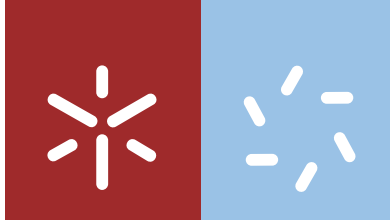


Universidade do Minho
Escola de Ciências

Lúcia Bernardete Azevedo de Sousa

**Sobre um método proposto por
Sebastião e Silva para o cálculo de raízes
de uma equação algébrica**

outubro de 2016



Universidade do Minho

Escola de Ciências

Lúcia Bernardete Azevedo de Sousa

**Sobre um método proposto por
Sebastião e Silva para o cálculo de raízes
de uma equação algébrica**

Dissertação de Mestrado

Mestrado em Ciências – Formação Contínua de Professores
Área de Especialização em Matemática

Trabalho realizado sob orientação do

Professor Doutor Rui Manuel da Silva Ralha

“Seria possível dizer o que é a Matemática se esta fosse uma ciência morta. Mas a Matemática é, pelo contrário, uma ciência viva, que se encontra hoje mais do que nunca, em rápido desenvolvimento, proliferando cada vez mais em novos ramos, que mudam não só a sua fisionomia, como até a sua essência.”

José Sebastião e Silva

AGRADECIMENTOS

A minha profunda gratidão ao Doutor Rui Ralha pela honra dada em ser meu orientador, de ter partilhado a sua sabedoria e experiência com uma atenção e disponibilidade inesgotáveis na concretização desta tese. Ao longo do ano acompanhou o meu trabalho de uma forma sempre crítica e muito construtiva através de uma capacidade singular de me ajudar a estudar e compreender conteúdos mais complexos de uma maneira simples, prática e entusiasta. Agradeço-lhe a paciência e o apoio incondicional que muito elevaram os meus conhecimentos científicos e pedagógicos e, sem dúvida, muito estimularam o meu desejo e a minha motivação de querer sempre saber mais e fazer melhor.

Agradeço à Doutora Elfrida Ralha não só pelo incentivo e orientação iniciais tão importantes para a tomada de decisão acerca do tema da tese mas também pela prestabilidade e simpatia sempre manifestadas ao longo deste percurso.

À minha família, em especial à Sónia, ao Marcel e ao Armando por todo o incentivo, perseverança no acompanhamento desta jornada.

A todos os que, direta ou indiretamente contribuíram para a realização desta tese.

Àqueles que um dia possam ler e explorar este trabalho tornando-o mais próspero e útil.

A todos um... Muito Obrigada!

SOBRE UM MÉTODO PROPOSTO POR JOSÉ SEBASTIÃO E SILVA PARA O CÁLCULO DAS RAÍZES DE UMA EQUAÇÃO ALGÉBRICA

RESUMO

Esta dissertação realiza-se no âmbito do Mestrado em Ciências - Formação Contínua de Professores - área de especialização em Matemática. É objetivo principal desta tese aprofundar conhecimentos e desenvolver competências com vista a uma melhoria do meu desempenho profissional.

Também gostaria que este trabalho viesse a ser lido por outros professores de Matemática do Ensino Básico e Secundário e que constituísse também para eles uma experiência de reflexão sobre o que se ensina (e como se ensina) no tema das equações algébricas.

Neste trabalho estudaremos um dos métodos apresentados por Sebastião e Silva em [SebSilva1940] e também os métodos de Graeffe e de Newton – Raphson que com aquele se relacionam. Incluído neste estudo está a análise da convergência baseada no “Teorema Fundamental” de Sebastião e Silva, resultado original e muito interessante.

As ferramentas matemáticas envolvidas no método de Sebastião e Silva são: a) as relações de Vieta para encontrar a aproximação u da menor das raízes de um polinómio; b) a transformação algébrica do polinómio decorrente da mudança de variável $y = x - u$; c) a transformada de Graeffe para a aceleração da convergência. Usadas de forma iterativa, a) e b) são afinal matematicamente equivalentes ao método de Newton- Raphson e a proposta de Sebastião e Silva difere daquele método clássico apenas na utilização da transformada de Graeffe para aumentar o afastamento entre as raízes, melhorando assim a rapidez de convergência.

Incluiremos secções onde tratamos, de forma autónoma, a transformada de Graeffe (secção 2) e o método de Newton-Raphson (secção 3). Em cada caso procurámos expôr de forma suficientemente detalhada a teoria básica subjacente e usámos exemplos ilustrativos do funcionamento dos métodos.

Reservámos uma secção própria (secção 4) para o “Teorema Fundamental” que aparece no primeiro parágrafo de [SebSilva1940] onde Sebastião e Silva apresenta uma condição necessária e suficiente para a convergência de um método iterativo. Optámos por não reproduzir a demonstração do Teorema Fundamental e dar ênfase à sua interpretação no caso dos exemplos incluídos.

ON A METHOD PROPOSED BY JOSÉ SEBASTIÃO E SILVA FOR THE CALCULATION OF THE ROOTS OF AN ALGEBRAIC EQUATION

ABSTRACT

This is a dissertation for obtaining the MSc degree “Mestrado em Ciências – Formação Contínua de Professores, área de especialização em Matemática”.

The main goal of this work is to enlarge my scientific background and to develop skills that make me a better secondary school teacher. Also, it is my wish that this thesis is read by other teachers of Mathematics and becomes for them an opportunity for rethinking what is taught (and how it is taught) in the topic of algebraic equations.

In this work we study one of the methods presented by Sebastião e Silva in [SebSilva1940], as well as the classic methods of Graeffe and Newton-Raphson, because these are related to the first one. We also devoted our attention to the convergence analysis of the method based upon an original and interesting theoretical result that Sebastião e Silva called “The Fundamental Theorem”.

The mathematical tools used in the method of Sebastião e Silva are: a) the Vieta relations to find an approximation u for the smallest root of a polynomial; b) the algebraic transformation of the polynomial that corresponds to a change of variable $y=x-u$ (roots displacement); c) the Graeffe transformation for accelerating the convergence. Used in an iterative way, a) and b) are mathematically equivalent to the Newton-Raphson method and the method of Sebastião e Silva differs from that classical method on using the Graeffe transformation to get a polynomial whose roots are the squares of the roots of the transformed polynomial, thus better separated, increasing the convergence speed.

We include sections dedicated to the Graeffe transformation (section 2) and the Newton-Raphson method (section 3). In both cases we tried to expound in a sufficiently detailed way the basic theory and we included many examples that illustrate how the methods work. We have reserved an autonomous section for the “Fundamental Theorem” in which Sebastião e Silva gives a necessary and sufficient condition for the convergence of an iterative method. We do not include the proof of the theorem (which may be found in [SebSilva1940]) but we include examples in order to make a complete interpretation of the theorem.

INDICE

1. INTRODUÇÃO	1
2. MÉTODO DE GRAEFFE.....	3
3. MÉTODO DE NEWTON – RAPHSON	11
4. TEOREMA FUNDAMENTAL	22
5. MÉTODO DE JOSÉ SEBASTIÃO E SILVA	29
6. CONCLUSÃO	40
BIBLIOGRAFIA	41

1. INTRODUÇÃO

José Sebastião e Silva (1914-1972), licenciado em Matemática pela Faculdade de Ciências de Lisboa, em 1937, é um dos maiores vultos portugueses na área da Matemática e do ensino da Matemática do século XX. De uma entrevista que concedeu ao Diário de Notícias em 23/01/1968 ficou escrito: (...) *a educação, na era científica, não pode continuar, de modo nenhum, a ser feita segundo os moldes do passado. Em todas as escolas o ensino das ciências tem que ser intensificado e remodelado desde as suas bases, não só quanto a programas mas ainda quanto a métodos. Compete à escola formar seres pensantes, dotados de imaginação criadora e de capacidade de adaptação em grau cada vez mais elevado. O que se pretende acima de tudo é levar o aluno a compreender o porquê dos processos matemáticos, em vez de lhe paralisar o espírito, automatizando-o desde logo.* (...)

Num tempo em que os computadores eram rudimentares e o seu uso muito limitado, Sebastião e Silva já reconhecia que o cálculo automático poderia (e deveria) ter um papel relevante no ensino da Matemática. Deu especial atenção aos métodos iterativos de resolução de equações algébricas e, em particular, um método original por ele proposto em [SebSilva1941] e [SebSilva1946] teve reconhecimento de reputados especialistas internacionais [Householder1973]. Este método que revela o conhecimento profundo de Sebastião e Silva sobre este assunto foi o culminar de outros trabalhos sobre o tema. Na sua primeira publicação [SebSilva1940], apenas trata de polinómios com raízes reais positivas e apresenta dois métodos para o cálculo da menor daquelas raízes.

Quero agora destacar as razões da escolha do tema da tese. Por um lado, o desafio de conhecer uma parte (embora pequena) do trabalho desenvolvido por um grande matemático português. Por outro lado, inspirar-me nas ideias do grande pedagogo que foi Sebastião e Silva e com isso tornar-me mais capaz de desenvolver nos meus alunos o gosto pela verdadeira matemática, a que desenvolve o espírito crítico e o raciocínio lógico. Por exemplo, no Ensino Secundário os alunos aprendem a calcular as raízes de uma equação do segundo grau automatizando a respetiva fórmula resolvente. Para além disto resolvem equações de grau superior que admitem uma fatorização simples (para

equações do terceiro grau ensina-se a regra de Ruffini quando uma das raízes é dada ou facilmente determinada).

Defendia Sebastião e Silva que os alunos, depois de aprenderem as regras mencionadas antes, devem ser confrontados com outras equações que mostrem a necessidade de outras ferramentas matemáticas, isto é, os métodos numéricos ligados à utilização de computadores cuja utilização o pedagogo já defendia então.

2. MÉTODO DE GRAEFFE

O método de Graeffe, também conhecido por método de Dandelin–Graeffe, pode ser usado para calcular todas as raízes de um polinómio. Foi desenvolvido de forma independente por Germinal Pierre Dandelin em 1826 e Karl Heinrich Gräffe in 1837. Lobachevsky em 1834 também descobriu a ideia principal deste método [Householder1959].

O método usa as fórmulas de Vieta¹ para aproximar as raízes e, uma vez que tais aproximações serão tanto melhores quanto mais afastadas estiverem as raízes, a ideia principal é a de produzir um polinómio do mesmo grau cujas raízes são os quadrados das raízes do polinómio inicial.

Seja

$$p(x) = (x-x_1)\dots(x-x_n).$$

O polinómio $q(x) = p(x) \cdot p(-x)$ é de grau $2n$ mas, por serem nulos os coeficientes dos termos de grau ímpar, uma simples mudança de variável $y = x^2$ produz um polinómio de grau n .

Com

$$p(x) = x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n,$$

tem-se

$$p(-x) = (-1)^n [x^n + \bar{a}_1x^{n-1} + \dots + \bar{a}_{n-1}x + \bar{a}_n]$$

onde

$$\bar{a}_i = \begin{cases} a_i, & \text{se } i \text{ é par} \\ -a_i, & \text{se } i \text{ é ímpar} \end{cases} \quad (2.1)$$

¹ Estas são as conhecidas fórmulas que relacionam os coeficientes de um polinómio com somas e produtos das raízes (ver à frente); tais relações foram descobertas no século XVI pelo matemático François Vieta para o caso de raízes positivas.

Resulta

$$q(x) = (-1)^n [x^{2n} + c_1 x^{2n-1} + c_2 x^{2n-2} + c_3 x^{2n-3} + \dots + c_{2n-1} x + c_{2n}]$$

onde, com $a_0 = \bar{a}_0 = 1$,

$$\begin{aligned} c_1 &= a_0 \bar{a}_1 + \bar{a}_0 a_1 \\ c_2 &= a_0 \bar{a}_2 + a_1 \bar{a}_1 + a_2 \bar{a}_0 \\ c_3 &= a_0 \bar{a}_3 + a_1 \bar{a}_2 + a_2 \bar{a}_1 + a_3 \bar{a}_0 \\ &\vdots \\ c_p &= \sum_{i+j=p} a_i \bar{a}_j, p = 1, \dots, 2n \end{aligned} \quad (2.2)$$

Para p ímpar, os índices i e j são de diferente paridade e de (2.1) resulta $a_i \bar{a}_j = -a_j \bar{a}_i$, logo $c_p = 0$. Para p par, tem-se

$$c_p = c_{2j} = (-1)^j a_j^2 + 2 \sum_{i \neq j} (-1)^i a_i a_{2j-i} \quad (2.3)$$

e com $b_j = c_{2j}$, podemos escrever

$$q(x) = (-1)^n (x^{2n} + b_1 x^{2n-2} + b_2 x^{2n-4} + \dots + b_{n-1} x^2 + b_n)$$

E, finalmente, substituindo x^2 por x ,

$$q(x) = (-1)^n (x^n + b_1 x^{n-1} + b_2 x^{n-2} + \dots + b_{n-1} x + b_n).$$

Graeffe observou que o polinómio q pode ser obtido escrevendo $p(x) = 0$ na forma

$$p_e(x) = x p_o(x)$$

onde p_e e p_o são polinómios só com termos de grau par. O grau destes polinómios é reduzido para metade substituindo x^2 por x e obtemos, com $\bar{p}_e(x^2) = p_e(x)$ e $\bar{p}_o(x^2) = p_o(x)$,

$$\bar{p}_e(x) = x^{\frac{1}{2}} \bar{p}_o(x).$$

O polinómio q é definido por

$$q(x) = (-1)^n (\bar{p}_e^2(x) - x\bar{p}_o^2(x)),$$

o que envolve o quadrado de dois polinómios de graus não superiores a $\frac{n}{2}$ e é a forma mais prática de implementar o método.

Como ilustração do que acaba de ser dito, consideremos o polinómio

$$p(x) = x^4 - 10x^3 + 35x^2 - 50x + 24,$$

cujas raízes são 4, 3, 2 e 1.

Seguindo o processo usado por Graeffe² para obter os coeficientes do polinómio q , escrevemos

$$x^4 + 35x^2 + 24 = x(-10x^2 - 50)$$

ou, substituindo x^2 por x ,

$$x^2 + 35x + 24 = x^{\frac{1}{2}}(-10x - 50)$$

e, quadrando ambos os membros, resulta

$$x^4 + 70x^3 + 1273x^2 + 1680x + 576 = x(100x^2 + 1000x + 2500)$$

ou seja

$$x^4 - 30x^3 + 273x^2 - 820x + 576 = 0,$$

² Este processo costuma designar-se por transformada de Graeffe.

cujas raízes são 1, 4, 9 e 16. Os coeficientes deste polinómio podem ser calculados diretamente a partir dos coeficientes do polinómio dado, usando as relações (2.3).

Com $a_0 = 1, a_1 = -10, a_2 = 35, a_3 = -50$ e $a_4 = 24$, tem-se

$$b_1 = 2a_0a_2 - a_1^2 = 70 - 100 = -30$$

$$b_2 = 2a_0a_4 - 2a_1a_3 + a_2^2 = 273$$

$$b_3 = 2a_2a_4 - a_3^2 = -820$$

$$b_4 = a_4^2 = 576$$

Para aumentar a separação das raízes, o processo anterior pode ser repetido. Iterando k vezes, obtemos o polinómio de grau n

$$q(y) = y^n + b_1y^{n-1} + \dots + b_{n-1}y + b_n$$

cujas raízes são

$$y_1 = x_1^{2^k}, y_2 = x_2^{2^k}, \dots, y_n = x_n^{2^k}.$$

Se os valores absolutos das raízes do polinómio original forem separados por um fator $\rho > 1$, isto é, $|x_k| \geq \rho |x_{k+1}|$ então $|y_i|$ e $|y_{i+1}|$, na k -ésima iteração são separados por um fator de crescimento rápido

$$\rho^{2^k} > 1 + 2^k(\rho - 1)$$

(a demonstração desta relação pode fazer-se por indução matemática).

As relações de Vieta

$$b_1 = -(y_1 + y_2 + \dots + y_n) = -y_1 \left(1 + \frac{y_2}{y_1} + \dots + \frac{y_n}{y_1} \right)$$

$$b_2 = y_1y_2 + y_1y_3 + \dots + y_{n-1}y_n = y_1y_2 \left(1 + \frac{y_1y_3}{y_1y_2} + \dots + \frac{y_{n-1}y_n}{y_1y_2} \right)$$

⋮

$$b_n = (-1)^n (y_1y_2 \cdots y_n)$$

podem agora ser usadas para obter aproximações para as raízes y_i . Desprezando as quantidades $\frac{y_i}{y_j}, j < i$, cujos valores absolutos tendem para zero à medida que k cresce,

temos

$$\begin{aligned} b_1 &\approx -y_1 \\ b_2 &\approx y_1 y_2 \\ &\vdots \\ b_n &= (-1)^n (y_1 y_2 \cdots y_n) \end{aligned}$$

e

$$\begin{aligned} y_1 &\approx -b_1 \\ y_2 &\approx -\frac{b_2}{b_1} \\ &\vdots \\ y_n &\approx -\frac{b_n}{b_{n-1}} \end{aligned} \quad (2.4)$$

De $y_i = x_i^{2^k}, i = 1, \dots, n$, podemos calcular os valores absolutos das raízes do polinómio original $|x_i| = y_i^{-2^k}$.

Em seguida, ilustraremos a aplicação repetida da transformada de Graeffe e a utilização das aproximações (2.4) dadas pelas fórmulas de Vieta.

Para equações de terceiro grau

$$x^3 + a_1 x^2 + a_2 x + a_3 = 0$$

a transformada de Graeffe dá

$$x^3 + (-a_1^2 + 2a_2)x^2 + (a_2^2 - 2a_1 a_3)x - a_3^2.$$

No caso do polinómio $x^3 - 4x^2 - 7x + 10$, cujas raízes são 5, -2 e 1, aplicações sucessivas da transformação anterior produz os polinómios cujos coeficientes, para valores de k até 8, estão listados na tabela seguinte, com valores arredondados a partir de $k = 3$, onde $p^{(2^k)}$ é o polinómio de grau 3 produzido na k -ésima transformada.

k	$p^{(2^k)}$	b_1	b_2	b_3
1	$p^{(2)}$	-30	129	-100
2	$p^{(4)}$	-642	10641	-10000
3	$p^{(8)}$	$-3,909.10^5$	$1,004.10^8$	$-1,000.10^8$
4	$p^{(16)}$	$-1,526.10^{11}$	$1,000.10^{16}$	$-1,000.10^{16}$
5	$p^{(32)}$	$-2,328.10^{22}$	$1,000.10^{32}$	$-1,000.10^{32}$
6	$p^{(64)}$	$-5,421.10^{44}$	$1,000.10^{64}$	$-1,000.10^{64}$
7	$p^{(128)}$	$-2,939.10^{89}$	$1,000.10^{128}$	$-1,000.10^{128}$
8	$p^{(256)}$	$-8,636.10^{178}$	$1,000.10^{256}$	$-1,000.10^{256}$

Tabela 2.1 Transformadas de Graeffe

Como se pode ver, os coeficientes crescem rapidamente; com efeito, $k = 8$ é o número máximo de transformadas que é possível fazer neste caso, uma vez que a nona transformada produz números com expoentes demasiadamente grandes para poderem ser representados na máquina (isto é, ocorre “overflow”).

Para apreciar a convergência para a raiz de maior valor absoluto, na tabela seguinte apresentamos para cada um dos polinómios $p^{(2^k)}(x)$, $k = 0, 1, 2, 3$, a razão

$$\frac{x_1^{2^k}}{x_1^{2^k} + x_2^{2^k} + x_3^{2^k}}, \text{ que converge para a unidade.}$$

Polinómio	Raízes	Soma das raízes do polinómio	Razão entre a raiz de maior valor absoluto e a soma das raízes do polinómio
$p(x)$	5, -2, 1	$x_1 + x_2 + x_3 = 1 + (-2) + 5 = 4$	$\frac{5}{4} = 1,25$
$p^{(2)}(x)$	25, 4, 1	$x_1^2 + x_2^2 + x_3^2 = 1 + 4 + 25 = 30$	$\frac{25}{30} \approx 0,833$
$p^{(4)}(x)$	625, 16, 1	$x_1^4 + x_2^4 + x_3^4 = 1 + 16 + 625 = 642$	$\frac{625}{642} \approx 0,974$
$p^{(8)}(x)$	390625, 256, 1	$x_1^8 + x_2^8 + x_3^8 = 1 + 256 + 390625 = 390882$	$\frac{390625}{390882} \approx 0,999343$
$p^{(16)}(x)$	152587890625, 65536, 1	$x_1^{16} + x_2^{16} + x_3^{16} = 152587956162$	$\frac{152587890625}{152587956162} \approx 0,999999$

Tabela 2.2

Neste caso, em que as raízes são todas reais e de valores absolutos distintos, as razões $\left(\frac{x_i}{x_j}\right)^{2^k}$, $j < i$, convergem todas para zero e podemos obter aproximações para os valores absolutos de todas as raízes. Na tabela 2.3 estão tais aproximações calculadas a partir dos coeficientes dos polinómios $p^{(2^k)}$ pelas fórmulas (2.4).

n	$\sqrt[n]{-b_1}$	$\sqrt[n]{-\frac{b_2}{b_1}}$	$\sqrt[n]{-\frac{b_3}{b_2}}$
2	5,477225575051660	2,073644135332770	0,880450906325624
4	5,033658601717030	2,017724436713330	0,984587661849564
8	5,000411081688840	2,000811036096890	0,999512470074940
16	5,000000134219750	2,000001853647860	0,999999046333005
32	5,000000000000030	2,000000000014540	0,999999999992724
64	5,000000000000000	2,000000000000000	1,000000000000000

Tabela 2.3 Aproximações para as raízes dadas por (2.4)

A partir do conhecimento de $|x_i|$, sendo $|x_i|$ real, a ambiguidade do sinal é resolvida através do cálculo do valor do polinómio original em $|x_i|$. No caso presente tem-se $p(5) = 0$, $p(2) = -12$ e $p(1) = 0$, donde se conclui que as raízes são $x_1 = 5$, $x_2 = -2$ e $x_3 = 1$.

Neste exemplo não foram necessárias muitas iterações para obter boas aproximações para as raízes porque

$$\left(\frac{x_2}{x_1}\right)^{2^k} = \left(\frac{2}{5}\right)^{2^k} \quad \text{e} \quad \left(\frac{x_3}{x_2}\right)^{2^k} = \left(\frac{1}{2}\right)^{2^k}$$

tendem rapidamente para zero quando k aumenta. Nos casos em que a convergência seja lenta não será prático usar o método para calcular boas aproximações (até porque, como se disse antes, os coeficientes dos polinómios produzidos nas transformadas de Graffe crescem muito rapidamente inviabilizando a continuação do processo). No entanto, as aproximações produzidas podem servir como pontos de partida para outros métodos de convergência mais rápida.

Como nota final sobre o método, observamos que ele também pode ajudar no cálculo de raízes reais múltiplas ou até raízes complexas; mas estas situações requerem outro estudo que é desnecessário para a análise que queremos fazer do trabalho de Sebastião e Silva sobre o cálculo de raízes de equações algébricas.

3. MÉTODO DE NEWTON – RAPHSON

O método de Newton é um dos mais conhecidos e utilizados métodos de resolução de equações. Também conhecido por método de Newton – Raphson, aparece na obra Isaac Newton *De Methodus Fluxionum et Serierum infinitorum*, escrita entre 1664 e 1671, para encontrar a solução da equação

$$y^3 - 2y - 5 = 0.$$

Joseph Raphson desenvolveu o método original de Newton mas, de acordo com [Nordgaard1922], a expressão funcional (3.1) é atribuída a Lagrange (Simpson usou-a antes na resolução de sistemas de equações).

Em linguagem moderna, escrevendo a equação na forma $f(x) = 0$, em que f é de classe C^2 num intervalo que contém a raiz procurada α , a partir de uma dada aproximação inicial x_0 calcula-se a aproximação seguinte

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

e, em geral,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

O método também é conhecido pelo método das tangentes. Faz-se em seguida a interpretação geométrica deste método, que justifica a designação.

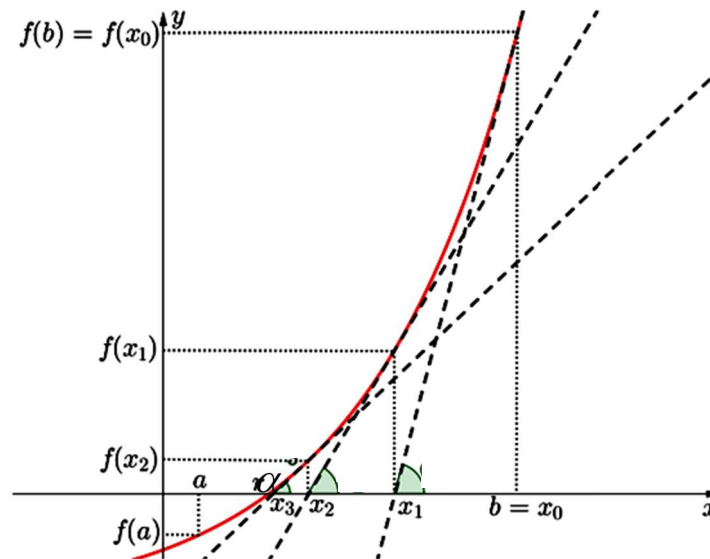


Figura 3.1

A partir da expansão de em série de Taylor em torno do ponto x_0

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!}f''(x_0) + \dots,$$

tomando apenas os dois primeiros termos da série, temos

$$y = f(x_0) + (x - x_0)f'(x_0)$$

que é a equação da reta que passa pelo ponto $(x_0, f(x_0))$ e tem inclinação $f'(x_0)$, isto é, ela é tangente à curva no ponto de abscissa x_0 . A ideia é usar a abscissa do ponto em que esta reta cruza o eixo dos x como aproximação da raiz da equação. Esta abscissa é dada por

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Este ponto x_1 será uma melhor aproximação, para a raiz α , do que x_0 se se cumprirem as condições de convergência que estudaremos mais adiante. Fazendo n iterações temos

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, f'(x_n) \neq 0 \quad (3.1)$$

que é a fórmula iterativa do método de Newton.

A partir do desenvolvimento em série dado antes escrevemos a fórmula de Taylor com resto de Lagrange

$$f(\alpha) = f(x_n) + f'(x_n)(\alpha - x_n) + \frac{f''(\xi)}{2!}(\alpha - x_n)^2,$$

para algum ξ que está entre x_n e α .

Por ser $f(\alpha) = 0$ e $f'(x) \neq 0$, resulta

$$0 = f(x_n) + f'(x_n)(\alpha - x_n) + \frac{f''(\xi)}{2!}(\alpha - x_n)^2$$

$$\alpha = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(\xi)}{2f'(x_n)}(\alpha - x_n)^2.$$

Sendo

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

resulta

$$\alpha - x_{n+1} = -\frac{f''(\xi)}{2f'(x_n)}(\alpha - x_n)^2$$

e, fazendo $\alpha_n = |\alpha - x_n|$, temos

$$\alpha_{n+1} = \left| \frac{f''(\xi)}{2f'(x_n)} \right| \alpha_n^2 \quad (3.2)$$

o que, de acordo com a definição usual da teoria dos métodos iterativos, nos permite concluir que a convergência do método de Newton é quadrática. Com efeito [Valença1993], p.46, sendo

$$\lim_{n \rightarrow +\infty} \frac{\alpha_{n+1}}{\alpha_n^p} = c, \quad c > 0, \quad (3.3)$$

então p é a chamada razão de convergência do método e C é a constante de convergência assintótica. No caso do método de Newton – Raphson é $p = 2$ e $c = \left| \frac{f''(\alpha)}{2f'(\alpha)} \right|$. Se este valor de C não for grande e $\alpha_n = 10^{-2}$, por exemplo, então será $\alpha_{n+1} = 10^{-4}$.

De seguida são apresentados exemplos de aplicação do método de Newton-Raphson. Consideremos assim a função f , utilizada na secção 2, definida por

$$f(x) = x^4 - 10x^3 + 35x^2 - 50x + 24$$

Exemplo 3.1

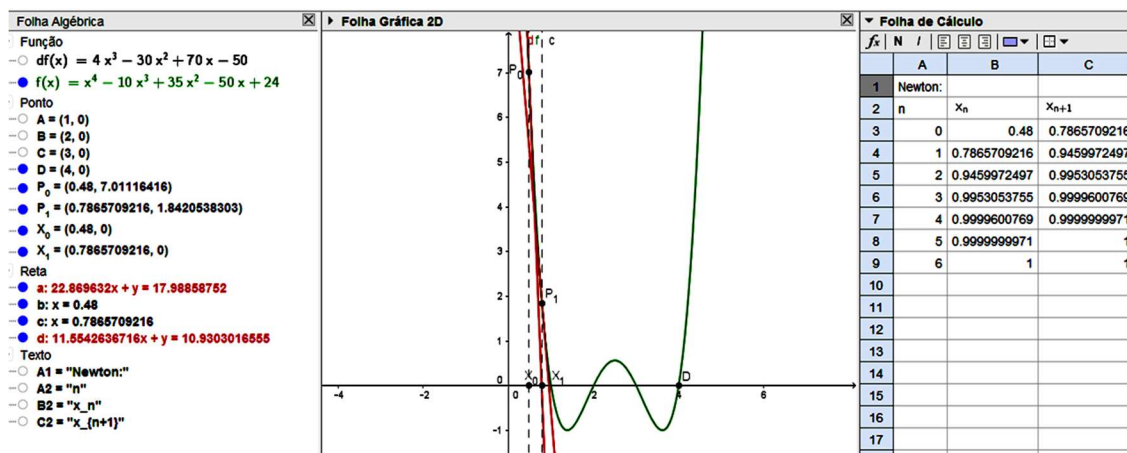


Figura 3.2

Com a aproximação inicial $x_0 = 0,48$, o método converge rapidamente para a menor das raízes $\alpha = 1$ do polinómio f , como se pode apreciar na parte direita da janela (produzida no GeoGebra).

Dependendo da equação a resolver e da aproximação inicial x_0 usada, o método de Newton – Raphson pode divergir. Tal acontece se $f'(x_n)=0$, o método falha, como resulta da expressão (3.1). Graficamente, tal significa que a reta tangente ao gráfico no ponto de abcissa x_n é horizontal, isto é, não intersesta o eixo dos x , conforme ilustra a figura 3.3.

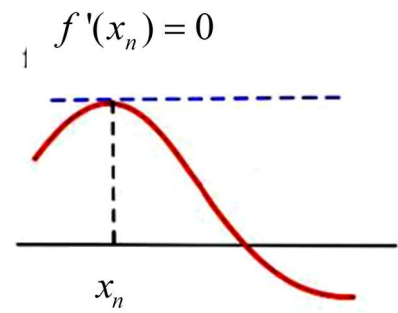


Figura 3.3

No exemplo que se segue vamos observar que para o valor inicial $x_0=0$ obtemos o valor $x_1=-1$ mas para a iteração seguinte temos $f'(x_1)=0$, o que nos impede de continuar. Na folha de cálculo da figura 3.4 o ponto de interrogação sinaliza a divisão por zero ocorrida.

Exemplo 3.2 $f(x)=-x^5-x^4+x+1$

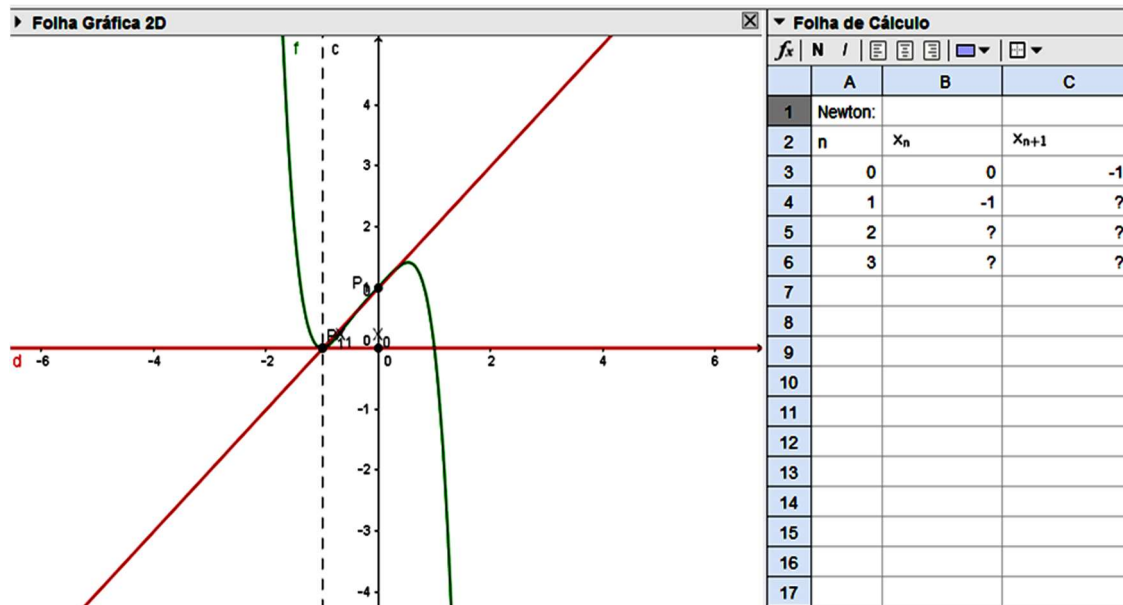


Figura 3.4

Conforme se ilustra na figura 3.5, pode verificar-se um comportamento oscilatório das iterações em que dois valores se repetem indefinidamente.

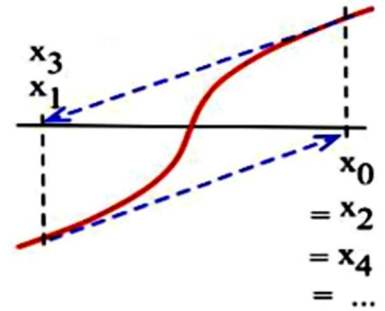


Figura 3.5

Exemplo 3.3 $f(x) = x^3 - 3x^2 + x - 1$

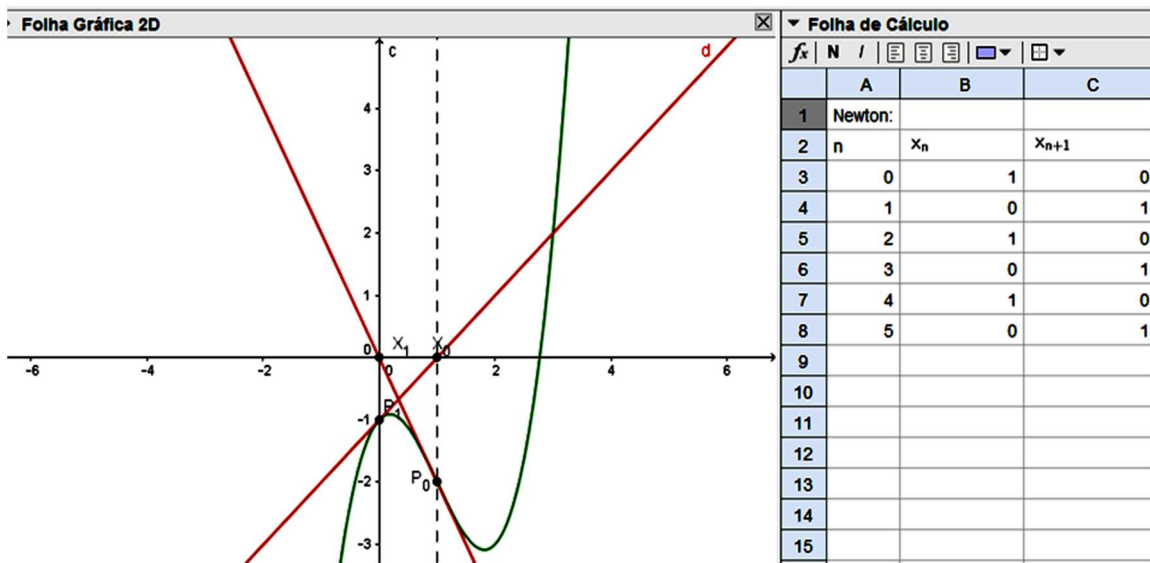


Figura 3.6

Se os valores das iterações x_n se afastarem cada vez mais da raiz como ilustramos na figura 3.7, temos outro caso de divergência.

No exemplo seguinte a raiz de f é 0 mas se tomarmos $x_0 = 1$, a sucessão de valores de x_{n+1} não converge.

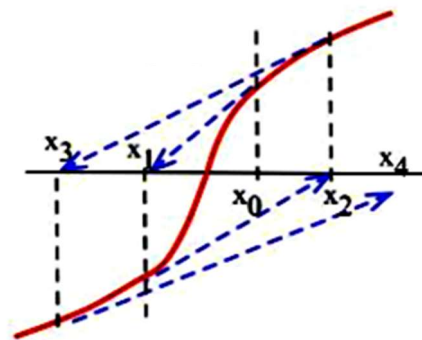


Figura 3.7

Exemplo 3.4 $f(x) = xe^{-x^2}$

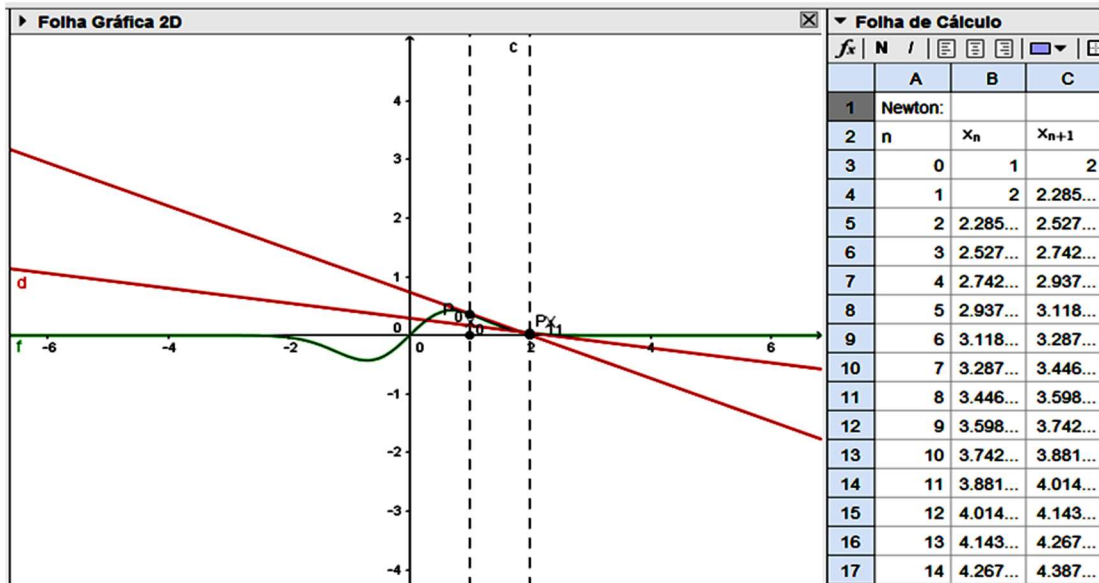


Figura 3.8

O método também pode falhar se existirem na vizinhança da raiz pontos onde f não é derivável.

Exemplo 3.5 $f(x) = |x-3| - 1$.

Os zeros da função são em $x = 2$ e $x = 4$ são suas raízes. Caso iniciemos o método de Newton-Raphson com $x_0 = 3$, o processo iterativo falhará porque não existe $f'(3)$.

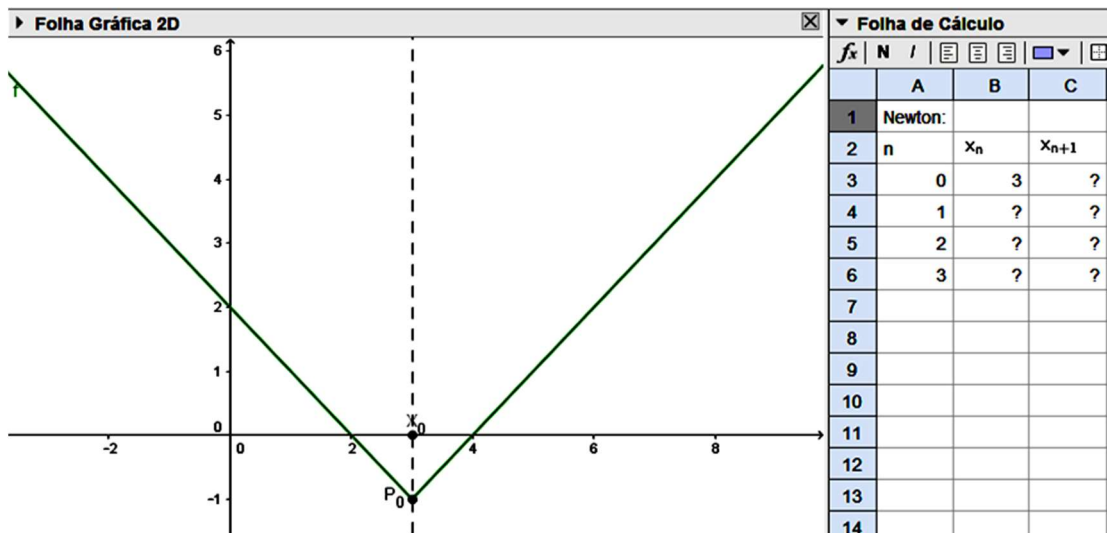


Figura 3.9

No exemplo seguinte, o método diverge seja qual for a aproximação inicial x_0 .

Exemplo 3.6 $f(x) = \sqrt[8]{x}$

Partindo de $x_0 = 1$, a fórmula iterativa $x_{n+1} = -7x_n$ é divergente.

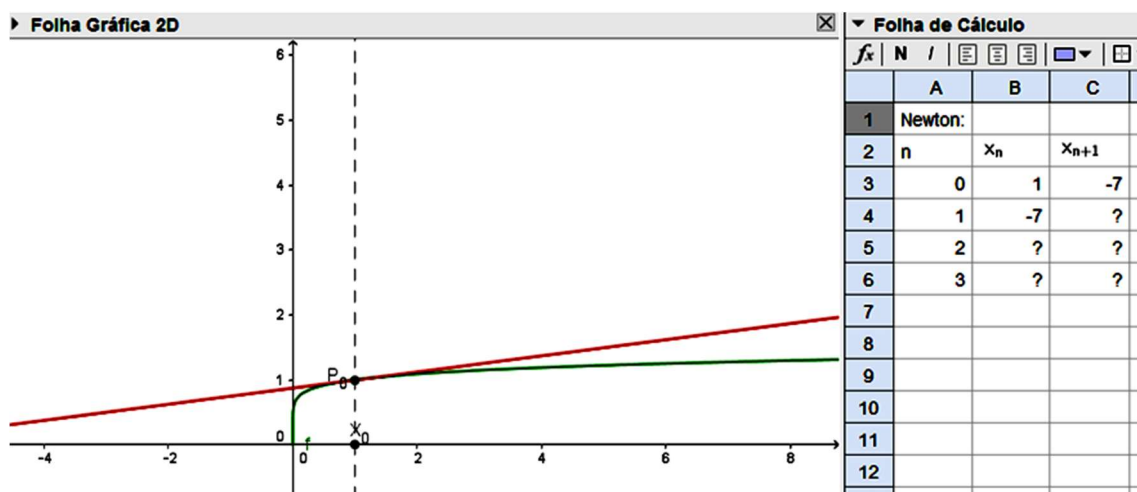


Figura 3.10

Mas pode-se garantir a convergência se se cumprem certas condições. Tem-se o seguinte [Valença1993], p.42.

Teorema 3.1 Seja $f \in C^2([a, b])$. Se

- (i) $f(a) \times f(b) < 0$;
- (ii) $f'(x) \neq 0, x \in [a, b]$;
- (iii) $f''(x)$ não muda de sinal em $[a, b]$;
- (iv) $\left| \frac{f(a)}{f'(a)} \right| < b - a$ e $\left| \frac{f(b)}{f'(b)} \right| < b - a$,

então para qualquer $x_0 \in [a, b]$, a sucessão $\{x_n\} = \{x_0, x_1, x_2, \dots\}$ gerada pelo método de Newton – Raphson converge para α , único zero de f em $[a, b]$.

As condições (i) e (ii) garantem a existência de uma só solução em $[a, b]$; a condição (ii) garante que a função é monótona; a condição (iii) garante que o sentido da concavidade da função não muda; e a condição (iv) garante que as tangentes à curva em $(a, f(a))$ e $(b, f(b))$ intersectam o eixo dos xx em (a, b) .

Analisemos o exemplo que se segue para ilustrar o teorema anterior, considerando $f(x) = e^{-x} - \log(x)$, $f \in C^2([1, 2])$. Note-se que:

- $f(1) = e^{-1} > 0, f(2) = e^{-2} - \log(2) < 0$;
- $f'(x) = e^{-x} - \frac{1}{x} < 0, \forall x \in [1, 2]$;
- $f''(x) = e^{-x} + \frac{1}{x^2} > 0, \forall x \in [1, 2]$;
- $\left| \frac{f(1)}{f'(1)} \right| = 0,2689... < |2 - 1| = 1$
 $\left| \frac{f(2)}{f'(2)} \right| = 0,8779... < |2 - 1| = 1$, $\forall x \in [1, 2]$;

Concluimos que a convergência ocorre para qualquer $x \in [1, 2]$.

Usando o GeoGebra vemos que a convergência é rápida, constatando-se que x_3 tem já 6 dígitos corretos.

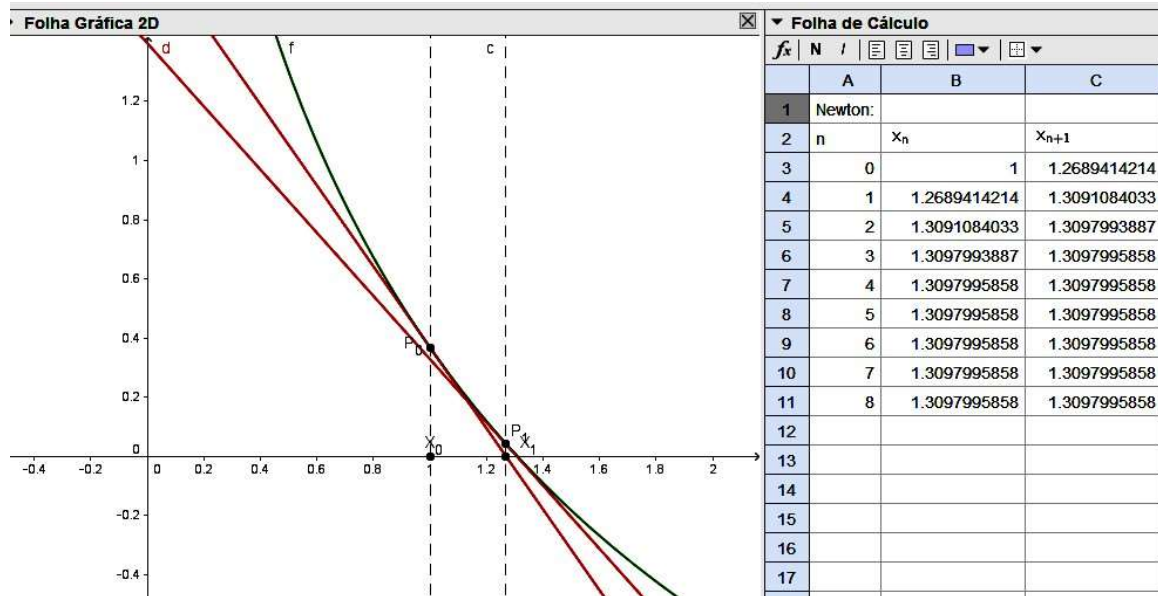


Figura 3.11

É fácil verificar que as condições do Teorema 3.1 também se verificam no caso do exemplo 3.1 qualquer que seja $x_0 \in [0; 1,3]$. Assim, seja $f(x) = x^4 - 10x^3 + 35x^2 - 50x + 24$, $f \in C^2([0; 1,3])$ verificamos que:

- $f(0) = 24 > 0, f(1,3) = -0,9639 < 0$;
- $f'(x) = 4x^3 - 30x^2 + 70x - 50 < 0, \forall x \in [0; 1,3]$;
- $f''(x) = 12x^2 - 60x + 70 > 0, \forall x \in [0; 1,3]$;
- $\left| \frac{f(0)}{f'(0)} \right| = 0,9230... < |1,3 - 0| = 1,3$
 $\left| \frac{f(1,3)}{f'(1,3)} \right| = 1,0569... < |1,3 - 0| = 1,3$, $\forall x \in [1, 2]$;

Concluimos que ocorre convergência para $\alpha = 1$, único zero de f em $[0; 1,3]$.

A partir de alguns exemplos ilustrativos do funcionamento do método várias dificuldades podem surgir e que a convergência do método de Newton-Raphson fica dependente de algumas condições. Também vimos que mesmo quando converge, não é garantido que a convergência seja rápida. Por exemplo, em certos casos é necessário que x_0 seja escolhido bastante próximo da solução (para garantir, em particular que a função derivada não se anule no intervalo definido por x_0 e α). O comportamento do método é tanto melhor quanto mais afastada de zero é o valor da função derivada nos valores iterados.

4. TEOREMA FUNDAMENTAL

Nesta secção apresentamos e ilustramos o resultado que Sebastião e Silva designa por teorema fundamental e usa para analisar a convergência dos métodos que propõe para o cálculo da menor das raízes reais positivas de equações polinomiais.

O teorema

Sejam α um número positivo qualquer, $\theta_0, \theta_1, \dots, \theta_k, \dots$ uma sucessão infinita de números tais que $0 < \theta_k < 1$, para todos os valores de k . Fazendo

$$u_1 = \theta_0 \alpha \quad e \quad \alpha_1 = \alpha - u_1;$$

em seguida

$$u_2 = \theta_1 \alpha_1 \quad e \quad \alpha_2 = \alpha_1 - u_2$$

e assim sucessivamente, fica determinada uma nova sucessão

$$u_1, u_2, \dots, u_k, \dots$$

cujos termos gerais são da forma

$$u_k = \theta_{k-1} \alpha_{k-1}, \quad \text{com} \quad \begin{aligned} \alpha_0 &= \alpha \\ \alpha_k &= \alpha_{k-1} - u_k. \end{aligned}$$

É condição necessária e suficiente para que se tenha $\sum_{k=0}^{\infty} u_k = \alpha$ que a série $\sum_{k=0}^{\infty} \theta_k$ seja divergente.

A demonstração, por redução ao absurdo, pode ser consultada no texto original [SebSilva1940]. Pela nossa parte, faremos uma interpretação do teorema fundamental no contexto da análise da convergência de um método iterativo para o cálculo da raiz α ; também apresentaremos alguns exemplos como ilustração. Finalmente, relacionaremos

as quantidades θ_k com a chamada constante de convergência assintótica, usual na teoria dos métodos iterativos.

No contexto dos métodos para o cálculo da menor das raízes positivas, o significado das quantidades presentes no enunciado anterior é o seguinte (e que se ilustra na figura 4.1):

α é a raiz da equação

u_1 é a aproximação inicial de α

u_2 é a correção feita a $S_1 = u_1$ para produzir a nova aproximação S_2 para α

u_3 é a correção feita a S_2 para produzir a nova aproximação S_3 para α

e assim sucessivamente.

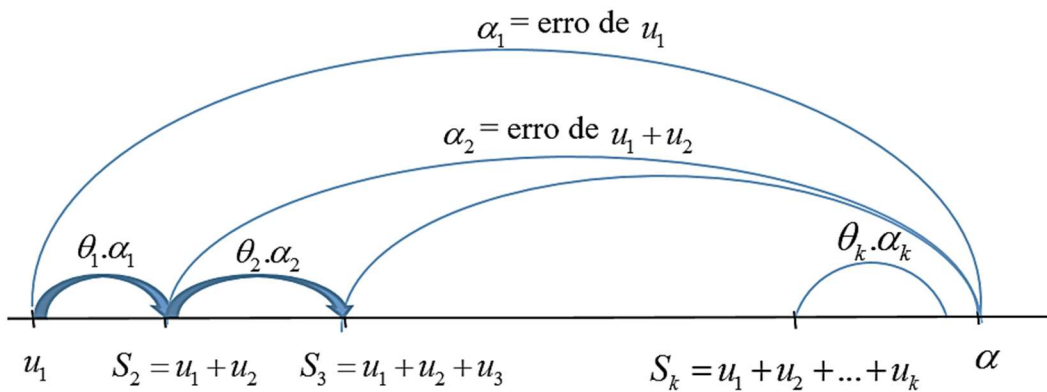


Figura 4.1

Sendo $\alpha_1 = \alpha - u_1$ o erro de u_1 como aproximação de α , calculamos a correção u_2 multiplicando este erro por θ_1 . Quanto mais próximo estiver θ_1 da unidade, melhor é a aproximação obtida.

Em geral, tendo em conta as relações

$$u_3 = \theta_2 \cdot \alpha_2$$

...

$$u_k = \theta_{k-1} \cdot \alpha_{k-1}$$

intui-se que a convergência depende da forma como evoluem as quantidades θ_k .

A k -ésima aproximação para a raiz α é $S_k = u_1 + u_2 + \dots + u_k$, logo

$$\lim_{k \rightarrow +\infty} S_k = \sum_{k=1}^{+\infty} u_k.$$

Que a série de termo geral u_k é convergente resulta do facto da sucessão de termo geral S_k (que é a sucessão associada da série) ser crescente e majorada. Com efeito, S_k é crescente por ser $S_k = S_{k-1} + u_k$ e $u_k > 0$; e é majorada por α por ser $u_k = \theta_{k-1}(\alpha - S_{k-1})$, com $\theta_{k-1} < 1$. Designemos tal limite por s . O teorema fundamental afirma que se $\sum \theta_k$ for convergente então $s < \alpha$ (isto é, o método converge para um valor inferior a α) tal como se ilustra no exemplo seguinte.

Exemplo 4.1
$$\sum_{n=1}^{+\infty} \theta_n = \sum_{n=1}^{+\infty} \left(\frac{1}{2}\right)^n$$

Neste exemplo a série $\sum_{n=1}^{+\infty} \theta_n$ converge, logo $\sum_{k=1}^{+\infty} u_k < \alpha$.

Com efeito tem-se

$$s = u_1 + u_2 + u_3 + \dots + u_n + \dots = u_1 + u_2 + \theta_2 \alpha_2 + \dots + \theta_{n-1} \alpha_{n-1} + \dots$$

e, por ser α_n decrescente

$$s < u_1 + u_2 + \alpha_2 (\theta_2 + \theta_3 + \dots + \theta_{n-1} + \dots).$$

Neste caso, sendo

$$\begin{aligned} & \theta_2 + \theta_3 + \dots + \theta_{n-1} + \dots \\ &= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \dots + \left(\frac{1}{2}\right)^{n-1} + \dots \\ &= \left(\frac{1}{2}\right)^2 \left/ \left(1 - \frac{1}{2}\right)\right. = \frac{1}{2}, \end{aligned}$$

resulta

$$s < u_1 + u_2 + \frac{1}{2} \alpha_2 < \alpha.$$

Na figura seguinte apresenta-se uma ilustração do exemplo.

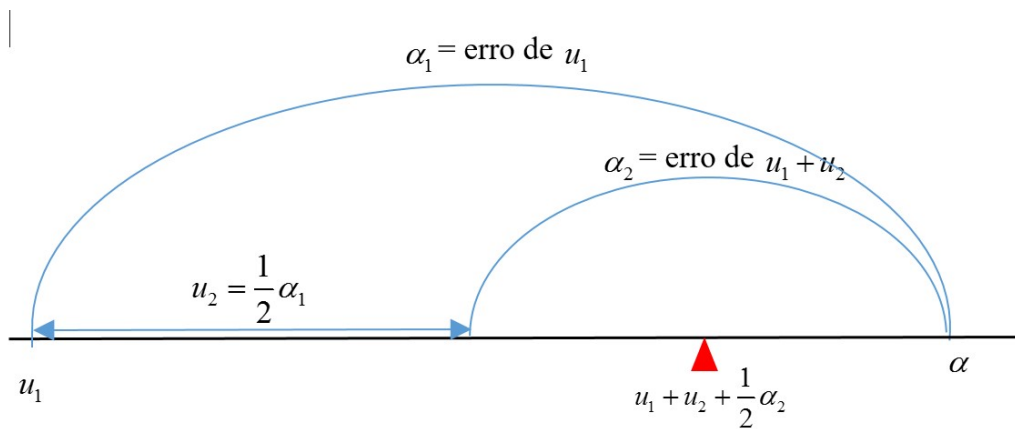


Figura 4.2

O teorema fundamental também afirma que se $\sum \theta_n$ for divergente então $S = \alpha$.

Uma condição suficiente, mas não necessária, para que tal aconteça é que θ_n não tenda para zero. Em particular, será $S = \alpha$ se existir $m < 1$, constante positiva, tal que $\theta_n \geq m$. Como ilustração disto, Sebastião e Silva refere o princípio em que se baseia a prova de convergência do método de exaustão utilizado pelos gregos e que corresponde a $m=1/2$.³

De acordo com o que se disse antes, será $S = \alpha$ ainda que $\lim_{n \rightarrow \infty} \theta_n = 0$ desde que $\sum_{n=1}^{+\infty} \theta_n$ seja divergente. No exemplo seguinte ilustramos este caso.

³ Os gregos procuraram determinar a área $a(S)$ de uma figura S (por exemplo, de um círculo) através de uma sequência P_1, P_2, \dots, P_n , de polígonos que preenchem ou "exaurem" S . O chamado método de exaustão, usualmente atribuído a Eudoxus (408–355 BC) e desenvolvido por Arquimedes (287 a.C. – 212 a.C.), foi usado para demonstrar com rigor que a área $a(S-P_n)$ compreendida entre S e P_n é tão pequena quanto se queira, bastando para tal tomar n suficientemente grande. Para este fim serviu o axioma de Arquimedes-Eudoxus que, em linguagem moderna, pode ser escrito da maneira seguinte: sejam M_0 e \mathcal{E} duas dadas quantidades e M_1, M_2, M_3, \dots uma sequência tal que $M_1 < M_0/2$, $M_2 < M_1/2$, $M_3 < M_2/2$, etc. Então $M_n < \mathcal{E}$ para algum n . Ver, por exemplo [Edwards1979], p.16-17, para uma aplicação deste princípio à exaustão do círculo.

Exemplo 4.2 Com $\theta_n = \frac{1}{n+2}$ obtém-se

$$\begin{cases} u_n = \frac{\alpha}{n(n+1)} \\ S_n = \alpha \left(1 - \frac{1}{n+1} \right) \end{cases} \quad (4.1)$$

logo $\lim_{n \rightarrow +\infty} S_n = \alpha$.

Provemos as relações (4.1) usando o método de indução. É fácil verificar que são

verdadeiras para $n = 1$. Admitindo (hipótese de indução) que se tem $\begin{cases} u_n = \frac{\alpha}{n(n+1)} \\ S_n = \alpha \left(1 - \frac{1}{n+1} \right) \end{cases}$

provemos que $\begin{cases} u_n = \frac{\alpha}{n(n+1)} \\ S_n = \alpha \left(1 - \frac{1}{n+1} \right) \end{cases}$.

Ora,

$$\begin{aligned} u_n &= \theta_{n-1} \cdot \alpha_{n-1} \\ &= \frac{1}{n+1} (\alpha - S_{n-1}) \\ &= \frac{1}{n+1} \left[\alpha - \alpha \left(1 - \frac{1}{n} \right) \right] \\ &= \frac{\alpha}{n+1} \left[1 - \left(1 - \frac{1}{n} \right) \right] \\ &= \frac{\alpha}{n(n+1)} \end{aligned}$$

e

$$\begin{aligned} S_n &= S_{n-1} + u_n = \alpha \left(1 - \frac{1}{n} + \frac{1}{n(n+1)} \right) \\ &= \alpha \left(1 - \left(\frac{n+1}{n(n+1)} - \frac{1}{n(n+1)} \right) \right) \\ &= \alpha \left(1 - \frac{1}{n+1} \right) \end{aligned}$$

Ficaram assim provadas, por indução, as relações (4.1). Neste exemplo a série $\sum_{n=0}^{+\infty} \theta_n$

diverge e $\sum_{k=1}^{+\infty} u_k$ converge para α (ilustração na figura 4.3).

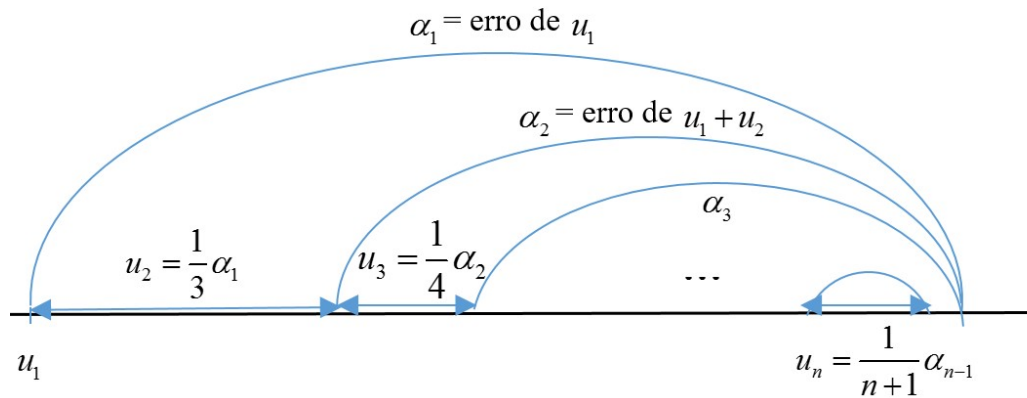


Figura 4.3

Sebastião e Silva usou este exemplo com $\alpha = 1$ para concluir, aplicando o seu teorema fundamental, que a série harmônica é divergente.

Parece-nos oportuno agora interpretar o significado das quantidades θ_n no contexto da definição dada em (3.3) Sendo

$$\alpha_n = \alpha_{n-1} - u_n = \alpha_{n-1} - \theta_{n-1} \alpha_{n-1} = \alpha_{n-1} (1 - \theta_{n-1})$$

resulta

$$\frac{\alpha_n}{\alpha_{n-1}} = 1 - \theta_{n-1},$$

o que mostra que com $\lim_{n \rightarrow +\infty} \theta_n < 1$ a convergência é apenas linear ($p = 1$) e a constante de convergência assintótica é afinal

$$c = 1 - \lim_{n \rightarrow +\infty} \theta_n. \quad (4.2)$$

Fica desta maneira estabelecida a relação entre as quantidades θ_n presentes no teorema fundamental de Sebastião e Silva e a constante de convergência assintótica. O

caso de ser $\lim_{n \rightarrow +\infty} \theta_n = 1$ tem interesse particular porque resulta $c = 0$. Como ilustração, temos:

Exemplo 4.3 $\theta_n = 1 - \frac{1}{n+2}$

Tem-se

$$\frac{\alpha_n}{\alpha_{n-1}} = \frac{1}{n+1},$$

ou seja

$$\alpha_n = \frac{\alpha_{n-1}}{n+1}.$$

Por exemplo, $\alpha_9 = \frac{\alpha_8}{10}$, isto é o erro na 9ª iteração é 10 vezes menor que o erro na iteração anterior. Embora também neste caso seja $p = 1$, a convergência (linear) é muito mais rápida do que no caso do Exemplo 4.2. Observe-se que a partir de (4.2) é imediato concluir que é $c = 0$ e $c = 1$ nos Exemplos 4.3 e 4.2, respetivamente.

5. MÉTODO DE JOSÉ SEBASTIÃO E SILVA

Nesta secção estudamos o método apresentado no parágrafo 3 de [SebSilva1940] para o cálculo da menor raiz α do polinómio p , mónico, de grau m ,

$$p_m(x) = x^m + b_1x^{m-1} + \dots + b_{m-1}x + b_m \quad (5.1)$$

cujas raízes, que se assumem serem todas reais e positivas, satisfazem

$$\alpha = a_1 < a_2 \leq \dots \leq a_{m-1} \leq a_m.$$

Definindo

$$S'_1 = \frac{1}{\alpha} + \frac{1}{a_2} + \dots + \frac{1}{a_m},$$

isto é, a soma dos inversos das raízes, resulta

$$\frac{\alpha}{m} < \frac{1}{S'_1} < \alpha. \quad (5.2)$$

Tem-se que

$$u_1 = \frac{1}{S'_1} = \frac{\alpha}{1 + \frac{\alpha}{a_2} + \dots + \frac{\alpha}{a_m}} \quad (5.3)$$

é uma aproximação de α , e usando (5.2) podemos escrever

$$u_1 = \theta_0 \alpha \quad (5.4)$$

com $\frac{1}{m} < \theta_0 < 1$. Observe-se que é

$$u_1 = -\frac{b_m}{b_{m-1}} \quad (5.5)$$

por ser S'_1 a soma das raízes do polinómio

$$x^m \cdot p_m\left(\frac{1}{x}\right) = b_m x^m + b_{m-1} x^{m-1} + \dots + b_1 x + 1$$

que são $\frac{1}{\alpha} > \frac{1}{a_2} \geq \frac{1}{a_3} \geq \dots \geq \frac{1}{a_m}$. Se for α muito menor do que a_2 , então $-\frac{b_{m-1}}{b_m}$ será uma

boa aproximação para $\frac{1}{\alpha}$ uma vez que

$$-\frac{b_{m-1}}{b_m} = \frac{1}{\alpha} \left(1 + \sum_{i=2}^m \frac{\alpha}{a_i} \right).$$

Da igualdade anterior resulta

$$\begin{aligned} -\frac{b_m}{b_{m-1}} &= \alpha \left(1 + \sum_{i=2}^m \frac{\alpha}{a_i} \right)^{-1} \\ &= \alpha (1 - \rho + \rho^2 - \rho^3 + \dots) \\ &= \alpha \frac{1}{1 + \rho} \end{aligned} \quad (5.6)$$

com

$$\rho = \sum_{i=2}^m \frac{\alpha}{a_i}.$$

Portanto, quando se toma $-\frac{b_m}{b_{m-1}}$ para aproximar α comete-se um erro relativo que é igual

a $\frac{\rho}{1+\rho}$. Parece-nos relevante relacionar o valor de ρ com o valor de θ_0 : de (5.4) - (5.6)

resulta $\theta_0 = \frac{1}{1+\rho}$.

Importa pois efetuar transformações sobre o polinómio inicial que torna pequenos os valores de ρ . Uma maneira efetiva de conseguir isto é usar mudanças de origem. Com efeito, se u_1 é uma boa aproximação de α , o polinómio $q(x) = p(x+u_1)$ tem as raízes $a_i - u_1$ e as razões $\frac{\alpha - u_1}{a_i - u_1}$ estarão próximas de zero.

Embora Sebastião e Silva não o faça, apresentaremos em detalhe um algoritmo para levar a cabo a transformação algébrica do polinómio correspondente à mudança de variável.

Dado o polinómio em (5.1) e $u_1 \neq 0$, pretendemos determinar os coeficientes c_1, c_2, \dots, c_m tais que o polinómio

$$q_m(x) = x^m + c_1 x^{m-1} + c_2 x^{m-2} + \dots + c_{m-1} x + c_m \quad (5.7)$$

satisfaz $q_m(x) = p_m(x+u)$. Escrevendo

$$p_m(x+u) = (x+u)^m + b_1(x+u)^{m-1} + b_2(x+u)^{m-2} + \dots + b_{m-1}(x+u) + b_m$$

na chamada “forma encaixada”

$$p_m(x+u) = \left(\dots \left(\left((x+u) + b_1 \right) (x+u) + b_2 \right) (x+u) + \dots + b_{m-1} \right) (x+u) + b_m$$

definimos de forma recursiva os polinómios $q_i(x)$, $i = 1, \dots, m$, com $q_0(x) = 1$ e

$$q_i(x) = q_{i-1}(x) \cdot (x+u) + b_i. \quad (5.8)$$

Com

$$q_{i-1}(x) = x^{i-1} + b_1^{(i-1)} x^{i-2} + \dots + b_{i-2}^{(i-1)} x + b_{i-1}^{(i-1)}$$

e

$$q_i(x) = x^i + b_1^{(i)} x^{i-1} + \dots + b_{i-2}^{(i)} x^2 + b_{i-1}^{(i)} x + b_i^{(i)},$$

de (5.8) resulta

$$\begin{aligned} b_1^{(i)} &= b_1^{(i-1)} + u \\ b_2^{(i)} &= b_2^{(i-1)} + u \cdot b_1^{(i-1)} \\ &\dots = \dots \\ b_{i-1}^{(i)} &= b_{i-1}^{(i-1)} + u \cdot b_{i-2}^{(i-1)} \\ b_i^{(i)} &= b_i + u \cdot b_{i-1}^{(i-1)} \end{aligned} \quad (5.9)$$

Efetuando estes cálculos, para cada i desde 1 até m , obtemos os coeficientes

$$c_k = b_k^{(m)}, \quad k = 1, \dots, m$$

do polinómio q dado em (5.7).

Na verdade, para obter a correção $u_2 = -\frac{c_m}{c_{m-1}}$ é necessário calcular, para cada

$i = 1, \dots, m$, apenas os coeficientes $b_{i-1}^{(i)}$ e $b_i^{(i)}$, de acordo com as duas últimas expressões em (5.9), o que reduz drasticamente o número de operações aritméticas envolvidas. Para obter os dois últimos termos do polinómio resultante da nova mudança de variável, usamos a transformação expressa em (5.9) a partir do polinómio original substituindo u_1 por $u_1 + u_2$.

Exemplo 5.1

Sendo

$$p_4(x) = x^4 - 10x^3 + 35x^2 - 50x + 24$$

de (5.5) temos a aproximação $u_1 = \frac{24}{50} = \frac{12}{25}$ para a menor das raízes⁴ de p_4 que resulta

de desprezar em (5.3), a quantidade

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{4}.$$

Substituindo x por $x + \frac{12}{25}$ (o que corresponde a fazer uma deslocação das raízes para a origem) e usando as relações (5.9), obtemos

$$\begin{aligned}q_1(x) &= x - \frac{238}{25} \\q_2(x) &= x^2 - \frac{226}{25}x + \frac{2404}{79} \\q_3(x) &= x^3 + \dots + \frac{8010}{307}x - \frac{2159}{61} \\q_4(x) &= x^4 + \dots - \frac{14911}{652}x + \frac{4396}{627}.\end{aligned}$$

Procedendo de forma análoga à usada para o cálculo de u_1 , aproximamos a menor das raízes de q_4 por

$$u_2 = \frac{4396}{627} \cdot \frac{652}{14911} \approx 0.31$$

Na tabela seguinte apresentamos os valores de u_2, \dots, u_8 e os últimos dois termos dos correspondentes polinómios $q_4(x)$ obtidos.

⁴ As raízes de p_4 são 1, 2, 3 e 4.

k	u_k	$q_k(x)$
1	0,48	$x^4 + \dots - \frac{14911}{652}x + \frac{4396}{627}$
2	0,30657092	$x^4 + \dots - 11,55426367x + 1,84205383$
3	0,15942632	$x^4 + \dots + 7,24118380x + 0,35704920$
4	0,04930812	$x^4 + \dots - 6,10367886x + 0,02841080$
5	0,00465470	$x^4 + \dots - 6,00087833x + 0,00023955$
6	0,00003992	$x^4 + \dots - 6,00000006x + 0,00000001$
7	0,0000000029	$x^4 + \dots - 5,99999999x + 0,00000000$
8	0,0000000000	$x^4 + \dots - 5,99999999x + 0,00000000$

Tabela 5.1

Na notação do teorema fundamental, a menor das raízes é aproximada por

$$S_8 = \sum_{k=1}^8 u_k = 1,0000000000\dots$$

Este método, como também o reconhece o seu autor, é afinal o método de Newton – Rapshon⁵. Com efeito, para o polinómio dado em (5.1), a aproximação (5.5) é afinal o valor $0 - \frac{p_m(0)}{p'_m(0)}$ que corresponde ao valor obtido pelo método de Newton a partir de uma aproximação inicial igual a zero. O método de Sebastião e Silva usa mudanças de variável que aproximam a menor das raízes positivas da origem das coordenadas e produz sucessivas correções u_1, u_2, \dots, u_k . Na k –ésima iteração, a aproximação obtida para a raiz é

$$S_k = u_1 + u_2 + \dots + u_k.$$

Conjugando a notação usada no método de Sebastião e Silva com a fórmula iterativa do método de Newton – Rapshon teremos

$$S_{k+1} = S_k - \frac{p_m(S_k)}{p'_m(S_k)}, k = 0, 1, \dots.$$

⁵ Embora Sebastião e Silva o designe por método de Newton – Fourier.

Consideremos novamente o polinómio $p_4(x) = x^4 - 10x^3 + 35x^2 - 50x + 24$ cuja representação gráfica se apresenta de seguida:

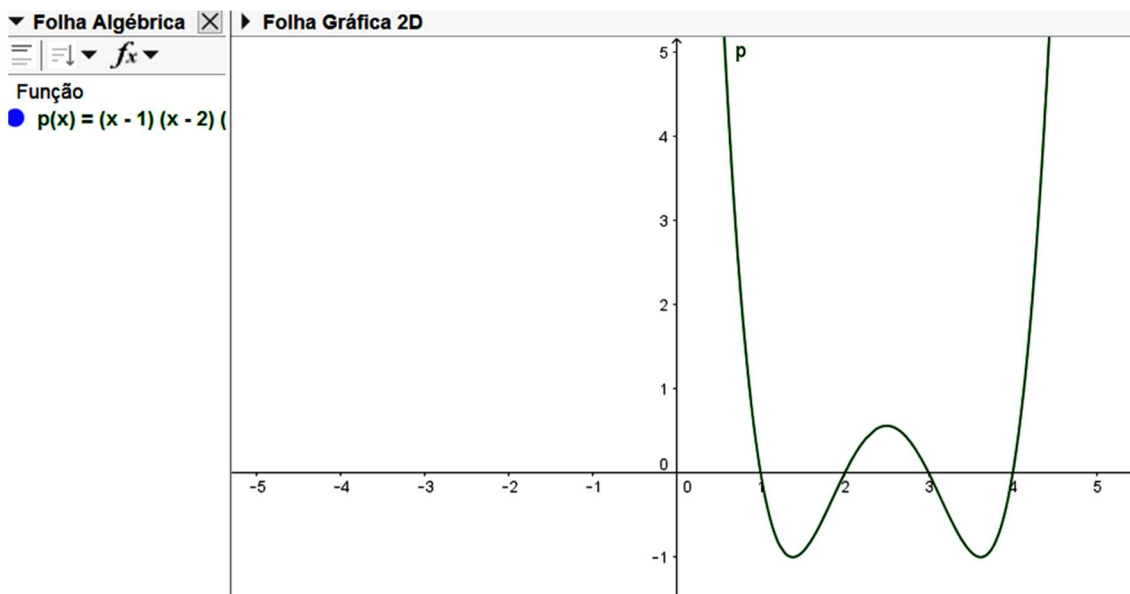


Gráfico 5.1

Tomando como aproximação inicial $u_1 = \frac{12}{25} = 0.48$ para a menor das raízes de

p_4 e efetuando a mudança de variável $y = x + 0.48$, obtemos um novo polinómio

$$P_4(x) = (x + 0.48)^4 - 10(x + 0.48)^3 + 35(x + 0.48)^2 - 50(x + 0.48) + 24$$

cujas representações gráficas são:

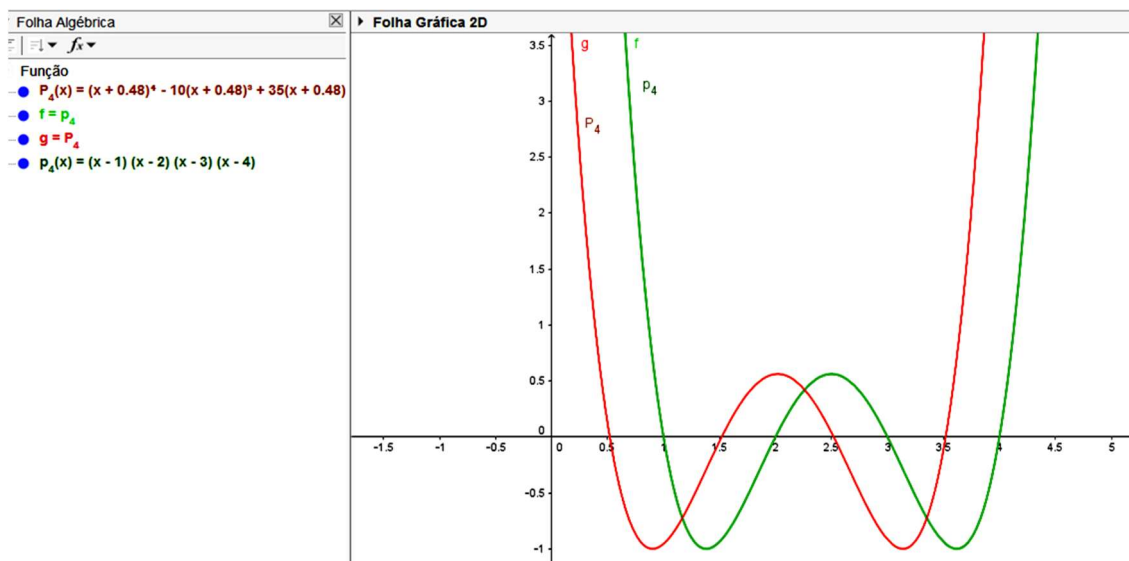


Gráfico 5.2

Observando os gráficos de $p_4(x)$ e $P_4(y)$ vemos que $p_4(0.48) = P_4(0)$ e $p'_4(0.48) = P'_4(0)$.

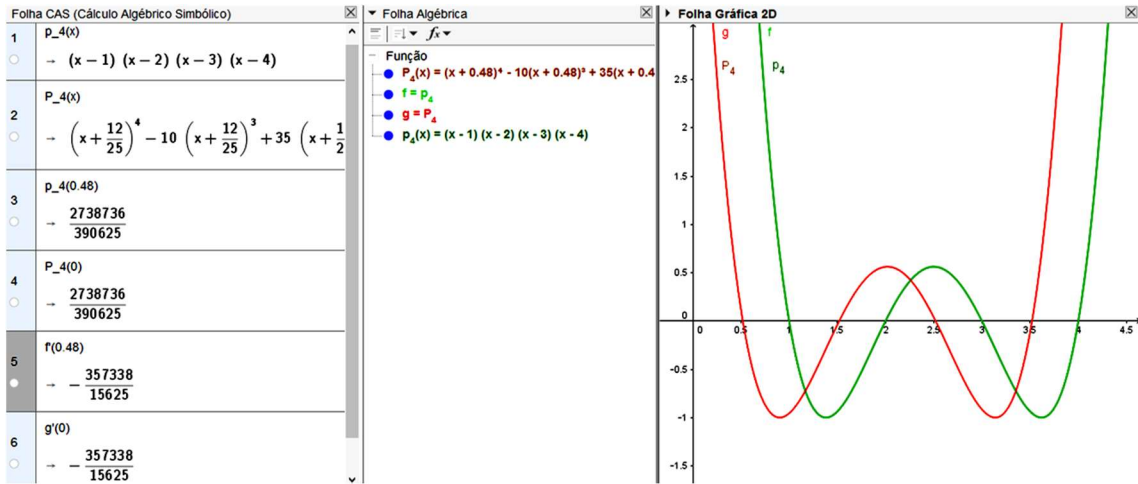


Gráfico 5.3

De seguida, após o cálculo de u_2 , tomando agora como aproximação inicial $u_1 + u_2$, efetuamos a mudança de variável

$$y = x + (0.48 + 0.31)$$

$$\Leftrightarrow y = x + 0.79$$

e obtemos um novo polinómio, seja

$$Q_4(x) = (x + 0.79)^4 - 10(x + 0.79)^3 + 35(x + 0.79)^2 - 50(x + 0.79) + 24$$

cuja representação gráfica é a que se segue. Note-se que

$$p_4(0.79) = Q_4(0) \text{ e } p'_4(0.79) = Q'_4(0).$$

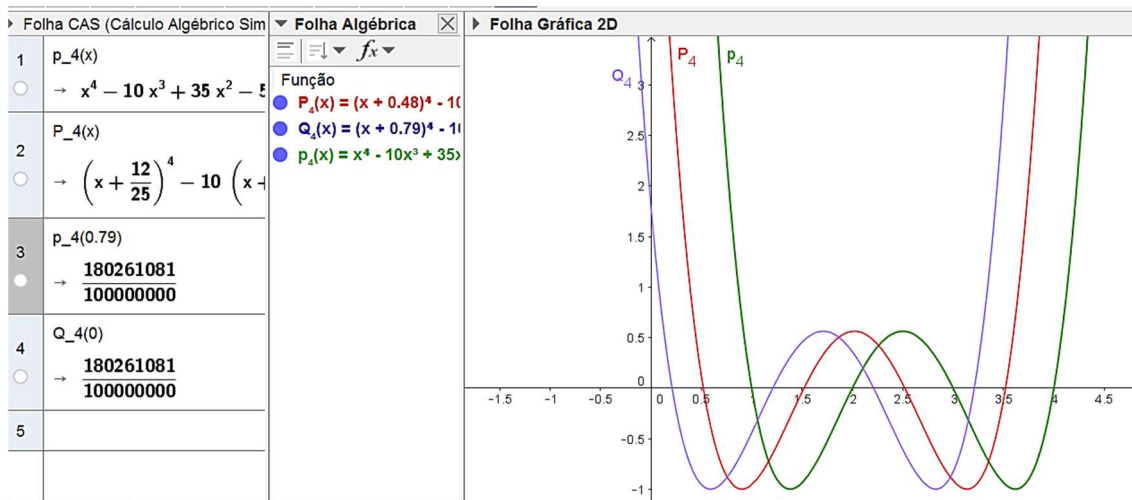


Gráfico 5.4

Continuando a efetuar mudanças de variável deslocando assim as raízes dos novos polinómios para a origem, verificamos que o método de Sebastião e Silva não é mais do que o método de Newton – Raphson usando sempre o zero como aproximação inicial.

Comparemos agora os métodos de Newton – Raphson e de Sebastião e Silva (sem a Transformada de Graeffe) em termos de eficiência computacional. Em (5.9) está indicada a forma de cálculo dos coeficientes $b_j^{(i)}$, $j=1, \dots, i$, do polinómio q_i de grau i dado em (5.8). São

$$1 + 2 + \dots + m = \frac{m+1}{2} \times m$$

coeficientes cuja determinação requer exatamente

$$\begin{aligned} & 1 + 3 + 5 + \dots + (2m - 1) \\ &= [1 + (2m - 1)] \times \frac{m}{2} \\ &= m^2 \end{aligned}$$

operações aritméticas.

Este é um preço elevado a pagar por cada iteração do método mas, como se observou antes, não é necessário calcular todos os coeficientes $b_j^{(i)}$ na sequência de polinómios

$$\begin{aligned}
q_1(x) &= x + b_1^{(1)} \\
q_2(x) &= x^2 + b_1^{(2)}x + b_2^{(2)} \\
q_3(x) &= x^3 + b_1^{(3)}x^2 + b_2^{(3)}x + b_3^{(3)} \\
&\dots = \dots \\
q_m(x) &= x^m + b_1^{(m)}x^{m-1} + b_2^{(m)}x^{m-2} + \dots + b_{m-1}^{(m)}x + b_m^{(m)}
\end{aligned}$$

mas apenas os coeficientes $b_1^{(1)}, b_2^{(2)}, \dots, b_m^{(m)}$ e $b_1^{(2)}, b_2^{(3)}, \dots, b_{m-1}^{(m)}$ através das relações

$$\begin{aligned}
b_1^{(1)} &= b_1 + u \\
b_2^{(2)} &= b_2 + ub_1^{(1)} \\
&\dots = \dots \\
b_m^{(m)} &= b_m + ub_{m-1}^{(m-1)}
\end{aligned} \tag{5.10}$$

e

$$\begin{aligned}
b_1^{(2)} &= b_1^{(1)} + u \\
b_2^{(3)} &= b_2^{(2)} + ub_1^{(2)} \\
&\dots = \dots \\
b_{m-1}^{(m)} &= b_{m-1}^{(m-1)} + ub_{m-2}^{(m-1)}
\end{aligned} \tag{5.11}$$

que resultam de (5.9) e requerem $4(m-1)$ operações aritméticas.

O cálculo de $p_m(u)$ faz-se usualmente através do conhecido método de Horner (ver, por exemplo, [Valença1993], p.53-54). De

$$\begin{aligned}
p_m(x) &= x^m + b_1x^{m-1} + \dots + b_{m-1}x + b_m \\
&= \left(\dots \left((x + b_1)x + b_2 \right) x + \dots + b_{m-1} \right) x + b_m
\end{aligned}$$

resulta, denotando por p_i , para cada $i = 1, \dots, m-1$, o polinómio de grau i envolvido na relação anterior,

$$\begin{aligned}
p_1(u) &= b_1 + u \\
p_2(u) &= u.p_1(u) + b_2 \\
&\dots = \dots \\
p_m(u) &= u.p_{m-1}(u) + b_m
\end{aligned}$$

Estas relações coincidem com as dadas em (5.10) com $p_i(u) = b_i^i, i = 1, \dots, m$. A forma mais eficiente para calcular o valor de $p_m'(u)$ (ver [Valença1993], p.54-55) consiste em escrever

$$p_m(x) = (x-u)t_{m-1}(x) + r \quad (5.12)$$

donde resulta

$$p_m'(u) = t_{m-1}(u) \quad (5.13)$$

Com

$$t_{m-1}(x) = x^{m-1} + d_1x^{m-2} + d_2x^{m-3} + \dots + d_{m-2}x + d_{m-1}$$

tem-se

$$\begin{aligned} d_1 &= b_1 + u \\ d_2 &= b_2 + u.d_1 \\ &\dots = \dots \\ d_{m-1} &= b_{m-1} + u.d_{m-2} \\ r &= b_m + u.d_{m-1}. \end{aligned}$$

Com esta formulação, é $d_i = p_i(u), i = 1, \dots, m-1$ e $r = p_m(u)$. O método de Horner pode agora ser usado para calcular $p_m'(u) = t_{m-1}(u)$. Escrevendo

$$\begin{aligned} t_{m-1}(x) &= x^{m-1} + d_1x^{m-2} + d_2x^{m-3} + \dots + d_{m-2}x + d_{m-1} \\ &= \left(\dots \left((x + d_1)x + d_2 \right) x + \dots + d_{m-2} \right) x + d_{m-1} \end{aligned}$$

resulta, denotando por t_i , para cada $i = 1, \dots, m-2$, o polinómio de grau i envolvido na relação anterior,

$$\begin{aligned} t_1(u) &= d_1 + u \\ t_2(u) &= u.t_1(u) + d_2 \\ &\dots = \dots \\ t_{m-1}(u) &= u.t_{m-2}(u) + d_{m-1}. \end{aligned}$$

Uma vez que é $d_i = b_i^{(i)}$, as relações anteriores coincidem com as dadas em (5.11) com $t_i(u) = b_i^{(i+1)}, i = 1, \dots, m-1$. Portanto, a formulação que encontramos para a implementação eficiente do método de Sebastião e Silva não é mais nem menos do que o método mais eficiente para calcular os valores de um dado polinómio e da sua derivada num ponto.

Com isto fica clara a completa equivalência, não apenas matemática mas também computacional, dos dois métodos.

Mas Sebastião e Silva também propõe que o seu método, que acabamos de concluir ser afinal equivalente ao método de Newton – Raphson, possa incluir a transformada de Graeffe para aceleração da convergência. Não nos alongaremos sobre a análise do método com esta transformação. Fazemos apenas notar que sendo quadrática a convergência do método à partida, é difícil perceber qual a vantagem da proposta aceleração.

6. CONCLUSÃO

José Sebastião e Silva foi um dos mais notáveis matemáticos portugueses do século XX. Além de deixar uma obra importantíssima como investigador, preocupou-se profundamente com o ensino da Matemática, escrevendo livros que se tornaram bem conhecidos e são hoje recordados por muitos professores. Sebastião e Silva, revelou ter preocupações pedagógicas e didáticas no que se refere à apresentação dos resultados e suas demonstrações, como se pode comprovar em [SebSilva1940].

Foi autor de importantes trabalhos a nível científico. Em particular, propôs um método para a resolução de equações algébricas que teve reconhecimento internacional. Este trabalho foi o culminar da investigação feita por Sebastião e Silva nesta temática. Com efeito, a sua primeira publicação em 1940 já tratava do método de resolução de equações algébricas. Neste artigo propôs um método que embora original nas ideias, é afinal, como ele próprio o reconhece, o popular método de Newton – Raphson.

Sebastião e Silva também considera a possibilidade de acelerar o método através de transformadas de Graeffe. Porém, atendendo à convergência quadrática do método de Newton, a utilização da transformada de Graeffe parece-nos a nós poder contribuir para a melhoria do método apenas nas iterações iniciais. Isto é interessante porque é bem sabido que o sucesso do método de Newton depende muitas vezes de uma boa escolha da aproximação inicial. O estudo de uma combinação eficiente dos dois métodos está, porém, para além do âmbito desta tese.

Terminamos com um pensamento que deverá acompanhar sempre um professor na sua vida profissional, uma vez que o seu papel principal como agente educativo é o de conduzir o aluno a pensar, a raciocinar, a investigar, a conjecturar, a experimentar e a criticar:

...

O que faço, aprendo.

Confúcio 41⁶

⁶ Filósofo chinês (c.551 a.C. – c. 479 a.C.).

BIBLIOGRAFIA

[Almeida1974] Costa, António Almeida, “*Elogio histórico de José Sebastião e Silva*”, In *Memórias da Academia das Ciências de Lisboa*, Classe de Ciências, Tomo XVII, Lisboa, 1974.

[CarvSilva2015] J. Carvalho e Silva, *O pensamento pedagógico de José Sebastião e Silva* <http://www.mat.uc.pt/~jaimecs/pessoal/sebsilva.html> (acedido em 30 outubro 2015).

[Edwards1979] Edwards, C. H., Jr. *The Historical Development of the Calculus*, Springer-Verlag, New York, 1979.

[HomSebSilva1997] *Homenagem a José Sebastião e Silva - Atas do colóquio de Homenagem*. Torre do Tombo : s.n., dezembro de 1997.

[Householder1959] Alston Scott Householder: *Dandelin, Lobačevskiĭ, or Graeffe?* Amer. Math. Monthly, 66 (1959), pp. 464–466.

[Householder1971] A S. Householder, “Generalizations of an algorithm of Sebastião e Silva”, *Numerische Mathematik*, Vol. 16 Fasc. 4, 375-382 (1971).

[Householder1973] A S. Householder, Generalizations of an algorithm of Sebastião e Silva, *Numerische Mathematik*, Vol. 20, 205-207 (1973).

[Nordgaard1922] M. A. Nordgaard, A Historical Survey of Algebraic Methods of Approximating the Roots of Numerical Higher Equations up to the year 1819, New York City: Teachers College, Columbia University, 1922.

[SebSilva1940] J. Sebastião e Silva, *Da resolução numérica de equações algébricas*, Portugaliae Mathematica, Vol. 1 Fasc. 4, 303-332 (1940).

[SebSilva1941] J. Sebastião e Silva, *Sur une méthode d'approximation semblable à celle de Gräffe*, Portugaliae Mathematica, Vol. 2, 271-279 (1941).

[SebSilva1946] J. Sebastião e Silva, *Complementi al metodo di Gräffe per la risoluzione delle equazioni algebriche, Nota I; Nota II*. Rendiconti, Accademia Nazionale dei Lincei, serie VIII, Vol. I Fasc. 3-4-5, 335-343; 547-551 (1946).

[SebSilva1997] Silva, José Sebastião e, *Guia para a utilização do compêndio da Matemática (2º e 3º volumes)*, Lisboa: GEP,ME, 1997.

[Valença1993] Valença, Maria Raquel, *Métodos Numéricos*, Livraria Minho, 3ª Edição, 1993.

[WeissteinMathWorld] Weisstein, Eric W. "*Graeffe's Method*." Resource: <http://mathworld.wolfram.com/GraeffesMethod.html>; acedido em 1 de abril de 2016.

