Sobre aproximações polinomiais de raízes reais de cúbicas

Edgar Rechtschaffen

UNIFESO

O problema de se obter as raízes de uma equação do terceiro grau, abreviadamente uma *cúbica*, foi resolvido no século XVI da era atual. Uma abordagem detalhada do problema, precedida de fascinante apanhado histórico dos desenvolvimentos pertinentes, está apresentada por Elon L. Lima em [2]. Esses desenvolvimentos culminam com as conhecidas fórmulas de Cardano-Tartaglia. Ocorre que estas fórmulas não são de cunho prático para o caso das cúbicas com três raízes reais, pois elas dependem do cálculo de funções trigonométricas inversas, tais como o arcocosseno ou a arcotangente, que requerem o emprego de métodos iterativos, como enfatizado em [1] e [5].

Em recente artigo ([4]), a partir de uma *forma canônica* comum às cúbicas, apresentei duas fórmulas polinomiais explícitas que determinam, com extraordinária precisão (módulo do erro relativo inferior a 3×10^{-6}), a raiz intermediária ou a raiz inferior de uma cúbica com três raízes reais. São as *quase-soluções*.

O objetivo desta nota é apresentar e justificar essas fórmulas polinomiais, com o uso do cálculo diferencial. Se os polinômios forem truncados para apenas ordem 3, sacrifica-se um pouco da precisão mas ganha-se com a simplicidade, permitindo a obtenção de quase-soluções de cúbicas com o uso de dois polinômios cúbicos. Essas quase-soluções podem ser usadas como condição inicial para algum método de aproximação numérica como o Método de Newton ou o Método de Newton Estendido, resultando em excelente aproximação já na primeira iteração.

Forma canônica

Estamos interessados na cúbica $Ay^3 + By^2 + Cy + D$, com $A \neq 0$, que, após divisão dos dois lados por A , passa à forma

$$y^3 + ay^2 + by + c\,, \tag{1}$$

que tem as mesmas raízes. Se em (1) aplicarmos a translação $y = x - \frac{a}{3}$ obteremos um novo polinômio,

$$h(x) = x^3 + px + q\,, \tag{2}$$

conhecido como *forma deprimida*, em que $p = b - 3s^2$, $q = c + s(p + s^2)$, para $s = -\frac{a}{3}$. A forma deprimida tem a propriedade de que o único ponto de inflexão da cúbica se localiza na origem. Para se obterem as raízes de (1), basta subtrair $\frac{a}{3}$ das raízes de (2).

Sendo $h'(x) = 3x^2 + p$, h não tem ponto crítico se $p > 0$, tem um ponto crítico (em $x = 0$) se $p = 0$ e, se $p < 0$, tem dois pontos críticos, em $x_M = -\sqrt{\frac{|p|}{3}}$ e $x_m = +\sqrt{\frac{|p|}{3}}$, que são, respectivamente, máximo e mínimo locais de h. Se $p \geq 0$ então h só tem uma raiz. Se $p < 0$ então h pode ter 1, 2 ou 3 raízes, dependendo se

$$f(x_M)f(x_m) = 4\left(\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3\right)$$

é positivo, nulo ou negativo, respectivamente.

Em todos os casos em que o chamado *discriminante* $\Delta = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$ é maior ou igual a zero (que inclui automaticamente o caso $p \geq 0$), pode-se empregar a fórmula de Cardano-Tartaglia para as raízes de h,

$$\sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}} \tag{3}$$

apenas usando-se algoritmos para o cálculo de raízes quadradas e cúbicas de números reais positivos (se $m < 0$ então toma-se $\sqrt[3]{m} = -\sqrt[3]{|m|}$). No caso em que $p < 0$ e $\Delta = 0$, em que se tem duas raízes reais, a fórmula dá a raiz simples, enquanto a raiz dupla é o ponto crítico de sinal contrário.

O maior problema no uso da fórmula de Cardano-Tartaglia ocorre quando $\Delta < 0$ (*casus irreducibilis*), em que $\sqrt{\Delta}$ é um número imaginário e necessita-se calcular raízes cúbicas de números complexos. Na prática, isso requer escrever o número complexo na sua forma polar

$$Re^{i\theta} = R(\cos(\theta) + i\mathrm{sen}(\theta))\,,$$

em que θ só se determina com o uso de funções trigonométricas inversas, o que limita seu uso e pode ter um custo computacional muito grande, para um grande volume de cálculos desse tipo (ver [1] e [5]).

Aqui proporemos fórmulas polinomiais para o cálculo das três raízes reais de $h(x)$ (e portanto do polinômio cúbico original) quando $\Delta < 0$, de forma aproximada mas direta, sem passagem aos números complexos. Para tanto, faremos a mudança de coordenadas

$$w = \frac{x}{x_m}\,,$$

que força a que o máximo e o mínimo locais se situem em -1 e $+1$, respectivamente. Dessa mudança de coordenadas, seguida da divisão da função por x_m^3, resulta a equação cúbica

$$f(w) = w^3 - 3w + 2\alpha = 0\,, \qquad (4)$$

em que

$$\alpha = \frac{q}{2}\left(\frac{|p|}{3}\right)^{-3/2}\,,$$

que em [4] cunhamos como *forma canônica*.

A forma canônica pode ser definida também para $p = 0$ (neste caso basta tomar $2\alpha = q$) ou para $p > 0$, definindo-se $x_m = \sqrt{\frac{|p|}{3}}$ (mesmo que x_m não seja ponto crítico). Neste caso, obtém-se $f(w) = w^3 + 3w + 2\alpha = 0$.

Quando $\Delta < 0$, tem-se $p < 0$ (logo $p = -|p|$), que implica

$$|\alpha| < 1\,.$$

Como consequência, o caso de três raízes reais fica não só reduzido a apenas um parâmetro, como esse parâmetro varia dentro de um intervalo limitado. É isso que permitirá a construção de fórmulas universais de aproximação das raízes.

Geometria da forma canônica

Como exposto acima, teceremos algumas observações sobre a forma canônica dada em (4), para $|\alpha| < 1$ (de fato, $\alpha \leq 1$, pois a análise se estende aos extremos). A equação (4) pode ser reescrita como

$$\alpha = \frac{1}{2}\left(-w^3 + 3w\right) \equiv F(w)\,.$$

A função F tem a seguinte interpretação: se quisermos que um determinado w seja uma raiz, basta fixarmos $\alpha = F(w)$. Por outro lado, dado um α, as raízes da forma canônica são os valores w tais que $F(w) = \alpha$.

O gráfico de F está mostrado na figura, no intervalo $[-2, +2]$. A função F tem raízes em $-\sqrt{3}$, 0 e $+\sqrt{3}$, pontos críticos -1 e $+1$ com valores críticos -1 e $+1$, respectivamente, e vale $+1$ e -1 em -2 e $+2$, respectivamente. Ela não tem inversa em $[-2, +2]$, mas tem os três ramos inversos

$$\begin{aligned} r_1 &: [-1, +1] \longrightarrow [-2, -1] \\ r_2 &: [-1, +1] \longrightarrow [-1, +1] \\ r_3 &: [-1, +1] \longrightarrow [+1, +2]\,. \end{aligned}$$

Para cada $\alpha \in [-1, +1]$, as três raízes da forma canônica são dadas por $r_1(\alpha)$, $r_2(\alpha)$ e $r_3(\alpha)$, e satisfazem

$$-2 \leq r_1(\alpha) \leq -1 \leq r_2(\alpha) \leq +1 \leq r_3(\alpha) \leq 2\,.$$

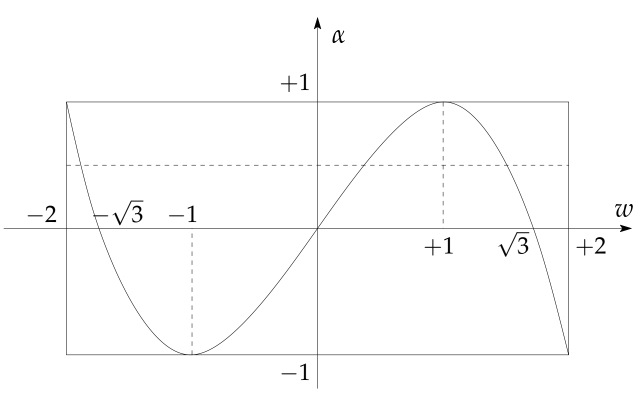

Gráfico de $\alpha = \frac{1}{2}\left(-w^3 + 3w\right)$ em $[-2, +2]$

Note também que se w é solução da forma canônica para um determinado α então $-w$ é solução da forma canônica para $-\alpha$. Portanto, caso α seja negativo, pode-se trocar seu sinal, para que fique positivo, resolver a equação, achando suas raízes, e finalmente trocar o sinal das raízes, para se obter as soluções originais. Em outras palavras, basta saber resolver a forma canônica para α entre 0 e 1.

Uma raiz determina as outras

Será útil também aplicar o princípio de que basta determinar uma das raízes da forma canônica para que as outras fiquem determinadas apenas com o uso de raízes quadradas, sem que se recorra à divisão explícita de polinômios.

Suponhamos que r seja uma raiz da forma canônica e reescrevamos a fórmula canônica expandindo o polinômio de Taylor de f de terceiro grau em torno de r, que neste caso é exato. Isso dá

$$f(r+d) = f(r) + f'(r)d + f''(r)\frac{d^2}{2} + f'''(r)\frac{d^3}{6}\,.$$

Se $r+d$ também for uma raiz, então

$$\begin{aligned} 0 &= f'(r)d + f''(r)\frac{d^2}{2} + f'''(r)\frac{d^3}{6} \\ &= 3(r^2-1)d + 3rd^2 + d^3 \\ &= d\left(3(r^2-1) + 3rd + d^2\right)\,. \end{aligned}$$

As soluções da equação são $d=0$, evidentemente, e as soluções da equação quadrática $3(r^2-1)+3rd+d^2=0$. Calculando

$$D = \sqrt{12-3r^2}\,,\; d^+ = \frac{-3r+D}{2}\,, \tag{5}$$

resulta que, se r é uma raiz da forma canônica, as demais raízes r^+ e r^- são dadas por

$$r^+ = r + d^+ = \frac{D-r}{2}\,,\; r^- = r^+ - D\,. \tag{6}$$

Os superescritos $+$ e $-$ são apenas para ressaltar que $r^+ > r^-$, nada tendo a ver com o sinal das raízes.

Estimativas da raiz intermediária e da raiz inferior: quase-soluções

As fórmulas de aproximação que apresentaremos aqui são truncamentos das expansões de Taylor dos ramos $r_1(\alpha)$ e $r_2(\alpha)$, com graus de truncamento previamente escolhidos (daremos duas sugestões para esses graus). O ramo r_1 é expandido em torno de $\alpha=1$, enquanto o ramo r_2 é expandido em torno de $\alpha=0$. Dependendo do valor de α, um ou outro é mais conveniente, pelo tamanho do erro obtido pelo truncamento. Usando um ou outro, as outras raízes são obtidas como exposto acima.

Para o uso combinado das expansões de $r_1(\alpha)$ e $r_2(\alpha)$ definiremos o *valor de transição* α_T: para $\alpha \le \alpha_T$ usa-se a expansão de $r_2(\alpha)$ em torno de $\alpha=0$, e para $\alpha > \alpha_T$ usa-se a expansão de $r_1(\alpha)$ em torno de $\alpha=1$. A escolha de α_T pode variar com a escolha dos graus das duas expansões, de forma que seja minimizado o erro de truncamento. Neste caso, influencia na escolha de α_T o erro de truncamento ser *absoluto* ou *relativo* ao tamanho de α. Chamaremos de *quase-soluções* aos resultados obtidos por combinação dessas expansões.

A primeira sugestão é expandir r_1 até ordem 6 e r_2 até ordem 9 (como as potências pares na expansão de r_2 são nulas, pode-se chegar a um grau maior com menos termos), usando $\alpha_T = 0.45$ como valor de transição.

Mais abaixo orientaremos o leitor para a obtenção das expansões. Por enquanto, restringiremo-nos a enunciá-las. Apresentaremos as fórmulas com mudanças de variáveis que simplificarão as expressões. Definindo-se

$$\gamma = \frac{2}{9}(1-\alpha)$$

a expansão $\psi_1(\alpha)$ de $r_1(\alpha)$ em torno de $\alpha=1$ até grau 6 é dada por

$$-2+\gamma\left(1+\frac{2}{3}\gamma+\frac{7}{9}\gamma^2+\frac{10}{9}\gamma^3+\frac{143}{81}\gamma^4+\frac{728}{243}\gamma^5\right) \tag{7}$$

e, definindo-se

$$\beta = \frac{2}{3}\alpha$$

a expansão $\psi_2(\alpha)$ de $r_2(\alpha)$ em torno de $\alpha=0$ até grau 9 é dada por

$$\beta\left(1+\frac{1}{3}\beta^2+\frac{1}{3}\beta^4+\frac{4}{9}\beta^6+\frac{55}{81}\beta^8\right)\,. \tag{8}$$

Pode-se constatar empiricamente, por meio de simulações computacionais, que esses truncamentos são de extraordinária qualidade: tanto

$$\frac{|r_2(\alpha)-\psi_2(\alpha)|}{|r_2(\alpha)|}\,,$$

para $\alpha \in (0, 0.45]$, quanto

$$\frac{|r_1(\alpha)-\psi_1(\alpha)|}{|r_1(\alpha)|}\,,$$

para $\alpha \in [0.45, 1]$, são menores do que 3×10^{-6}.

Se estivéssemos na era pré-computador, as fórmulas de $\psi_1(\alpha)$ e $\psi_2(\alpha)$ poderiam ser o receituário para a confecção de uma tabela, com poucas páginas, que forneceria as raízes reais de qualquer cúbica com três raízes reais.

Raízes de cúbicas por meio de cúbicas

Para uma boa parte das situações, a determinação das raízes da forma canônica com precisão melhor do que 5×10^{-4} talvez seja bastante aceitável. Nestes casos, podemos usar as expansões de r_1 e r_2 apenas até grau 3:

$$r_1 \simeq -2 + \gamma + \frac{2}{3}\gamma^2 + \frac{7}{9}\gamma^3 \tag{9}$$

$$r_2 \simeq \beta + \frac{1}{3}\beta^3 , \tag{10}$$

em que $\beta = \frac{2}{3}\alpha$ e $\gamma = \frac{2}{9}(1-\alpha)$. Pode-se dizer então que "dois polinômios cúbicos determinam com boa precisão as raízes das cúbicas" (no caso em que são três as raízes reais). O valor de transição e a precisão obtida serão mostrados adiante.

Expansões de Taylor dos ramos inversos

Para cada $i = 1, 2, 3$, r_i é a inversa de um ramo monótono de F, que é uma função cúbica. A ideia é obter aproximações das funções r_i truncando-se suas expansões de Taylor em pontos apropriados. Para as expansões de Taylor é preciso ter expressões para as derivadas (de ordens variadas) das r_i's. A regra da cadeia aplicada a $F(r_i(\alpha)) = \alpha$ implica

$$r_i'(\alpha) = \frac{1}{F'(r_i(\alpha))} .$$

Daí se obtêm as derivadas de r_i de ordem mais alta, por indução. É fácil mostrar que

$$r_i^{(k)} = \frac{X_k \circ r_i}{(F' \circ r_i)^{2k-1}} ,$$

em que X_k é uma expressão polinomial nas derivadas de F, que pode ser calculada pela recorrência

$$X_{k+1} = X_k' F' - (2k-1) X_k F'' . \tag{11}$$

Como F é cúbica, todas as suas derivadas de ordem maior do que 3 são nulas. Então, na expressão polinomial de X_k, todos os termos que envolvem essas derivadas podem ser ignorados. Chamemos de $\widetilde{X}_k$ a expressão de X_k sem os termos envolvendo derivadas de ordem maior do que 3. Os primeiros termos dessa recorrência são

$$\begin{aligned} \widetilde{X}_1 &= 1 , \\ \widetilde{X}_2 &= -F'' , \\ \widetilde{X}_3 &= 3(F'')^2 - F'F''' , \\ \widetilde{X}_4 &= -5\left(3(F'')^3 - 2F'F''F'''\right) . \end{aligned}$$

Também por indução, pode-se detalhar mais um pouco a forma dos $\widetilde{X}_k$'s. Se n é par, então $\widetilde{X}_{n-1}$ e $\widetilde{X}_n$ têm $\frac{n}{2}$ termos e são da forma

$$\begin{aligned} \widetilde{X}_{n-1} &= \sum_{i=0}^{\frac{n}{2}-1} \sigma_i^{(n)} (F'F''')^i (F'')^{n-2-2i} , \\ \widetilde{X}_n &= F'' \sum_{i=0}^{\frac{n}{2}-1} \tau_i^{(n)} (F'F''')^i (F'')^{n-2-2i} , \end{aligned}$$

em que $\sigma_i^{(n)}$ e $\tau_i^{(n)}$, $i = 0, 1, \ldots, \frac{n}{2} - 1$, são inteiros. Usando-se (11) pode-se deduzir a transformação linear, na forma matricial, que leva o vetor de coeficientes

$$(\sigma_0^{(n)}, \sigma_1^{(n)}, \ldots, \sigma_{\frac{n}{2}-1}^{(n)})$$

no outro vetor de coeficientes

$$(\tau_0^{(n)}, \tau_1^{(n)}, \ldots, \tau_{\frac{n}{2}-1}^{(n)}) ,$$

e também este último no vetor seguinte

$$(\sigma_0^{(n+1)}, \sigma_1^{(n+1)}, \ldots, \sigma_{\frac{n}{2}}^{(n+1)}) .$$

Deixamos para o leitor completar os detalhes que levam às expansões apresentadas em (7) e (8).

Informações técnicas

Dependendo das fórmulas utilizadas e se o objetivo é controlar o erro absoluto ou relativo, pode variar o valor de transição que minimiza o erro. A tabela mostra, para os pares de funções considerados nesse texto (isto é, primeiro ψ_1 e ψ_2, que são truncamentos de graus 6 e 9 das séries de Taylor de r_1 e r_2, e os truncamentos de grau 3 dessas mesmas séries), o valor de transição e uma limitação para o erro cometido obtida numericamente.

graus	rel/abs	transição	erro
6 e 9	absoluto	0.454	2.7×10^{-6}
6 e 9	relativo	0.406	2.7×10^{-6}
3 e 3	absoluto	0.395	4.7×10^{-4}
3 e 3	relativo	0.290	5.1×10^{-4}

Sendo necessária uma maior precisão, esses valores podem ser utilizados como 'semente' para uma iteração de algum método de aproximação de raízes, como o Método de Newton ou o Método de Newton Estendido. Como as sementes já são bastante próximas da raiz, basta fazer uma iteração desses métodos para se ter uma aproximação melhor ainda.

Por exemplo, o Método de Newton é a iteração de

$$\phi(w) = w - \frac{f(w)}{f'(w)},$$

se f é a função da qual se procura a raiz. Neste caso, sendo f a forma canônica, a iteração fica

$$\phi(w) = \frac{2(w^3 - \alpha)}{3(w^2 - 1)}.$$

Pode-se então combinar as fórmulas (9) e (10), em que se obtém r (a aproximação da raiz que essas fórmulas dão), e depois calcular $\phi(r)$. Cada uma dessas composições é uma função racional de α. Usando 0.38 como valor de transição para o uso de (9) ou (10), isso dá um erro relativo menor do que 1.6×10^{-7}.

Melhor ainda é aplicar uma iteração do Método de Newton Estendido, dada por

$$\phi(w) = w - \frac{f(w)}{f'(w)}\left(1 - \frac{f(w)f''(w)}{f'(w)^2}\right)^{-1/2}.$$

Para a forma canônica, isso dá a fórmula

$$\phi(w) = w - \frac{w^3 - 3w + 2\alpha}{2(w^2 - 1)}\left(1 - \frac{6w(w^3 - 3w + 2\alpha)}{9(w^2 - 1)^2}\right)^{-1/2}.$$

Neste caso, usando 0.35 como valor de transição, da composição com (9) ou (10) obtém-se um erro relativo menor do que 2.5×10^{-11}.

Um exemplo

A título de exemplo, estudemos as raízes da cúbica $2y^3 - 6y^2 - 6y + 2$. A divisão por dois, para deixar o coeficiente do maior grau igual a 1, não altera o conjunto de raízes, portanto basta-nos estudar a cúbica $y^3 - 3y^2 - 3y + 1$. Derivando duas vezes e igualando a zero encontramos o ponto de inflexão $y = 1$. Fazemos então a mudança de coordenadas $x = y - 1$, que leva à forma deprimida $x^3 - 6x - 4 = 0$ (exemplo que aparece em [2] de forma detalhada). O discriminante negativo mostra que a equação tem três raízes reais. Para passar à forma canônica, calculamos o mínimo local $x_m = \sqrt{2}$, e fazemos a mudança de coordenadas $w = -x/x_m$, que leva à forma canônica $w^3 - 3w + \sqrt{2} = 0$, isto é, $\alpha = \sqrt{2}/2$ (o sinal negativo na mudança de coordenadas, escolhido *a posteriori*, é conveniente para se obter α positivo na forma canônica).

Usaremos (7) ou (8), com valor de transição $\alpha_T = 0.45$. Como $\alpha > \alpha_T$, usamos a primeira delas. Como esse método, de acordo com a tabela, tem precisão máxima da ordem de 10^{-6}, usaremos 8 casas decimais (neste caso, $\alpha = 0.70710678$). Com (7), obtemos $r_1 = -1.93185168$. Com as fórmulas (5) e (6) conseguimos as outras duas raízes: $r_2 = 0.51763820$ e $r_3 = 1.41421349$. Para chegarmos às raízes de $x^3 - 6x - 4$ basta multiplicar esses três números por $-x_m = -\sqrt{2}$, do que resultam -1.99999989, -0.73205093 e 2.73205084, que estão, respectivamente, a menos de 1.1×10^{-7}, 1.3×10^{-7} e 3.3×10^{-8} das raízes exatas -2, $1 - \sqrt{3}$ e $1 + \sqrt{3}$. Para obtermos as raízes da cúbica original, basta somarmos 1 a esses valores, em vista da mudança de coordenadas $y = x + 1$.

Agradecimento. O autor é muitíssimo grato aos editores pelos diversos comentários e sugestões propostas para a versão final destas notas.

Referências

[1] KUROSCH, A. M. *Curso de álgebra superior*. Moscou: Mir, 1968.

[2] LIMA, E. L. Equação do terceiro grau. *Matemática Universitária*, n. 5, p. 10–23, 1987.

[3] OSTROWSKY, A. M. *Solution of equations and systems of equations*. New York: Academic Press, 1960.

[4] RECHTSCHAFFEN, E. E. M. Real roots of cubics: explicit formula for quasi-solutions. *Mathematical Gazette*, v. 92, n. 524, p. 268–276, 2008.

[5] USPENSKY, J. V. *Theory of equations*. New York: McGraw-Hill, 1948.

Edgar E.M. Rechtschaffen - edgarxrecht@terra.com.br
Centro Universitário Serra dos Órgãos – UNIFESO