Capítulo 1

Fundamentos

1.1 Introdução

Sempre que se pretende tratar algum problema cuja solução toma a forma do cálculo de um valor numérico é habitual ter de considerar não só conceitos de carácter mais abstracto (que fornecem um modelo consistente para a análise do problema) mas também questões de natureza mais prática relacionadas com os cálculos a efectuar ou com os números necessários à realização de tais cálculos.

Exemplo 1.1.1. *Suponha-se que se pretende determinar o volume V de um paralelipípedo a partir dos comprimentos de três arestas a, b e c, perpendiculares entre si. Neste caso, o modelo abstracto consiste na expressão $V = abc$, que permite calcular o volume a partir dos comprimentos das três arestas. Para aplicar esta expressão é então necessário começar por medir cada uma das arestas. Ora, à medição de cada uma das arestas estar associado um erro (erro de medida). Ou seja, o processo de medição fornecerá valores aproximados dos comprimentos das arestas, sendo eventualmente possível obter alguma caracterização dos erros de medida. Ao efectuar, em seguida, o produto das medidas dos três comprimentos ir-se-á obter um valor que apenas poderá ser considerado uma aproximação do volume do paralelipípedo. Obviamente que este valor aproximado terá associado um erro que dependerá dos erros cometidos nos processos de medida.*

A situação descrita neste exemplo de não se conseguir obter um valor numérico exacto para muitos problemas é a mais comum. Esta impossibilidade pode ter origens diversas, de que são exemplos erros associados a processos de medida, modelos abstractos aproximados, ou cálculos efectuados de forma aproximada. Contudo esta situação não é necessariamente má, pois na grande maioria (ou até talvez na totalidade) dos problemas bastará obter um valor numérico suficientemente próximo do valor exacto.

De uma forma simples, pode dizer-se que a Análise Numérica abrange o estudo de métodos

e técnicas que permitam obter soluções aproximadas de problemas numéricos de uma forma eficiente. É por natureza uma disciplina que se situa na fronteira entre a Matemática e a Ciência de Computadores.

Neste capítulo apresentam-se os conceitos fundamentais necessários à compreensão e utilização dos métodos numéricos que irão ser estudados nos capítulos subsequentes.

1.2 Valores exactos e aproximados: erros

Consideremos um problema cuja solução é um número real. Este número é designado por **valor exacto** do problema. Habitualmente, este valor é caracterizado por um modelo abstracto que define o problema, por exemplo uma equação. No que se segue o valor exacto de um problema será representado por x.

Designa-se por **valor aproximado** ou **aproximação**, e representa-se por x^*, qualquer valor que se pretende utilizar como solução do problema. Associado a um dado valor aproximado x^* define-se o **erro de aproximação** como sendo a quantidade

$$\Delta x^* = x - x^*.$$

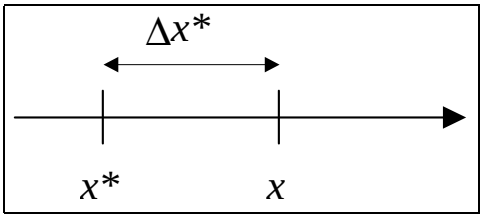

Figura 1.1: Valor exacto e aproximação.

No caso de $x^* < x$, a aproximação diz-se ser **por defeito**, verificando-se então que $\Delta x^* > 0$. No caso de $x^* > x$, a aproximação diz-se ser **por excesso**, tendo-se então que $\Delta x^* < 0$.

Exemplo 1.2.1. *É sabido que* $\pi \simeq 3.14159265359$. *Então,*

$$3 \qquad 3.1 \qquad 3.14 \qquad 3.141 \qquad \ldots$$

são aproximações de π *por defeito e*

$$4 \qquad 3.2 \qquad 3.15 \qquad 3.142 \qquad \ldots$$

são aproximações de π *por excesso.*

Quando o sinal do erro não é conhecido é usual considerar o seu valor absoluto, isto é, $|\Delta x^*| = |x - x^*|$ que se designa por **erro absoluto**.

É de notar que na maioria ou até mesmo em todas as situações, não é conhecido o erro Δx^* associado a uma dada aproximação x^*. De facto, se ambos fossem conhecidos, o valor exacto x poder-se-ia calcular por intermédio da expressão $x = x^* + \Delta x^*$, e então não se utilizaria tal aproximação!

Assim, a situação mais comum é aquela em que se conhece um determinado valor aproximado x^* e um intervalo para o erro de aproximação Δx^*. Este intervalo é muitas vezes caracterizado a partir de majorantes do erro absoluto. A expressão **erro máximo absoluto** é utilizada para designar um majorante do erro absoluto. É usual indicar o erro máximo absoluto por ε. Então, se x^* for um valor aproximado de x com um erro máximo absoluto ε, verifica-se que $x \in [x^* - \varepsilon, x^* + \varepsilon]$.

Outra forma de caracterizar uma aproximação x^* é através do **erro relativo**, que se define por

$$\frac{\Delta x^*}{|x|},$$

para valores de x diferentes de zero. Muitas vezes é também considerado o erro relativo aproximado definido por

$$\frac{\Delta x^*}{|x^*|}.$$

A noção de erro relativo advém do facto de o mesmo erro absoluto poder ter significados reais diferentes conforme o valor exacto em causa. Os erros relativos exprimem-se muitas vezes em termos percentuais. Por exemplo, um erro relativo de 0.02 é referido habitualmente como um erro de 2%.

Define-se também **erro máximo relativo**, normalmente indicado por ε', como sendo um majorante do erro relativo em valor absoluto, isto é,

$$\varepsilon' = \frac{\varepsilon}{|x|},$$

onde ε representa um erro máximo absoluto. Também aqui é normal trabalhar com a aproximação do erro máximo relativo dada por

$$\varepsilon' \simeq \frac{\varepsilon}{|x^*|},$$

valor que é possível calcular com base na aproximação x^* e no erro máximo absoluto ε conhecido.

Para uma dada aproximação x^*, o erro máximo relativo pode ser calculado a partir do erro máximo absoluto conhecido e vice-versa, ainda que de uma forma aproximada. Habitualmente, os erros máximos quer absolutos quer relativos são indicados com um número reduzido de casas decimais (raramente mais do que duas).

Exemplo 1.2.2.
Seja $x^* = 3.45$ *com* $\varepsilon = 0.01$. *Então* $\varepsilon' \simeq \frac{0.01}{3.45} \simeq 3 \times 10^{-3}$.
Seja $x^* = -2.7$ *com* $\varepsilon' = 0.07$. *Então* $\varepsilon \simeq 0.07 \times 2.7 \simeq 0.19$.

1.3 Representação de números reais

Uma das formas possíveis de representar números reais é a designada **notação científica**. Na notação científica decimal (base 10), que é a mais utilizada, o número real x é representado por

$$x = \pm 0.d_1 d_2 d_3 \ldots \times 10^e.$$

Cada número é representado pelo sinal ($+$ ou $-$), pela **mantissa** ($0.d_1 d_2 d_3 \ldots$) e pelo **expoente** e na base 10. Os dígitos da mantissa são elementos de $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ e o expoente é um número inteiro. Refira-se que a mantissa de um número real poderá ter comprimento infinito.

Um sistema de representação em **vírgula flutuante** é um sistema semelhante à notação científica mas que apenas permite representar um conjunto finito de números racionais. Cada sistema de vírgula flutuante é caracterizado por 4 parâmetros: a sua base (β), o número de dígitos da mantissa (n) e os valores mínimos e máximos do expoente (m e M, respectivamente). Tal sistema é habitualmente representado por $\mathrm{FP}(\beta, n, m, M)$. Assim, dizer que $x \in \mathrm{FP}(\beta, n, m, M)$ é equivalente a ter

$$x = 0 \quad \text{ou} \quad x = \pm(0.d_1 d_2 \ldots d_n) \times \beta^e$$

onde e é um inteiro tal que $m \leq e \leq M$, e d_i, para $i = 1, \ldots, n$, são dígitos na base β, com $d_1 \neq 0$.

A representação mais comum de números reais em computadores é feita em vírgula flutuante, sendo normalmente utilizada a base 2.

Numa representação em vírgula flutuante, o conjunto de expoentes permitidos limita a gama de valores representáveis e o número de dígitos da mantissa caracteriza a precisão com que se podem aproximar números que não tenham representação exacta.

Consideremos então um número $x \neq 0$, descrito em notação científica por

$$x = \pm 0.d_1 d_2 d_3 \ldots d_n d_{n+1} \ldots \times 10^e$$

com $d_1 \neq 0$, que se pretende representar (de forma aproximada) no sistema $\mathrm{FP}(10, n, m, M)$.

Se o expoente e for superior a M dá-se a designada situação de **overflow**. Se o expoente e for inferior a m dá-se a designada situação de **underflow**. Em qualquer dos casos não é aconselhável utilizar um elemento de $\mathrm{FP}(10, n, m, M)$ para representar x, pois o erro relativo de aproximação pode ser arbitrariamente elevado. Por tal motivo, é habitual os sistemas computacionais tratarem as situações de overflow e underflow como situações de erro.

Quando se verifica que $m \leq e < M$ e a mantissa tem mais dígitos do que os permitidos numa dada representação, podem usar-se dois processos para representar o número: **truncatura** ou **arredondamento**. No primeiro caso ignoram-se os dígitos da mantissa a partir do índice $n+1$. No segundo caso é escolhido o número representável mais próximo do número que se pretende aproximar de acordo com as seguintes regras

- se $0.d_{n+1}d_{n+2}\ldots > 0.5$ arredonda-se a casa decimal n para **cima**;
- se $0.d_{n+1}d_{n+2}\ldots < 0.5$ arredonda-se a casa decimal n para **baixo**;
- se $0.d_{n+1}d_{n+2}\ldots = 0.5$ arredonda-se a casa decimal n para o algarismo **par** mais próximo.

Esta regra de arredondamento garante que toda aproximação com uma mantissa de n casas decimais terá um erro absoluto inferior a 5 unidades da primeira casa não representada.

Quando se pretende indicar que x^* é uma aproximação de um valor exacto x com um dado erro máximo absoluto ε, ter-se-á de indicar x^* com todas as casas decimais conhecidas bem como o valor de ε. Por forma a tornar a escrita mais simples, é habitual considerar erros absolutos da forma 0.5×10^n e representar a aproximação até à casa decimal correspondente a 10^n. Quando se utiliza esta convenção, os algarismos da mantissa de uma representação, com excepção dos zeros à esquerda, designam-se **algarismos significativos**.

Exemplo 1.3.1. *A tabela seguinte mostra alguns exemplos de aproximações, utilizando esta convenção. Em cada caso, é apresentado o erro máximo absoluto, o menor intervalo em que se garante estar o valor exacto, o número de algarismos significativos, bem como o erro máximo relativo.*

x^*	ε	*Intervalo*	*Algarismos significativos*	ε'
2.24	0.005	$[2.235, 2.245]$	*3*	2.2×10^{-3}
2.240	0.0005	$[2.2395, 2.2405]$	*4*	2.2×10^{-4}
1.5×10^2	5	$[145, 155]$	*2*	3.3×10^{-2}
0.1×10^3	50	$[50, 150]$	*1*	5×10^{-1}

O teorema seguinte relaciona o número de algarismos significativos de uma aproximação com o seu erro relativo.

Teorema 1.3.1. *Uma aproximação com n algarismos significativos tem um erro relativo aproximado inferior a 5×10^{-n}.*

Demonstração. Se x^* é uma aproximação com n algarismos significativos, então x^* é da forma

$$x^* = \pm d_1 d_2 \cdots d_n \times 10^k,$$

para algum $k \in \mathbb{Z}$ e com $d_1 \neq 0$. De acordo com a convenção utilizada, esta aproximação terá um erro máximo absoluto $\varepsilon = 5 \times 10^{k-1}$.

O erro máximo relativo (aproximado) ε' satisfaz

$$\varepsilon' = \frac{\varepsilon}{|x^*|} = \frac{5 \times 10^{k-1}}{d_1 d_2 \cdots d_n \times 10^k} = \frac{5 \times 10^{-1}}{d_1 d_2 \cdots d_n}.$$

Como $d_1 \neq 0$ tem-se que $10^{n-1} \leq d_1 d_2 \cdots d_n < 10^n$, concluindo-se finalmente que

$$\varepsilon' \leq \frac{5 \times 10^{-1}}{10^{n-1}} = 5 \times 10^{-n}. \qquad \square$$

Esta relação entre algarismos significativos e erro relativo fornece uma regra prática para o número de algarismos a utilizar quando representar uma dada aproximação. Devem apenas representar-se os algarismos significativos, podendo também considerar-se um algarismo adicional.

1.4 Aritmética em representações finitas

O cálculo de uma expressão envolvendo múltiplas operações aritméticas realizadas utilizando representações finitas deve ser efectuado com algum cuidado. De facto, a necessidade de guardar resultados intermédios, obviamente utilizando uma representação finita, faz com que se cometam diversos erros de arredondamento desses resultados intermédios, erros esses que se podem ir acumulando à medida que os cálculos progridem, resultando em elevados erros no resultado final.

Um dos pontos a considerar advém do facto de operações aritméticas que habitualmente gozam de associatividade (como a soma e a multiplicação) poderem perder essa propriedade quando se trabalha em representações finitas. O exemplo seguinte ilustra este efeito.

Exemplo 1.4.1. *Calcular* $0.5 + 0.024 + 0.012$ *utilizando* 2 *dígitos em vírgula flutuante.*

a) Somando da esquerda para a direita

$$\begin{aligned}(0.50 \times 10^0 + 0.24 \times 10^{-1}) + 0.12 \times 10^{-1} &= (0.50 \times 10^0 + 0.02 \times 10^0) + 0.12 \times 10^{-1} \\ &= 0.52 \times 10^0 + 0.01 \times 10^0 \\ &= 0.53 \times 10^0\end{aligned}$$

b) Somando da direita para a esquerda

$$\begin{aligned}0.50 \times 10^0 + (0.24 \times 10^{-1} + 0.12 \times 10^{-1}) &= 0.50 \times 10^0 + 0.36 \times 10^{-1} \\ &= 0.50 \times 10^0 + 0.04 \times 10^0 \\ &= 0.54 \times 10^0\end{aligned}$$

Utilizando aritmética exacta o resultado seria sempre 0.536.

Outro caso que é necessário ter em atenção é a subtracção de dois números quase iguais. Aqui, o resultado poderá ter um erro máximo absoluto da sua ordem de grandeza, originando um erro relativo elevado. Este fenómeno de perda de algarismos significativos é designado por **cancelamento subtractivo**.

Exemplo 1.4.2. *Efectuar a subtracção* $2.034 - 2.016$ *utilizando* 3 *dígitos em vírgula flutuante.*

Resolução

Arredondando os dois números dados para 3 *algarismos obtém-se* 2.03 *e* 2.02*, respectivamente. O resultado aproximado da subtracção, utilizando os números arredondados é* $x^* = 0.01$*.*

O valor exacto da subtracção é 0.018*, pelo que o erro absoluto de* x^* *é* 0.008 *e o seu erro relativo é* 44%*, aproximadamente.*

O cancelamento subtractivo pode levar a resultados com elevados erros relativos que são sempre indesejáveis. Por vezes, é possível dispor os cálculos de forma a evitar tal cancelamento.

Exemplo 1.4.3. *Seja* $x \gg 1$ *e* $y = \sqrt{x+1} - \sqrt{x}$*. O cálculo de* y *pela expressão dada pode originar um erro relativo elevado devido ao cancelamento subtractivo. Contudo, a expressão equivalente*

$$y = \frac{1}{\sqrt{x+1} + \sqrt{x}}$$

permite calcular y*, evitando tal fenómeno.*

1.5 Propagação de erros no cálculo de funções

Nesta secção iremos analisar como se propagam os erros de aproximação no cálculo de funções. Abordaremos primeiro o caso de uma função real de variável real e posteriormente o caso de uma função real de variável vectorial.

Seja então $f : \mathbb{R} \to \mathbb{R}$. A situação que iremos tratar pode descrever-se do seguinte modo: conhecendo uma aproximação x^* de x, que valor y^* considerar para aproximar $y = f(x)$ e como relacionar os erros de aproximação de x^* e de y^*?

No caso de a função f ser contínua verifica-se que à medida que x^* se aproxima de x mais o valor $f(x^*)$ se aproxima de $f(x)$. Nesta situação, que é a mais usual, pode utilizar-se o valor $y^* = f(x^*)$ como aproximação de $y = f(x)$.

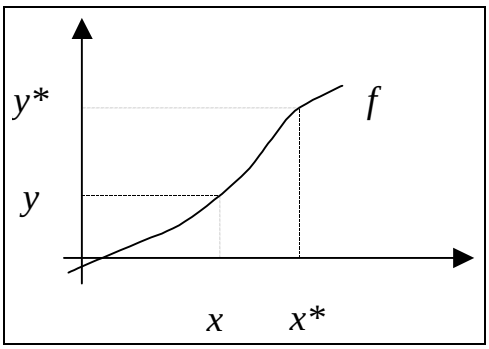

Figura 1.2: $f(x^*)$ aproximação de $f(x)$.

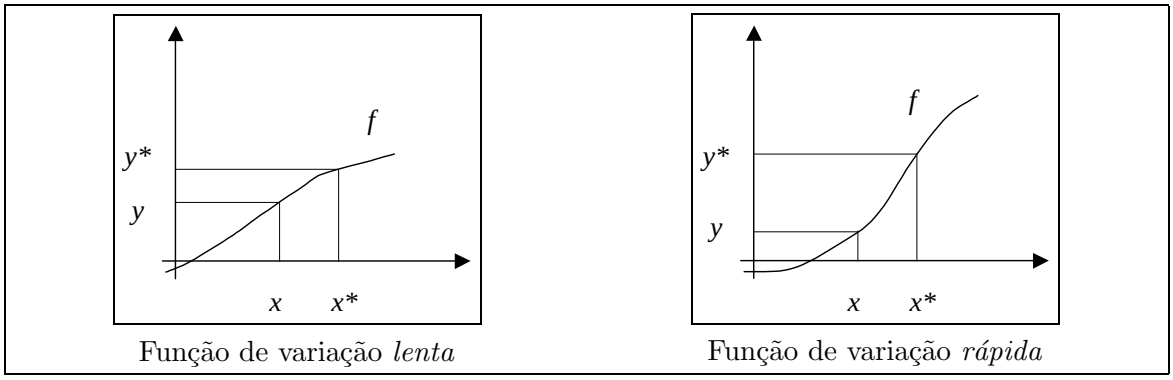

Figura 1.3: Influência de f na propagação de erros.

Para além da determinação do valor aproximado de $y^* = f(x^*)$, interessa também caracterizar o erro cometido nesta aproximação, ou melhor, relacionar este erro com o erro de aproximação de x por x^*. É claro que o erro $\Delta y^* = y - y^*$ dependerá do erro $\Delta x^* = x - x^*$ e também da função f em questão. De facto, o erro de aproximação Δy^* é obtido pela expressão

$$\Delta y^* = y - y^* = f(x) - f(x^*) = f(x^* + \Delta x^*) - f(x^*).$$

Se a função f for continuamente diferenciável, a aplicação do teorema do valor médio permite escrever

$$f(x^* + \Delta x^*) - f(x^*) = f'(\bar{x}) \cdot \Delta x^*$$

para algum $\bar{x}$ entre x^* e $x^* + \Delta x^*$, obtendo-se então a seguinte expressão

$$\Delta y^* = f'(\bar{x}) \cdot \Delta x^*.$$

Habitualmente, apenas se conhece um majorante ε_x para $|\Delta x^*|$. Então, o valor

$$\varepsilon_y = |f'|_{|\max} \cdot \varepsilon_x$$

é um majorante para o erro absoluto da aproximação y^* de y. Nesta expressão, o valor máximo de $|f'|$ é determinado no intervalo $[x^* - \varepsilon_x, x^* + \varepsilon_x]$.

Exemplo 1.5.1. *Calcular um valor aproximado de $y = \sin x$ e o correspondente erro máximo absoluto quando $x \approx 0.57$ (isto é, $x = 0.57 \pm 0.005$).*

Resolução

Um valor aproximado será $\bar{y} = \sin \bar{x} = \sin 0.57 \simeq 0.5396$.

O erro máximo absoluto será

$$\varepsilon_y \leq \max_x \left| \frac{dy}{dx} \right| \cdot \varepsilon_x = \max_x |\cos x| \cdot \varepsilon_x$$

No intervalo em questão, a função cos *é positiva e decrescente. Então*

$$\varepsilon_y \leq \cos(0.57 - 0.005) \cdot 0.005 \simeq 4.2 \times 10^{-3}$$

Finalmente tem-se que $y = 0.5396 \pm 4.2 \times 10^{-3}$, ou ainda, $y \approx 0.54 \pm 5 \times 10^{-3}$.

Sendo x^* um valor aproximado de x com erro máximo absoluto ε_x, a expressão $|\Delta y^*| \leq |f'|_{\max} \cdot \varepsilon_x$ permite obter o majorante para o erro relativo de $y^* = f(x^*)$ dado por

$$\varepsilon'_y = \left| f'(x) \cdot \frac{x}{f(x)} \right|_{\max} \cdot \varepsilon'_x$$

onde $\varepsilon'_x = \frac{\varepsilon_x}{|x|}$, e o máximo de $\left| \frac{xf'(x)}{f(x)} \right|$ é determinado no intervalo $[x^* - \varepsilon_x, x^* + \varepsilon_x]$.

Dados $x \in \mathbb{R}$ e uma função f, o **número de condição** de f em x é definido como sendo

$$\left| \frac{xf'(x)}{f(x)} \right|.$$

Este valor pode ser utilizado para avaliar a perda ou o ganho de algarismos significativos no cálculo de uma função, uma vez que caracteriza a ampliação ou redução do erro relativo. Quando o número de condição for reduzido a função diz-se **bem condicionada**. Quando o número de condição for elevado a função diz-se **mal condicionada** e o erro relativo é amplificado.

Exemplo 1.5.2. *Quantos dígitos significativos se podem perder no cálculo da função $y = \tan(x)$ quando x está próximo de 1?*

Resolução
Como $\frac{dy}{dx} = 1 + \tan^2(x)$ tem-se que

$$\left| \frac{dy}{dx} \cdot \frac{x}{y} \right|_{x=1} = \left| \frac{(1 + \tan^2(x)) \cdot x}{\tan(x)} \right|_{x=1} = \frac{1 + \tan^2(1)}{\tan(1)} \approx 2.2 > 1$$

podendo perder-se um dígito significativo.

Repetindo os cálculos para $x = 1.5$, obter-se-ia $\left| \frac{dy}{dx} \cdot \frac{x}{y} \right| \approx 21$, concluindo-se que em tal caso se poderiam perder até 2 dígitos significativos.

Passemos agora a analisar o caso em que y depende de diversas variáveis, isto é, $y = f(x_1, x_2, \ldots, x_n)$, onde f é uma função de $\mathbb{R}$ em $\mathbb{R}^n$, que se considera continuamente diferenciável.

Para cada $i = 1, \ldots, n$, seja x_i^*, um valor aproximado de x_i, com erro máximo absoluto ε_{x_i}. Nestas condições verifica-se que

$$y^* = f(x_1^*, x_2^*, \ldots, x_n^*)$$

será um valor aproximado de $y = f(x_1, x_2, \ldots, x_n)$ com erro máximo absoluto

$$\varepsilon_y = \sum_{i=1}^{n} \left(\left| \frac{\partial f}{\partial x_i} \right|_{\max} \cdot \varepsilon_{x_i} \right),$$

onde cada um dos máximos das derivadas parciais de f em relação às diversas variáveis independentes é determinado em $\prod_{i=1}^{n} [x_i - \varepsilon_{x_i}, x_i + \varepsilon_{x_i}]$.

É também possível obter o erro relativo máximo para y^* dado por

$$\varepsilon'_y = \sum_{i=1}^{n} \left(\left| \frac{\partial f}{\partial x_i} \cdot \frac{x_i}{f} \right|_{\max} \cdot \varepsilon'_{x_i} \right).$$

Nesta expressão, considera-se que ε'_{x_i} é um majorante do erro relativo de x_i^*, para $i = 1, \ldots, n$. As maximizações são ainda realizadas no conjunto indicado acima, tomando-se agora $\varepsilon_{x_i} = \varepsilon'_{x_i}|x_i|$.

Exemplo 1.5.3. *Sendo $\tilde{a}$ e $\tilde{b}$ valores aproximados de a e b, com erros máximas absolutos ε_a e ε_b, então $\tilde{s} = \tilde{a} + \tilde{b}$, é um valor aproximado de $s = a + b$ com erro máximo absoluto*

$$\varepsilon_s = \varepsilon_a + \varepsilon_b.$$

Exemplo 1.5.4. *O erro máximo relativo no cálculo de $w = xyz$, pode ser obtido a partir dos erros máximos relativos em x, y e z da seguinte forma*

$$\begin{aligned}
\varepsilon'_w &= \left| \frac{\partial w}{\partial x} \cdot \frac{x}{w} \right| \cdot \varepsilon'_x + \left| \frac{\partial w}{\partial y} \cdot \frac{y}{w} \right| \cdot \varepsilon'_y + \left| \frac{\partial w}{\partial z} \cdot \frac{z}{w} \right| \cdot \varepsilon'_z \\
&= \left| yz \cdot \frac{x}{xyz} \right| \cdot \varepsilon'_x + \left| xz \cdot \frac{y}{xyz} \right| \cdot \varepsilon'_y + \left| xy \cdot \frac{z}{xyz} \right| \cdot \varepsilon'_z \\
&= \varepsilon'_x + \varepsilon'_y + \varepsilon'_z.
\end{aligned}$$

1.6 Erro de truncatura

Por vezes a determinação de um certo valor envolve a realização de uma sucessão infinita de operações. O erro cometido quando se toma uma aproximação resultante da realização de um número finito de operações designa-se **erro de truncatura**.

Um dos casos em que se surge o erro de truncatura é no caso da aproximação da soma S de uma série convergente $\sum_{i=0}^{\infty} a_i$ pela soma parcial $S_n = \sum_{i=0}^{n} a_i$. Neste caso, o erro de truncatura será $R_n = S - S_n$.

O erro de truncatura é particularmente importante quando se efectua a aproximação de uma função por polinómios de Taylor, reduzindo assim o seu cálculo à realização de operações de soma, subtracção, multiplicação e divisão, que são as operações aritméticas elementares à custa das quais todos os cálculos numéricos são realizados.

O desenvolvimento de Taylor de uma função f em torno do ponto x_0 permite escrever

$$f(x) \quad = \quad \underbrace{f(x_0) + f'(x_0)(x - x_0) + \cdots + f^{(n)}(x_0)\frac{(x - x_0)^n}{n!}}_{P_{x_0,n}(x)} \quad + \quad R_{x_0,n}(x)$$

onde $R_{x_0,n}(x) = f^{(n+1)}(x_0 + (x - x_0)\theta)\frac{(x-x_0)^{n+1}}{(n+1)!}$ para algum $\theta \in [0, 1]$.

O erro de truncatura na aproximação $f(x) \approx P_{x_0,n}(x)$ é dado pelo resto de Taylor $R_{x_0,n}(x)$. Se se verificar que $R_{x_0,n}(x) \xrightarrow[n \to +\infty]{} 0$ a aproximação por polinómios de Taylor pode ser tão boa quanto se queira, bastando para tal considerar um número suficientemente elevado de termos.

Exemplo 1.6.1. *Considere aproximações da função e^x no intervalo $[-2, 2]$ dadas por polinómios de Taylor. Qual deverá ser o grau do polinómio a utilizar se se pretender que o erro absoluto devido à truncatura da série seja inferior a 5×10^{-5}?*

Resolução

O desenvolvimento de Taylor em torno de 0 é

$$e^x = 1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} + R_n(x),$$

onde $R_n(x) = e^{\theta x} \frac{x^{n+1}}{(n+1)!}$, para $\theta \in [0, 1]$.

O erro absoluto devido à truncatura pode ser majorado da seguinte forma

$$\varepsilon_{trunc} = |R_n(x)| = \left| e^{\theta x} \frac{x^{n+1}}{(n+1)!} \right| \leq 8 \frac{2^{n+1}}{(n+1)!}$$

uma vez que $\theta \in [0, 1]$ e $x \in [-2, 2]$.

Calculando estes majorantes para alguns valores de n, obtêm-se os seguintes valores

n	$8\frac{2^{n+1}}{(n+1)!}$
10	4.1×10^{-4}
11	6.8×10^{-5}
12	1.1×10^{-6}
13	1.5×10^{-7}

Conclui-se então que para $n = 12$ se tem $\varepsilon_{trunc} \leq 1.0 \times 10^{-5}$, devendo-se portanto utilizar um polinómio de grau 12.

1.7 Exercícios

1. Sendo x^* uma aproximação de x por defeito com erro máximo absoluto ε, verificar que $x^* + \frac{\varepsilon}{2}$ é uma aproximação de x com erro máximo absoluto $\frac{\varepsilon}{2}$.

2. Arredondar os seguintes números como indicado.

   (a) 3.567 às centésimas.

   (b) -5.78 às décimas.

   (c) 4.5675 às milésimas.

   (d) 4.5685 às milésimas.

   (e) 8764 para 3 casas decimais.

   (f) 8.34×10^{-1} para 1 casa decimal.

3. Sendo $y \approx 0$, como calcular $\cos(x + y) - \cos(x)$ evitando erros de cancelamento elevados?

4. As duas soluções da equação $x^2 - bx + 1 = 0$ podem ser obtidas por

$$x_1 = \frac{b + \sqrt{b^2 - 4}}{2} \qquad\qquad x_2 = \frac{b - \sqrt{b^2 - 4}}{2}$$

   Se $b \gg 2$ então b será muito próximo de $\sqrt{b^2 - 4}$ e logo o cálculo de x_2 poderá ser afectado por erros de cancelamento elevados. Será possível evitar tais erros, calculando x_2 de uma forma alternativa?

5. Determinar um majorante para o erro absoluto no cálculo de $d = x - y$, conhecendo majorantes para os erros absolutos em x e y.

6. Determinar um majorante para o erro relativo de $z = \frac{x}{y}$, a partir de majorantes para os erros absolutos em x e y.

7. A frequência de ressonância de um circuito LC série é dada for $f = \frac{1}{2\pi\sqrt{LC}}$. Calcular o valor de f e um majorante para o erro relativo quando

$$L = 10\,\mu\text{H} \pm 5\% \qquad \text{e} \qquad C = 10\,\text{pF} \pm 1\,\text{pF}.$$

8. Sendo $R_1 = 6.7 \times 10^2\,\Omega$ e $R_2 = 1\,\text{k}\Omega$, com erros relativos de 5% e 10%, respectivamente, calcular $R_1//R_2$ e um majorante para o seu erro relativo.

9. Pretende-se calcular a função $\sin(x)$ no intervalo $[-\frac{\pi}{4}, \frac{\pi}{4}]$ recorrendo a um polinómio de Taylor. Qual deverá ser o polinómio a utilizar de modo a que o erro relativo, devido à truncatura da série de Taylor, seja inferior a 5×10^{-8} ?