

Minimización de una forma cuadrática sobre una recta (ejercicios)

Objetivos. Dada una forma cuadrática positiva definida, encontrar su mínimo sobre una recta dada. Mostrar que en el punto mínimo el gradiente es ortogonal a la recta.

El gradiente de formas lineales y cuadráticas (repaso)

1. El gradiente de una función (repaso). Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ y sea $x \in \mathbb{R}^n$. Supongamos que f tiene todas las derivadas parciales en el punto x , esto es, para cada $j \in \{1, \dots, n\}$ existe y es finito el límite

$$(D_j f)(x) := \lim_{h \rightarrow 0} \frac{f(x + h e_j) - f(x)}{h}.$$

Entonces el gradiente de la función f en el punto x se define como el vector

$$(\nabla f)(x) := \left[\begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \right]_{j=1}^n.$$

En otras palabras, $(\nabla f)(x) \in \text{ } y$ para cada $j \in \{1, \dots, n\}$

$$((\nabla f)(x))_j = \text{ }.$$

2. Ejemplo (el gradiente de una forma lineal). Consideremos la función $\ell: \mathbb{R}^3 \rightarrow \mathbb{R}$ definida mediante la regla

$$\ell(x) = 7x_1 - 5x_2 + 4x_3.$$

Calculemos sus derivadas parciales:

$$(D_1 \ell)(x) = 7, \quad (D_2 \ell)(x) = \text{ }, \quad (D_3 \ell)(x) = \text{ }.$$

El gradiente de ℓ es

$$(\nabla \ell)(x) = \left[\begin{array}{c} (D_1 \ell)(x) \\ \text{ } \\ \text{ } \end{array} \right] = \left[\begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \right].$$

3. Fórmula general para el gradiente de una forma lineal (repaso). Sea $a \in \mathbb{R}^n$. Consideremos la función $\ell: \mathbb{R}^n \rightarrow \mathbb{R}$ definida mediante la regla

$$\ell(x) = a^\top x.$$

Entonces para cada $j \in \{1, \dots, n\}$ tenemos $(D_j \ell)(x) = \text{ },$ así que

$$(\nabla \ell)(x) = \left[\begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \right]_{j=1}^n = \text{ }.$$

4. Ejemplo (el gradiente de una forma cuadrática). Consideremos la función $q: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida mediante la regla

$$q(x) := x^\top \begin{bmatrix} 7 & -3 \\ -3 & 4 \end{bmatrix} x.$$

Escribimos $q(x)$ en coordenadas:

$$\begin{aligned} q(x) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 7 & -3 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \text{[]} \\ \text{[]} \end{bmatrix} \\ &= \text{[]} - 3 \text{[]} - \text{[]} + \text{[]} = \text{[]} - 6 \text{[]} + \text{[]}. \end{aligned}$$

Calculamos las derivadas parciales:

$$(D_1q)(x) = \text{[]}, \quad (D_2q)(x) = \text{[]}.$$

Formamos el vector gradiente y notamos que este vector se escribe como el producto de una matriz por un vector:

$$(\nabla q)(x) = \begin{bmatrix} \text{[]} \\ \text{[]} \end{bmatrix} = 2 \begin{bmatrix} \text{[]} & \text{[]} \\ \text{[]} & \text{[]} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

5. Fórmula general para el gradiente de una forma cuadrática asociada a una matriz simétrica (repaso). Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica: $A^\top = A$. Definimos $q: \mathbb{R}^n \rightarrow \mathbb{R}$ mediante la regla

$$q(x) = x^\top Ax.$$

Entonces

$$(\nabla q)(x) = \text{[]}.$$

6. Fórmula para el gradiente de una forma cuadrática con términos lineales.

Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica ($A^\top = \text{[]}$) y sea $b \in \mathbb{R}^n$. Consideremos la función $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(x) = \frac{1}{2} x^\top Ax - b^\top x. \tag{1}$$

Su gradiente es

$$(\nabla f)(x) = \text{[]}. \tag{2}$$

Restricción de una forma cuadrática a una recta

7. La regla de la cadena en el caso vectorial (repaso). Sea $\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$ una función vectorial continuamente derivable, y sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una función que tiene derivadas parciales continuas. Notamos que en cada punto del dominio la derivada de φ y el gradiente de f son vectores:

$$\varphi'(\alpha) \in \mathbb{R}^n, \quad (\nabla f)(x) \in \mathbb{R}^n. \quad (3)$$

Formamos la composición $h: \mathbb{R} \rightarrow \mathbb{R}$,

$$g(\alpha) := f(\varphi(\alpha)).$$

Entonces la derivada de la función g en cada punto $\alpha \in \mathbb{R}$ se puede calcular como el producto punto de los vectores (3), con x sustituido por $\varphi(\alpha)$:

$$g'(\alpha) = (\nabla f)(\varphi(\alpha))^\top \varphi'(\alpha) = \varphi'(\alpha)^\top (\nabla f)(\varphi(\alpha)). \quad (4)$$

8. Ejemplo (la derivada de una función restringida a una recta). Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ una función que tiene derivadas parciales continuas, y sean $y, p \in \mathbb{R}^2$. Definimos $\varphi: \mathbb{R} \rightarrow \mathbb{R}^2$ y $g: \mathbb{R} \rightarrow \mathbb{R}$ mediante las reglas

$$\varphi(\alpha) := y + \alpha p = \begin{bmatrix} y_1 + \alpha p_1 \\ \end{bmatrix},$$

$$g(\alpha) := f(\varphi(\alpha)) = f(y + \alpha p) = f(y_1 + \alpha p_1,).$$

Entonces φ es una función derivable, y en cada punto $\alpha \in \mathbb{R}$ la derivada de φ es el vector

$$\varphi'(\alpha) = \begin{bmatrix} \\ \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}.$$

La derivada de g se calcula por la regla de la cadena vectorial y se puede escribir como cierto producto interno:

$$\begin{aligned} g'(\alpha) &= (D_1 f)(y + \alpha p) \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + D_2 f(y + \alpha p) \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \\ &= \begin{bmatrix} & \end{bmatrix} \begin{bmatrix} \\ D_2 f(y + \alpha p) \end{bmatrix} = \begin{bmatrix} \end{bmatrix}^\top (\nabla f)(y + \alpha p). \end{aligned}$$

9. Fórmula general para la derivada de una función restringida a una recta.

Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una función que tiene derivadas parciales continuas, y sean $y, p \in \mathbb{R}^n$. Definimos $\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$ y $g: \mathbb{R} \rightarrow \mathbb{R}$ mediante las reglas

$$\varphi(\alpha) := y + \alpha p, \quad g(\alpha) := f(\varphi(\alpha)) = f(\text{[]}).$$

Entonces φ es una función derivable, y en cada punto $\alpha \in \mathbb{R}$ la derivada de φ es el vector

$$\varphi'(\alpha) = \text{[]}.$$

La derivada de g se calcula por la fórmula (4):

$$g'(\alpha) = \varphi'(\alpha)^\top \text{[]} = \text{[]} f(\text{[]}). \quad (5)$$

10. Restricción a una recta de una forma cuadrática con términos lineales.

Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz real simétrica estrictamente positiva definida, y sea $b \in \mathbb{R}^n$. Definamos la función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mediante la regla (1):

$$f(x) := \text{[]}.$$

Luego fijemos un punto $y \in \mathbb{R}^n$ y un vector $p \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, y consideremos la función $g: \mathbb{R} \rightarrow \mathbb{R}$,

$$g(\alpha) := f(y + \alpha p).$$

Calculemos $g(\alpha)$ en forma explícita:

$$g(\alpha) =$$

Respuesta final:

$$g(\alpha) = \left(\text{[]} \right) \alpha^2 - \left(\text{[]} \right) \alpha + f(y). \quad (6)$$

11. Derivada de una forma cuadrática con términos lineales, restringida a una recta. Sustituyendo $x = y + \alpha p$ en (2) obtenemos

$$(\nabla f)(y + \alpha p) = \text{[blanco]} . \quad (7)$$

Combinamos (5) y (7):

$$g'(\alpha) = \text{[blanco]}^\top (\text{[blanco]}), \quad (8)$$

esto es,

$$g'(\alpha) = \text{[blanco]} \alpha - \text{[blanco]}^\top (b - Ay). \quad (9)$$

La misma fórmula se puede deducir de (6):

$$g'(\alpha) =$$

Calculemos también la segunda derivada de g :

$$g''(\alpha) = \text{[blanco]} . \quad (10)$$

12. Minimización sobre una recta de una forma cuadrática con términos lineales. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica positiva definida:

$$A^\top = \text{[blanco]}, \quad \forall v \in \mathbb{R}^n \setminus \{\text{[blanco]}\} \quad v^\top Av > 0.$$

Sea $b \in \mathbb{R}^n$ y sea $p \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$. Definimos f y g como en los ejercicios anteriores:

$$f(x) := \frac{1}{2}x^\top Ax - b^\top x, \quad g(\alpha) := f(y + \alpha p).$$

La condición que A es positiva definida, la condición $p \neq \mathbf{0}_n$, y la fórmula (10) nos dan:

$$p^\top Ap \text{ [blanco]}, \quad \text{esto es,} \quad g''(\alpha) \text{ [blanco]} .$$

En otras palabras, por la fórmula (6), g es una función cuadrática con coeficiente mayor

$$\underbrace{\text{[blanco]}}_{\text{positivo/negativo}},$$

su gráfica es una parábola que se abre $\underbrace{\text{[blanco]}}_{\text{hacia arriba/hacia abajo}},$

y la función g tiene un único punto $\underbrace{\text{[blanco]}}_{\text{máximo/mínimo}}$ global estricto,

el cual denotemos por α_{\min} y que se puede encontrar de la ecuación $g'(\alpha) = 0$. Usando (9) y resolviendo la ecuación $g'(\alpha) = 0$ obtenemos

$$\alpha_{\min} = \frac{\text{[blanco]}}{\text{[blanco]}} . \quad (11)$$

13. Observación sobre el gradiente de la forma cuadrática con términos lineales, en el punto mínimo sobre una recta. En condiciones del ejercicio anterior, recordamos que

$$(\nabla f)(y + \alpha p) = \text{[]},$$

recordamos (8):

$$g'(\alpha) = \text{[]}^\top (\text{[]}),$$

y concluimos que la condición $g'(\alpha) = 0$ es equivalente a la ortogonalidad de dos vectores:

$$g'(\alpha) = 0 \iff \text{[]} \perp \text{[]}.$$

El punto α_{\min} satisface $g'(\alpha_{\min}) = 0$, así que

$$\text{[]} \perp b - A(y + \alpha_{\min}p). \tag{12}$$

Verifiquemos (12) de otra manera, usando (11):

$$\text{[]}^\top (b - A(y + \alpha_{\min}p)) =$$

14. Resumen. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica ($A^\top = \text{[]}$) y positiva definida:

$$\forall v \in$$

Además, sean $b, y \in \text{[]}$ y $p \in \text{[]}$. Definimos

$$f: \text{[]}, \quad f(x) := \text{[]},$$

y consideremos la restricción de la función f a la recta que pasa por el punto y en la dirección p :

$$g: \text{[]}, \quad g(\alpha) := f(\text{[]}).$$

Entonces la función g alcanza su [] global estricto en el punto

$$\alpha_{\min} = \text{[]}.$$

El gradiente de la función f en el punto correspondiente $y + \alpha_{\min}p$ es ortogonal al vector [], esto es, a la dirección de la recta. En otras palabras,

$$\text{[]} \perp b - A(\text{[]}).$$