Notas sobre a Fórmula de Taylor e o estudo de extremos

O Teorema de Taylor estabelece que (sob certas condições) uma função pode ser aproximada (na proximidade de algum ponto dado) por um polinómio, de modo que o erro que se comete ao substituir a função pelo polinómio seja *pequeno*. Começamos por considerar o caso de funções de uma só variável.

1 Funções de uma variável

1.1

É sabido que a função exponencial é definida por meio da série:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots \tag{1}$$

Quando $|x| < 1$, os valores de x^2, x^3, etc. são menores (em módulo) que $|x|$. É de esperar que, para valores de x *pequenos*, se tenha

$$e^x \approx 1 + x \ , \tag{2}$$

onde o símbolo '$\approx$' se lê "é aproximadamente igual a", e deve ser entendido de modo informal. (Note que a validade da aproximação (2) não é evidente, uma vez que a quantidade desprezada é a soma de parcelas pequenas mas em número infinito.)

Do ponto de vista geométrico, trata-se de aproximar o gráfico da função exponencial pela recta de equação $y = 1 + x$ (observe que esta recta é a tangente ao gráfico no ponto de abcissa 0). Definindo a função $p_1 : \mathbb{R} \to \mathbb{R}$ por

$$p_1(x) = x + 1 \ ,$$

o gráfico desta função é precisamente a recta em causa.

Suponhamos agora que pretendemos saber se a função exponencial tem um extremo no ponto $x = 0$ (sabemos que não é verdade, mas trata-se de tentar estudar a questão com base na aproximação (2)). Se houver um mínimo local em $x = 0$, teremos

$$e^x \geq e^0$$

para todo x pertencente a alguma vizinhança de 0; isto é, em alguma vizinhança de 0 a diferença $e^x - e^0$ será sempre positiva ou nula. Assim, o que se pretende é saber se o sinal daquela diferença é sempre o mesmo (e qual) ou se depende de x estar à direira ou à esquerda de 0.

Admitindo que a função pode de facto ser aproximada pela função p_1 em alguma vizinhança de 0, é de esperar que o sinal da diferença seja o mesmo quando se substitui a função exponencial pela função p_1. Para esta função, vemos que:

$$p_1(x) - p_1(0) = x \begin{cases} < 0 & \text{se } x < 0 \\ > 0 & \text{se } x > 0 \end{cases} \tag{3}$$

Por conseguinte, a função p_1 não tem extremo local em 0 (era evidente, uma vez que o seu gráfico é uma recta de declive não nulo). Para ver se podemos estender esta conclusão à função exponencial, há que substituir a *igualdade aproximada* (2) por uma igualdade rigorosa:

$$e^x = 1 + x + r_1(x) \tag{4}$$

A função r_1 (resto de $1^{\underline{a}}$ ordem) é definida precisamente como

$$r_1(x) = e^x - p_1(x) \ .$$

(Note que a igualdade (4), por si só, não nos dá nenhuma informação! O importante é saber como se comporta a função r_1.) Estudemos agora o sinal da diferença $e^x - e^0$ (quando $x \neq 0$):

$$\begin{aligned} e^x &= 1 + x + r_1(x) \\ e^x - e^0 &= x + r_1(x) \\ &= x\left(1 + \frac{r_1(x)}{x}\right) \end{aligned} \tag{5}$$

A diferença entre este caso e o de (3) é que o factor x vem multiplicado por $1 + \frac{r_1(x)}{x}$. Ora, se o valor de $\frac{r_1(x)}{x}$ for pequeno comparado com 1, o sinal de $1 + \frac{r_1(x)}{x}$ será igual ao sinal de 1, e portanto o sinal de (5) será igual ao sinal de x. Deste modo poderemos reduzir o estudo da existência de extremo da função exponencial ao da existência de extremo da função p_1, e concluiremos que não existe extremo.

A questão será, então: que condição deveria o resto r_1 satisfazer para garantir que o valor de $\frac{r_1(x)}{x}$ é pequeno comparado com 1? Observe que estamos a supor que $r_1(x)$ é *pequeno* (para que a aproximação (2) seja válida), mas é preciso que seja pequeno mesmo quando comparado com x. Podemos formalizar este requisito exigindo que:

$$\lim_{x \to 0} \frac{r_1(x)}{x} = 0 \tag{6}$$

De facto, se isto for verdade, então o valor de $\frac{r_1(x)}{x}$ será menor que 1/2 (por exemplo) para todo x pertencente a alguma vizinhança de zero, e para estes valores de x virá

$$1 - \frac{1}{2} < 1 + \frac{r_1(x)}{x} < 1 + \frac{1}{2} \ ,$$

isto é, o valor de $1 + \frac{r_1(x)}{x}$ será positivo.

É fácil provar que a condição (6) é efectivamente verificada (demonstre este facto, recorrendo a uma majoração da série e atendendo a que para cálculo do limite só interessam os valores de x tais que $|x| < 1$).

1.2

Consideremos a função f definida em $\mathbb{R}$ pela expressão $f(x) = 2\cos x$, e vejamos o que se passa no ponto $x = 0$ (sabemos que tem um máximo local, mas o que se pretende é chegar a este resultado com base nas considerações anteriores).

A função f é dada por:

$$f(x) = 2\left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots\right) \tag{7}$$

Isto sugere que se tenha

$$f(x) \approx 2 \tag{8}$$

para x *pequeno*. Em rigor, fazendo $p_1(x) = 2$ e $r_1(x) = f(x) - p_1(x)$ teremos $f(x) = p_1(x) + r_1(x)$; o importante é que

$$\lim_{x \to 0} \frac{r_1(x)}{x} = 0 \ ,$$

isto é, para valores de x próximos (mas diferentes) de zero o resto é pequeno quando comparado com o próprio x.

Para verificar se existe extremo em $x = 0$, estudemos o sinal da diferença $f(x) - f(0)$:

$$\begin{aligned} f(x) &= 2 + r_1(x) \\ f(x) - f(0) &= r_1(x) \end{aligned}$$

Sabemos que $r_1(x)$ é muito pequeno (em módulo) quando x é muito próximo de zero, mas não sabemos qual o seu sinal; por conseguinte, esta aproximação (isto é, a aproximação (8)) de nada serviu.

No entanto, a expressão (7) sugere que se aproxime a função por um polinómio de grau 2, o que significa que o gráfico de f é aproximadamente uma parábola na vizinhança do ponto de abcissa 0:

$$\begin{aligned} f(x) &\approx 2 - x^2 \\ f(x) &= 2 - x^2 + r_2(x) \end{aligned}$$

Temos agora

$$r_2(x) = 2\left(\frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots\right) \ , \tag{9}$$

e prova-se (recorrendo à definição de r_2) que:

$$\lim_{x \to 0} \frac{r_2(x)}{x^2} = 0 \tag{10}$$

Virá então:

$$\begin{aligned} f(x) &= 2 - x^2 + r_2(x) \\ f(x) - f(0) &= -x^2 + r_2(x) \\ &= x^2\left(-1 + \frac{r_2(x)}{x^2}\right) \end{aligned} \tag{11}$$

O valor de $\frac{r_2(x)}{x^2}$ não afecta o sinal da soma $-1 + \frac{r_2(x)}{x^2}$, desde que x esteja numa vizinhança suficientemente pequena para que $\left|\frac{r_2(x)}{x^2}\right| < 1$ (o que é possível devido a (10)). Portanto, para $x \neq 0$ (nessa vizinhança) o valor de (11) será sempre negativo, o que significa que f tem um máximo local em $x = 0$.

Note-se que chegámos a este resultado sem utilizar a expressão (9)—apenas a propriedade (10) foi necessária. Repare ainda que as funções f e p_2 têm o mesmo valor na origem, e que o mesmo sucede às suas derivadas de $1^{\underline{a}}$ e $2^{\underline{a}}$ ordem.

1.3

Os procedimentos anteriores generalizam-se a uma classe vasta de funções (não necessariamente definidas por séries de potências). O Teorema de Taylor estabelece que se uma função f for diferenciável n vezes num ponto a então é válida (numa vizinhança de a) a aproximação

$$f(x) \approx p_n(x) \ ,$$

ou, em rigor,

$$f(x) = p_n(x) + r_n(x) \ ,$$

onde p_n é o polinómio de Taylor (de ordem n) da função f relativo ao ponto a, definido por

$$p_n(x) = f(a) + \frac{1}{1!}f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n \ ,$$

e a função r_n (resto de Taylor de ordem n da função f relativo ao ponto a) satisfaz a condição:

$$\lim_{x \to a} \frac{r_n(x)}{(x-a)^n} = 0$$

Dados f, n e a, o polinómio p_n anterior é o único polinómio para o qual o resto r_n satisfaz aquela condição. Tem a propriedade de todas as suas derivadas (até à ordem n) no ponto a coincidirem com as respectivas derivadas de f no mesmo ponto.

Como usar a fórmula de Taylor para estudar a existência de um extremo de f no ponto a? Comecemos por recorrer à fórmula de 1ª ordem:

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + r_1(x) \\ f(x) - f(a) &= (x-a)\left(f'(a) + \frac{r_1(x)}{(x-a)}\right) \end{aligned} \tag{12}$$

Supondo que $f'(a) \neq 0$, o sinal de $f'(a) + \frac{r_1(x)}{(x-a)}$ será o mesmo de $f'(a)$ (numa vizinhança de a suficientemente pequena); logo, o produto por $x - a$ terá sinais diferentes consoante $x > a$ ou $x < a$, pelo que não existe extremo no ponto a. Conclui-se assim que para existir extremo num ponto (em que a função seja diferenciável!) a derivada deve ser nula (facto que já era conhecido de Análise I).

Supondo agora que $f'(a) = 0$, a fórmula (12) nada permite concluir. Se a função f tiver segunda derivada finita, podemos recorrer à aproximação de 2ª ordem:

$$\begin{aligned} f(x) - f(a) &= f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + r_2(x) \\ &= (x-a)^2\left(\frac{1}{2}f''(a) + \frac{r_2(x)}{(x-a)^2}\right) \end{aligned}$$

Se $f''(a) \neq 0$, a diferença $f(x) - f(a)$ tem sempre (numa vizinhança de a) o mesmo sinal, uma vez que é igual ao produto de $(x-a)^2$ (sempre não-negativo) por um valor que tem um mesmo sinal (o sinal de $f''(a)$). Neste caso conclui-se que existe extremo em a (de que tipo?).

Se $f''(a) = 0$, estamos numa situação semelhante à anterior, e há que tentar uma aproximação de maior ordem. (Que acontece se $f'''(0) \neq 0$? E se $f'''(0) = 0$ mas $f^{(4)}(0) \neq 0$?)

> Note que pode acontecer que todas as derivadas de f no ponto a existam mas sejam nulas (sem que a função seja nula!); também pode suceder que todas as derivadas se anulem até certa ordem e não exista derivada de ordem maior—nestes casos, o método é inconclusivo.

2 Funções de várias variáveis

Para funções reais f definidas em subconjuntos de $\mathbb{R}^m$, o Teorema de Taylor tem enunciado semelhante ao do caso $m = 1$, embora a notação possa complicar a sua leitura.

2.1

Comecemos por retomar o caso $m = 1$ com a mudança de variável $h = x - a$; a fórmula de 3ª ordem (por exemplo) assume a forma:

$$f(a+h) = f(a) + hf'(a) + \frac{1}{2!}h^2 f''(a) + \frac{1}{3!}h^3 f'''(a) + r_3(h)$$

(Repare-se que continuamos a designar o resto pelo mesmo símbolo r_3—em rigor, deveríamos escrever $r_3(a+h)$, ou então substituir 'r' por outra letra!)

Utilizando a notação $(Df)(a)$, $(D^2 f)(a)$, etc., obtemos:

$$f(a+h) = f(a) + h(Df)(a) + \frac{1}{2!}h^2(D^2 f)(a) + \frac{1}{3!}h^3(D^3 f)(a) + r_3(h)$$

O valor de $f(a+h) - r_3(h)$ (que podemos designar por $p_3(h)$, com o mesmo abuso de linguagem já mencionado a propósito do resto) pode ser escrito na forma:

$$p_3(h) = \left(f + (hD)f + \frac{1}{2!}(h^2D^2)f + \frac{1}{3!}(h^3D^3)f\right)(a)$$

Tente compreender cuidadosamente o significado do 2º membro! Por exemplo, a expressão 'h^2D^2' não designa um número, nem uma função, mas sim uma operação—a operação que consiste em "derivar duas vezes e multiplicar pelo número h^2"; uma expressão como '$(h^2D^2)f$' designa uma função—a função que se obtém aplicando a operação anterior à função f; por fim, '$((h^2D^2)f)(a)$' já designa um número—o valor da função anterior no ponto a.

> No que se segue, escrevemos $(h^2D^2)f$ em vez de $\big((h^2D^2)f\big)(a)$—em geral, subentende-se que todas as funções estão aplicadas a um dado ponto a.

Na mesma ordem de ideias, a expressão '$(hD)^2$' designa a operação que consiste em "derivar e multiplicar por h, duas vezes" (ou seja, $(hD)^2 = (hD) \circ (hD)$). Como é claramente o mesmo que "derivar duas vezes e multiplicar por h^2," obteremos ainda:

$$p_3(h) = \left(f + (hD)f + \frac{1}{2!}(hD)^2 f + \frac{1}{3!}(hD)^3 f\right) = \left(1 + (hD) + \frac{1}{2!}(hD)^2 + \frac{1}{3!}(hD)^3\right)f \quad (13)$$

A última expressão obteve-se "pondo em evidência" f. Isto é legitimado pela definição de *soma de* duas (ou mais) *aplicações*: sendo T, S aplicações (como hD, $(hD)^2$, etc.), a soma $T + S$ associa a cada f a função $Tf + Sf$.

2.2

Passando ao caso das funções de m variáveis, a fórmula de Taylor (de ordem n) será

$$f(\mathbf{a} + \mathbf{h}) = p_n(\mathbf{h}) + r_n(\mathbf{h}) \ , \quad (14)$$

onde $\mathbf{a} = (a_1, \ldots, a_m)$ e $\mathbf{h} = (h_1, \ldots, h_m)$. A função p_n é ainda um polinómio (nas variáveis $h_1, \ldots, h_m$) e o resto verifica a condição:

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{r_n(\mathbf{h})}{\|\mathbf{h}\|^n} = 0$$

A expressão de p_n é:[1]

$$p_n(\mathbf{h}) = \left(1 + (\mathbf{h}\cdot\nabla) + \frac{1}{2!}(\mathbf{h}\cdot\nabla)^2 + \cdots + \frac{1}{n!}(\mathbf{h}\cdot\nabla)^n\right) f \tag{15}$$

Para melhor compreender o significado da expressão do 2º membro, compare com (13) e recorde que para funções de várias variáveis o gradiente desempenha o papel que cabe à derivada em diversas fórmulas relativas a funções de uma só variável. Em vez do escalar h, temos agora o vector $\mathbf{h}$, e o "produto" hD é substituído pelo "produto interno" $\mathbf{h}\cdot\nabla$. A expressão '$\mathbf{h}\cdot\nabla$' designa a operação que à função f faz corresponder a função $(\mathbf{h}\cdot\nabla)f$, definida como:

$$(\mathbf{h}\cdot\nabla)f = (\mathbf{h}\cdot\nabla f) = (h_1, \ldots, h_m)\cdot\left(\tfrac{\partial f}{\partial x_1}, \ldots, \tfrac{\partial f}{\partial x_m}\right) = h_1\tfrac{\partial f}{\partial x_1} + \cdots + h_m\tfrac{\partial f}{\partial x_m}$$

Por sua vez, o valor desta função no ponto $\mathbf{a}$ será:

$$\big((\mathbf{h}\cdot\nabla)f\big)(\mathbf{a}) = (h_1\tfrac{\partial f}{\partial x_1} + \cdots + h_m\tfrac{\partial f}{\partial x_m})(\mathbf{a}) = h_1\tfrac{\partial f}{\partial x_1}(\mathbf{a}) + \cdots + h_m\tfrac{\partial f}{\partial x_m}(\mathbf{a})$$

Considerando $m = 2$ (como exemplo), podemos ainda escrever:

$$\mathbf{h}\cdot\nabla = (h_1, h_2)\cdot(\tfrac{\partial}{\partial x}, \tfrac{\partial}{\partial y}) = h_1\tfrac{\partial}{\partial x} + h_2\tfrac{\partial}{\partial y} \tag{16}$$

O segundo membro representa a operação tal que:

$$\big(h_1\tfrac{\partial}{\partial x} + h_2\tfrac{\partial}{\partial y}\big)f = h_1\tfrac{\partial f}{\partial x} + h_2\tfrac{\partial f}{\partial y}$$

Que significará $(\mathbf{h}\cdot\nabla)^2$? Tal como acontecia com $(hD)^2$, trata-se de aplicar duas vezes a mesma operação à função f:

$$\begin{aligned}(\mathbf{h}\cdot\nabla)^2 f &= \big((\mathbf{h}\cdot\nabla)\circ(\mathbf{h}\cdot\nabla)\big)f = (\mathbf{h}\cdot\nabla)\big((\mathbf{h}\cdot\nabla)f\big) = (h_1\tfrac{\partial}{\partial x} + h_2\tfrac{\partial}{\partial y})\big((h_1\tfrac{\partial}{\partial x} + h_2\tfrac{\partial}{\partial y})f\big)\\ &= (h_1\tfrac{\partial}{\partial x} + h_2\tfrac{\partial}{\partial y})\big(h_1\tfrac{\partial f}{\partial x} + h_2\tfrac{\partial f}{\partial y}\big) = h_1\tfrac{\partial}{\partial x}\big(h_1\tfrac{\partial f}{\partial x} + h_2\tfrac{\partial f}{\partial y}\big) + h_2\tfrac{\partial}{\partial y}\big(h_1\tfrac{\partial f}{\partial x} + h_2\tfrac{\partial f}{\partial y}\big)\\ &= h_1^2\tfrac{\partial^2 f}{\partial x^2} + h_1h_2\tfrac{\partial^2 f}{\partial x\partial y} + h_2h_1\tfrac{\partial^2 f}{\partial y\partial x} + h_2^2\tfrac{\partial^2 f}{\partial y^2}\end{aligned} \tag{17}$$

A fórmula de Taylor de ordem n é válida quando a função f é n vezes diferenciável no ponto $\mathbf{a}$. (A função é duas vezes diferenciável em $\mathbf{a}$ sse as suas derivadas de 1ª ordem são diferenciáveis em $\mathbf{a}$, é três vezes diferenciável sse as suas derivadas de 2ª ordem são diferenciáveis, etc.) Nestas condições, prova-se (**Teorema de Young**) que nas derivadas mistas de ordem não superior a n é possível trocar à vontade a ordem da derivação—$\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x}$, etc.

Assim, (17) equivale a:

$$\begin{aligned}(\mathbf{h}\cdot\nabla)^2 f &= h_1^2\frac{\partial^2 f}{\partial x^2} + 2h_1h_2\frac{\partial^2 f}{\partial x\partial y} + h_2^2\frac{\partial^2 f}{\partial y^2}\\ &= \left(h_1^2\frac{\partial^2}{\partial x^2} + 2h_1h_2\frac{\partial^2}{\partial x\partial y} + h_2^2\frac{\partial^2}{\partial y^2}\right)f\end{aligned} \tag{18}$$

[1]Não se pretende demonstrar aqui a fórmula.

Por outro lado, retomando a igualdade (16) podemos escrever:

$$(\mathbf{h}\cdot\nabla)^2 = \left(h_1\frac{\partial}{\partial x} + h_2\frac{\partial}{\partial y}\right)^2 \tag{19}$$

Comparando com (18), vemos que o quadrado no 2º membro de (19) pode ser desenvolvido como se se tratasse do quadrado de uma soma de números reais:

$$\begin{aligned}\left(h_1\frac{\partial}{\partial x} + h_2\frac{\partial}{\partial y}\right)^2 &= \left(h_1\frac{\partial}{\partial x}\right)^2 + \left(h_2\frac{\partial}{\partial y}\right)^2 + 2h_1h_2\frac{\partial}{\partial x}\frac{\partial}{\partial y}\\ &= h_1^2\frac{\partial^2}{\partial x^2} + h_2^2\frac{\partial^2}{\partial y^2} + 2h_1h_2\frac{\partial^2}{\partial x\partial y}\end{aligned}$$

Para entender estas igualdades, há que recordar que os "produtos" representam de facto composições de aplicações; ora, as propriedades dos números reais utilizadas na dedução da fórmula do binómio de Newton são ainda válidas para as aplicações em causa, desde que se substitua o produto de números pela composição de aplicações. Por conseguinte, a fórmula do binómio permanece válida, para qualquer potência; por exemplo:

$$\begin{aligned}(\mathbf{h}\cdot\nabla)^3 &= \left(h_1\frac{\partial}{\partial x} + h_2\frac{\partial}{\partial y}\right)^3\\ &= h_1^3\frac{\partial^3}{\partial x^3} + 3h_1^2h_2\frac{\partial^3}{\partial x^2\partial y} + 3h_1h_2^2\frac{\partial^3}{\partial x\partial y^2} + h_2^3\frac{\partial^3}{\partial y^3}\end{aligned}$$

Para calcular $(\mathbf{h}\cdot\nabla)^3 f$, podemos aplicar o 2º membro anterior a f, directamente; se for conveniente, também podemos atender a que

$$(\mathbf{h}\cdot\nabla)^3 f = (\mathbf{h}\cdot\nabla)\big((\mathbf{h}\cdot\nabla)^2 f\big)\ ,$$

ou seja:

$$\begin{aligned}(\mathbf{h}\cdot\nabla)^3 f &= \left(h_1\frac{\partial}{\partial x} + h_2\frac{\partial}{\partial y}\right)\left(h_1^2\frac{\partial^2 f}{\partial x^2} + 2h_1h_2\frac{\partial^2 f}{\partial x\partial y} + h_2^2\frac{\partial^2 f}{\partial y^2}\right)\\ &= h_1\frac{\partial}{\partial x}\Big(\dots\Big) + h_2\frac{\partial}{\partial y}\Big(\dots\Big)\\ &= \cdots\end{aligned}$$

2.3

O estudo dos pontos de extremo local de uma função de m variáveis faz-se de modo semelhante ao do caso $m = 1$. Em primeiro lugar, para que f tenha um extremo no ponto $\mathbf{a}$ é necessário que o gradiente de f se anule nesse ponto, desde que a função seja diferenciável no ponto:[2]

> Pensemos no caso $m = 2$, para fixar ideias. Se f tem extremo local em (a, b), então a função g tal que $g(x) = f(x, b)$ também tem extremo local em $x = a$—se a função f aumenta (por exemplo) quando nos afastamos um pouco do ponto (a, b), aumenta, em particular, quando nos afastamos de (a, b) ao longo da recta $y = b$. Logo, $g'(a) = 0$, ou seja, $\frac{\partial f}{\partial x}(a, b) = 0$. Pela mesma razão, $\frac{\partial f}{\partial y}(a, b) = 0$.

[2] O argumento seguinte mostra que não é necessário que a função seja diferenciável no ponto—basta que existam as derivadas parciais.

Por conseguinte, para além dos pontos em que o gradiente não exista, os únicos "candidatos" a pontos de extremo são os pontos em que o gradiente se anula—pontos de estacionaridade. Para cada um destes pontos, recorre-se à fórmula de Taylor para tentar estudar o sinal de $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})$.

Os exemplos seguintes ilustram as diversas situações que podemos encontrar. (Todos os exemplos se referem a funções $f : \mathbb{R}^2 \to \mathbb{R}$. O caso geral é essencialmente idêntico.)

Exemplo 1

Função: $f(x,y) = \dfrac{1}{1+x^2+y^2}$

Pontos de estacionaridade?

$$\frac{\partial f}{\partial x} = -\frac{2x}{(1+x^2+y^2)^2} \qquad \frac{\partial f}{\partial y} = -\frac{2y}{(1+x^2+y^2)^2}$$
$$\nabla f = \mathbf{0} \Leftrightarrow (x,y) = (0,0)$$

Fórmula de Taylor:

$$\begin{aligned} f(h_1,h_2) &= f(0,0) + \underbrace{(\mathbf{h}\cdot\nabla)f}_{=0 \text{ no ponto } (0,0)} + \frac{1}{2!}(\mathbf{h}\cdot\nabla)^2 f + r_2(h_1,h_2) \\ &= f(0,0) + \frac{1}{2!}\left(h_1^2\frac{\partial^2 f}{\partial x^2} + h_2^2\frac{\partial^2 f}{\partial y^2} + 2h_1h_2\frac{\partial^2 f}{\partial x\partial y}\right) + r_2(h_1,h_2) \end{aligned}$$

Termo de 2ª ordem?

$$\frac{\partial^2 f}{\partial x^2}(0,0) = -2$$
$$\frac{\partial^2 f}{\partial y^2}(0,0) = -2$$
$$\frac{\partial^2 f}{\partial y\partial x}(0,0) = 0$$

Fórmula de Taylor:

$$\begin{aligned} f(h_1,h_2) - f(0,0) &= \frac{1}{2!}(-2h_1^2 - 2h_2^2) + r_2(h_1,h_2) \\ &= (h_1^2+h_2^2)\left(-\frac{2}{2!} + \frac{r_2(h_1,h_2)}{h_1^2+h_2^2}\right) \\ &= \|\mathbf{h}\|^2 \underbrace{\left(-1 + \underbrace{\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}}_{\to 0 \text{ qd. } \mathbf{h}\to\mathbf{0}}\right)}_{\leq 0 \text{ para } \|\mathbf{h}\| \text{ pequena}} \end{aligned}$$

Logo, $(0,0)$ é ponto de máximo local. (Para esta função, era fácil ver imediatamente que a função tem um único extremo, na origem, e que se trata de um máximo global.)

Exemplo 2

Função: $f(x,y) = xy + x + 1$

Pontos de estacionaridade?

$$\frac{\partial f}{\partial x} = y + 1 \qquad \frac{\partial f}{\partial y} = x$$
$$\nabla f = \mathbf{0} \Leftrightarrow (x,y) = (0,-1)$$

Derivadas de $2^{\underline{a}}$ ordem:

$$\frac{\partial^2 f}{\partial x^2}(0,-1) = 0 \qquad \frac{\partial^2 f}{\partial y^2}(0,-1) = 0 \qquad \frac{\partial^2 f}{\partial x \partial y}(0,-1) = \frac{\partial^2 f}{\partial y \partial x}(0,-1) = 1$$

Fórmula de Taylor:

$$\begin{aligned} f(0+h_1,-1+h_2) - f(0,-1) &= \frac{1}{2!}(h_1^2 \cdot 0 + h_2^2 \cdot 0 + 2h_1h_2) + r_2(h_1,h_2) \\ &= \|\mathbf{h}\|^2 \underbrace{\left(\frac{h_1h_2}{h_1^2+h_2^2} + \underbrace{\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}}_{\to 0}\right)}_{??} \end{aligned}$$

A expressão $\frac{h_1h_2}{h_1^2+h_2^2}$ não tem sempre o mesmo sinal (compare com o Exemplo 1, onde havia uma constante no lugar desta expressão). Por exemplo, para $(h_1,h_2) = (r,r)$, com $r > 0$, vem $\frac{h_1h_2}{h_1^2+h_2^2} > 0$, e para $(h_1,h_2) = (r,-r)$ vem $\frac{h_1h_2}{h_1^2+h_2^2} < 0$. No caso $(h_1,h_2) = (r,r)$ teremos:

$$f(h_1,-1+h_2) - f(0,-1) = \|\mathbf{h}\|^2 \left(\underbrace{\frac{r^2}{r^2+r^2}}_{1/2} + \underbrace{\frac{r_2(\mathbf{h})}{r^2+r^2}}_{\to 0 \text{ qd. } r \to 0} \right)$$

Note-se que $\lim_{r\to 0} \frac{r_2(r,r)}{\|(r,r)\|^2} = 0$, por se tratar de um limite relativo de $\lim_{\mathbf{h}\to 0} \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$. Por conseguinte, quando $\mathbf{h}$ é da forma $\mathbf{h} = (r,r)$ a diferença $f((0,-1)+\mathbf{h}) - f(0,-1)$ é positiva (para $\mathbf{h}$ suficientemente pequeno). De modo semelhante se vê que aquela diferença é negativa quando $\mathbf{h}$ é da forma $\mathbf{h} = (r,-r)$. Assim, a função não tem extremo na origem.

Exemplo 3

Função: $f(x,y) = x^2 + 2y^2 + xy$

Pontos de estacionaridade?

$$\frac{\partial f}{\partial x} = 2x + y \qquad \frac{\partial f}{\partial y} = 4y + x$$
$$\nabla f = \mathbf{0} \Leftrightarrow \begin{cases} 2x + y = 0 \\ 4y + x = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases} \Leftrightarrow (x,y) = (0,0)$$

Derivadas de $2^{\underline{a}}$ ordem:

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2 \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = 4 \qquad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1$$

Fórmula de Taylor:

$$\begin{aligned} f(h_1, h_2) - f(0,0) &= \frac{1}{2!}(\mathbf{h}\cdot\nabla)^2 f + r_2(\mathbf{h}) \\ &= \frac{1}{2!}\|\mathbf{h}\|^2\left(\frac{1}{\|\mathbf{h}\|^2}(\mathbf{h}\cdot\nabla)^2 f + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}\right) \qquad (20) \\ &= \frac{1}{2!}\|\mathbf{h}\|^2\left(\frac{2h_1^2 + 2h_1h_2 + 4h_2^2}{h_1^2 + h_2^2} + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}\right) \qquad (21) \end{aligned}$$

Tal como nos exemplos anteriores, a questão que agora se põe é a de saber o sinal de (21). Supondo que $\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$ não tem influência (tal como acontecia nos exemplos anteriores), há que estudar o sinal de $2h_1^2 + 2h_1h_2 + 4h_2^2$:

$$\begin{aligned} 2h_1^2 + 2h_1h_2 + 4h_2^2 &= 2(h_1^2 + h_1h_2 + 2h_2^2) = 2\big((h_1^2 + h_1h_2 + \tfrac{1}{4}h_2^2) + \tfrac{7}{4}h_2^2\big) \\ &= 2\big((h_1 + \tfrac{1}{2}h_2)^2 + \tfrac{7}{4}h_2^2\big) \geq 0 \end{aligned}$$

Como se vê, o valor de $(2h_1^2 + 2h_1h_1 + 4h_2^2)/(h_1^2 + h_2^2)$ é sempre não-negativo, e só é nulo quando $(h_1, h_2) = (0,0)$. Isto parece indicar que $(0,0)$ seja ponto de mínimo local; no entanto, há que saber se é possível desprezar o termo $r_2(\mathbf{h})/\|\mathbf{h}\|^2$. O ponto essencial desta questão é que o conjunto dos valores de $(2h_1^2 + 2h_1h_1 + 4h_2^2)/(h_1^2 + h_2^2)$ (para $\mathbf{h} \neq \mathbf{0}$) tem um valor mínimo M. Como M é necessariamente positivo, o termo $M + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$ é positivo para $\mathbf{h}$ numa vizinhança de zero apropriada; ora, se $\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$ não afecta o sinal de $M + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$, também não afecta o sinal de $\frac{2h_1^2+2h_1h_1+4h_2^2}{h_1^2+h_2^2} + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$, uma vez que se $\left|\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}\right| < M$ também $\left|\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}\right| < \frac{2h_1^2+2h_1h_1+4h_2^2}{h_1^2+h_2^2}$, pois $M \leq \frac{2h_1^2+2h_1h_1+4h_2^2}{h_1^2+h_2^2}$.

Quanto à existência do mínimo M: Façamos $\mathbf{v} = \mathbf{h}/\|\mathbf{h}\|$, para $\mathbf{h} \neq 0$. Então:

$$\frac{2h_1^2 + 2h_1h_2 + 4h_2^2}{h_1^2 + h_2^2} = 2v_1^2 + 2v_1v_2 + 4v_2^2$$

Quando $\mathbf{h}$ passa por todos os valores possíveis (excepto $\mathbf{0}$), o vector $\mathbf{v}$ correspondente passa por todos os pontos da circunferência de raio 1 e centro em $\mathbf{0}$. Ora, a função $v_1 \mapsto 2v_1^2 + 2v_1v_2 + 4v_2^2$ é contínua; pelo **Teorema de Weierstrass**, a função tem mínimo quando (v_1, v_2) varia num conjunto (não vazio) limitado e fechado (como é o caso da circunferência $\|\mathbf{v}\| = 1$).

Que outros tipos de casos serão possíveis, além dos vistos nos exemplos anteriores? Se o valor de $(\mathbf{h}\cdot\nabla)f$ (no ponto de estacionaridade subentendido) for sempre negativo (nulo apenas quando $\mathbf{h} = \mathbf{0}$), é claro que f terá máximo local no ponto em questão (como exemplo, basta considerar a função do Exemplo 3 com o sinal trocado).

Assim, para tentar classificar um dado ponto de estacionaridade, estudamos o termo de $2^{\underline{a}}$ ordem da fórmula de Taylor. Sabemos o que concluir nos casos em que:

- $(\mathbf{h}\cdot\nabla)^2 f > 0$ para todo $\mathbf{h} \neq \mathbf{0}$

- $(\mathbf{h}\cdot\nabla)^2 f < 0$ para todo $\mathbf{h} \neq \mathbf{0}$
- $(\mathbf{h}\cdot\nabla)^2 f > 0$ para algum valor de $\mathbf{h}$ e $(\mathbf{h}\cdot\nabla)^2 f < 0$ para algum outro valor de $\mathbf{h}$

Pode acontecer que:

- $(\mathbf{h}\cdot\nabla)^2 f = 0$ para todo $\mathbf{h}$

Neste caso, haverá que estudar a fórmula de $3^{\underline{a}}$ ordem, tal como acontecia com as funções de uma só variável. Porém, podemos encontrar duas outras situações:

- $(\mathbf{h}\cdot\nabla)^2 f \geq 0$ para todo $\mathbf{h}$, com $(\mathbf{h}\cdot\nabla)^2 f = 0$ para algum $\mathbf{h} \neq \mathbf{0}$
- $(\mathbf{h}\cdot\nabla)^2 f \leq 0$ para todo $\mathbf{h}$, com $(\mathbf{h}\cdot\nabla)^2 f = 0$ para algum $\mathbf{h} \neq \mathbf{0}$

Note que o caso "$(\mathbf{h}\cdot\nabla)^2 f = 0$ para todo $\mathbf{h}$" é um caso particular dos dois últimos.

Exemplo 4

Função: $f(x,y) = x^2 \operatorname{sen} x + y^2 \operatorname{sh} y$

Pontos de estacionaridade?

$$\nabla f = (x^2\cos x + 2x\operatorname{sen} x, y^2 \operatorname{ch} y + 2y \operatorname{sh} y)$$

Para $(x,y) = (0,0)$ vem $\nabla f = (0,0)$. (Há uma infinidade de pontos de estacionaridade, mas vamos estudar apenas o ponto $(0,0)$.)

Termo de $2^{\underline{a}}$ ordem:

$$\frac{\partial^2 f}{\partial x^2} = 4x\cos x - x^2 \operatorname{sen} x + 2\operatorname{sen} x \qquad \frac{\partial^2 f}{\partial y^2} = 4y\operatorname{ch} y + y^2\operatorname{sh} y + 2\operatorname{sh} y \qquad \frac{\partial^2 f}{\partial x \partial y} = 0$$

No ponto $(0,0)$ vem $(\mathbf{h}\cdot\nabla)^2 f = 0$ para todo $\mathbf{h}$.

Termo de $3^{\underline{a}}$ ordem:

$$\begin{aligned}(\mathbf{h}\cdot\nabla)^3 f &= (\mathbf{h}\cdot\nabla)\big((4x\cos x - x^2\operatorname{sen} x + 2\operatorname{sen} x)h_1^2 + (4y\operatorname{ch} y + y^2\operatorname{sh} y + 2\operatorname{sh} y)h_2^2 \\ &\qquad + 2\cdot 0 \cdot h_1 h_2\big) \\ &= \big(h_1 \tfrac{\partial}{\partial x} + h_2 \tfrac{\partial}{\partial y}\big)(\dots) \\ &= h_1^3(-6x\operatorname{sen} x + 6\cos x - x^2\cos x) + h_2^3(6\operatorname{ch} y + 6y\operatorname{sh} y + y^2\operatorname{ch} y)\end{aligned}$$

No ponto $(0,0)$:

$$\big((\mathbf{h}\cdot\nabla)^3 f\big)(0,0) = 6h_1^3 + 6h_2^3$$

Fórmula de Taylor:

$$\begin{aligned}f(h_1,h_2) - f(0,0) &= \frac{1}{3!}(6h_1^3 + 6h_2^3) + r_3(\mathbf{h}) \\ &= \|\mathbf{h}\|^3\left(\frac{6}{3!}\frac{h_1^3 + h_2^3}{\|\mathbf{h}\|^3} + \frac{r_3(\mathbf{h})}{\|\mathbf{h}\|^3}\right)\end{aligned}$$

Estamos numa situação semelhante à do Exemplo 2. Para $(h_1, h_2) = (r, r)$ (com $r > 0$), vem $h_1^3 + h_2^3 = 2r^3 > 0$. Isto é, quando o vector $\mathbf{h}$ tem coordenadas iguais (por exemplo) e positivas, o valor de $\frac{6}{3!}\frac{h_1^3+h_2^3}{\|\mathbf{h}\|^3}$ é positivo (independentemente de r ser ou não pequeno); para r suficientemente pequeno, o valor de $\frac{r_3(r,r)}{\|\mathbf{h}\|^3}$ não afecta o sinal de $\frac{6}{3!}\frac{h_1^3+h_2^3}{\|\mathbf{h}\|^3}$. Logo, a função aumenta quando se passa de $(0, 0)$ para um ponto (r, r) suficientemente próximo (desde que $r > 0$).

Analogamente, considerando $(h_1, h_2) = (r, r)$ com $r < 0$ vemos que a função diminui. Por conseguinte, a função não tem extremo no ponto $(0, 0)$.

A conclusão deste exemplo estende-se a qualquer função para a qual o primeiro termo da fórmula de Taylor que não se anule para todo $\mathbf{h}$ seja o de 3ª ordem. De facto, em vez de $h_1^3 + h_2^3$ teremos uma expressão da forma[3]

$$\sum_{\substack{i,j=0 \\ (j+k=3)}}^{3} a_{ij} h_1^j h_2^k \ ,$$

onde pelo menos algum dos coeficientes a_{ij} não é nulo. Escolha-se (h_1, h_2) de modo que o somatório não seja nulo. Multiplicando $\mathbf{h}$ por r, o somatório vem multiplicado por r^3. Logo, escolhendo $r > 0$ e $r < 0$ obtemos valores de $(r\mathbf{h} \cdot \nabla)^3 f$ de sinais diferentes.

Mais geralmente, a conclusão mantém-se em situações deste tipo, quando o primeiro termo da fórmula de Taylor que não se anule para todo $\mathbf{h}$ seja de ordem ímpar (não necessariamente de 3ª ordem).

Exemplo 5

Função: $f(x, y) = x^4 + y^4$

O único ponto de estacionaridade é a origem. Neste caso, o primeiro termo da fórmula de Taylor que não é identicamente nulo é o de 4ª ordem:

$$f(h_1, h_2) - f(0, 0) = \frac{1}{4!}(24h_1^4 + 24h_2^4) + r_4(h_1, h_2) \tag{22}$$

Como se vê, existe um mínimo local na origem.

Neste exemplo, acontece que o resto de 4ª ordem é idênticamente nulo. Por outras palavras, estudar o sinal do termo de 4ª ordem é equivalente em dificuldade (ou facilidade) a estudar directamente o sinal do 1º membro de (22). Isto deve-se ao facto de a função ser ela própria um polinómio, o que, obviamente, não é o caso geral.

[3]Se o espaço em questão for $\mathbb{R}^n$, obtemos uma expressão da forma $\sum_{i,j,k=1}^{n} h_i h_j h_k a_{ijk}$, e o resto do argumento é semelhante.

Exemplo 6

Função:

$$\begin{aligned} f(x,y) &= (x-y)^2 - x^4 - y^4 \\ \nabla f &= \big(2(x-y) - 4x^3, -2(x-y) - 4y^3\big) \\ \nabla f = \mathbf{0} &\Leftrightarrow (x,y) = (0,0) \vee (x,y) = (1,-1) \vee (x,y) = (-1,1) \end{aligned}$$

Termo de $2^{\underline{a}}$ ordem, no ponto (x,y):

$$(\mathbf{h}\cdot\nabla)^2 f = (2-12x^2)h_1^2 + 2(-2)h_1h_2 + h_2^2(2-12y^2)$$

Para $(x,y) = (1,-1)$ ou $(x,y) = (-1,1)$, obtemos:

$$(\mathbf{h}\cdot\nabla)^2 f = -10h_1^2 - 4h_1h_2 - 10h_2^2$$

Fórmula de Taylor, para $(x,y) = (1,-1)$:

$$\begin{aligned} f(1+h_1,-1+h_2) - f(1,-1) &= \frac{1}{2!}(-10h_1^2 - 4h_1h_2 - 10h_2^2) + r_2(h_1,h_2) \\ -10h_1^2 - 4h_1h_2 - 10h_2^2 &= -10\big(h_1^2 + \tfrac{2}{5}h_1h_2 + h_2^2\big) \\ &= -10\big(h_1^2 + \tfrac{2}{5}h_1h_2 + \tfrac{1}{25}h_2^2 - \tfrac{1}{25}h_2^2 + h_2^2\big) \\ &= -10\Big(\big(h_1 + \tfrac{1}{5}h_2\big)^2 + \tfrac{24}{25}h_2^2\Big) \end{aligned}$$

Conclui-se que $(1,-1)$ é ponto de máximo local.

A conclusão é a mesma para o ponto $(-1,1)$.

Fórmula de Taylor, para o ponto $(0,0)$:

$$\begin{aligned} f(h_1,h_2) - f(0,0) &= \frac{1}{2!}(2h_1^2 - 4h_1h_2 + 2h_2^2) + r_2(h_1,h_2) \\ &= \|\mathbf{h}\|^2\left(\frac{(h_1-h_2)^2}{\|\mathbf{h}\|^2} + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}\right) \end{aligned}$$

O valor de $(h_1-h_2)^2$ é sempre não-negativo, mas pode ser nulo sem que $(h_1,h_2) = (0,0)$ (por exemplo, quando $(h_1,h_2) = (1,1)$).

Ao contrário do que poderia parecer à primeira vista, *não se pode concluir que exista um mínimo* local na origem. Comparando com o que acontecia no Exemplo 2, e fazendo $\mathbf{v} = \mathbf{h}/\|\mathbf{h}\|$, vemos que o mínimo de $(v_1-v_2)^2$ na circunferência $\|\mathbf{v}\| = 1$ é zero, pelo que não podemos desprezar $\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$.

No entanto, para uma dada direcção de $\mathbf{h}$, diferente da direcção de $(1,1)$, o valor de $(h_1-h_2)^2/\|\mathbf{h}\|^2$ será positivo, e a parcela $\frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2}$ não afectará o sinal daquele valor desde que $\|\mathbf{h}\|$ seja suficientemente pequena. Por exemplo, para $\mathbf{h} = r(1,2)$ virá $f(\mathbf{h}) > f(\mathbf{0})$ para todo $r > 0$ suficientemente pequeno. Se encontrarmos outra direcção de $\mathbf{h}$ para a qual $f(\mathbf{h}) < f(\mathbf{0})$ desde que $\mathbf{h}$ seja suficientemente pequeno, poderemos concluir que não temos um ponto de extremo. Ora, a única direcção para a qual há possibilidade de isto acontecer é a direcção do vector $\mathbf{h} = (1,1)$, pois para qualquer outra a situação seria idêntica à do caso $\mathbf{h} = (1,2)$.

Assim, há que recorrer a uma fórmula de Taylor de ordem superior, a qual só nos interessará para vectores $\mathbf{h}$ colineares com $(1,1)$:

$$(\mathbf{h}\cdot\nabla)^3 f = (-24x)h_1^3 + (-24y)h_2^3$$

Para $(x,y) = (0,0)$, vem $(\mathbf{h}\cdot\nabla)^3 f = 0$.

$$(\mathbf{h}\cdot\nabla)^4 f = -24h_1^4 - 24h_2^4$$

Fórmula de Taylor de 4ª ordem:

$$f(h_1,h_2) - f(0,0) = \frac{1}{2!}(2h_1^2 - 4h_1h_2 + 2h_2^2) + \frac{1}{4!}(-24h_1^4 - 24h_2^4) + r_4(h_1,h_2)$$

Quando $h_1 = h_2$, vem:

$$\begin{aligned} f(h_1,h_2) - f(0,0) &= \frac{1}{4!}(-24h_1^4 - 24h_2^4) + r_4(h_1,h_2) \\ &= \|\mathbf{h}\|^4 \left(\frac{1}{4!}\frac{-24(h_1^4 + h_2^4)}{\|\mathbf{h}\|^4} + \frac{r_4(\mathbf{h})}{\|\mathbf{h}\|^4} \right) \\ &= \|\mathbf{h}\|^4 \left(\frac{-48}{4!}\underbrace{\frac{h_1^4}{|h_1|^4(\sqrt{2})^4}}_{1/4} + \frac{r_4(\mathbf{h})}{\|\mathbf{h}\|^4} \right) \end{aligned}$$

É claro que a parcela $r_4(\mathbf{h})/\|\mathbf{h}\|^4$ não afecta o sinal da soma (desde que $\|\mathbf{h}\|$ seja suficientemente pequena).

Concluindo: Para $\mathbf{h} = r(1,2)$ e todo $r < \varepsilon_1$ (com $\varepsilon_1 > 0$ apropriado), temos $f(\mathbf{h}) - f(\mathbf{0}) > 0$. Para $\mathbf{h} = r(1,1)$ e todo $r < \varepsilon_2$ (com $\varepsilon_2 > 0$ apropriado) temos $f(\mathbf{h}) - f(\mathbf{0}) < 0$. Logo, $\mathbf{0}$ não é ponto de extremo: se fosse ponto de máximo (por exemplo), existiria $\varepsilon > 0$ tal que $f(\mathbf{h}) - f(\mathbf{0}) < 0$ sempre que $\|\mathbf{h}\| < \varepsilon$; mas para $\mathbf{h} = (r,r)$ e r apropriado (mais precisamente, $r < \min\{\varepsilon_2, \varepsilon/\sqrt{2}\}$) viria $f(\mathbf{h}) - f(\mathbf{0}) < 0$ (porque $r < \varepsilon_2$) e $f(\mathbf{h}) - f(\mathbf{0}) > 0$ (porque $\|(r,r)\| < \varepsilon$), o que é uma contradição.

Exemplo 7

Função: $f(x,y) = y^2 - 4x^2y + 3x^4$

Ponto de estacionaridade: $(0,0)$

Termo de 2ª ordem: $(\mathbf{h}\cdot\nabla)^2 f = 2h_2^2$

Fórmula de Taylor:

$$\begin{aligned} f(h_1,h_2) - f(0,0) &= \frac{1}{2!}2h_2^2 + r_2(h_1,h_2) \\ &= \|\mathbf{h}\|^2 \left(\frac{h_2^2}{\|\mathbf{h}\|^2} + \frac{r_2(\mathbf{h})}{\|\mathbf{h}\|^2} \right) \end{aligned}$$

O valor de h_2^2 é sempre não-negativo, mas pode ser nulo sem que $\mathbf{h}$ seja nulo: caso de $h_2 = 0$ ($\mathbf{h} = (1,0)$, por exemplo).

A situação é semelhante à do Exemplo 6 (caso do ponto de estacionaridade $(0,0)$). Escolhida uma direcção que não seja a do vector $(1,0)$, a diferença $f(\mathbf{h}) - f(\mathbf{0})$ é positiva desde que $\|\mathbf{h}\|$ suficientemente pequena. Será que ao longo da direcção de $(1,0)$ aquela diferença é negativa?

Fórmula de Taylor para $\mathbf{h}$ colinear com $(1,0)$:

$$\begin{aligned}\underbrace{f(h_1,h_2) - f(0,0)}_{f(h_1,0)} &= \frac{1}{4!}72h_1^4 + r_4(h_1,0)\\ &= \|\mathbf{h}\|^4\left(\frac{1}{4!}\frac{72h_1^4}{\|\mathbf{h}\|^4} + \frac{r_4(h_1,0)}{\|\mathbf{h}\|^4}\right) \qquad (23)\end{aligned}$$

O valor de $72h_1^4$ é positivo (para ser nulo, teria de ser $h_1 = h_2 = 0$, e este caso está excluído). Para $\|\mathbf{h}\|$—ou seja, $|h_1|$—suficientemente pequeno, a parcela $r_4(\mathbf{h})/\|\mathbf{h}\|^4$ não afecta o sinal de (23). Logo, $f(h_1,0) - f(0,0) > 0$.

> Poderemos concluir que a função tem um mínimo local na origem? *Tal conclusão estaria errada*!

O que se provou é que para cada direcção do plano a diferença $f(\mathbf{h}) - f(\mathbf{0})$ é positiva desde que $\mathbf{h}$ tenha essa direcção e tenha norma suficientemente pequena. Mas o grau de "pequenês" a exigir de $\|\mathbf{h}\|$ (isto é, o valor de ε tal que $\|\mathbf{h}\| < \varepsilon$ implique $f(\mathbf{h}) > f(\mathbf{0})$) pode depender da direcção; como o número de direcções distintas é infinito, pode ser impossível obter um valor de ε que seja apropriado para todas as direcções (se estivéssemos interessados em apenas duas direcções, bastaria escolher o menor dos dois valores de ε correspondentes).

Em situações deste tipo, o recurso à fórmula de Taylor não permite classificar o ponto de estacionaridade.

No caso presente, é possível resolver o problema por inspecção:

$$\begin{aligned}f(x,y) - f(0,0) &= y^2 - 4x^2y + 3x^4\\ &= y^2 - 4x^2y + 4x^4 - 4x^4 + 3x^4\\ &= (y-2x^2)^2 - x^4\\ &= (y-2x^2-x^2)(y-2x^2+x^2)\\ &= (y-3x^2)(y-x^2)\end{aligned}$$

Para pontos (x,y) situados entre as parábolas de equações $y = 3x^2$ e $y = x^2$, vem $y - 3x^2 < 0$ e $y - x^2 > 0$ (isto é, $f(x,y) - f(0,0) < 0$). Para (x,y) acima da parábola de equação $y = 3x^2$, virá $f(x,y) > f(0,0)$. Logo, em qualquer vizinhança da origem existem pontos em que a diferença $f(x,y) - f(0,0)$ é positiva e pontos em que a diferença é negativa; por conseguinte, a função não tem extremo na origem.