Aula Laboratorial II
Erros e sua propagação

Quando se representa um número real x num sistema de vírgula flutuante, de base b, isso significa que x é aproximado por um certo número $\tilde{x}$, em geral diferente de x, da forma

$$\tilde{x} = a_0.a_1a_2\ldots a_n \times b^t,$$

onde a sucessão de dígitos

$$m = a_0.a_1a_2\ldots a_n$$

é chamada a *mantissa* e t é um número inteiro (o expoente). Cada sistema de vírgula flutuante caracteriza-se pela sua base (b), pelo número de dígitos da mantissa ($n+1$) e pelos valores mínimo e máximo que o expoente t pode tomar (t_1 e t_2).

Por exemplo, no *Mathematica* usa-se um sistema de base $b = 10$ e o número de dígitos da mantissa depende do computador em que se trabalhar. Num PC, temos 16 dígitos, embora normalmente só surjam 6 dígitos no ecran. Pode-se, porém elevar a precisão , aumentando o número de dígitos da mantissa, até ao limite máximo de 308. Para isso, usa-se a instrução $N[x,k]$, onde x é o número real que se pretende representar e k é o número de dígitos.

Os valores máximo e mínimo do expoente no *Mathematica* são , respectivamente, $t_1 = -333328015$ e $t_2 = 323228010$. Isto significa que não é possível representar números reais muito superiores a 10^{t_2} ou muito inferiores a 10^{t_1}. Verifique o que acontece quando tais números surgem nos cálculos.

Uma vez que, como vimos, a representação de um número real x num sistema de vírgula flutuante não coincide, regra geral, com o próprio número, a representação tem um erro, o chamado **erro de arredondamento**, cujo valor é

$$e_{ar} = x - \tilde{x}.$$

Ao módulo deste valor chamamos o erro de arredondamento absoluto. O valor do erro de arredondamento absoluto depende, naturalmente, do número x que queremos representar (admitindo que ele tem representação no sistema considerado) e do número de dígitos da mantissa. Note-se que, quanto maior é o número que se pretende representar, maior tende a ser o erro de arredondamento absoluto, no mesmo sistema numérico. Isso, no entanto, não significa, em geral, que a *precisão* seja mais baixa para números grandes. Ou seja, a precisão depende não do erro de arredondamento absoluto, mas sim do erro de arredondamento relativo, que se define como

$$\delta_{ar} = \frac{e_{ar}}{x}$$

Na tabela 1 apresentam-se, como exemplos, as representações de alguns números reais no sistema de vírgula flutuante do *Mathematica*, com precisão normal (16 dígitos) ou aumentada (20 dígitos), e os valores aproximados dos respectivos erros de arredondamento, absoluto e relativo.

Tabela 1

x	$N[x]$	$N[x,20]$	$\lvert N[x]-N[x,20]\rvert$	$\frac{N[x]-N[x,20]}{x}$
$\sqrt{2}$	1.414213562373095	1.4142135623730950488	2.2204×10^{-16}	1.57009×10^{-16}
$\sqrt{\pi}$	1.772453850905516	1.7724538509055160273	2.2204×10^{-16}	1.57009×10^{-16}
e^{20}	$4.851651954097899\times10^{8}$	$4.851651954097902780\times10^{8}$	4.1723×10^{-7}	-8.5998×10^{-16}
e^{-20}	$2.061153622438559\times10^{-9}$	$2.0611536224385578280\times10^{-9}$	1.6544×10^{-24}	8.02638×10^{-16}

Assim, vemos que, para cada número real x, que tem representação exacta num sistema de vírgula flutuante, ou seja, $\tilde{x} = x$, existe um intervalo de números que têm a mesma representação: $[\tilde{x}-\delta, \tilde{x}+\delta]$, onde o valor de δ depende de x e do número de dígitos da mantissa. Em *Mathematica*,

o valor de δ, para um dado $\tilde{x}$, pode ser obtido através do comando $Interval[\tilde{x}]$, cujo *output* é o intervalo de números reais que são representados por $\tilde{x}$. Por exemplo, no caso de $x = \sqrt{2}$, temos

$$In : Interval[N[Sqrt[2]]] - Sqrt[2]$$
$$Out : Interval[\{-2.2045 \times 10^{-16}, 2.2045 \times 10^{-16}\}],$$

de onde se conclui que o valor de δ , para $\tilde{x} = N[\sqrt{2}]$, é 2.2045×10^{-16}. Quando aproximamos qualquer número do intervalo de $\tilde{x}$, cometemos um erro de arredondamento absoluto que não excede δ. Como exercício, verifique, para cada um dos valores de x da tabela, qual é o valor correspondente de δ.

Influência dos erros de arredondamento nos cálculos.

Quando os dados dos cálculos são afectados por erros de arredondamento, estes erros não podem deixar de afectar os resultados. Neste parágrafo, vamos debruçar-nos sobre esta questão e ver exemplos de como pequenos erros de arredondamento nos dados podem ter uma repercussão desastrosa no resultado. Na realidade, quando nós efectuamos uma certa operação envolvendo os números x e y , num certo sistema de cálculo, a operação vai efectuar-se sobre os números $\tilde{x}$ e $\tilde{y}$ que os representam no sistema de cálculo considerado. Isto faz com que o resultado da operação seja afectado por um certo erro, dependente da propagação dos erros de $\tilde{x}$ e $\tilde{y}$. Esse erro pode ter mais ou menos peso no resultado, conforme os dados e a operação efectuada. O caso mais desfavorável é aquele em que se subtraem dois números muito próximos um do outro. De facto, quando tal acontece, nós substituímos $x - y$ por $\tilde{x} - \tilde{y}$. Mas, se os erros absolutos de $\tilde{x}$ e $\tilde{y}$ forem da mesma ordem de grandeza que $x - y$, naturalmente $\tilde{x} - \tilde{y}$ vai ser uma má aproximação da diferença. Este fenómeno é conhecido como **cancelamento subtractivo**.

Exemplo. Consideremos a função real

$$f(x) = \frac{x}{1 - \sqrt{1 - x}}, \tag{1}$$

como $x \in [0, 1]$. É sabido que esta função é infinitamente diferenciável numa vizinhança de zero e que

$$\lim_{x \to 0} f(x) = 2,$$

pelo que, para valores de x, muito próximos de zero, ela toma valores muito próximos de 2. No entanto, se calcularmos a função usando a precisão habitual do *Mathematica*, obtemos os valores que constam da segunda coluna da tabela 2, muito diferentes daquilo que se poderia esperar. A explicação é a seguinte: quando efectuamos a subtracção no denominador da fracção , o resultado é muito próximo de 0 , uma vez que, quando $x \to 0$, $\sqrt{1 - x}$ toma valores próximos de $1 - x/2$. Logo, se $x/2$ for da mesma ordem de grandeza que δ, dá-se o cancelamento subtractivo. Aumentando a precisão do argumento da função para 20 dígitos, obtêm-se os resultados da terceira coluna, que já correspondem ao valor esperado. Conclui-se então que os primeiros resultados estão afectados de erros, cujos valores absolutos estão representados na quarta coluna, e que resultam dos erros de arredondamento do argumento.

Por vezes, consegue-se evitar o cancelamento subtractivo recorrendo a transformações algébricas da função . Neste caso, por exemplo, verifica-se facilmente que

$$f(x) = 1 + \sqrt{1 - x}. \tag{2}$$

Se utilizarmos a fórmula 2 para calcular a função , deixamos de ter cancelamento subtractivo, pelo que, mesmo com precisão simples se obtêm resultados satisfatórios (ver quinta coluna). O estudo deste exemplo poderá ser complementado com a ajuda do *Mathematica*, traçando o gráfico de $f(10^{-x})$, primeiro com uma e depois com a outra fórmula. Para que os resultados sejam elucidativos, sugere-se que considere valores de x superiores a 10.

Tabela 2

x	$f[N[x]]$	$f[N[x,20]]$	$\lvert f[N[x]] - f[N[x,20]]\rvert$	$f[N[x]](form.2)$
10^{-11}	2.	1.999999999995000000	5×10^{-12}	2.
10^{-12}	1.99982	1.999999999999500000	1.7×10^{-4}	2.
10^{-13}	1.99274	1.999999999999950000	7.3×10^{-3}	2.
10^{-14}	1.95809	1.999999999999995000	0.042	2.
10^{-15}	1.5012	1.999999999999999500	0.499	2.
10^{-16}	0.45036	1.999999999999999950	1.55	2.

Problemas

Para consolidar os conhecimentos adquiridos e treinar o uso do *Mathematica*, resolva os problemas abaixo formulados, bem como os problemas das aulas práticas que envolvam o uso deste software, momeadamente o 3-e e o 4-b do I capítulo.

1. Se introduzir no *Mathematica* as instruções

$$x = 10.^{16}; \quad i = 100; \quad y = x + i; \quad y == x$$

   o resultado é $True$. Como explica isso? Qual o mínimo valor (inteiro) que deve dar a i de modo a que o resultado seja $False$?

2. Considere a função $g(x) = \frac{cosx-1}{x^2}$

   (a) Recorrendo ao *Mathematica*, construa para esta função uma tabela, semelhante à tabela 2, que mostre a influência dos erros de arredondamento para valores de x próximos de 0. Na escolha dos valores de x a considerar, tenha em conta que, quando $x \to 0$, a função $cosx$ toma valores próximos de $1 - x^2/2$. Como explica os resultados obtidos?

   (b) Recorrendo a uma transformação trigonométrica, proponha uma fórmula para calcular g que evite o cancelamento subtractivo. Verifique o resultado do cálculo através desta fórmula e complete a tabela.

   (c) Use o *Mathematica* para traçar o gráfico da função $g\,(10^{-x})$, aplicando cada uma das fórmulas. Comente o resultado.