) + \cos(\theta_u) & \sin(\theta_v) + \sin(\theta_u) \\ \cos(\theta_v) - \cos(\theta_u) & \sin(\theta_v) - \sin(\theta_u) \end{bmatrix}$$

$$= \sqrt{2} \begin{bmatrix} \cos(b) \cos(a) & \cos(b) \sin(a) \\ -\sin(b) \sin(a) & \sin(b) \cos(a) \end{bmatrix}.$$

这正是 ΣV^T 。

利用这一结果，可以得出 \mathcal{D}_2 椭圆 (5.1) 具有 45 度倾角，半轴为 $|\sigma_1|$ 和 $|\sigma_2|$ ，其中

$$(5.2) \quad \sigma_1 = \frac{2}{\sqrt{n}} \cos\left(\frac{\theta_v - \theta_u}{2}\right), \quad \sigma_2 = \frac{2}{\sqrt{n}} \sin\left(\frac{\theta_v - \theta_u}{2}\right).$$

这验证了猜想 1。

5.3 通过 $x^{(0)}$ 和 $y^{(0)}$ 进行规范化

现在我们通过初始顶点向量 $x^{(0)}$ 和 $y^{(0)}$ 来指定与算法 2 相关联的极限椭圆。根据 (4.1) 和正交性 $\{c, s, z_4, \dots, z_n\}$ ，我们得知

$$x^{(0)} = (c^T x^{(0)})c + (s^T x^{(0)})s + \text{span}\{z_4, \dots, z_n\} \text{ 中的向量,}$$

$$y^{(0)} = (c^T y^{(0)})c + (s^T y^{(0)})s + \text{span}\{z_4, \dots, z_n\} \text{ 中的向量.}$$

参考 (4.2) 和 (4.3)，算法 2 的极限椭圆与算法 3 的极限椭圆相同，其起始单位向量为

$$u^{(0)} = \cos(\theta_u)c + \sin(\theta_u)s,$$

$$v^{(0)} = \cos(\theta_v)c + \sin(\theta_v)s,$$

其中

$$\cos(\theta_u) = \frac{c^T x^{(0)}}{\sqrt{(c^T x^{(0)})^2 + (s^T x^{(0)})^2}}, \quad \sin(\theta_u) = \frac{s^T x^{(0)}}{\sqrt{(c^T x^{(0)})^2 + (s^T x^{(0)})^2}},$$

$$\cos(\theta_v) = \frac{c^T y^{(0)}}{\sqrt{(c^T y^{(0)})^2 + (s^T y^{(0)})^2}}, \quad \sin(\theta_v) = \frac{s^T y^{(0)}}{\sqrt{(c^T y^{(0)})^2 + (s^T y^{(0)})^2}}.$$

这四个数与 (5.2) 一起完全指定了与算法 2 相关联的 45 度倾斜椭圆。

6 理论证明过程与结果总结

本文的核心目标是寻找多边形边的中点迭代后图形的规律并确定该图形的性质。研究通过将几何迭代问题转化为线性代数问题，利用平均矩阵 M_n 进行分析。 M_n 的特征值为 $\lambda_j = \cos(\pi j/n)$ ，其中特别关注特征值 1 对应的特征向量 c 和 s ，它们张成一个重要的不变子空间 D_2 。

在迭代算法方面，算法 2 通过归一化迭代在每次迭代后对顶点向量进行归一化，确保其保持单位长度，利用幂方法的原理使顶点向量逐渐收敛到 D_2 。而算法 3 直接使用特征向量 c 和 s 的线性组合初始化顶点向量，使得迭代过程更快地收敛到极限图形。

研究证明了迭代的极限图形是一个与坐标轴成 45 度倾斜的椭圆，并给出了椭圆半轴长的精确计算公式： $|\sigma_1|$ 和 $|\sigma_2|$ ，其中

$$\sigma_1 = \frac{2}{\sqrt{n}} \cos\left(\frac{\theta_v - \theta_u}{2}\right), \quad \sigma_2 = \frac{2}{\sqrt{n}} \sin\left(\frac{\theta_v - \theta_u}{2}\right).$$

参数 θ_u 和 θ_v 可通过初始顶点向量计算得出：

$$\cos(\theta_u) = \frac{c^T \cdot x^{(0)}}{\sqrt{(c^T \cdot x^{(0)})^2 + (s^T \cdot x^{(0)})^2}}, \quad \sin(\theta_u) = \frac{s^T \cdot x^{(0)}}{\sqrt{(c^T \cdot x^{(0)})^2 + (s^T \cdot x^{(0)})^2}},$$

$$\cos(\theta_v) = \frac{c^T \cdot y^{(0)}}{\sqrt{(c^T \cdot y^{(0)})^2 + (s^T \cdot y^{(0)})^2}}, \quad \sin(\theta_v) = \frac{s^T \cdot y^{(0)}}{\sqrt{(c^T \cdot y^{(0)})^2 + (s^T \cdot y^{(0)})^2}}.$$

通过数学推导和实验观察，研究验证了两个关键猜想：猜想 1：无论初始多边形如何，迭代后的多边形边最终不相交，并且极限情况下顶点围绕一个 45 度倾斜的椭圆排列；猜想 2：在算法 3 中，偶数次和奇数次迭代得到的多边形分别完全一致。

研究成功地将几何迭代问题转化为线性代数问题，利用矩阵特征值和特征向量进行分析，证明了多边形中点迭代的极限图形是一个 45 度倾斜的椭圆。这一研究不仅解决了一个有趣的几何问题，还展示了线性代数在分析复杂几何系统行为中的强大应用能力。