Aproximación por Polinomios

(Polinomio de Taylor en $\mathbb{R}^n$)

María del Carmen Calvo

En este trabajo nos ocuparemos de extender a funciones de varias variables los resultados obtenidos para funciones de una variable.

I. *Introducción*

Antes de ocuparnos del tema específico necesitamos tratar algunos asuntos esencialmente algebraicos que facilitarán luego la comprensión de los conceptos principales.

A partir de una función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ y de un punto $\mathbf{a} \in \mathbb{R}^n$ vamos a encontrar la forma en que se puede escribir un polinomio de grado 1 y otro de grado 2 que cumplen la condición

$$P(\mathbf{a}) = f(\mathbf{a})$$

Nos interesan especialmente porque el polinomio de Taylor, al tener que aproximar a la función cerca de **a**, deberá comenzar por coincidir con ella en ese punto.

❒ Forma general de un polinomio de grado 1 en $\mathbb{R}^n$

Sea $P : \mathbb{R}^n \longrightarrow \mathbb{R}$ un polinomio de grado menor o igual que 1 tal que $P(\mathbf{a}) = f(\mathbf{a})$. La condición sobre su grado obliga a que sólo contenga monomios de grado 1 o cero. Por lo tanto,

$$P(\mathbf{x}) = b_1x_1 + \cdots + b_nx_n + c$$

donde $b_1, \ldots, b_n, c \in \mathbb{R}$ están fijos. Llamando

$$\mathbf{b} = (b_1, \ldots, b_n)$$

se puede escribir en la forma

$$P(\mathbf{x}) = \mathbf{b} \cdot \mathbf{x} + c$$

La condición $P(\mathbf{a}) = f(\mathbf{a})$ dice que

$$\mathbf{b} \cdot \mathbf{a} + c = f(\mathbf{a})$$

O sea,

$$c = f(\mathbf{a}) - \mathbf{b} \cdot \mathbf{a}$$

Llegamos finalmente a nuestro objetivo

$$\boxed{P(\mathbf{x}) = \mathbf{b} \cdot (\mathbf{x} - \mathbf{a}) + f(\mathbf{a})} \tag{1.1}$$

donde $\mathbf{b} \in \mathbb{R}^n$.

Ejemplo

Consideremos la función $f(x, y, z) = 3\cos(x+y) - xyz^2$, el punto $\mathbf{a} = (0, \pi, 2)$ y P el polinomio de grado 1

$$P(x, y, z) = 3x - 2y + 4z + 2\pi - 11 \qquad \text{que satisface} \qquad P(0, \pi, 2) = f(0, \pi, 2) = -3$$

Es inmediato ver que se puede escribir en la forma

$$P(x, y, z) = (3, -2, 4) \cdot (x, y, z) + 2\pi - 11$$

Siguiendo el procedimiento del desarrollo anterior, reemplazamos (x, y, z) por

$$(x, y - \pi, z - 2) + (0, \pi, 2)$$

y obtenemos

$$\begin{aligned} P(x, y, z) &= (3, -2, 4) \cdot [(x, y - \pi, z - 2) + (0, \pi, 2)] + 2\pi - 11 \\ &= (3, -2, 4) \cdot (x, y - \pi, z - 2) + (3, -2, 4) \cdot (0, \pi, 2) + 2\pi - 11 \\ &= (3, -2, 4) \cdot (x, y - \pi, z - 2) - 2\pi + 8 + 2\pi - 11 \\ &= (3, -2, 4) \cdot (x, y - \pi, z - 2) - 3 \end{aligned}$$

Lo hemos escrito entonces en la forma (1.1) con

$$\mathbf{b} = (3, -2, 4)$$

❒ FORMA GENERAL DE UN POLINOMIO DE GRADO 2 EN $\mathbb{R}^n$

Sea $P : \mathbb{R}^n \longrightarrow \mathbb{R}$ un polinomio de grado menor o igual que 2 tal que $P(\mathbf{a}) = f(\mathbf{a})$. La condición sobre su grado obliga a que sólo contenga monomios de grado 2, 1 o cero. Por lo tanto,

$$P(\mathbf{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j + d_1x_1 + \cdots + d_nx_n + c$$

donde a_{ij} , d_k y c son números reales fijos y la matriz (a_{ij}) simétrica.

Recordando que los polinomios de grado 2 —con todos sus monomios de grado 2— representan formas cuadráticas, podemos escribir a este polinomio en la forma

$$P(\mathbf{x}) = \mathbf{x}\, A\, \mathbf{x}^t \ + \ \mathbf{d} \cdot \mathbf{x} \ + \ c$$

siendo $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ simétrica , $\mathbf{d} \in \mathbb{R}^n$ y $c \in \mathbb{R}$.

Nos proponemos aquí hacer algo similar al caso anterior; i.e., expresarlo en términos de las diferencias $\mathbf{x} - \mathbf{a}$. Con ese propósito, en la ecuación anterior escribimos

$$\mathbf{x} = \mathbf{x} - \mathbf{a} + \mathbf{a}$$

con lo cual, recordando las propiedades del producto de matrices y del producto escalar [1], obtenemos

$$\begin{aligned} P(\mathbf{x}) &= (\mathbf{x} - \mathbf{a} + \mathbf{a})\, A\, (\mathbf{x} - \mathbf{a} + \mathbf{a})^t \ + \ \mathbf{d} \cdot (\mathbf{x} - \mathbf{a} + \mathbf{a}) \ + \ c \\ &= (\mathbf{x} - \mathbf{a})\, A\, (\mathbf{x} - \mathbf{a})^t + \mathbf{a}\, A\, (\mathbf{x} - \mathbf{a})^t + (\mathbf{x} - \mathbf{a})\, A\, \mathbf{a}^t + \mathbf{a}\, A\, \mathbf{a}^t \ + \ \mathbf{d} \cdot (\mathbf{x} - \mathbf{a}) + \mathbf{d} \cdot \mathbf{a} \ + \ c \\ &= (\mathbf{x} - \mathbf{a})\, A\, (\mathbf{x} - \mathbf{a})^t \ + \ 2(\mathbf{a}A) \cdot (\mathbf{x} - \mathbf{a}) + \mathbf{d} \cdot (\mathbf{x} - \mathbf{a}) + \mathbf{a}\, A\, \mathbf{a}^t + \mathbf{d} \cdot \mathbf{a} + c \\ &= (\mathbf{x} - \mathbf{a})\, A\, (\mathbf{x} - \mathbf{a})^t \ + \ [2(\mathbf{a}A) + \mathbf{d}] \cdot (\mathbf{x} - \mathbf{a}) + [\mathbf{a}\, A\, \mathbf{a}^t + \mathbf{d} \cdot \mathbf{a} + c] \end{aligned}$$

Llamando

$$\mathbf{b} = \mathbf{d} + 2\, \mathbf{a}A$$

y teniendo en cuenta que

$$f(\mathbf{a}) = P(\mathbf{a}) = \mathbf{a}\, A\, \mathbf{a}^t + \mathbf{d} \cdot \mathbf{a} + c$$

Llegamos finalmente a la forma en que queríamos escribir a P,

$$\boxed{P(\mathbf{x}) = (\mathbf{x} - \mathbf{a})\, A\, (\mathbf{x} - \mathbf{a})^t \ + \ \mathbf{b} \cdot (\mathbf{x} - \mathbf{a}) \ + \ f(\mathbf{a})} \tag{1.2}$$

donde $A \in \mathbb{R}^{n\times n}$ es simétrica y $\mathbf{b} \in \mathbb{R}^n$.

Ejemplo

Sean $f(x, y, z) = xy^2 + 2xz - e^{yz}$, $\mathbf{a} = (1, 0, -2)$ y $P(x, y, z) = 4x^2 - 3y^2 + z^2 + xy - yz + 3x - y + z - 14$ que satisface

$$P(1, 0, -2) = -5 = f(1, 0, -2)$$

[1] y además que: $\mathbf{y}\, A\, \mathbf{y}^t = (\mathbf{y}\, A\, \mathbf{y}^t)^t = ((\mathbf{y})^t)^t(\mathbf{y}A)^t = \mathbf{y}(\mathbf{y}A)^t = \mathbf{y} \cdot (A\mathbf{y})$

Este polinomio se puede escribir en la forma,

$$P(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + (3,-1,1)\cdot(x,y,z) - 14$$

Reemplazamos $(x,y,z) = (x-1,y,z+2) + (1,0,-2)$

$$\begin{aligned}
P(x,y,z) &= \left[\begin{pmatrix} x-1 & y & z+2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & -2 \end{pmatrix}\right] \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \left[\begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}\right] \\
&\quad + (3,-1,1)\cdot[(x-1,y,z+2)+(1,0,-2)] - 14 \\
&= \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} - \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \\
&\quad - \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \\
&\quad + (3,-1,1)\cdot(x-1,y,z+2) + (3,-1,1)\cdot(1,0,-2) - 14 \\
&= \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} - \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 \\ -\frac{1}{2} \\ -2 \end{pmatrix} \\
&\quad - \begin{pmatrix} 4 & -\frac{1}{2} & -2 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} + \begin{pmatrix} 4 & -\frac{1}{2} & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \\
&\quad + (3,-1,1)\cdot(x-1,y,z+2) + 1 - 14 \\
&= \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} - 2(4,-\frac{1}{2},-2)\cdot(x-1,y,z+2) + 8 \\
&\quad + (3,-1,1)\cdot(x-1,y,z+2) - 13 \\
&= \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} \\
&\quad + [-2(4,-\frac{1}{2},-2) + (3,-1,1)]\cdot(x-1,y,z+2) - 5 \\
&= \begin{pmatrix} x-1 & y & z+2 \end{pmatrix} \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z+2 \end{pmatrix} + (-1,-\frac{1}{2},5)\cdot(x-1,y,z+2) - 5
\end{aligned}$$

Esto responde a la forma (1.2) con

$$A = \begin{pmatrix} 4 & \frac{1}{2} & 0 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \qquad \text{y} \qquad \mathbf{b} = (-1,-\frac{1}{2},5)$$

II. *Polinonio de Taylor de orden* 1 *en* $\mathbb{R}^n$

Vamos a probar que para cada función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ diferenciable en el punto $\mathbf{a}$ existe un único polinomio $P_1 : \mathbb{R}^n \longrightarrow \mathbb{R}$ que satisface

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - P_1(\mathbf{x})}{\|\mathbf{x}-\mathbf{a}\|} = 0 \tag{2.1}$$

Comencemos por notar que un polinomio que cumple esta condición, en particular satisface

$$P_1(\mathbf{a}) = f(\mathbf{a})$$

dado que, siendo que el cociente en (2.1) tiende a cero y el demoninador también, no queda otra posibilidad que también lo haga el numerador; con lo cual, por ser ambas funciones continuas implica que

$$f(\mathbf{a}) - P_1(\mathbf{a}) = \lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) - P_1(\mathbf{x}) = 0$$

Lo que diferencia a P_1 del resto de los polinomios que coinciden con f justo en $\mathbf{a}$ es que su gráfico es el que está más cerca del gráfico de f para los $\mathbf{x}$ suficiente cercanos a $\mathbf{a}$.

Recordando la forma que tienen todos los polinomios que coinciden con f en $\mathbf{a}$ [2] podemos escribir

$$P_1(\mathbf{x}) = \mathbf{b}\cdot(\mathbf{x}-\mathbf{a}) + f(\mathbf{a})$$

Se trata entonces de ver quién es ese vector $\mathbf{b}$. Escribamos la condición (2.1) usando esta expresión de P_1,

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \mathbf{b}\cdot(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = 0$$

Pero esto, por ser f diferenciable, sólo se cumple cuando

$$\mathbf{b} = \nabla f(\mathbf{a})$$

Concluimos de esta forma que

$$P_1(\mathbf{x}) = \nabla f(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a}) + f(\mathbf{a})$$

es el único polinomio de grado menor o igual que 1 *que cumple* (2.1).

Además, cuando $n = 2$*, su gráfico es el plano tangente al gráfico de* f *en* $(\mathbf{a}, f(\mathbf{a}))$.

III. *Expresión del término complementario de orden* 1 *en* $\mathbb{R}^n$

El término complementario de Taylor de orden 1 alrededor del punto $\mathbf{a}$ se define por

$$R_1(\mathbf{x}) = f(\mathbf{x}) - P_1(\mathbf{x})$$

[2] cf. la ecuación (1.1)

es decir, indica la diferencia entre la función y el polinomio de Taylor de orden 1 y nos da por lo tanto el *error* que cometemos al aproximar a f por P_1 en cada punto $\mathbf{x}$. Teniendo en cuenta que

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{R_1(\mathbf{x})}{\mathbf{x}-\mathbf{a}} = \lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x})-P_1(\mathbf{x})}{\mathbf{x}-\mathbf{a}} = 0$$

podemos lograr que el error que cometemos en esa aproximación sea tan chico como se quiera con tal de tomar a $\mathbf{x}$ suficientemente cerca de $\mathbf{a}$.

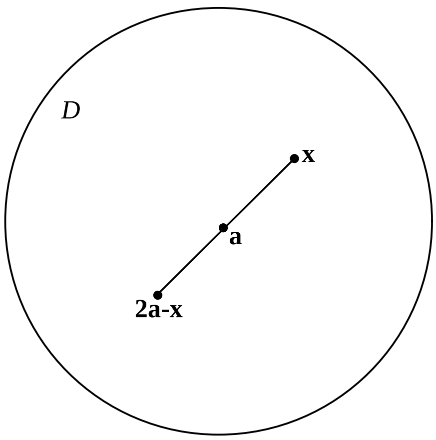
Nos proponemos ahora encontrar una forma de escribir a este resto que recuerde lo sabido para funciones de una variable.

Como pretendemos utilizar la forma Hessiana de f, vamos a pedir que esta función sea de clase C^2 en un entorno de $\mathbf{a}$.

Para facilitar los cálculos definiremos –a partir de f– una función de una variable para la cual ya sabemos escribir su fórmula de Taylor de orden 1. Más precisamente, supongamos que f es de clase C^2 en el disco

$$D: \quad \|\mathbf{x}-\mathbf{a}\| < r$$

sabemos que el segmento $[\mathbf{a},\mathbf{x}] \subset D$ para todo $\mathbf{x} \in D$. Pero D también incluye al segmento $[2\mathbf{a}-\mathbf{x},\mathbf{a}]$ [3]; luego, el segmento $[2\mathbf{a}-\mathbf{x},\mathbf{x}] \subset D$. Esto se aprecia en el siguiente esquema gráfico,



Comenzamos parametrizando este segmento,

$$\alpha : [-1,1] \longrightarrow D \qquad , \qquad \alpha(t) = \mathbf{a} + t(\mathbf{x}-\mathbf{a})$$

cuyo vector tangente en todo instante es

$$\mathbf{x}-\mathbf{a}$$

y a continuación definimos la función

$$\varphi : [-1,1] \longrightarrow \mathbb{R} \qquad , \qquad \varphi(t) = f(\alpha(t))$$

[3] simétrico de $[\mathbf{a},\mathbf{x}]$ respecto de $\mathbf{a}$ sobre la recta que pasa por $\mathbf{a}$ y $\mathbf{x}$

que resulta de clase C^2 y además satisface

$$\varphi(0) = f(\mathbf{a}) \quad , \quad \varphi(1) = f(\mathbf{x}) \quad , \quad \varphi'(t) = \nabla f(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a})$$

La fórmula de Taylor de orden 1 de φ alrededor de $t = 0$ es

$$\varphi(t) = p_1(t) + r_1(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(t_0)}{2}t^2$$

para un cierto t_0 entre 0 y t. Escrito en términos de f, para $t = 1$, nos da

$$f(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + \frac{\varphi''(t_0)}{2} \qquad \text{(para un } 0 < t_0 < 1)$$

Necesitamos encontrar la expresión del resto en términos de f para lo cual deberemos derivar

$$\varphi'(t) = \nabla f(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a}) = \frac{\partial f}{\partial x_1}(\boldsymbol{\alpha}(t))(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(\boldsymbol{\alpha}(t))(x_n - a_n)$$

Aplicando la regla de la cadena en cada sumando obtenemos

$$\begin{aligned}
\varphi''(t) &= \nabla(\tfrac{\partial f}{\partial x_1})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a})\,(x_1 - a_1) \;+\; \ldots \;+\; \nabla(\tfrac{\partial f}{\partial x_n})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a})\,(x_n - a_n) \\
&= \Big(\nabla(\tfrac{\partial f}{\partial x_1})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a}),\ \ldots\ , \nabla(\tfrac{\partial f}{\partial x_n})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a})\Big) \cdot (\mathbf{x} - \mathbf{a}) \\
&= (\mathbf{x} - \mathbf{a}) \cdot \Big(\nabla(\tfrac{\partial f}{\partial x_1})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a}),\ \ldots\ , \nabla(\tfrac{\partial f}{\partial x_n})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a})\Big)
\end{aligned}$$

Recordando que cada fila de la matriz Hessiana es el gradiente de cada derivada parcial primera de f vemos que

$$\begin{pmatrix} \nabla(\frac{\partial f}{\partial x_1})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a}) \\ \vdots \\ \nabla(\frac{\partial f}{\partial x_n})(\boldsymbol{\alpha}(t)) \cdot (\mathbf{x} - \mathbf{a}) \end{pmatrix} = MHf(\boldsymbol{\alpha}(t))(\mathbf{x} - \mathbf{a})^t$$

de donde resulta que

$$\varphi''(t) = (\mathbf{x} - \mathbf{a})MHf(\boldsymbol{\alpha}(t))(\mathbf{x} - \mathbf{a})^t = 2Hf(\boldsymbol{\alpha}(t))(\mathbf{x} - \mathbf{a})$$

Tomando $t = t_0$ obtenemos

$$\frac{\varphi''(t_0)}{2} = Hf(\boldsymbol{\alpha}(t_0))(\mathbf{x} - \mathbf{a})$$

Si llamamos $\mathbf{c} = \boldsymbol{\alpha}(t_0)$, hemos probado que la fórmula de Taylor de orden 1 de f alrededor de $\mathbf{a}$ se escribe

$$f(\mathbf{x}) = P_1(\mathbf{x}) + Hf(\mathbf{c})(\mathbf{x} - \mathbf{a})$$

donde $\mathbf{c} = \boldsymbol{\alpha}(t_0) \in [\mathbf{a}, \mathbf{x}]$ [4].

[4] por ser $0 < t_0 < 1$

IV. *Polinonio de Taylor de orden* 2 *en* $\mathbb{R}^n$

Supongamos que la función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es de clase C^2 en el punto $\mathbf{a}$. Mostraremos que existe un único polinomio $P_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ de grado menor o igual que 2 tal que

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - P_2(\mathbf{x})}{||\mathbf{x}-\mathbf{a}||^2} = 0 \tag{4.1}$$

Comencemos por notar que esta condición implica que, en particular,

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - P_2(\mathbf{x})}{||\mathbf{x}-\mathbf{a}||} = 0 \tag{4.2}$$

Esto se deduce inmediatamente de

$$\frac{f(\mathbf{x}) - P_2(\mathbf{x})}{||\mathbf{x}-\mathbf{a}||} = \frac{f(\mathbf{x}) - P_2(\mathbf{x})}{||\mathbf{x}-\mathbf{a}||^2}\, ||\mathbf{x}-\mathbf{a}||$$

Como observamos en una sección anterior, la condición (4.1) obliga a que $P_2(\mathbf{a}) = f(\mathbf{a})$ y por lo tanto es posible escribirlo en la forma (1.2)

$$P_2(\mathbf{x}) = (\mathbf{x}-\mathbf{a})A(\mathbf{x}-\mathbf{a})^t + \mathbf{b}\cdot(\mathbf{x}-\mathbf{a}) + f(\mathbf{a})$$

Reescribimos la condición (4.2) usando esta expresión de P_2

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - \mathbf{b}\cdot(\mathbf{x}-\mathbf{a}) - f(\mathbf{a})}{||\mathbf{x}-\mathbf{a}||} + \frac{\mathbf{x}-\mathbf{a}}{||\mathbf{x}-\mathbf{a}||}A(\mathbf{x}-\mathbf{a})^t = 0$$

Ahora bien, la expresión $\dfrac{\mathbf{x}-\mathbf{a}}{||\mathbf{x}-\mathbf{a}||}$ está acotada dado que tiene norma 1 mientras que $A(\mathbf{x}-\mathbf{a})^t$ tiende a cero cuando $\mathbf{x}$ tiende a $\mathbf{a}$. Esto nos dice que el segundo sumando tiende a cero cuando $\mathbf{x}\to\mathbf{a}$ y en consecuencia —dado que la suma también converge a cero— lo propio debe suceder con el primer sumando; i.e.,

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - \mathbf{b}\cdot(\mathbf{x}-\mathbf{a}) - f(\mathbf{a})}{||\mathbf{x}-\mathbf{a}||} = 0$$

Pero entonces $\mathbf{b}\cdot(\mathbf{x}-\mathbf{a}) + f(\mathbf{a})$ tiene que ser el polinomio de Taylor de orden 1 de f en $\mathbf{a}$ por lo cual

$$\mathbf{b} = \nabla f(\mathbf{a})$$

Tenemos hasta ahora que el polinomio que buscamos debe tener la forma

$$P_2(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a}) + (\mathbf{x}-\mathbf{a})A(\mathbf{x}-\mathbf{a})^t$$

sólo nos resta ver quiénes son los elementos de la matriz A.

Para lograr este objetivo vamos a usar el hecho que la existencia del límite múltiple de la condición (4.1) asegura que se obtiene el mismo resultado cuando uno se aproxima a $\mathbf{a}$ por cualquier camino en particular.

Consideremos en primera instancia

$$\mathbf{c}(t) = (a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n)$$

que tiende a $\mathbf{a}$ cuando $t \to a_i$. Conviene notar además que

$$\mathbf{d}(t) = \mathbf{c}(t) - \mathbf{a} = (0, \ldots, 0, t - a_i, 0, \ldots, 0) \quad \text{y} \quad \|\mathbf{c}(t) - \mathbf{a}\| = \|\mathbf{d}(t)\| = |t - a_i|$$

y también que

$$\mathbf{d}(t)A(\mathbf{d}(t))^t = \mathbf{d}(t)\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & t - a_i & 0 & \cdots & 0 \end{pmatrix}\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} = (t - a_i)^2 a_{ii}$$

La condición (4.1) nos permite afirmar entonces que

$$\frac{f(\mathbf{c}(t)) - f(\mathbf{a}) - \nabla f(\mathbf{a})(\mathbf{d}(t)) - (t - a_i)^2 a_{ii}}{(t - a_i)^2} \longrightarrow 0$$

cuando $t \to a_i$. De aquí podemos concluir que

$$a_{ii} = \lim_{t \to a_i} \frac{f(\mathbf{c}(t)) - f(\mathbf{a}) - \nabla f(\mathbf{a})(\mathbf{d}(t))}{(t - a_i)^2}$$

es decir,

$$a_{ii} = \lim_{t \to a_i} \frac{f(a_1, \ldots, t, \ldots, a_n) - f(a_1, \ldots, a_n) - \frac{\partial f}{\partial x_i}(\mathbf{a})(t - a_i)}{(t - a_i)^2}$$

Siendo éste un límite de una variable en el que tanto numerador como denominador tienden a cero cuando $t \to a_i$ podemos aplicar l'Hospital para concluir que

$$\begin{aligned} a_{ii} &= \lim_{t \to a_i} \frac{\frac{\partial f}{\partial x_i}(a_1, \ldots, t, \ldots, a_n) - \frac{\partial f}{\partial x_i}(\mathbf{a})}{2(t - a_i)} \\ &= \frac{1}{2}\lim_{t \to a_i} \frac{\frac{\partial f}{\partial x_i}(a_1, \ldots, t, \ldots, a_n) - \frac{\partial f}{\partial x_i}(\mathbf{a})}{t - a_i} \\ &= \frac{1}{2}\frac{\partial^2 f}{\partial x_i^2}(\mathbf{a}) \end{aligned}$$

A esta altura podemos decir que la matriz A coincide con $\frac{1}{2}MHf(\mathbf{a})$, al menos, en su diagonal principal.

Para calcular el resto de los elementos de A y comprobar finalmente que

$$A = \frac{1}{2}MHf(\mathbf{a})$$

bastaría con reproducir lo hecho con los elementos de la diagonal pero esta vez usando el camino

$$\mathbf{g}(t) = (a_1, \ldots, a_{i-1},\ a_i + t\ , a_{i+1}, \ldots, a_{j-1},\ a_j + t\ , a_{j+1}, \ldots, a_n) \longrightarrow \mathbf{a}$$

cuando $t \to 0$.

Llegamos entonces a que hay a lo sumo un polinomio que puede cumplir la condición (4.1)

$$P_2(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + Hf(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

Sólo nos resta comprobar que este polinomio efectivamente satisface la condición que lo caracteriza. Para ello utilizaremos un argumento muy similar al empleado en la sección anterior para llegar a la expresión del resto de orden 1. Más precisamente, para cada $\mathbf{x}$ llamamos

$$\mathbf{u} = \frac{\mathbf{x} - \mathbf{a}}{||\mathbf{x} - \mathbf{a}||}$$

y definimos

$$\varphi(t) = f(\mathbf{a} + t\mathbf{u})$$

Cuentas similares a las hechas en esa ocasión nos permiten ahora decir que

$$\varphi(0) = f(\mathbf{a}) \quad , \quad \varphi(||\mathbf{x} - \mathbf{a}||) = f(\mathbf{x}) \quad , \quad \varphi'(0) = \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) \quad , \quad \varphi''(0) = Hf(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

Como para φ sabemos que

$$\lim_{t \to 0} \frac{\varphi(t) - \varphi(0) - \varphi'(0)t - \frac{1}{2}\varphi''(0)t^2}{t^2} = 0$$

escribiéndolo en términos de f llegamos a que

$$\begin{aligned}
&\lim_{\mathbf{x} - \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) - Hf(\mathbf{a})(\mathbf{x} - \mathbf{a})}{||\mathbf{x} - \mathbf{a}||^2} \\
&\qquad = \lim_{\mathbf{x} - \mathbf{a}} \frac{\varphi(||\mathbf{x} - \mathbf{a}||) - \varphi(0) - \varphi'(0)||\mathbf{x} - \mathbf{a}|| - \frac{1}{2}\varphi''(0)||\mathbf{x} - \mathbf{a}||^2}{||\mathbf{x} - \mathbf{a}||^2} \\
&\qquad = 0
\end{aligned}$$

dado que cuando $\mathbf{x} \longrightarrow \mathbf{a}$, $||\mathbf{x} - \mathbf{a}|| \longrightarrow 0$.

Queda probado entonces que

$$P_2(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + Hf(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

es el único polinomio de grado menor o igual que 2 *que satisface la condición* (4.1).

Observaciones

1. El polinomio de Taylor de orden 2 de f alrededor de $\mathbf{a}$ se puede expresar en la forma

$$P_2(\mathbf{x}) = P_1(\mathbf{x}) + Hf(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

2. El polinomio de Taylor de orden 2 de f alrededor de $\mathbf{a}$ aproxima mejor a f —cerca de $\mathbf{a}$— que el polinomio de orden 1. En efecto, P_2 satisface

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{f(\mathbf{x}) - P_2(\mathbf{x})}{||\mathbf{x} - \mathbf{a}||^2} = 0$$

y como $||\mathbf{x}-\mathbf{a}||^2$ tiende a cero *más rápido* que $||\mathbf{x}-\mathbf{a}||$, esto nos dice que para cada $\mathbf{x}$ suficientemente cercano a $\mathbf{a}$, la diferencia

$$f(\mathbf{x}) - P_2(\mathbf{x})$$

es *menor* que la diferencia

$$f(\mathbf{x}) - P_1(\mathbf{x})$$

3. Esta propiedad vale en general: cuanto más alto es el grado del polinomio de Taylor mayor es su aproximación a la función cerca del punto.