Fórmula para el gradiente
de una forma cuadrática

Objetivos. Dada una matriz $A \in \mathcal{M}_n(\mathbb{R})$, deducir la fórmula para el gradiente de la forma cuadrática $q(x) \coloneqq x^\top Ax$.

Requisitos. Formas lineales y cuadráticas, las derivadas parciales de una función en un punto, el gradiente de una función en un punto, la derivada del producto, sumas con la delta de Kronecker, la matriz transpuesta, matrices simétricas.

Fórmula para el gradiente de una forma lineal

1. Definición del gradiente de una función (repaso). Sea $f\colon \mathbb{R}^n \to \mathbb{R}$ una función differenciable en cada punto. Denotamos por $(D_p f)(x)$ a la p-ésima derivada parcial de la función f en el punto x. Entonces el gradiente de f en el punto x se define como el vector de las derivadas parciales y se denota por $(\operatorname{grad} f)(x)$ o por $(\nabla f)(x)$:

$$(\operatorname{grad} f)(x) = (\nabla f)(x) = \big[(D_p f)(x)\big]_{p=1}^n.$$

En otras palabras, para cada $x \in \mathbb{R}^n$ tenemos $(\operatorname{grad} f)(x) \in \mathbb{R}^n$, y para cada $p \in \{1,\ldots,n\}$ la p-ésima componente del vector $(\operatorname{grad} f)(x)$ es

$$((\operatorname{grad} f)(x))_p = (D_p f)(x).$$

2. Fórmula para la derivada parcial de una función monomial. Sean $j, p \in \{1,\ldots,n\}$. Entonces

$$D_p x_j = \delta_{j,p}. \tag{1}$$

Demostración. Si $j = p$, entonces ambos lados de (1) son 1. En otro caso, si $j \neq p$, ambos lados son 0. □

3. Fórmula para el gradiente de una forma lineal. Sea $b \in \mathbb{R}^n$. Definimos $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ mediante la regla

$$\varphi(x) = b^\top x = \sum_{j=1}^n b_j x_j.$$

Entonces para cada $x \in \mathbb{R}^n$

$$(\operatorname{grad} \varphi)(x) = b.$$

Demostración. Sea $x \in \mathbb{R}^n$. Para cada $p \in \{1,\ldots,n\}$, calculemos la p-ésima derivada parcial de φ en el punto x, usando la linealidad del operador D_p y la fórmula (1):

$$((\operatorname{grad} \varphi)(x))_p = (D_p \varphi)(x) = \sum_{j=1}^n b_j \delta_{j,p} = b_p.$$

Formando un vector de estas derivadas parciales obtenemos el vector b. □

Fórmula para el gradiente de una forma cuadrática

4. Forma cuadrática con la notación matricial y en coordenadas. Sea $A \in \mathcal{M}_n(\mathbb{R})$ tal que $A^\top = A$. Definimos $q\colon \mathbb{R}^n \to \mathbb{R}$ mediante la fórmula

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

Entonces

$$q(x) = \sum_{j=1}^{n} \sum_{k=1}^{n} A_{j,k} x_j x_k. \tag{2}$$

Demostración. Aplicamos dos veces la definición del producto de matrices (para el caso especial de vectores-renglones y vectores-columnas):

$$x^\top A x = \sum_{j=1}^{n} x_j (Ax)_j = \sum_{j=1}^{n} \left(x_j \sum_{k=1}^{n} A_{j,k} x_k \right).$$

Aplicando propiedades del producto en $\mathbb{R}$ obtenemos (2). □

5. Fórmula para la derivada parcial del producto de dos variables. Sean $j, k, p \in \{1, \ldots, n\}$. Entonces

$$D_p(x_j\, x_k) = \delta_{p,j} x_k + \delta_{p,k} x_j. \tag{3}$$

Demostración. Primero aplicamos la fórmula para la derivada del producto:

$$D_p(x_j x_k) = x_j\, D_p(x_k) + x_k\, D_p(x_j),$$

luego utilizamos (1). □

6. Comprobación. Para comprobar (3) consideremos todos los casos posibles:

1. Si $p = j = k$, entonces ambos lados de (3) son iguales a $2x_j$.
2. Si $p = j$, $p \neq k$, entonces ambos lados son iguales a x_k.
3. Si $p = k$, $p \neq j$, entonces ambos lados son iguales a x_j.
4. Si $p \neq j$ y $p \neq k$, entonces ambos lados son 0.

7. Fórmula para el gradiente de una forma cuadrática. Sea $A \in \mathcal{M}_n(\mathbb{R})$. Definimos $q\colon \mathbb{R}^n \to \mathbb{R}$ mediante la fórmula

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

Entonces q es diferenciable, y para cada $x \in \mathbb{R}^n$ el gradiente de la función q en el punto x se calcula por la fórmula

$$(\operatorname{grad} q)(x) = (A + A^\top) x. \tag{4}$$

En particular, si la matriz A es simétrica, entonces

$$(\operatorname{grad} q)(x) = 2Ax. \tag{5}$$

Demostración. Sea $x \in \mathbb{R}^n$. Para cada $p \in \{1, \ldots, n\}$ calculemos la p-ésima componente del vector $(\operatorname{grad} q)(x)$ aplicando (2) y (3):

$$
\begin{aligned}
((\operatorname{grad} q)(x))_p = D_p q(x) &= \sum_{j=1}^{n} \sum_{k=1}^{n} A_{j,k} (\delta_{p,j} x_k + \delta_{p,k} x_j) \\
&= \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} \delta_{p,j} x_k + \sum_{j=1}^{n} \sum_{k=1}^{n} A_{j,k} \delta_{p,k} x_j \\
&= \sum_{k=1}^{n} A_{p,k} x_k + \sum_{j=1}^{n} A_{j,p} x_j \\
&= \sum_{k=1}^{n} A_{p,k} x_k + \sum_{j=1}^{n} (A^\top)_{p,j} x_j \\
&= (Ax)_p + (A^\top x)_p = ((A + A^\top)x)_p.
\end{aligned}
$$

Como los vectores $(\operatorname{grad} q)(x)$ y $(A+A^\top)x$ son de la misma longitud n y tienen las mismas componentes en posiciones correspondientes, son iguales. □