Содержание

1 Введение

В данной статье приведены две модификации метода Лагерра. Они определяют методы одновременной аппроксимации всех нулей данного многочлена. Метод Лагерра для аппроксимации нулей многочлена считается одним из стандартов, по которому следует оценивать другие алгоритмы нахождения нуля. Он характеризуется отличным начальным поведением, глобальной сходимостью для многочлены только с вещественными корнями, локальная кубическая сходимость к простому корню и локальная линейная сходимость к кратному корню (см. Ostrowski [1]).
В первом разделе мы рассмотрим вывод метода Лагерра, найденный в Parlett [2].

Во втором разделе мы описываем две модификации для одновременной аппроксимации всех нулей данного многочлена. Одна из этих модификаций особенно подходит для реализации на компьютерах с параллельной обработкой.

Третий раздел посвящен изучению скорости сходимости модифицированных методов.

2 Теоретическое обоснование

2.1 Вывод метода

Пусть $f(z)$ - комплексный многочлен с корнями $r_1, r_2 \dots r_n$ и пусть z аппроксимирует корень r_j для некоторого фиксированного j. Определим

$$S_1(z) = \frac{f'(z)}{f(z)} = \sum_{i=1}^{n} \frac{1}{z - r_j} \tag{1}$$

и

$$S_2(z) = -\frac{dS_1}{dz} = \frac{(f'(z))^2 - f(z)f''(z)}{(f(z))^2} = \sum_{i=1}^{n} \frac{1}{(z - r_j)^2}. \tag{2}$$

Пусть $\alpha(z) = \frac{1}{z-r_j}$ и $\beta(z) + \gamma_i(z) = \frac{1}{z-r_j}$(для $i = 1, 2, \dots, n, i \neq j$), где β это среднее значение коллекции $\frac{1}{z-r_j}, i \neq j$. Ясно, что $\sum_{i=1, i\neq j}^{n} \gamma_i = 0$.
Определим

$$\delta^2 = \sum_{i=1, i\neq j}^{n} \gamma_i^2. \tag{3}$$

Используя определения $\alpha(z), \beta(z)$ и δ^2, перепишем (1) и (2) в виде

$$S_1 = \alpha + (n-1)\beta \tag{4}$$

и

$$S_2 = \alpha^2 + (n-1)\beta^2 + \delta^2 \tag{5}$$

Исключая β из (4) и (5) и решая для α, получим

$$\alpha = \frac{S_1 \pm \sqrt{(n-1)[nS_2 - S_1^2 - n\delta^2]}}{n}. \tag{6}$$

Поскольку $\alpha = \frac{1}{z-r_j}$ и (6) следует

$$r_j = z - \frac{n}{S_1 \pm \sqrt{(n-1)[nS_2 - S_1^2 - n\delta^2]}}, \tag{7}$$

где S_1, S_2 и δ^2 оцениваются по z.

Это точное выражение для r_j в терминах аппроксимации z. Задавая $\delta^2 = 0$ в (7), получаем итерационную формулу Лагерра

$$z_j^{(k+1)} = z_j^{(k)} - \frac{n}{S_1 \pm \sqrt{(n-1)[nS_2 - S_1^2]}}, \tag{8}$$

где S_1 и S_2 вычислены при $z_j^{(k)}$, текущая аппроксимация к r_j, и они определяется (1) и (2), соответственно.

2.2 Две модификации метода Лагерра.

1. Подставляя

$$\beta = \frac{1}{n-1} \sum_{i=1, i\neq j}^{n} \frac{1}{z - r_i}$$

в (5) и решая для δ^2, мы получаем, используя (2)

$$\delta^2 = \sum_{i=1, i\neq j}^{n} \frac{1}{(z-r_i)^2} - \frac{1}{n-1}\left[\sum_{i=1, i\neq j}^{n} \frac{1}{z-r_i}\right]^2 = \sum_{i\neq j}\left[\frac{1}{z-r_i} - \beta\right]^2. \tag{9}$$

Таким образом, δ^2 всегда неотрицательна для действительных z и r_i.

Если бы корни $r_1, \ldots, r_{j-1}, r_{j+1}, \ldots, r_n$ были известны, то для произвольного z, δ^2 можно было бы вычислить из (9) и использовать (7) для получения точного выражения для корня r_j в терминах S_1 и S_2, вычисленные при z и других корнях. В общем случае это невозможно, поскольку мы не знаем оставшихся $n-1$ корней.

Сформулируем алгоритм, позволяющий аппроксимировать δ^2, одновременно аппроксимируя все корни заданного многочлена $f(z)$. Пусть $z_1^{(k)}, z_2^{(k)}, \ldots, z_n^{(k)}$ аппроксимации нулей $r_1, r_2, \ldots, r_n$ многочлена $f(z)$, соответственно, и определим

$$\delta_j^2 = \sum_{i=1, i\neq j}^{n} \left[\frac{1}{z_j^{(k)} - z_i^{(k)}} - \beta_j\right]^2 \tag{10}$$

где

$$\beta_j = \frac{1}{n-1} \sum_{i=1, i\neq j}^{n} \frac{1}{z_j^{(k)} - z_i^{(k)}}.$$

Знаем, что $\delta_j^2 \approx \delta^2$ вычисляется при $z = z_j^{(k)}$, где δ^2 определяется в (9). Точность этой аппроксимации зависит от точности аппроксимации $z_i^{(k)}$ к r_i $(i = 1, \ldots, n; i \neq j)$.

Рассмотрим метод Лаггера (8). На практике мы выбираем знак перед квадратным корнем, чтобы сделать меньший шаг $z_j^{(k+1)} - z_j^{(k)}$. Если $f(z)$ вещественное со всеми вещественными корнями, то этот шаг "подрезает"корень r_j. А также, $nS_2 - S_1^2 \geq 0$ в вещественном случае.

Если заменим $nS_2 - S_1^2$ под квадратным корнем на $nS_2 - S_1^2 - n^2$, то получим больший шаг, который, согласно (7), и необходим для достижения корня r_j.

Аппроксимируем δ^2 величиной δ_j^2, заданной в (10), и используем итерационную функцию

$$z_j^{(k+1)} = z_j^{(k)} - \frac{n}{S_1 \pm \sqrt{(n-1)[nS_2 - S_1^2 - n\delta_j^2]}}. \tag{11}$$

Этот больший шаг (в общем случае) обеспечивает более близкую аппроксимацию к r_j, чем это делал метод Лаггера (8).

Если у нас имеются аппроксимации $z_j^{(k)}$ для r_j для всех $j = 1, \ldots, n$, мы можем вычислить δ^2 для всех j, используя (10). Следовательно, мы можем определить алгоритм для одновременной аппроксимации всех корней. Таким образом, алгоритм является подходящим для реализации на параллельной обработке данных.

2. Рассмотрим теперь модификацию типа "Гаусс-Зейдель" , которая является возможным улучшением как (8), так и (11). Она использует любую новую аппроксимацию, как только она будет найдена. То есть, предположим, что мы использовали (11) с $z_1^{(k)}, z_2^{(k)}, \ldots, z_n^{(k)}$ для получения новой аппроксимации $z_1^{(k+1)}$. Затем мы можем использовать $z_1^{(k+1)}, z_2^{(k)}, \ldots, z_n^{(k)}$ вместо $z_1^{(k)}, z_2^{(k)}, \ldots, z_n^{(k)}$ для получения новой аппроксимации $z_1^{(k+1)}$. Для обобщения пусть

$$\delta_j^* = \sum_{i=1}^{j-1} \left[\frac{1}{z_j^{(k)} - z_i^{(k+1)}} - \beta_j^* \right]^2 + \sum_{i=j+1}^{n} \left[\frac{1}{z_j^{(k)} - z_i^{(k)}} - \beta_j^* \right]^2 \tag{12}$$

где

$$\beta_j^* = \frac{1}{n-1} \left[\sum_{i=1}^{j-1} \frac{1}{z_j^{(k)} - z_i^{(k+1)}} + \sum_{i=j+1}^{n} \frac{1}{z_j^{(k)} - z_i^{(k)}} \right].$$

Если заменим в (11) δ_j^2 на δ_j^*, то сможем определить алгоритм, который использует новые аппроксимации $z_i^{(k+1)}$ к r_i по мере их поступления вместо того, чтобы ждать пока все $z_j^{(k+1)}$, $j = 1, 2, \ldots, n$ будут определены. Поскольку новые $z_j^{(k+1)}$, предположительно, являются более точны, чем старые $z_j^{(k)}$, то результирующий алгоритм должен сходиться быстрее.

Итерационная формула, полученная в результате этой модификации, имеет вид

$$z_j^{(k+1)} = z_j^{(k)} - \frac{n}{S_1 \pm \sqrt{(n-1)[nS_2 - S_1^2 - n\delta_j^*]}} \tag{13}$$

где $S1$ и S_2, заданные в (1) и (2), соответственно, вычисляются в $z_j^{(k)}$ и δ_j^*, определенный в (12).

2.3 Исследование скорости сходимости

Пусть $\varepsilon_j^{(k+1)} = z_j^{(k+1)} - r_j$ и $\varepsilon_j^{(k)} = z_j^{(k)} - r_j$. Предположим, что δ_j^2 в (11) - константа и, что r_j - простой корень. Ранее было показано, что $S_1(z) = \frac{f'(z)}{f(z)}$ и $S_2(z) = \frac{(f'(z))^2 - f(z)f''(z)}{(f(z))^2}$. С помощью несложных действий можем показать, что для $j = 1, 2, \ldots, n$

$$\varepsilon_j^{(k+1)} = (\varepsilon_j^{(k)})^3 \left[\frac{3(f'')^2 - 4f'f'''}{24(f')^2} - \frac{1}{8}\frac{1}{n-1}\left(\frac{f''}{f'}\right)^2 - \frac{\delta_j^2}{2} \right] + O\Big((\varepsilon_j^{(k)})^4\Big) \tag{14}$$

где f', f'' и f''' вычислены при $z = r_j$. Это показывает, что, когда δ_j^2 любая постоянная величина, (11) определяет итерационную формулу с кубической скоростью сходимости к простым корням. При нулевом значении δ_j^2 в (14), полученная асимптотическая константа ошибки согласуется с константой метода Лаггера[6].

Если $\delta_j^2 > 0$, что имеет место в вещественном случае, рассмотренный во втором разделе, и если оно не слишком велико, то есть, если $\delta_j^2 < 2L(r_j)$, где

$$L(r_j) = \frac{3(f'')^2 - 4f'f'''}{24(f')^2} - \frac{1}{8}\frac{1}{n-1}\left(\frac{f''}{f'}\right)^2, \tag{15}$$

тогда последовательность $< z_j^{(k)} >$, порожденная (11), имеет меньшую асимптотическую константу ошибки, чем чем соответствующая последовательность, порожденная (8). Это означает, по крайней мере асимптотически, что итерационная функция (11) с δ_j^2, определенный в (10), имеет улучшенную сходимость по сравнению с методом Лагерра. То же самое справедливо и для δ_j^2, замененого на δ_j^*, определенный в (12), если $\delta_j^2 < 2L(r_j)$.

Рассмотрим скорость сходимости (11), рассматривая δ_j^2, как функцию от z. Знаем, что $\varepsilon_j^{(k)} = z_j^{(k)} - r_j$, $j = 1, 2, \ldots, n$ и пусть $r_{ij} = r_i - r_j$, тогда можем показать, что

$$\varepsilon_j^{(k+1)} = \left(\varepsilon_j^{(k)}\right)^3 \left[\sum_{i=1, i\neq j}^{n} \frac{\varepsilon_i^{(k)}}{r_{ij}^3} - \frac{i}{n-1}\left(\sum_{i=1, i\neq j}^{n} \frac{1}{r_{ij}}\right) \sum_{i=1, i\neq j}^{n} \frac{\varepsilon_i^{(k)}}{r_{ij}^2} \right] \tag{16}$$

$$+ \text{ члены высших порядков в } \varepsilon_j^{(k)} \text{ и } \varepsilon_i^{(k)}.$$

Если возьмем $\varepsilon_i^{(k)} = \eta_i \varepsilon^{(k)}$, то убедимся в том, что метод, определенный в (11), с δ_j^2, заданный в (10), имеет по крайней мере сходимость 4-го порядка к простым корням. Исходя из определения (12) о δ_j^*, следует, что дальнейшее улучшение сходимости будет достигнуто с помощью итерации, определенной в (13).

2.4 Параллельный алгоритм

Итерационная формула, определенная (11), кажется подходящей в качестве параллельного алгоритма для одновременного определения всех корней заданного многочлена степени n.

Известно, что δ_j^2, определенный (10), сходится к δ^2, определенный (9), вычисленный при $z = r_j$ как $z_i^{(k)}$ сходящийся к r_i, $i = 1, \ldots, n$. Поэтому (11) как функция от $z_i^{(k)}$, $i = 1, 2, \ldots, n$ сходится к точному выражению для корня r_j, $j = 1, 2, \ldots, n$ и имеет отличное асимптотическое поведение. Кроме того, метод Лагерра имеет хорошее начальное поведение в следующем смысле: даже если начальная оценка корня достаточно велика, то один шаг метода Лагерра дает новую оценку, "достаточно близкую"к корню[9].

Одним из способов описания такого поведения состоит в следующем. Предположим, что в (8) допускаем, что $z_j^{(k)}$ становится бесконечно большой величиной. Поскольку $z_j^{(k)} \to \infty$, $z_j^{(k+1)}$ приближается к конечному пределу. В частности, если

$$f(z) = \sum_{i=0}^{n} a_i z^i,$$

тогда

$$\lim_{z_j^{(k)} \to \infty} z_j^{(k+1)} = \frac{1}{n a_n} \{[(n-1)^2 a_{n-1}^2 - 2n(n-1) a_n a_{n-2}]^{\frac{1}{2}} - a_{n-1}\}.$$

Наша модификация (11) обладает тем свойством, что все $z_i^{(k)}$ для $i = 1, \ldots, n$ $(i \neq j)$ остаются конечными при $z_j^{(k)} \to \infty$.