Fórmula para el gradiente de una forma cuadrática

http://esfm.egormaximenko.com/numlinalg/gradient_of_quadratic_form_es.pdf

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

Fórmula para el gradiente de una forma lineal

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

$$\frac{\partial x_j}{\partial x_p} = \delta_{j,p}.$$

Demostración. En realidad, si $j = p$, entonces ambos lados son 1, y si $j \neq p$, entonces ambos lados son 0. $\square$

2. Fórmula para las derivadas parciales 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 $p \in \{1, \ldots, n\}$

$$(D_p\varphi)(x) = b_p. \tag{1}$$

Demostración.

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

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

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

Demostración. Usamos la fórmula (1) y la definición del gradiente:

$$(\operatorname{grad}\varphi)(x) = \big[D_p\varphi(x)\big]_{j=1}^n = \big[b_p\big]_{j=1}^n = b. \qquad \square$$

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

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

Demostración. 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.

□

6. Fórmula para las derivadas parciales de una forma cuadrática. 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 para cada $p \in \{1, \ldots, n\}$ y cada $x \in \mathbb{R}^n$, la p-ésima derivada parcial de la función q en el punto x se calcula por la fórmula

$$D_p q(x) = 2 \sum_{k=1}^{n} A_{p,k} x_k. \tag{4}$$

Demostración. Aplicamos (2) y (3):

$$\begin{aligned} D_p q(x) &= \sum_{j=1}^n \sum_{k=1}^n A_{j,k}(\delta_{p,j} x_k + \delta_{p,k} x_j) = \sum_{k=1}^n \sum_{j=1}^n A_{j,k}\delta_{p,j} x_k + \sum_{j=1}^n \sum_{k=1}^n A_{j,k}\delta_{p,k} x_j \\ &= \sum_{k=1}^n A_{p,k} x_k + \sum_{j=1}^n A_{j,p} x_j = 2\sum_{k=1}^n A_{p,k} x_k. \end{aligned} \qquad \square$$

7. Fórmula para el gradiente de una forma cuadrática. 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 es diferenciable, y para cada $x \in \mathbb{R}^n$ la matriz Jacobiana de la función q en el punto x se calcula por la fórmula

$$q'(x) = 2x^\top A.$$

En otras palabras, el vector gradiente de la función q en el punto x es

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

Demostración. Aplicar (4) y la definición del producto de una matriz por un vector:

$$(\operatorname{grad} q)(x) = \Big[D_p q(x)\Big]_{p=1}^n = \left[2\sum_{k=1}^n A_{p,k} x_k\right]_{p=1}^n = 2Ax. \qquad \square$$

8. El caso de una matriz no simétrica. Si $A \in \mathcal{M}_n(\mathbb{R})$ sin suponer que $A^\top = A$, y definimos $q\colon \mathbb{R}^n \to \mathbb{R}$ mediante la fórmula $q(x) = x^\top A x$, entonces

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

La demostración es muy similar a la de escrita arriba.