Teorema do Ponto Fixo a uma Variável Real e Aplicações à Resolução de Equações Não Lineares

Luís Camilo do Canto de Loura
Faculdade de Motricidade Humana - Universidade Técnica de Lisboa

I. O Zelmiro e a máquina de calcular oferecida pela tia Zagolinda.

Suponhamos que o aluno Zelmiro do 12º ano do ensino secundário, para experimentar a sua nova calculadora, oferecida no "Dia Mundial do Cálculo sem Esforço" pela tia Zagolinda, decide carregar nas teclas "1" e "EXE", obtendo o valor 1. Depois carrega nas teclas "cos", "Ans" e "EXE"; obtém o valor 0,5403023059 (por acaso a calculadora do Zelmiro estava a trabalhar em radianos). O Zelmiro gostou do resultado obtido e, num ímpeto experimentalista raro nele, resolve voltar a carregar sucessivamente na tecla "cos" seguida das teclas "Ans" e "EXE", por forma a calcular o coseno do número previamente calculado; designando por n o número de vezes que cada uma das teclas "cos" e "Ans" foram premidas e por x_n valor obtido, tem-se:

n	x_n	n	x_n
1	0,5403023059	12	0,7414250866
2	0,8575532158	13	0,7375068905
3	0,6542897905	14	0,7401473356
4	0,7934803587	15	0,7383692041
5	0,7013687736	16	0,7395672022
6	0,7639596829	17	0,7387603199
7	0,7221022425	18	0,7393038924
8	0,7504177618	19	0,7389377567
9	0,7314040424	20	0,7391843998
10	0,7442373549	21	0,7390182624
11	0,7356047404	22	0,7391301765

n	x_n	n	x_n
23	0,7390547907	42	0,7390851499
24	0,7391055719	43	0,739085122
25	0,7390713653	44	0,7390851408
26	0,7390944074	45	0,7390851281
27	0,739078886	46	0,7390851366
28	0,7390893414	47	0,7391301765
29	0,7390822985	48	0,7390851309
30	0,7390870427	49	0,7390851348
31	0,739083847	50	0,7390851322
32	0,7390859996	51	0,7390851339
33	0,7390845496	52	0,7390851327
34	0,7390855264	53	0,7390851335
35	0,7390848684	54	0,739085133
36	0,7390853116	55	0,7390851334
37	0,739085013	56	0,7390851331
38	0,7390852142	57	0,7390851333
39	0,7390850787	58	0,7390851332
40	0,7390851699	59	0,7390851332
41	0,7390851085	60	0,7390851332

Chegado a este ponto, depois de carregar 60 vezes na tecla "cos", o Zelmiro assustou-se (arrependendo-se até do impulso experimentalista, raro nele, como aliás foi referido). Ter-se-ia a calculadora avariado? Por que é que, depois de carregar 58 vezes nas teclas "cos", "Ans" e "EXE", por mais que continuasse a carregar nessas teclas, obtinha sempre o mesmo valor? Aflito (e não encontrando o número de telefone da "Reparadora Triturante de Calculadoras"), resolve fazer uma experiência análoga, mas partindo do valor 335 em vez do valor 1. Obteve:

n	x_n	n	x_n
1	–0,4081095818	31	0,739083183
2	0,917872722	32	0,7390864469
3	0,6075112505	33	0,7390842483
4	0,8210712016	34	0,7390857293
5	0,6814376114	35	0,7390847317
6	0,7766679555	36	0,7390854037
7	0,7132529455	37	0,739084951
8	0,7562374876	38	0,7390852559
9	0,7274229495	39	0,7390850505
10	0,7468904827	40	0,7390851889
11	0,7338048956	41	0,7390850957
12	0,7426316451	42	0,7390851585
13	0,7366915171	43	0,7390851162
14	0,740695383	44	0,7390851447
15	0,7379994919	45	0,7390851255
16	0,7398159986	46	0,7391301384
17	0,7385926162	47	0,7390851297
18	0,739416809	48	0,7390851356
19	0,7388616718	49	0,7390851316
20	0,7392356411	50	0,7390851343
21	0,7389837409	51	0,7390851325
22	0,7391534285	52	0,7390851337
23	0,739039127	53	0,7390851329
24	0,7391161228	54	0,7390851334
25	0,7390642579	55	0,7390851331
26	0,7390991949	56	0,7390851333
27	0,739075661	57	0,7390851332
28	0,7390915138	58	0,7390851332
29	0,7390808352	59	0,7390851332
30	0,7390880284	60	0,7390851332

O Zelmiro pensou que era bruxedo! Ao fim de premir 57 vezes as teclas "cos", "Ans" e "EXE" obtinha o mesmo valor de há pouco. Desalentado, resolveu fazer uma experiência análoga com a tecla "$\tan^{-1}$" (correspondente à função arctg), partindo do valor 1; eis os resultados:

n	x_n	n	x_n
1	0,7853981934	6	0,4541714733
2	0,66577375	7	0,4263175061
3	0,5873841757	8	0,4029860575
4	0,5310915102	9	0,3830779111
5	0,4882103378	10	0,3658337989

n	x_n	n	x_n
11	0,3507104329	36	0,2012661543
12	0,3373075815	37	0,198612719
13	0,325323103	38	0,1960612806
14	0,314524107	39	0,193605477
15	0,3047278305	40	0,1912394875
16	0,2957885899	41	0,1889579751
17	0,2875886511	42	0,1867560362
18	0,2800317036	43	0,1846291561
19	0,2730381016	44	0,1825731707
20	0,2665413363	45	0,180584238
21	0,2604853784	46	0,1786587773
22	0,2548226492	47	0,1767935042
23	0,2495124513	48	0,1749853452
24	0,2445197411	49	0,1732314471
25	0,2398141588	50	0,1715291523
26	0,2353692536	51	0,1698759817
27	0,2311618604	52	0,1682696198
28	0,2271715935	53	0,166707901
29	0,2233804326	54	0,165188797
30	0,2197723811	55	0,1637104065
31	0,2163331839	56	0,1622709445
32	0,2130500906	57	0,1608687339
33	0,2099116578	58	0,1595021966
34	0,2069075819	59	0,1581698468
35	0,204028557	60	0,1568702837

O Zelmiro não entendeu o que se passava. Obteve uma sucessão de números positivos que aparentemente é decrescente; mas sê-lo-á de facto? Será que tende para zero? Será que converge para outro limite? Por que é que a máquina se comporta de maneira diferente neste último caso, não estabilizando o resultado como nos casos anteriores?

O Zelmiro resolveu pensar um pouco e, ao chegar a conclusão nenhuma, decidiu escrever ao primo Zagalote, professor na "Universidade de Todas as Matemáticas", em Freixo-de-Espada-à-Cinta. O que se segue é a resposta do dito primo Zagalote. Para comodidade do leitor, refiro que essa resposta tem uma primeira secção onde o problema é equacionado e onde são introduzidos os conceitos necessários ao seu tratamento e que não fazem parte do programa do 12º ano do ensino secundário, uma segunda secção onde o problema é tratado e uma última secção onde a teoria estudada é aplicada à resolução de algumas equações não lineares. No final é apresentada uma bibliografia sucinta.

II. Uma lição magistral e suave do primo Zagalote.

1. Introdução.

Tivesse o Zelmiro conseguido analisar o que tinha feito, teria verificado que, em qualquer dos cálculos efectuados, tinha fixado um número real x_0 (no primeiro caso x_0=1, no segundo x_0=351 e no terceiro x_0=1) e tinha construído os primeiros 61 termos de uma sucessão $(x_n)_{n\in\mathbb{N}}$ assim definida por recorrência (designando por $\mathbb{N}$ o conjunto dos inteiros não nulos):

$$(1) \qquad \begin{cases} x_0 \in IR \\ x_{n+1} = f(x_n) \qquad \forall n \in IN \end{cases},$$

com f(x)=cos x nos dois primeiros casos e $f(x)=\mathrm{arctg}x$ no último caso. O problema que se põe é então estudar a convergência de uma tal sucessão e saber se essa convergência, e o valor do limite no caso de haver convergência, dependem ou não do valor x_0. Neste momento observaremos apenas que, se a referida sucessão for convergente, digamos $x_n \to \mathbf{a}$ com $\mathbf{a}\in\mathbb{R}$, e se f for contínua ter-se-á:

$$x_{n+1} = f(x_n) \quad \Rightarrow \quad \lim x_{n+1} = \lim f(x_n) \Rightarrow$$
$$\Rightarrow \quad \mathbf{a} = f(\lim x_n) \quad \Rightarrow \quad \mathbf{a} = f(\mathbf{a}) .$$

Vemos assim que **a** deverá ser uma solução da equação x=f(x). Atendendo aos gráficos da função identidade e da função cos facilmente concluímos que a equação x=cos x tem uma e uma única solução, pelo que o limite da sucessão $(x_n)_{n\in\mathbb{N}}$ definida por (1) com f(x)=cos x, se existir, deverá ser independente do valor x_0; já não nos devemos espantar por obter o mesmo valor quando partimos de x_0=1 ou de x_0=351 no caso de f ser a função cos. No caso da função arctg sabemos que a equação $x=\mathrm{arctg}\ x$ tem uma e uma única solução, nomeadamente x=0; logo a sucessão do terceiro exemplo, se convergir, converge para zero; será que tende de facto para zero?

Pelas considerações anteriores somos levados a considerar equações do tipo x=f(x), onde f é um função real de variável real. Para esse estudo necessitamos de alguns conceitos sobre sucessões que, ou não fazem parte dos programas do ensino secundário, ou não foram aí tratados com a profundidade conveniente.

Seja $(u_n)_{n\in\mathbb{N}}$ uma sucessão de números reais; relembramos que a sucessão $(u_n)_{n\in\mathbb{N}}$ se diz convergente se e só se existir um número real **a** tal que

$$(2) \qquad \forall\varepsilon>0 \;\; \exists p\in\mathbb{N} \;\; \forall n\in\mathbb{N} \;\; (n \geq p \;\Rightarrow\; |u_n - \mathbf{a}| \leq \varepsilon).$$

Nesse caso escreveremos $u_n \to \mathbf{a}$ e dizemos que a sucessão $(u_n)_{n\in IN}$ converge (ou tende) para **a**. É fácil provar que, se $u_n \to \mathbf{a}$ e $u_n \to \mathbf{b}$ (com **a** e **b** números reais quaisquer), então $\mathbf{a}=\mathbf{b}$; este resultado permite-nos definir o limite de uma sucessão convergente como sendo aquele único número real para o qual ela converge e escrever $\mathbf{a} = \lim u_n$ com o sentido de $u_n \to \mathbf{a}$.

Diremos que a sucessão $(u_n)_{n\in\mathbb{N}}$ de números reais é de Cauchy se e só se verificar a condição seguinte:

$$(3) \qquad \forall\varepsilon>0 \;\; \exists p\in\mathbb{N} \;\; \forall m,n\in\mathbb{N} \;\; (m,n \geq p$$
$$\Rightarrow |u_n - u_m| \leq \varepsilon).$$

Tendo em conta a desigualdade

$$|u_n - u_m| \;\leq\; |u_n - \mathbf{a}| + |\mathbf{a} - u_m|$$

é imediato provar que toda a sucessão convergente é de Cauchy. Será a recíproca verdadeira? Por outras palavras, será que toda a sucessão de Cauchy (em $\mathbb{R}$) é convergente? De facto assim é, mas a demonstração não se enquadra numa exposição deste tipo (para uma demonstração, ver [1]). Podemos pois enunciar o seguinte teorema:

Teorema 1. Uma sucessão de números reais é de Cauchy se e só se for convergente em $\mathbb{R}$.

♦♦♦

Podem os mais incautos perguntar-se: para quê então a noção de sucessão de Cauchy? A resposta não é difícil: pela simples razão que, por vezes, é muito mais fácil provar que uma sucessão é de Cauchy do que provar que é convergente. Isto porque, para provar por definição de limite a convergência de uma sucessão, é necessário conhecer o valor do limite, tal como se depreende de (2); para provar que ela é de Cauchy não é necessário conhecer o valor do limite, visto ele não intervir em (3).

Seja $\varepsilon>0$ e seja $x\in\mathbb{R}$; poremos, por definição, $V_\varepsilon(x)=[x-\varepsilon,x+\varepsilon]$ e diremos que $V_\varepsilon(x)$ é a vizinhança-ε de x. Diremos que um subconjunto X do conjunto dos números reais é fechado se e só se, para todo o y em $\mathbb{R}\backslash X$, existir $\varepsilon>0$ tal que a vizinhança-ε de y ainda esteja contida em $\mathbb{R}\backslash X$. São fechados, por exemplo, os seguintes subconjuntos de $\mathbb{R}$: $[0,3]$, $[1,+\infty[$, $\mathbb{N}$, $V_\varepsilon(x)$; já $[0,3[$, $]1,+\infty[$ e $\mathbb{Q}$ não são subconjuntos fechados de $\mathbb{R}$.

Teorema 2. Seja $(u_n)_{n\in\mathbb{N}}$ uma sucessão convergente de números reais e seja X um subconjunto fechado de $\mathbb{R}$. Então, se u_n pertencer a X para todo o n natural, também o limite de u_n pertence a X.

Demonstração. Ponhamos $\mathbf{a} = \lim u_n$. Queremos mostrar que, na hipótese de se ter $u_n\in X$ para todo o n natural, tem-se ainda $\mathbf{a}\in X$. Suponhamos, por absurdo, que $\mathbf{a}\notin X$; então, como X é fechado, deverá existir $\varepsilon_0>0$ tal que a vizinhança-ε_0 de $\mathbf{a}$ não possua pontos de X: $[\mathbf{a}-\varepsilon_0, \mathbf{a}+\varepsilon_0]\cap X = \emptyset$. Consequentemente tem-se, para todo o n natural, a desigualdade $|u_n-\mathbf{a}| > \varepsilon_0$, pelo que a condição (2) não é verificada com $\varepsilon=\varepsilon_0$. Mas isto é um absurdo porque, por hipótese, $u_n\to\mathbf{a}$.

♦♦♦

2. O teorema do ponto fixo de Banach.

Seja D um subconjunto não vazio de $\mathbb{R}$ e seja f uma aplicação de D em $\mathbb{R}$. Diremos que $x\in D$ é um *ponto fixo* de f se e só se

(4) $\quad f(x) = x$.

Nesta secção preocupar-nos-emos com a existência, unicidade e método de cálculo de pontos fixos de aplicações obedecendo a certas condições.

Começaremos com uma definição. Diremos que a função $f:D\to\mathbb{R}$ é *lipschitziana* se e só se existir uma constante $L>0$ tal que

(5) $\quad \forall x,y\in D \quad |f(x) - f(y)| \leq L\,|x - y|$.

Se f for lipschitziana, chamaremos *constante de Lipschitz* de f ao ínfimo das constantes $L>0$ que verificam (5). Se supusermos que o conjunto D é um intervalo aberto e que f é derivável em D, é imediato ver que f é lipschitziana se e só se a sua derivada for limitada em D. De facto, se f verificar (5) tem-se, para $x,\mathbf{a}\in D$, com $x\neq\mathbf{a}$,

$$\frac{|f(x)-f(a)|}{|x-a|}\leq L$$

pelo que, passando ao limite quando x tende para $\mathbf{a}$, se obtém $|f'(\mathbf{a})|\leq L$. Reciprocamente, se f' for limitada em D, sabemos que existe $C>0$ tal que $|f'(z)|\leq C$ para todo o z em D; pelo teorema de Lagrange sabemos que, para todo o $x,y\in D$, com $x\neq y$, existe ξ no intervalo de extremos x e y tal que $f(x)-f(y) = (x-y)f'(\xi)$ pelo que $|f(x)-f(y)| = |x-y|\,|f'(\xi)| \leq C|x-y|$: a função f verifica (5) com $L=C$. Mostrámos assim que, num intervalo aberto, uma função derivável f é lipschitziana se e só se a sua derivada for limitada e é fácil ver que a constante de Lipschitz de f é precisamente igual ao supremo do módulo da derivada de f no intervalo considerado.

Chamaremos *contracção* a toda a função real f, definida num subconjunto D não vazio de $\mathbb{R}$, lipschitziana em D, e tal que a sua constante de Lipschitz L seja inferior a 1. Vemos assim que f é uma contracção se e só se

(6) $\exists L\in]0,1[\quad \forall x,y\in D \quad |f(x) - f(y)| \leq L\,|x - y|$.

Repare-se que toda a contracção verifica a condição

(7) $\forall x,y\in D \quad |f(x) - f(y)| \leq |x - y|$;

no entanto esta condição não é suficiente para que f seja uma contracção. Por exemplo a função $f:]1,+\infty[\to\mathbb{R}$ definida por $f(x) = x + \frac{1}{x}$ verifica (7) e não é uma contracção porque o supremo do módulo da sua derivada em $]1,+\infty[$ (que coincide com a sua constante de Lipschitz) é 1 e não um número estritamente inferior a 1.

Teorema 3. (Teorema do ponto fixo de Banach). Seja D um subconjunto fechado não vazio de $\mathbb{R}$ e seja f uma aplicação de D em $\mathbb{R}$. Então, se f for uma contracção e se $f(D)\subset D$, f tem um e um só ponto fixo $\mathbf{a}\in D$ e, para todo o $x_0\in D$, a sucessão $(x_n)_{n\in\mathbb{N}}$ definida por recorrência por

(8) $x_{n+1}= f(x_n)$,

onde $n\in\mathbb{N}$, converge para **a**. Tem-se ainda

(9) $|x_n - a| \leq \frac{L^n}{1-L}|x_0 - x_1| \quad \forall n\in\mathbb{N}$;

(10) $|x_n - a| \leq \frac{L}{1-L}|x_n - x_{n-1}| \quad \forall n\in\mathbb{N}\setminus\{0\}$,

onde L é a constante de Lipschitz de f.

Demonstração. Seja $x_0\in D$ e seja $(x_n)_{n\in\mathbb{N}}$ a sucessão definida por (8); é evidente, devido à hipótese $f(D)\subset D$, que $x_n\in D$ para todo o $n\in\mathbb{N}$. Vamos mostrar que se trata de uma sucessão de Cauchy em $\mathbb{R}$; atendendo a (8) vê-se que, sendo L a constante de Lipschitz de f,

$$|x_{n+1} - x_n| = |f(x_n) - f(x_{n-1})| \leq L\,|x_n - x_{n-1}| =$$

$$= L\,|f(x_{n-1}) - f(x_{n-2})| \leq L^2|x_{n-1} - x_{n-2}| \leq\ldots\leq L^n|x_1 - x_0| \ .$$

Sejam agora $m,n\in\mathbb{N}$, com $m>n$; tem-se

$$|x_m - x_n| = |x_m - x_{m-1} + x_{m-1} - x_{m-2} + \ldots + x_{n+1} - x_n| \leq$$

$$\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \ldots + |x_{n+1} - x_n| \leq$$

$$\leq (L^{m-1} + \ldots + L^n)\,|x_1 - x_0|$$

$$= L^n\,\frac{1-L^{m-n}}{1-L}|x_1 - x_0| \leq \frac{L^n}{1-L}|x_1 - x_0| \ .$$

Obtivémos assim as majorações:

(11) $|x_{n+1} - x_n| \leq L^n|x_1 - x_0| \quad \forall n\in\mathbb{N}$;

(12) $|x_m - x_n| \leq \frac{L^n}{1-L}|x_1 - x_0| \quad \forall m,n\in\mathbb{N},\ m>n$.

Como, por hipótese, $0\leq L<1$, tem-se

$$\lim_{n\to+\infty}\frac{L^n}{1-L} = 0$$

e portanto, dado $\varepsilon>0$, existe $p\in\mathbb{N}$ tal que $\frac{L^n}{1-L}\leq\varepsilon$ para todo o $n\geq p$. Vemos assim que, para todo o $m,n\in\mathbb{N}$, $m\geq n\geq p$, se tem $|x_m - x_n| \leq \varepsilon\,|x_1 - x_0|$, o que mostra que a sucessão $(x_n)_{n\in\mathbb{N}}$ é de Cauchy. Sabemos então, pelo teorema 1, que $(x_n)_{n\in\mathbb{N}}$ é convergente em $\mathbb{R}$, ou seja existe $\mathbf{a}\in\mathbb{R}$ tal que $x_n\to\mathbf{a}$; como D é fechado, tem-se, pelo teorema 2, $\mathbf{a}\in D$.

Vamos demonstrar que **a** é o único ponto fixo de f. Comecemos por mostrar que **a** é de facto um ponto fixo de f; para cada $n\in\mathbb{N}$ tem-se

$$0 \leq |\mathbf{a} - f(\mathbf{a})| \leq |\mathbf{a} - x_{n+1}| + |x_{n+1} - f(\mathbf{a})| =$$

$$= |\mathbf{a} - x_{n+1}| + |f(x_n) - f(\mathbf{a})| \leq |\mathbf{a} - x_{n+1}| + L\,|x_n - \mathbf{a}| \ ;$$

fazendo $n\to+\infty$ vem, pelo teorema das sucessões enquadradas, $|\mathbf{a} - f(\mathbf{a})| = 0$, ou seja $\mathbf{a} = f(\mathbf{a})$, o que prova que **a** é um ponto fixo de f.

Seja agora **b** um ponto fixo de f; vamos mostrar que **b**=**a**. Atendendo a que $\mathbf{a} = f(\mathbf{a})$ e $\mathbf{b} = f(\mathbf{b})$ vem $|\mathbf{a} - \mathbf{b}| = |f(\mathbf{a}) - f(\mathbf{b})| \leq L\,|\mathbf{a} - \mathbf{b}|$, ou seja $(1 - L)\,|\mathbf{a} - \mathbf{b}| \leq 0$. Como $1 - L > 0$ conclui-se que $|\mathbf{a} - \mathbf{b}| \leq 0$, o que implica **a**=**b**.

Resta-nos mostrar as desigualdades (9) e (10). A desigualdade (9) decorre imediatamente de (12), passando ao limite quando $m \to +\infty$. Para demonstrarmos (10) fixamos $p \in \mathbb{N}$ e definimos a sucessão $(y_m)_{m\in\mathbb{N}}$ por recorrência da seguinte forma:

$$y_0 = x_p \quad ; \quad y_{m+1} = f(y_m) \quad .$$

Aplicando (9) a esta sucessão vem

$|y_1 - a| \leq \frac{L}{1-L}|y_0 - y_1|$ ou, atendendo a que $y_1 = f(y_0) = f(x_p) = x_{p+1}$,

$$|x_{p+1} - a| \leq \frac{L}{1-L}|x_p - x_{p+1}| \quad .$$

Fazendo $n = p+1$ obtemos (10).

♦♦♦

O teorema do ponto fixo de Banach indica-nos um algoritmo para o cálculo aproximado do ponto fixo de uma contracção $f: D \to \mathbb{R}$, sendo D um subconjunto fechado não vazio de $\mathbb{R}$ e aplicando f de D em D. Esse algoritmo, conhecido por método das aproximações sucessivas, escreve-se:

$$(13) \qquad \begin{cases} x_0 \in D \\ x_{n+1} = f(x_n) \end{cases} \quad .$$

Diremos que a sucessão $(x_n)_{n\in\mathbb{N}}$ definida por (13) é a sucessão de aproximações sucessivas da função f com origem no ponto x_0. Repare-se que, sendo **a** o ponto fixo de f, a sucessão de termo geral $|x_n - \mathbf{a}|$ é decrescente porque

$$|x_{n+1} - \mathbf{a}| = |f(x_n) - f(\mathbf{a})| \leq L\,|x_n - \mathbf{a}| \leq |x_n - \mathbf{a}| \,,$$

o que justifica o nome de método das aproximações sucessivas.

Diremos que (9) é uma majoração a priori e que (10) é uma majoração a posteriori. Repare-se ainda que as majorações (9) e (10) são tanto melhores quanto menor for L; de facto, de uma forma geral, o método das aproximações sucessivas (13) é tanto mais eficaz para o cálculo aproximado do ponto fixo de f quanto menor for L.

Suponhamos agora que queremos utilisar o algoritmo (13) para o cálculo aproximado do ponto fixo de uma contracção $f: D \to \mathbb{R}$ (D fechado não vazio de $\mathbb{R}$ e $f(D) \subset D$), de constante de Lipschitz $L \in [0,1[$. Por razões de ordem prática, normalmente relacionadas com erros de arredondamento, inevitáveis em cálculos com computadores (a aritmética de um computador é finita!), pode acontecer que não determinemos, para cada $n \in \mathbb{N}$, o valor exacto de $f(x_n)$, mas apenas uma sua aproximação. Qual a importância destes erros no cálculo do ponto fixo **a** de f? É o problema que vamos tratar em seguida.

Para cada $n \in \mathbb{N}$ seja f uma aplicação de D em $\mathbb{R}$ (que nem sequer supomos contínua), que aplique D em D e tal que:

$$(14) \qquad \exists M \geq 0 \quad \forall \xi \in D \quad \forall n \in \mathbb{N} \qquad |f_n(\xi) - f(\xi)| \leq M \quad .$$

Suponhamos que construímos a sucessão $(y_n)_{n\in\mathbb{N}}$ de números reais da forma seguinte:

$$(15) \qquad \begin{cases} y_0 = x_0 \\ y_{n+1} = f_n(y_n) \end{cases} \quad .$$

Será que poderemos controlar o erro cometido ao aproximarmos **a** por y_n? Claro que, se f_n fosse igual a f para todo o $n \in \mathbb{N}$, (15) não seria mais do que (13), tendo-se então a majoração do erro (9). Caso contrário tem-se, para todo o $n \in \mathbb{N}\backslash\{0\}$,

$$|\mathbf{a} - y_n| = |f(\mathbf{a}) - f_{n-1}(y_{n-1})| \leq$$

$$\leq |f(\mathbf{a}) - f(y_n)| + |f(y_n) - f(y_{n-1})| + |f(y_{n-1}) - f_{n-1}(y_{n-1})| \leq$$

$$\leq L\,|\mathbf{a} - y_n| + L\,|y_n - y_{n-1}| + M \quad ,$$

donde se conclui que

$$(16) \qquad |\mathbf{a}-y_n| \leq \frac{M}{1-L}+\frac{L}{1-L}|y_n-y_{n-1}| \, .$$

Resta-nos majorar $|y_n - y_{n-1}|$; tem-se:

$$|y_n - y_{n-1}| = |f_{n-1}(y_{n-1}) - f_{n-2}(y_{n-2})| \leq$$

$$\leq |f_{n-1}(y_{n-1}) - f(y_{n-1})| + |f(y_{n-1}) - f(y_{n-2})| +$$

$$+ |f(y_{n-2}) - f_{n-2}(y_{n-2})| \leq 2M + L\,|y_{n-1} - y_{n-2}| \quad ,$$

qualquer que seja $n \in \mathbb{N}$, $n \geq 2$. Reiterando as majorações precedentes concluímos que

$$|y_n - y_{n-1}| \leq 2M + 2ML + 2ML^2 + \ldots + 2ML^{n-2} + L^{n-1}|y_1 - y_0| =$$

$$= 2M\frac{1-L^{n-1}}{1-L}+L^{n-1}|y_1-y_0| \quad .$$

Introduzindo esta majoração em (16) vemos que

$$|\mathbf{a}-y_n| \leq \frac{M}{1-L}+\frac{L}{1-L}\left(2M\frac{1-L^{n-1}}{1-L}+L^{n-1}|y_1-y_0|\right)$$

$$= M\frac{1+L-2L^n}{(1-L)^2}+\frac{L^n}{1-L}|y_1-y_0| \leq$$

$$\leq M\frac{1+L}{(1-L)^2}+\frac{L^n}{1-L}|y_1-y_0| \quad .$$

Obtivémos assim a majoração

$$(17) \qquad |\mathbf{a}-y_n| \leq M\frac{1+L}{(1-L)^2}+\frac{L^n}{1-L}|f_0(y_0)-y_0| \quad ,$$

válida para todo o $n \in \mathbb{N}$. Nesta majoração a segunda parcela da soma que aparece no segundo membro é semelhante ao da majoração (9) e é um erro de método, enquanto a primeira parcela tem em conta a aproximação de f por f_n, sendo um erro de arredondamento. Claro que este primeiro termo é nulo quando $f_n = f$, pois nesse caso pode tomar-se $M = 0$; este termo é tanto menor quanto menor for L ou seja a influência de L em cada um dos termos do segundo membro de (17) é do mesmo tipo.

Vamos agora ver que a equação (4) pode ser tratada localmente, sem supormos que f é uma contracção em todo o domínio D. Na verdade basta supormos que f é uma contracção numa vizinhança-ε de uma solução de (4), como se prova no teorema que apresentamos de seguida.

Teorema 4. Seja D um subconjunto não vazio de $\mathbb{R}$ e seja f uma função de D em $\mathbb{R}$. Seja **a** um ponto de D, solução da equação (4), e suponhamos que existe $\varepsilon>0$ tal que $V_\varepsilon(\mathbf{a})$ esteja contida em D e f seja uma contracção em $V_\varepsilon(\mathbf{a})$ (queremos com isto dizer que a restrição de f a $V_\varepsilon(\mathbf{a})$ é uma contracção). Então **a** é a única solução da equação (4) em $V_\varepsilon(\mathbf{a})$. Para além disso, o método de aproximações sucessivas (13) converge para **a** qualquer que seja $x_0 \in V_\varepsilon(\mathbf{a})$.

Demonstração. Fixemos $\varepsilon>0$ tal que $V_\varepsilon(\mathbf{a}) \subset D$ e f seja uma contracção em $V_\varepsilon(\mathbf{a})$. É imediato ver que f aplica $V_\varepsilon(\mathbf{a})$ em $V_\varepsilon(\mathbf{a})$ porque, para todo o $y \in V_\varepsilon(\mathbf{a})$, se tem (sendo L a constante de Lipschitz de f restrita a $V_\varepsilon(\mathbf{a})$) $|f(y) - \mathbf{a}| = |f(y) - f(\mathbf{a})| \leq L\,|y - \mathbf{a}| \leq |y - \mathbf{a}|$. O resultado é então uma consequência directa do teorema do ponto fixo de Banach.

♦♦♦

Claro que, nas condições do teorema anterior, temos a garantia de existência e unicidade de solução de (4) em $V_\varepsilon(\mathbf{a})$, mas pode haver outras soluções de (4) fora de $V_\varepsilon(\mathbf{a})$. Também não há qualquer garantia de que a sucessão definida por (13) convirja para **a** se x_0 não pertencer a $V_\varepsilon(\mathbf{a})$. Para x_0 pertencente a $V_\varepsilon(\mathbf{a})$ a sucessão de aproximações sucessivas definida por (13) converge para **a** e verifica as majorações (9) e (10). No caso de, em cada iteração n, aproximarmos f por uma função f_n que aplique $V_\varepsilon(\mathbf{a})$ em $V_\varepsilon(\mathbf{a})$ e que obedeça à majoração (14) em $V_\varepsilon(\mathbf{a})$, são verificadas a majoração a posteriori (16) e a majoração a priori (17).

Retomemos agora os exemplos do Zelmiro (secção I). Nos primeiros dois exemplos a função f era o coseno. Esta função não é uma contracção em $\mathbb{R}$ visto a sua constante de Lipschitz ser 1 (a derivada de cos é –sen e o máximo do módulo do seno em $\mathbb{R}$ é 1). No entanto trata-se de um contracção de [0,1] em [0,1] visto que aplica o intervalo [0,1] em si próprio e, para todo o $x \in [0,1]$, verifica $|\cos'(x)|=|\operatorname{sen}x| \leq \operatorname{sen}1 < 1$ (porque $1< \frac{\pi}{2}$). Logo, partindo de $x_0=1$, a sucessão de aproximações sucessivas deve convergir para a solução de $x=\cos x$ em [0,1] que, neste caso particular, é a única solução de $x=\cos x$ em $\mathbb{R}$; foi o que se passou no primeiro exemplo. Sendo x_0 um valor real arbitrário, tem-se $x_1=\cos x_0$ e portanto x_1 pertence a [-1,1]; então $x_2=\cos x_1$ pertence a [0,1], pelo que a sucessão de aproximações sucessivas também converge para a solução de $x=\cos x$; foi o que se passou no segundo exemplo. No terceiro exemplo, relativo à função arctg, o único ponto fixo é o zero e em nenhuma vizinhança de zero a função arctg é uma contracção porque a sua derivada no ponto zero é 1. No entanto a sucessão construída no exemplo converge de facto para zero porque, por um lado é uma sucessão decrescente e minorada de números reais, logo convergente; por outro, já sabemos que, sendo convergente, só pode convergir para uma solução de $x=\operatorname{arctg} x$; mas a única solução desta equação é 0. No entanto o número de iterações a efectuar é substancialmente maior do que no caso do coseno.

3. A equação f(x)=b.

Seja b um número real qualquer e seja $f:\mathbb{R}\to\mathbb{R}$ uma função dada. Nesta secção vamos considerar a equação

(18) $\quad f(x) = b$.

Veremos que, sob certas hipóteses para a função f, a equação (18) tem uma e uma única solução, que pode ser aproximada através de um método de aproximações sucessivas.

Seja r um número real diferente de zero; a equação (18) é equivalente à equação

(19) $\quad x = x - r\,(f(x) - b)$.

A equação (19) pode escrever-se

(20) $\quad x = T_r(x)$,

onde, para cada $r \neq 0$, T_r é a aplicação de $\mathbb{R}$ em $\mathbb{R}$ definida por

(21) $\quad T_r(x) = x - r\,(f(x) - b)$.

As equações (18) e (20) são equivalentes; pelo teorema do ponto fixo de Banach podemos garantir a existência e unicidade de solução de (20) sempre que T_r seja uma contracção, para algum $r \neq 0$. Queremos agora dar condições sobre a função f que garantam que, para algum $r \neq 0$, T_r seja uma contracção.

Teorema 5. Seja $f:\mathbb{R}\to\mathbb{R}$ uma função derivável e seja $b\in\mathbb{R}$. Então, se existirem números reais $c,L>0$ tais que $c \leq f'(x) \leq L$ para todo o x em $\mathbb{R}$, a função T_r definida por (21) é uma contracção em $\mathbb{R}$ para todo o r tal que

(22) $\quad r \in \left]\, 0\, , \frac{2}{L} \,\right[$.

Demonstração. Atendendo a (21) tem-se, para todo o $x,y\in\mathbb{R}$:

$$|T_r(x)-T_r(y)| = |x - r(f(x) - b) - (y - r(f(y) - b))| = \\ = |x - y - r(f(x) - f(y))| \; .$$

Se $x\neq y$ sabemos, pelo teorema de Lagrange, que existe um número real ξ pertencente ao intervalo de extremos x e y tal que $f(x) - f(y) = (x - y)\, f'(\xi)$ e portanto $|T_r(x)-T_r(y)| = |x - y - r(x - y)\, f'(\xi)| = |x - y|\; |1 - r\, f'(\xi)|$.

Utilizando as desigualdades $c \leq f'(\xi) \leq L$ vemos que $1 - rL \leq 1 - r\, f'(\xi) \leq 1 - rc$, pelo que $|1 - r\, f'(\xi)| \leq \max\{|1 - rL| , |1 - rc|\}$. Portanto

$$| T_r(x) - T_r(y) | \leq |x - y| \max \{|1 - rL| , |1 - rc|\}$$

e T_r será uma contracção sempre que $\max \{|1 - rL| , |1 - rc|\} < 1$. Na figura 1 esboçamos o gráfico da função $r \to \max \{|1 - rL| , |1 - rc|\}$ e vemos que ela é estritamente inferior a 1 se e só se r verifica (22). Então, para todo o r verificando (22), a aplicação T_r é uma contracção.

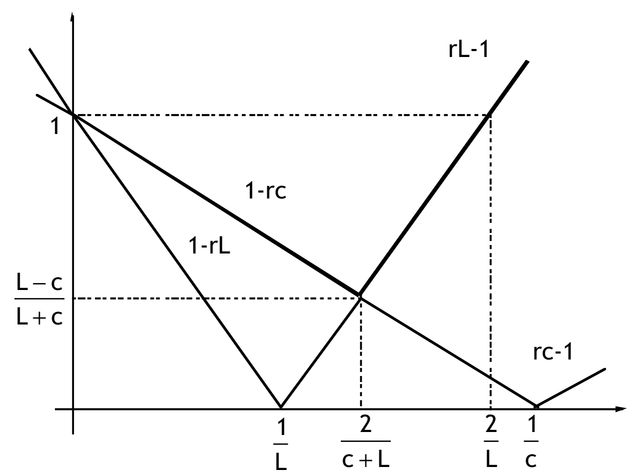

Figura 1.

♦♦♦

O valor óptimo para o parâmetro r é, como se depreende da figura 1, $\frac{2}{c+L}$, sendo c e L respectivamente o ínfimo e o supremo da derivada de f. No entanto, de forma análoga ao que fizemos no teorema 4, podemos restringir f a um intervalo que contenha a única solução **a** de (18), pelo que o importante são de facto os valores do ínfimo e do supremo da derivada de f numa vizinhança de **a**. Como a vizinhança pode ser arbitrariamente pequena e como ambos aqueles valores tendem para f'(**a**), o valor ideal para o parâmetro r será $\frac{1}{f'(a)}$. Claro que este valor não nos é acessível a priori, visto que desconhecemos o ponto **a**.

O número de iterações na resolução de (18) pelo método das aproximações sucessivas aplicado à função definida por (21) varia consideravelmente com r. Na tabela que se segue apresentamos o número de iterações como função de r para a resolução da equação x + arctg x = 1, cuja solução aproximada (arredondada até às dez casas decimais) é **a**=0,5202689927 ; partimos sempre da aproximação inicial x_0=1.

Resolução de x + arctg x = 1 com x_0= 1

r	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
Iterações	120	54	31	19	11	9	17	27	47	97

r	0,55	0,56	0,57	0,58	0,59
Iterações	7	5	6	8	8

Repare-se que o número de iterações varia fortemente com r e que os melhores resultados são obtidos para r = 0,56.

Como já dissemos, a solução arredondada de f(x)=1, sendo f a função definida por f(x)=x+arctg x, é **a**=0,5202689927.

Tem-se $f'(x) = 1 + \frac{1}{1+x^2}$ e portanto $\frac{1}{f'(a)}$ é aproximadamente igual (com arredondamento até á décima casa decimal) a 0,5596032567. Escolhendo este valor para r chega-se à solução em apenas 4 iterações, partindo de x_0=1. Mesmo partindo de valores afastados da solução o número de iterações é pequeno como se vê na tabela seguinte:

Resolução de x + arctg x = 1 com r = 0,5596032567

valor inicial	0,1	1	10	100	1 000	10 000	100 000
Iterações	4	4	6	9	12	15	18

III. Bibliografia.

[1] Jaime Campos Ferreira, "Introdução à Análise Matemática", Fundação Calouste Gulbenkian, Lisboa, 1985

[2] Erwin Kreyszig, "Introductory Functional Analysis with Applications", John Wiley & Sons, New York, 1978

[3] Moïse Sibony, Jean-Claude Mardon, "Analyse Numérique I - Systèmes linéaires et non linéaires", Hermann, Paris, 1982