1. Ejercicios

Definición 1. *A una función* $f\colon \mathbb{R}^n \to \mathbb{R}$ *es cuadrática si es de la forma*

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c$$

donde $Q \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$.

En muchas de las tareas abajo, la función f es cuadrática como en Definición 1 .

1. Aplicación 1: *El problema de mínimos cuadrados (lineales) regularizados.*
Sea $A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$, entonces el problema es encontrar $\boldsymbol{x}_\star \in \mathbb{R}^n$ tal que

$$f(\boldsymbol{x}_\star) = \min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) \qquad \text{donde} \qquad f(\boldsymbol{x}) \stackrel{\text{def}}{=\!=} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \mu\|\boldsymbol{x}\|_2^2$$

y $\mu > 0$ es un parámetro. La función f es una función cuadrática.
Encuentre, $Q, \boldsymbol{b}, c$ y muestre que Q es simétrica positiva definida.

Nota: Sin el termino azul, la solución puede ser no única (ver SVD, Cálculo Numérico).

2. Supón que f es una función cuadrática con una matriz simétrica positiva definida $Q \in \mathbb{R}^{n\times n}$. Demuestre que $\phi(\alpha) \stackrel{\text{def}}{=\!=} f(\boldsymbol{x}+\alpha\boldsymbol{d})$ es un polinomio cuadrático en α que siempre tiene un mínimo en el parámetro $\alpha_\star$. Además, si $\boldsymbol{d}$ es dirección de descenso, entonces el mínimo de ϕ esta en $\alpha_\star > 0$.

Resultado: $\alpha_\star = \frac{-\boldsymbol{d}^T\nabla f(\boldsymbol{x})}{\boldsymbol{d}^T Q\boldsymbol{d}}$

3. Supón que $f(\boldsymbol{x}) = \boldsymbol{x}^T Q\boldsymbol{x}$ donde $Q \in \mathbb{R}^{n\times n}$ es una matriz simétrica positiva semi-definida. Muestra que f es convexa en $\mathbb{R}^n$. Ayuda, para $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ muestre que

$$f(\boldsymbol{y} + \alpha(\boldsymbol{x} - \boldsymbol{y})) - \alpha f(\boldsymbol{x}) - (1-\alpha)f(\boldsymbol{y}) \le 0$$

para todo $\alpha \in [0, 1]$.

Ayuda: Considerar en el apartado anterior.

4. Supón que $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x}$ donde $\boldsymbol{b} \in \mathbb{R}^n$ y $Q \in \mathbb{R}^{n\times n}$ es una matriz simétrica positiva definida. Sea $\boldsymbol{x}_\star$ el mínimo (global) único. Demuestre que

$$f(\boldsymbol{x}) - f(\boldsymbol{x}_\star) = \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}_\star\|_Q^2 \qquad \text{donde} \qquad \|\boldsymbol{x} - \boldsymbol{x}_\star\|_Q^2 \stackrel{\text{def}}{=\!=} (\boldsymbol{x} - \boldsymbol{x}_\star)^T Q(\boldsymbol{x} - \boldsymbol{x}_\star)\,.$$

Ayuda: Use primero $\nabla f(\boldsymbol{x}_\star) = \boldsymbol{0}$ *para encontrar una expresión para* $\boldsymbol{b}^T(\boldsymbol{x} - \boldsymbol{x}_\star)$.

5. Supón que $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x}$ donde $\boldsymbol{b} \in \mathbb{R}^n$ y $Q \in \mathbb{R}^{n\times n}$ es una matriz simétrica positiva definida. Sean $\boldsymbol{x} \in \mathbb{R}^n$ fijo, $\boldsymbol{d} \overset{\text{def}}{=\!=} -\nabla f(\boldsymbol{x})$ y $\phi(\alpha) \overset{\text{def}}{=\!=} f(\boldsymbol{x} + \alpha\boldsymbol{d})$.

a) Muestre que
$$\phi(\alpha) = \phi(0) + \frac{1}{2}\alpha^2\boldsymbol{d}^T Q\boldsymbol{d} - \alpha\boldsymbol{d}^T\boldsymbol{d}\,.$$

b) Encuentre el paso óptimo $\alpha_\star$ que minimiza a la función $\phi(\alpha)$.

c) Usando el paso óptimo $\alpha_\star = \frac{\boldsymbol{d}^T\boldsymbol{d}}{\boldsymbol{d}^T Q\boldsymbol{d}}$, muestre que
$$\phi(\alpha_\star) = \phi(0) - \frac{1}{2}\frac{(\boldsymbol{d}^T\boldsymbol{d})^2}{\boldsymbol{d}^T Q\boldsymbol{d}}$$

d) Demuestre que $\boldsymbol{d}^T Q^{-1}\boldsymbol{d} = \|\boldsymbol{x} - \boldsymbol{x}_\star\|_Q^2$. Además, usando la tarea anterior
$$\|\boldsymbol{x} + \alpha_\star\boldsymbol{d} - \boldsymbol{x}_\star\|_Q^2 = \left(1 - \frac{(\boldsymbol{d}^T\boldsymbol{d})^2}{(\boldsymbol{d}^T Q\boldsymbol{d})(\boldsymbol{d}^T Q^{-1}\boldsymbol{d})}\right)\|\boldsymbol{x} - \boldsymbol{x}_\star\|_Q^2$$

e) La identidad en d) es poderosa.
Demuestre, que si $\boldsymbol{x} - \boldsymbol{x}_\star$ es un eigenvector de Q, entonces $\boldsymbol{x} + \alpha_\star\boldsymbol{d} = \boldsymbol{x}_\star$.

Ayuda: Para a) se usa $\boldsymbol{d} = -\nabla f(\boldsymbol{x}) = -(Q\boldsymbol{x} + \boldsymbol{b})$ *(similar a* ***2.****),*
para d) considere c), ***5.*** *y* $\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{x}_\star) = Q(\boldsymbol{x} - \boldsymbol{x}_\star)$*,*
para e) considere d) y $\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{x}_\star) = Q(\boldsymbol{x} - \boldsymbol{x}_\star)$.

2. Otra aplicación

El problema de Fermat-Weber.

El objetivo de este problema es encontrar una posición óptima $\boldsymbol{x}_\star \in \mathbb{R}^n$ cuya distancia a distintos puntos $\boldsymbol{a}^i \in \mathbb{R}^n$, $i = 1, \dots, m$ con pesos w_i (de Ingles: *weights*) es mínima. El vector $\boldsymbol{x}_\star$ debe ser solución del problema:
$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \left(f(\boldsymbol{x}) \overset{\text{def}}{=\!=} \sum_{i=1}^{m} w_i \big\|\boldsymbol{a}^i - \boldsymbol{x}\big\|_2 \right).$$

Ejemplo: El vector $\boldsymbol{x}_\star$ *puede ser una posición de un almacén, con el fin de reducir gastos de transporte. Los pesos* w_i *pueden contener información sobre como se convierte la norma de taxi a la norma euclidiana y gastos de usar una calle o un medio de transporte.*

El problema se puede resolver con una iteración de punto fijo diseñada por *Weiszfeld (1937)*. Esta iteración construye (de manera explícita) una sucesión $\{\boldsymbol{x}^k\}_{k\geq 0}$ empezando en $\boldsymbol{x}^0 \notin \{\boldsymbol{a}^1, \dots, \boldsymbol{a}^m\}$. Cada punto $\boldsymbol{x}^{k+1}$ de la sucesión coincide con la solución del problema
$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \left(g(\boldsymbol{x}) \overset{\text{def}}{=\!=} \sum_{i=1}^{m} w_i \frac{\big\|\boldsymbol{a}^i - \boldsymbol{x}\big\|_2^2}{\big\|\boldsymbol{a}^i - \boldsymbol{x}^k\big\|_2} \right).$$

En caso que $f(\boldsymbol{x}^0) < \min\big\{|f(\boldsymbol{a}^j)|\colon j = 1, \dots, m\big\}$, todos los puntos de acumulación de la sucesión son puntos estacionarios.

Referencia: Introduction to non-linear optimization, Beck, p. 68-72.