Departamento de Ciencias de la Computación
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
(FCEIA-UNR)

Optimización Continua

Prof. Alejandro G. Marchetti

Optimización Convexa

Abril de 2025

1. Conjuntos Convexos

1.1. Conjuntos Afines y Conjuntos Convexos

Líneas y Segmentos

Suponga que $\mathbf{x}_1$ y $\mathbf{x}_2$ son dos puntos en $\mathbb{R}^n$. Los puntos de la forma

$$\mathbf{x} = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2,$$

donde $\theta \in \mathbb{R}$, forman la *línea* que pasa por $\mathbf{x}_1$ y $\mathbf{x}_2$. El valor del parámetro $\theta = 0$ corresponde a $\mathbf{x} = \mathbf{x}_2$ y el valor del parámetro $\theta = 1$ corresponde a $\mathbf{x} = \mathbf{x}_1$. A medida que θ se incrementa de 0 a 1, el punto $\mathbf{x}$ se mueve de $\mathbf{x}_2$ a $\mathbf{x}_1$. Para $\theta > 1$, el punto $\mathbf{x}$ se mueve más alla de $\mathbf{x}_1$. Los valores de θ entre 0 y 1 corresponden al segmento entre $\mathbf{x}_1$ y $\mathbf{x}_2$. Esto se ilustra en la Figura 1.

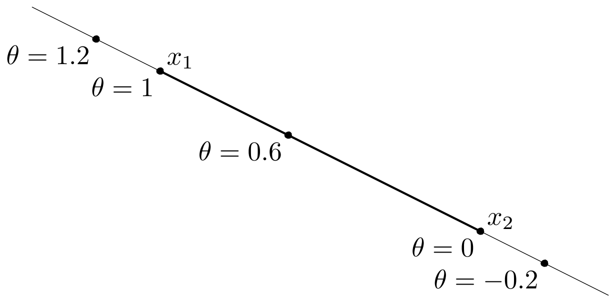

Figura 1: Línea que pasa por $\mathbf{x}_1$ y $\mathbf{x}_2$.

Conjunto Afín

Un conjunto $\mathcal{C} \subseteq \mathbb{R}^n$ es *afín* si la línea que pasa por dos puntos cualesquiera de $\mathcal{C}$ pertence a $\mathcal{C}$, es decir, si para cualquier $\mathbf{x}_1,\ \mathbf{x}_2\ \in \mathcal{C}$ y $\theta \in \mathbb{R}$, tenemos que $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{C}$. En otras palabras, $\mathcal{C}$ contiene la combinación lineal de cualquier par de puntos distintos en $\mathcal{C}$, provisto que la suma de los coeficientes de la combinación lineal sea igual a uno.

Esta idea se puede generalizar a más de dos puntos. Nos referimos a un punto de la forma $\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 + \cdots + \theta_k\mathbf{x}_k$, donde $\theta_1 + \theta_2 + \cdots + \theta_k = 1$, como una *combinación afín* de los puntos $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Por inducción se puede demostrar que un conjunto afín contiene toda combinación afín de sus puntos: Si $\mathcal{C}$ es un conjunto afín, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathcal{C}$, y $\theta_1 + \theta_2 + \cdots + \theta_k = 1$, luego el punto $\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 + \cdots + \theta_k\mathbf{x}_k$ también pertence a $\mathcal{C}$.

Si $\mathcal{C}$ es un conjunto afín y $\mathbf{x}_0 \in \mathcal{C}$, luego el conjunto

$$\mathcal{V} = \mathcal{C} - \mathbf{x}_0 = \{\mathbf{x} - \mathbf{x}_0 | \mathbf{x} \in \mathcal{C}\}$$

es un subespacio. De ello se desprende que el conjunto afín $\mathcal{C}$ se puede expresar como

$$\mathcal{C} = \mathcal{V} + \mathbf{x}_0 = \{\mathbf{v} + \mathbf{x}_0 | \mathbf{v} \in \mathcal{V}\},$$

es decir, un subespacio más una traslación. El subespacio $\mathcal{V}$ asociado al conjunto afín $\mathcal{C}$ no depende de la elección de $\mathbf{x}_0$, por lo cual $\mathbf{x}_0$ se puede elegir como cualquier punto en $\mathcal{C}$.

Definimos la *dimensión* de un conjunto afín $\mathcal{C}$ como la dimensión del subespacio $\mathcal{V} = \mathcal{C} - \mathbf{x}_0$, donde $\mathbf{x}_0$ es cualquier elemento de $\mathcal{C}$.

Ejemplo 1 (Solución de un sistema de ecuaciones lineales) *La solución de un sistema de ecuaciones lineales,* $\mathcal{C} = \{\mathbf{x} | \mathsf{A}\mathbf{x} = \mathbf{b}\}$*, donde* $\mathsf{A} \in \mathbb{R}^{m\times n}$ *y* $\mathbf{b} \in \mathbb{R}^m$*, es un conjunto afín. Para* *demostrar esto, suponga que* $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$*, es decir,* $\mathsf{A}\mathbf{x}_1 = \mathbf{b}$ *y* $\mathsf{A}\mathbf{x}_2 = \mathbf{b}$*. Luego, para cualquier* θ *tenemos*

$$\begin{aligned}\mathsf{A}\left(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right) &= \theta\mathsf{A}\mathbf{x}_1 + (1-\theta)\mathsf{A}\mathbf{x}_2 \\ &= \theta\mathbf{b} + (1-\theta)\mathbf{b} \\ &= \mathbf{b},\end{aligned}$$

lo cual demuestra que la combinación afín $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2$ *también pertenece a* $\mathcal{C}$*. El subespacio asociado con el conjunto afín* $\mathcal{C}$ *es el espacio nulo de* A*.*

También se cumple el contrario: todo conjunto afín se puede expresar como el conjunto solución de un sistema de ecuaciones lineales.

Conjunto Convexo

Un conjunto $\mathcal{S} \subset \mathbb{R}^n$ es convexo si contiene el segmento entre dos puntos distintos pertenecientes al conjunto, es decir, si para todo $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ tenemos

$$\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in \mathcal{S}, \quad \forall\ \lambda \in [0,1]$$

Todo conjunto afín es también convexo, ya que contiene la línea completa entre cualquier par de puntos distintos pertenecientes al conjunto. La Figura 2 ilustra ejemplos de conjuntos convexos y no convexos en $\mathbb{R}^2$.

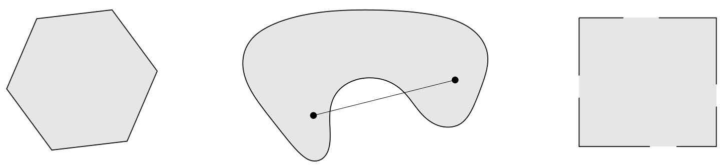

Figura 2: Ejemplos de conjuntos convexos y no convexos. *Izquierda*: El hexágono que incluye su frontera es convexo. *Centro*: El conjunto con forma de riñon no es convexo, ya que existen segmentos entre pares de puntos del conjunto que no están contenidos en el conjunto. *Derecha*: El cuadrado contiene algunos de sus puntos frontera y otros no, y por lo tanto no es convexo.

Combinación convexa. Decimos que un punto $\mathbf{x}$ tal que

$$\mathbf{x} = \theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 + \cdots + \theta_k\mathbf{x}_k, \quad \text{con} \quad \sum_{i=1}^{k}\theta_i = 1, \text{ y } \ \theta_i \geq 0, \quad i = 1, 2, \ldots, k$$

es una *combinacion convexa* de los puntos $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Se puede demostrar que un conjunto es convexo si y solo si contiene todas las combinaciones convexas de sus puntos.

Envoltura convexa (o cáscara convexa). Sea un conjunto $\mathcal{S} \in \mathbb{R}^n$. La envoltura convexa (denominada en inglés *convex hull*) de $\mathcal{S}$, denotada $\operatorname{conv}(\mathcal{S})$, es el conjunto de todas las combinaciones convexas de puntos es $\mathcal{S}$:

$$\operatorname{conv}(\mathcal{S}) = \left\{\mathbf{x} = \sum_{i=1}^{k}\theta_i\mathbf{x}_i : \ \mathbf{x}_i \in \mathcal{S}, \ \theta_i \geq 0, \ i = 1, \ldots, k, \ \sum_{i=1}^{k}\theta_i = 1\right\}$$

Como el nombre lo indica, la envoltura convexa de un conjunto $\mathcal{S}$ es siempre un conjunto convexo. Es el menor conjunto convexo que contiene $\mathcal{S}$. La Figura 3 ilustra la definición de envoltura convexa.

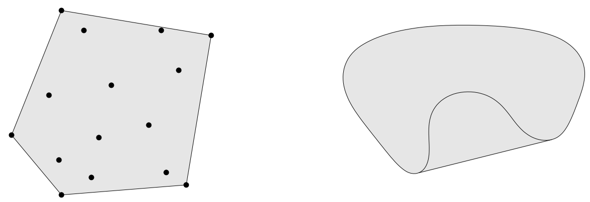

Figura 3: Envoltura convexa de dos conjuntos en $\mathbb{R}^2$. *Izquierda*: La envolutura convexa del conjunto de quince puntos es un pentágono. *Derecha*: La envoltura convexa del conjunto con forma de riñon de la Figura 2 es el conjunto sombreado.

El siguiente lema es consecuencia inmediata de la definición de convexidad. Establece que la intersección de dos conjuntos convexos es convexo y que la suma algebraica de dos conjuntos convexos es también un conjunto convexo.

Lema 1 *Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ dos conjuntos convexos en $\mathbb{R}^n$. Luego*

1) $\mathcal{S}_1 \cap \mathcal{S}_2$ *es convexo.*

2) $\mathcal{S}_1 \oplus \mathcal{S}_2 = \{\mathbf{x}_1 + \mathbf{x}_2 : \ \mathbf{x}_1 \in \mathcal{S}_1, \ \mathbf{x}_2 \in \mathcal{S}_2\}$ *es convexo*

3) $\mathcal{S}_1 \ominus \mathcal{S}_2 = \{\mathbf{x}_1 - \mathbf{x}_2 : \ \mathbf{x}_1 \in \mathcal{S}_1, \ \mathbf{x}_2 \in \mathcal{S}_2\}$ *es convexo*

Polítopo. La envoltura convexa de un número finito de puntos $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ en $\mathbb{R}^n$ es un polítopo (poliedro acotado). Si $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ son *afínmente independientes*, lo cual significa que $\mathbf{x}_2 - \mathbf{x}_1, \mathbf{x}_3 - \mathbf{x}_1, \mathbf{x}_{k+1} - \mathbf{x}_1$ son linealmente independientes, luego la envoltura convexa de $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ es un *simplex* de vértices $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$. Un simplex en $\mathbb{R}^n$ es un polítopo de $n+1$ vértices.

Por definición, un punto en $\mathrm{conv}(\mathcal{S})$ puede ser representado como combinación convexa de un número finito de puntos en $\mathcal{S}$. El siguiente teorema establece que cualquier punto $\mathbf{x} \in \mathrm{conv}(\mathcal{S})$ puede ser representado por combinación convexa de a lo sumo $n+1$ puntos de $\mathcal{S}$.

Teorema 1 (Teorema de Carathéodory) *Sea $\mathcal{S}$ un conjunto arbitrario en $\mathbb{R}^n$. Si $\mathbf{x} \in \mathrm{conv}(\mathcal{S})$, luego $\mathbf{x} \in \mathrm{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_{n+1})$ donde $\mathbf{x}_j \in \mathcal{S}$ para $j = 1, \ldots, n+1$. En otras palabras, $\mathbf{x}$ puede ser representado como*

$$
\begin{aligned}
&\mathbf{x} = \sum_{j=1}^{n+1} \lambda_j \mathbf{x}_j \\
&\sum_{j=1}^{n+1} \lambda_j = 1 \\
&\lambda_j \geq 0 \quad \text{para} \quad j = 1, \ldots, n+1 \\
&\mathbf{x}_j \in \mathcal{S} \quad \text{para} \quad j = 1, \ldots, n+1
\end{aligned}
$$

Demostración. No lo vemos (Teorema 2.1.6 de Bazaraa) $\square$

1.2. Separación y Soporte de Conjuntos

Dado un conjunto $\mathcal{S}$ convexo y cerrado, y un punto $\mathbf{y} \notin \mathcal{S}$, existe un único punto $\bar{\mathbf{x}} \in \mathcal{S}$ a distancia mínima de $\mathbf{y}$, y un hiperplano que separa $\mathbf{y}$ de $\mathcal{S}$.

Para establecer este importante resultado, haremos uso de la siguiente ley del paralelogramo:

Ley del paralelogramo: Sean $\mathbf{a}$ y $\mathbf{b}$ dos vectores en $\mathbb{R}^n$. Luego

$$\|\mathbf{a}+\mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 + 2\mathbf{a}^\mathsf{T}\mathbf{b}$$
$$\|\mathbf{a}-\mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\mathbf{a}^\mathsf{T}\mathbf{b}$$

Sumando obtenemos

$$\|\mathbf{a}+\mathbf{b}\|^2 + \|\mathbf{a}-\mathbf{b}\|^2 = 2\|\mathbf{a}\|^2 + 2\|\mathbf{b}\|^2$$

El resultado se ilustra en la Figura 4 y se puede iterpretar como sigue: La suma de las normas al cuadrado de las diagonales de un paralelogramo es igual a la suma de las normas al cuadrado de todos sus lados.

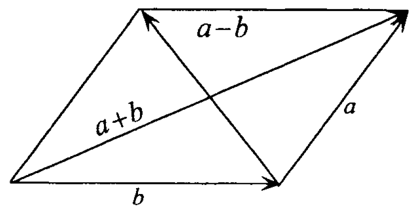

Figura 4: Ley del paralelogramo

Teorema 2 *Sea $\mathcal{S} \neq \emptyset$, convexo y cerrado en $\mathbb{R}^n$, y $\mathbf{y} \notin \mathcal{S}$. Luego, existe un único punto $\bar{\mathbf{x}} \in \mathcal{S}$ a mínima distanca de $\mathbf{y}$. Además, $\bar{\mathbf{x}}$ minimiza la distancia si y solo si $(\mathbf{y}-\bar{\mathbf{x}})^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) \leq 0$ para todo $\mathbf{x} \in \mathcal{S}$.*

Demostración. Demostraremos la existencia, unicidad, suficiencia y necesidad.
Existencia: $\mathcal{S} \neq \emptyset$, luego existe $\hat{\mathbf{x}} \in \mathcal{S}$, y podemos buscar el mínimo en

$$\bar{\mathcal{S}} = \mathcal{S} \cap \{\mathbf{x} : \ \|\mathbf{y}-\mathbf{x}\| \leq \|\mathbf{y}-\hat{\mathbf{x}}\|\}$$

En otras palabras, el problema de hallar el punto más cercano

$$\inf\{\|\mathbf{y}-\mathbf{x}\| : \ \mathbf{x} \in \mathcal{S}\}$$

es equivalente a

$$\inf\{\|\mathbf{y}-\mathbf{x}\| : \ \mathbf{x} \in \bar{\mathcal{S}}\}$$

Siendo $f(\mathbf{x}) = \|\mathbf{y}-\mathbf{x}\|$ una función continua y $\bar{\mathcal{S}}$ un conjunto compacto no vacío, sabemos que existe un mínimo $\bar{\mathbf{x}} \in \mathcal{S}$ por el Teorema de Weierstrass.
Unicidad: Suponga que existe $\bar{\mathbf{x}}' \in \mathcal{S}$ tal que $\|\mathbf{y}-\bar{\mathbf{x}}\| = \|\mathbf{y}-\bar{\mathbf{x}}'\| = \gamma$. Por la convexidad de $\mathcal{S}$, tenemos que $\frac{\bar{\mathbf{x}}+\bar{\mathbf{x}}'}{2} \in \mathcal{S}$. Por la desigualdad del triángulo, tenemos

$$\left\|\mathbf{y} - \frac{\bar{\mathbf{x}}+\bar{\mathbf{x}}'}{2}\right\| \leq \frac{1}{2}\|\mathbf{y}-\bar{\mathbf{x}}\| + \frac{1}{2}\|\mathbf{y}-\bar{\mathbf{x}}'\| = \gamma.$$

Si la desigualdad fuera estricta, tendríamos una contradicción ya que el punto $\frac{\bar{\mathbf{x}}+\bar{\mathbf{x}}'}{2}$ estaría más cercano a $\mathbf{y}$ que $\bar{\mathbf{x}}$. Por lo tanto se cumple la igualdad:

$$\left\|\mathbf{y}-\frac{\bar{\mathbf{x}}+\bar{\mathbf{x}}'}{2}\right\| = \frac{1}{2}\|(\mathbf{y}-\bar{\mathbf{x}})+(\mathbf{y}-\bar{\mathbf{x}}')\| = \frac{1}{2}\|\mathbf{y}-\bar{\mathbf{x}}\| + \frac{1}{2}\|\mathbf{y}-\bar{\mathbf{x}}'\| = \gamma.$$

Por lo tanto, debemos tener $\mathbf{y}-\bar{\mathbf{x}} = \lambda(\mathbf{y}-\bar{\mathbf{x}}')$ para algún λ. Como $\|\mathbf{y}-\bar{\mathbf{x}}\| = \|\mathbf{y}-\bar{\mathbf{x}}'\| = \gamma$, debemos tener $|\lambda| = 1$. Claramente, $\lambda \neq -1$, ya que de lo contrario obtenemos $\mathbf{y} = (\bar{\mathbf{x}}+\bar{\mathbf{x}}')/2 \in \mathcal{S}$, contradiciendo la hipótesis de que $\mathbf{y} \notin \mathcal{S}$. Luego, $\lambda = 1$, y $\bar{\mathbf{x}}' = \bar{\mathbf{x}}$, estableciéndose la unicidad de la solución óptima.
Suficiencia: Sea $\mathbf{x} \in \mathcal{S}$. Luego

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|\mathbf{y}-\bar{\mathbf{x}}+\bar{\mathbf{x}}-\mathbf{x}\|^2 = \|\mathbf{y}-\bar{\mathbf{x}}\|^2 + \|\bar{\mathbf{x}}-\mathbf{x}\|^2 + 2(\bar{\mathbf{x}}-\mathbf{x})^\mathsf{T}(\mathbf{y}-\bar{\mathbf{x}})$$

Como $\|\bar{\mathbf{x}}-\mathbf{x}\|^2 \geq 0$, y $(\bar{\mathbf{x}}-\mathbf{x})^\mathsf{T}(\mathbf{y}-\bar{\mathbf{x}}) \geq 0$ por hipótesis, luego $\|\mathbf{y}-\mathbf{x}\|^2 \geq \|\mathbf{y}-\bar{\mathbf{x}}\|^2$, y por lo tanto $\bar{\mathbf{x}}$ es un minimizador.
Necesidad: Suponga que $\|\mathbf{y}-\mathbf{x}\|^2 \geq \|\mathbf{y}-\bar{\mathbf{x}}\|^2$ para todo $\mathbf{x} \in \mathcal{S}$. Sea $\mathbf{x} \in \mathcal{S}$ y note que $\bar{\mathbf{x}}+\lambda(\mathbf{x}-\bar{\mathbf{x}}) \in \mathcal{S}$ para $0 \leq \lambda \leq 1$ por la convexidad de $\mathcal{S}$. Luego

$$\|\mathbf{y}-\bar{\mathbf{x}}-\lambda(\mathbf{x}-\bar{\mathbf{x}})\|^2 \geq \|\mathbf{y}-\bar{\mathbf{x}}\|^2 \tag{1}$$

Además,

$$\|\mathbf{y}-\bar{\mathbf{x}}-\lambda(\mathbf{x}-\bar{\mathbf{x}})\|^2 = \|\mathbf{y}-\bar{\mathbf{x}}\|^2 + \lambda^2\|\mathbf{x}-\bar{\mathbf{x}}\|^2 - 2\lambda(\mathbf{y}-\bar{\mathbf{x}})^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) \tag{2}$$

De (1) y (2) obtenemos

$$2\lambda(\mathbf{y}-\bar{\mathbf{x}})^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) \leq \lambda^2\|\mathbf{x}-\bar{\mathbf{x}}\|^2 \tag{3}$$

para todo $\lambda \in [0,1]$. Dividiendo (3) por $\lambda > 0$, y para $\lambda \to 0^+$, obtenemos $(\mathbf{y}-\bar{\mathbf{x}})^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) \leq 0$. □

La interpretación del Teorema 5 se ilustra en la Figura 5a. El ángulo entre $(\mathbf{y}-\bar{\mathbf{x}})$ y $(\mathbf{x}-\bar{\mathbf{x}})$ es mayor o igual a 90° para cualquier $\mathbf{x} \in \mathcal{S}$, y por lo tanto $(\mathbf{y}-\bar{\mathbf{x}})^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) \leq 0$ para todo $\mathbf{x} \in \mathcal{S}$. Esto implica que $\mathcal{S}$ se halla en el hemiespacio $\mathbf{a}^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) \leq 0$ relativo al hiperplano $\mathbf{a}^\mathsf{T}(\mathbf{x}-\bar{\mathbf{x}}) = 0$ que pasa por el punto $\bar{\mathbf{x}}$ y cuyo vector normal es $\mathbf{a} = (\mathbf{y}-\bar{\mathbf{x}})$. La Figura 5b muestra que esto no se cumple necesariamente si $\mathcal{S}$ no es convexo.

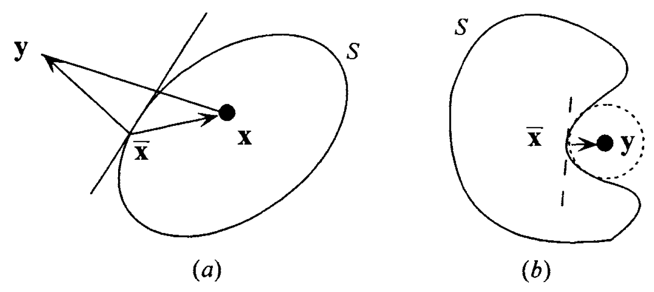

Figura 5: Distancia mínima a un conjunto convexo cerrado.

Separación por Hiperplanos

Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ conjuntos no vacios en $\mathbb{R}^n$.

- Un hiperplano $H = \{\mathbf{x} : \mathbf{p}^\mathsf{T}\mathbf{x} = \alpha\}$ *separa* $\mathcal{S}_1$ y $\mathcal{S}_2$ si $\mathbf{p}^\mathsf{T}\mathbf{x} \geq \alpha$ para todo $\mathbf{x} \in \mathcal{S}_1$ y $\mathbf{p}^\mathsf{T}\mathbf{x} \leq \alpha$ para todo $\mathbf{x} \in \mathcal{S}_2$.

- Si además $\mathcal{S}_1 \cup \mathcal{S}_2 \not\subseteq H$, entonces H *separa propiamente* $\mathcal{S}_1$ y $\mathcal{S}_2$.
- Si $\mathbf{p}^\mathsf{T}\mathbf{x} > \alpha$ para todo $\mathbf{x} \in \mathcal{S}_1$ y $\mathbf{p}^\mathsf{T}\mathbf{x} < \alpha$ para todo $\mathbf{x} \in \mathcal{S}_2$, entonces H *separa estrictamente* $\mathcal{S}_1$ y $\mathcal{S}_2$.
- Si $\mathbf{p}^\mathsf{T}\mathbf{x} > \alpha + \varepsilon$ para todo $\mathbf{x} \in \mathcal{S}_1$ y $\mathbf{p}^\mathsf{T}\mathbf{x} < \alpha$ para todo $\mathbf{x} \in \mathcal{S}_2$, donde $\varepsilon > 0$, entonces H *separa fuertemente* $\mathcal{S}_1$ y $\mathcal{S}_2$.

La Figura 6 muestra los distintos tipos de separación. Por supuesto, separación fuerte implica separación estricta, que a su vez implica separación propia, que a su vez implica separación.

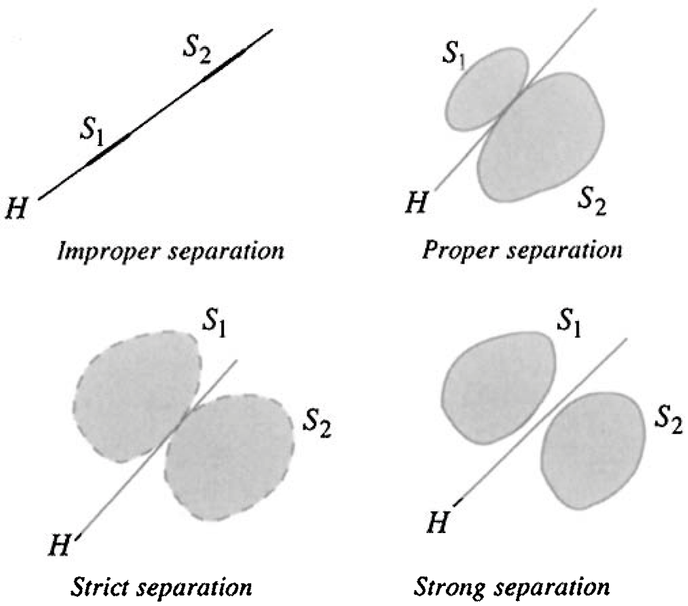

Figura 6: Varios tipos de separación.

Separación de un Conjunto Convexo y un Punto

Teorema 3 (Teorema de Separación) *Sea $\mathcal{S} \neq \emptyset$ cerrado y convexo en $\mathbb{R}^n$ e $\mathbf{y} \notin \mathcal{S}$. Luego, existe un vector $\mathbf{p} \neq \mathbf{0}$ y un escalar α tal que $\mathbf{p}^\mathsf{T}\mathbf{y} > \alpha$ y $\mathbf{p}^\mathsf{T}\mathbf{x} \leq \alpha$ para todo $\mathbf{x} \in \mathcal{S}$.*

Demostración. El conjunto $\mathcal{S} \neq \emptyset$ es cerrado y convexo en $\mathbb{R}^n$ e $\mathbf{y} \notin \mathcal{S}$. Luego, existe un único mínimo $\bar{\mathbf{x}} \in \mathcal{S}$ tal que $(\mathbf{x} - \bar{\mathbf{x}})^\mathsf{T}(\mathbf{y} - \bar{\mathbf{x}}) \leq 0$ para todo $\mathbf{x} \in \mathcal{S}$. Tomando $\mathbf{p} = \mathbf{y} - \bar{\mathbf{x}} \neq 0$ y $\alpha = \bar{\mathbf{x}}^\mathsf{T}(\mathbf{y} - \bar{\mathbf{x}}) = \mathbf{p}^\mathsf{T}\bar{\mathbf{x}}$, tenemos que $\mathbf{p}^\mathsf{T}\mathbf{x} \leq \alpha$ para todo $\mathbf{x} \in \mathcal{S}$. Además, $\mathbf{p}^\mathsf{T}\mathbf{y} - \alpha = (\mathbf{y} - \bar{\mathbf{x}})^\mathsf{T}(\mathbf{y} - \bar{\mathbf{x}}) = \|\mathbf{y} - \bar{\mathbf{x}}\|^2 > 0$. □

Teorema de Farkas

Teorema 4 (Teorema de Farkas) *Sea $\mathsf{A} \in \mathbb{R}^{m \times n}$ y $\mathbf{c} \in \mathbb{R}^n$. Luego, exactamente uno de los sistemas siguientes tiene solución:*

$$\text{S1:} \quad \mathsf{A}\mathbf{x} \leq \mathbf{0} \ \text{ y } \ \mathbf{c}^\mathsf{T}\mathbf{x} > 0 \ \text{ para algún } \mathbf{x} \in \mathbb{R}^n.$$
$$\text{S2:} \quad \mathsf{A}^\mathsf{T}\mathbf{y} = \mathbf{c} \ \text{ e } \ \mathbf{y} \geq \mathbf{0} \ \text{ para algún } \mathbf{y} \in \mathbb{R}^m.$$

Demostración. Ver demostración en Teorema 2.4.5 de Bazaraa. □

Interpretación del Teorema de Farkas: Sean $\mathbf{a}_1, \ldots, \mathbf{a}_m$ las columnas de A^T. Luego

- S2 tiene solución si $\mathbf{c}$ pertenece al cono convexo generado por $\mathbf{a}_1, \ldots, \mathbf{a}_m$.
- S1 tiene solución si el cono convexo cerrado $\{\mathbf{x} : \mathsf{A}\mathbf{x} \leq \mathbf{0}\}$ intersecta con el hemiespacio abierto $\{\mathbf{x} : \mathbf{c}^\mathsf{T}\mathbf{x} > 0\}$.

Estos dos casos se ilustran geométricamente en la Figura 7.

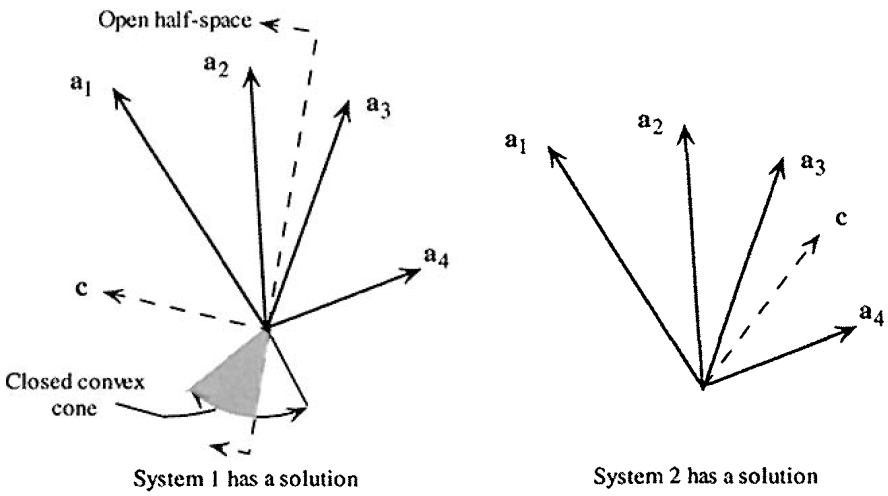

Figura 7: Teorema de Farkas.

Corolario 1 (Teorema de Gordan) *Sea* $\mathsf{A} \in \mathbb{R}^{m \times n}$. *Exactamente uno de los sistemas siguientes tiene solución:*

$$\text{S1:} \quad \mathsf{A}\mathbf{x} < \mathbf{0} \quad \text{para algún } \mathbf{x} \in \mathbb{R}^n.$$

$$\text{S2:} \quad \mathsf{A}^\mathsf{T}\mathbf{y} = \mathbf{0} \quad \text{e} \quad \mathbf{y} \geq \mathbf{0} \quad \text{para algún } \mathbf{y} \neq \mathbf{0},\ \mathbf{y} \in \mathbb{R}^m.$$

Demostración. Ver demostración en Corolario 1 de Bazaraa. □

Soporte de Conjuntos en Puntos Frontera

Definición. Sea $\mathcal{S} \neq \emptyset$ un conjunto en $\mathbb{R}^n$, y sea $\bar{\mathbf{x}} \in \partial\mathcal{S}$. $H = \{\mathbf{x} : \mathbf{p}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}) = \mathbf{0}\}$ es un *hiperplano soporte* de $\mathcal{S}$ en $\bar{\mathbf{x}}$ si:

$$\mathcal{S} \subseteq H^+, \quad \text{es decir,} \quad \mathbf{p}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}) \geq 0, \quad \forall\, \mathbf{x} \in \mathcal{S}.$$

o bien

$$\mathcal{S} \subseteq H^-, \quad \text{es decir,} \quad \mathbf{p}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}) \leq 0, \quad \forall\, \mathbf{x} \in \mathcal{S}.$$

donde H^+ es el hemiespacio positivo y H^- es el hemiespacio negativo. Si además $\mathcal{S} \nsubseteq H$, luego H es un *hiperplano soporte propio* de $\mathcal{S}$ en $\bar{\mathbf{x}}$.

La Figura 8 muestra algunos ejemplos de hiperplanos soporte. Se muestra el caso de un hiperplano soporte único en un punto frontera, un número infinito de hiperplanos soporte en un punto frontera, y un hiperplano soporte en más de un punto.

El siguiente teorema establece que un conjunto convexo tiene un hiperplano soporte en cada punto de su frontera.

Teorema 5 *Sea* $\mathcal{S} \neq \emptyset$ *un conjunto convexo en* $\mathbb{R}^n$, *y sea* $\bar{\mathbf{x}} \in \partial\mathcal{S}$. *Luego, existe un hiperplano que soporta* $\mathcal{S}$ *en* $\bar{\mathbf{x}}$, *es decir, existe* $\mathbf{p} \neq \mathbf{0}$ *tal que* $\mathbf{p}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}) \leq 0$ *para todo* $\mathbf{x} \in \mathrm{cl}(\mathcal{S})$.

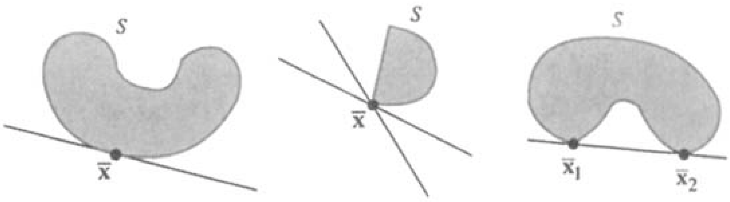

Figura 8: Hiperplanos soporte.

Demostración. Ver demostración en Teorema 2.4.7 de Bazaraa. □

Corolario 2 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$ y $\bar{\mathbf{x}} \notin \operatorname{int} \mathcal{S}$. Luego existe un vector $\mathbf{p} \neq \mathbf{0}$ tal que $\mathbf{p}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}) \leq 0$ para todo $\mathbf{x} \in \operatorname{cl} \mathcal{S}$.*

Demostración. Si $\bar{\mathbf{x}} \in \operatorname{cl} \mathcal{S}$, el corolario es cierto por el Teorema 3. Por otro lado, si $\bar{\mathbf{x}} \in \partial\mathcal{S}$, el corolario se reduce al Teorema 5. □

Separación de Dos Conjuntos Convexos Disjuntos

Dos conjuntos convexos disjuntos pueden ser separados por un hiperplano H tal que uno de los conjuntos pertenezca a H^+ y el otro pertenezca a H^-. De hecho, este resultado es válido aún si los conjuntos comparten algunos puntos, siempre y cuando sus interiores sean disjuntos.

Teorema 6 *Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ conjuntos convexos disjuntos no vacios en $\mathbb{R}^n$, tales que $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$. Luego, existe un hiperplano que separa $\mathcal{S}_1$ y $\mathcal{S}_2$, es decir, existe $\mathbf{p} \in \mathbb{R}^n$, $\mathbf{p} \neq \mathbf{0}$, tal que*

$$\inf\{\mathbf{p}^\mathsf{T}\mathbf{x} : \ \mathbf{x} \in \mathcal{S}_1\} \geq \sup\{\mathbf{p}^\mathsf{T}\mathbf{x} : \ \mathbf{x} \in \mathcal{S}_2\}$$

Demostración. Sea $\mathcal{S} = \mathcal{S}_1 \ominus \mathcal{S}_2 = \{\mathbf{x}_1 - \mathbf{x}_2 : \ \mathbf{x}_1 \in \mathcal{S}_1; \ \mathbf{x}_2 \in \mathcal{S}_2\}$. Notar que $\mathcal{S}$ es un conjunto convexo. Además, $\mathbf{0} \notin \mathcal{S}$ porque de lo contrario $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$. Por el Corolario 2 del Teorema 5, existe $\mathbf{p} \neq \mathbf{0}$ tal que $\mathbf{p}^\mathsf{T}\mathbf{x} \geq 0$ para todo $\mathbf{x} \in \mathcal{S}$. Esto significa que $\mathbf{p}^\mathsf{T}\mathbf{x}_1 \geq \mathbf{p}^\mathsf{T}\mathbf{x}_2$ para todo $\mathbf{x}_1 \in \mathcal{S}_1$ y $\mathbf{x}_2 \in \mathcal{S}_2$. □

Definición 1 (Conos) *$\mathcal{C} \neq \emptyset$ en $\mathbb{R}^n$ es un cono con vértice en cero si $\bar{\mathbf{x}} \in \mathcal{C}$ implica que $\lambda\bar{\mathbf{x}} \in \mathcal{C}$ para todo $\lambda \geq 0$. Si además $\mathcal{C}$ es convexo, entonces $\mathcal{C}$ es un cono convexo. Ver Figura 9.*

Definición 2 (Punto Extremo) *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$. Un vector $\mathbf{x} \in \mathcal{S}$ es un punto extremo de $\mathcal{S}$ si $\mathbf{x} = \alpha\mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2$ con $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ y $\alpha \in (0, 1)$ implica que $\mathbf{x} = \mathbf{x}_1 = \mathbf{x}_2$.*

2. Funciones Convexas

2.1. Definición de Funciones Convexas y Propiedades

Definición. Sea $f : \ \mathcal{S} \to \mathbb{R}$, $\mathcal{S} \neq \emptyset$ convexo en $\mathbb{R}^n$. La función f es *convexa* en $\mathcal{S}$ si

$$f\left(\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2\right) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2), \quad \forall\ \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}, \quad \text{y} \quad \forall\ \lambda \in (0, 1) \tag{4}$$

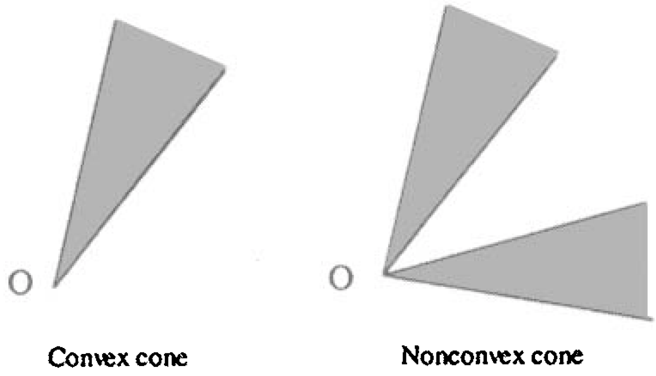

Figura 9: Conos.

Se dice que f es *estrictamente convexa* en $\mathcal{S}$ si

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) < \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2), \quad \forall\, \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S},\ \mathbf{x}_1 \neq \mathbf{x}_2, \quad \text{y} \quad \forall\, \lambda \in (0,1)$$

Se dice que f es (estrictamente) *cóncava* en $\mathcal{S}$ si $-f$ es (estrictamente) convexa en $\mathcal{S}$.

La Figura 10 illustra ejemplos de una función convexa, una función cóncava, y una función que no es ni convexa ni cóncava.

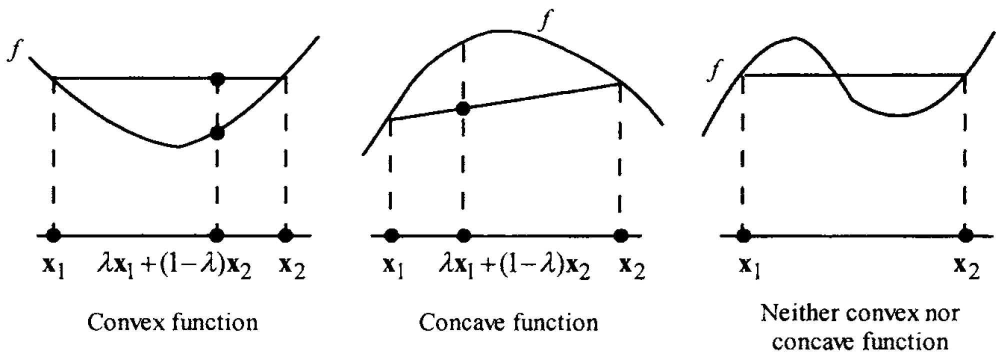

Figura 10: Función convexa, función cóncava, y función que no es convexa ni cóncava.

Algunas funciones convexas importantes:

1) Sean $f_1, f_2, \ldots, f_k :\ \mathbb{R}^n \to \mathbb{R}$ funciones convexas. Luego:

a) $f(\mathbf{x}) = \sum_{j=1}^{k} \alpha_j f_j(\mathbf{x})$, con $\alpha_j > 0$ para $j = 1, \ldots, k$, es una función convexa.

b) $f(\mathbf{x}) = \text{máx}\,\{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\}$ es una función convexa.

2) Sea $g : \mathbb{R}^n \to \mathbb{R}$ una función convexa. Sea $\mathcal{S} = \{\mathbf{x} :\ g(\mathbf{x}) > 0\}$. Defina $f : \mathcal{S} \to \mathbb{R}$ como $f(\mathbf{x}) = \frac{1}{g(\mathbf{x})}$. Luego f es convexa en $\mathcal{S}$.

3) Sea $g : \mathbb{R} \to \mathbb{R}$ una función no decreciente, univariada y convexa, y sea $h : \mathbb{R}^n \to \mathbb{R}$ una función convexa. Luego, la función compuesta $f(\mathbf{x}) = g(h(\mathbf{x}))$ es convexa.

Conjunto de nivel de $f(\mathbf{x})$**:** Sea $f : \mathbb{R}^n \to \mathbb{R}$ convexa. Luego $\mathcal{S}_\alpha = \{\mathbf{x} \in \mathcal{S} :\ f(\mathbf{x}) \leq \alpha\}$ es un conjunto convexo.

Lema 2 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa. Luego, el conjunto de nivel $\mathcal{S}_\alpha = \{\mathbf{x} \in \mathcal{S} :\ f(\mathbf{x}) \leq \alpha\}$, donde $\alpha \in \mathbb{R}$, es un conjunto convexo.*

Demostración. Sean $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}_\alpha$. Luego $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ y $f(\mathbf{x}_1) \leq \alpha$, $f(\mathbf{x}_2) \leq \alpha$. Sea $\lambda \in (0,1)$ y $\mathbf{x} = \lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$. Por convexidad de $\mathcal{S}$, $\mathbf{x} \in \mathcal{S}$. Por convexidad de f,

$$f(\mathbf{x}) \leq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) \leq \lambda\alpha + (1-\lambda)\alpha = \alpha$$

Luego, $\mathbf{x} \in \mathcal{S}_\alpha$ y $\mathcal{S}_\alpha$ es convexo. □

Continuidad de Funciones Convexas

Las funciones cóncavas y convexas son contínuas en el interior de su dominio.

Teorema 7 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa. Luego, f es continua en* $\mathrm{int}(\mathcal{S})$.

Demostración. La demostración se encuentra en el Teorema 3.1.3 de Bazaraa. □

Derivada Direccional de Funciones Convexas

Definición 3 *Sea $\mathcal{S} \neq \emptyset$ un conjunto en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$. Sea $\bar{\mathbf{x}} \in \mathcal{S}$ y el vector $\mathbf{d} \neq \mathbf{0}$ tal que $\bar{\mathbf{x}} + \lambda\mathbf{d} \in \mathcal{S}$ para $\lambda > 0$ suficientemente pequeño. Luego, la derivada direccional de f en la dirección de $\mathbf{d}$, en caso de existir, está dada por:*

$$f'(\bar{\mathbf{x}}, \mathbf{d}) = \nabla_{\mathbf{d}} f(\bar{\mathbf{x}}) = \lim_{\lambda \to 0^+} \frac{f(\bar{\mathbf{x}} + \lambda\mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda}$$

En particular, el límite de la Definición 3 existe para funciones cóncavas y convexas definidas globalmente. Si $f : \mathcal{S} \to \mathbb{R}$ es convexa en $\mathcal{S}$, el límite existe si $\bar{\mathbf{x}} \in \mathrm{int}(\mathcal{S})$, pero puede no existir si $\bar{\mathbf{x}} \in \partial\mathcal{S}$.

2.2. Epígrafo e Hipógrafo de una Función

Una función f en $\mathcal{S}$ puede describirse por medio del conjunto:

$$\{[\mathbf{x}, f(\mathbf{x})] : \ \mathbf{x} \in \mathcal{S}\} \subset \mathbb{R}^{n+1}$$

que se conoce como *grafo* de una función.

Definición 4 (Epígrafo) *Sea $\mathcal{S} \neq \emptyset$ un conjunto en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$. El epígrafo de f, denotado* $\mathrm{epi}f$ *es un subconjunto de $\mathbb{R}^{n+1}$ definido como*

$$\{(\mathbf{x}, y) : \ \mathbf{x} \in \mathcal{S} : \ y \in \mathbb{R}, \ y \geq f(\mathbf{x})\}$$

El hipógrafo de f, denotado $\mathrm{hyp}f$ *es un subconjunto de $\mathbb{R}^{n+1}$ dado por*

$$\{(\mathbf{x}, y) : \ \mathbf{x} \in \mathcal{S} : \ y \in \mathbb{R}, \ y \leq f(\mathbf{x})\}$$

La Figura 11 ilustra el epígrafo y el hipógrafo de varias funciones. En la Figura 11a, ni el epígrafo ni el hipógrafo de f son conjuntos convexos. En la Figura 11b, el epígrafo de f es un conjunto convexo, mientras que en la Figura 11c, el hipógrafo de f es convexo.

Una función resulta ser convexa si y solo si su epígrafo es un conjunto convexo, y en forma equivalente, una función es cóncava si y solo si su hipógrafo es un conjunto convexo.

Teorema 8 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$. Luego, f es convexa si y solo si* $\mathrm{epi}f$ *es un conjunto convexo.*

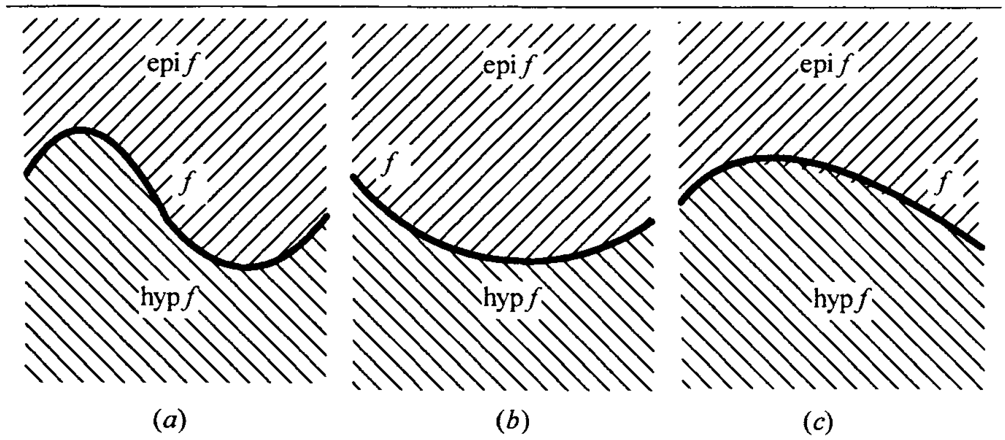

Figura 11: Epígrafos e Hipógrafos.

Demostración. Suponga que f es convexa, y sea $(\mathbf{x}_1, y_1)$ y $(\mathbf{x}_2, y_2) \in \text{epi}f$, es decir, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$, $y_1 \geq f(\mathbf{x}_1)$ e $y_2 \geq f(\mathbf{x}_2)$. Sea $\lambda \in (0,1)$. Luego

$$\lambda y_1 + (1-\lambda) y_2 \geq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) \geq f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2),$$

donde la última desigualdad vale por la convexidad de f. Notar que $\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in \mathcal{S}$. Luego, $[\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2, \lambda y_1 + (1-\lambda) y_2] \in \text{epi}f$, y por lo tanto epi$f$ es convexo.

A la inversa, suponga que epi$f$ es convexo, y sean $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$. Luego $[\mathbf{x}_1, f(\mathbf{x}_1)]$ y $[\mathbf{x}_2, f(\mathbf{x}_2)]$ pertenecen a epi$f$, y por la convexidad de epi$f$, debemos tener

$$[\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2,\ \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2)] \in \text{epi}f, \quad \text{para } \lambda \in (0,1).$$

En otras palabras, $\lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) \geq f[\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2]$ para todo $\lambda \in (0,1)$, es decir, f es convexa. Esto completa la demostración. $\square$

Dado que el epígrafo de una función convexa, y el hipógrafo de una función cóncava son conjuntos convexos, tienen hiperplanos soporte en sus puntos frontera.

2.3. Subgradientes

Definición 5 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa. Luego $\boldsymbol{\xi}$ es un subgradiente de f en $\bar{\mathbf{x}} \in \mathcal{S}$ si*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \boldsymbol{\xi}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall\ \mathbf{x} \in \mathcal{S}$$

Similarmente, sea $f : \mathcal{S} \to \mathbb{R}$ una función cóncava. Luego $\boldsymbol{\xi}$ es un subgradiente de f en $\bar{\mathbf{x}} \in \mathcal{S}$ si

$$f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) + \boldsymbol{\xi}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall\ \mathbf{x} \in \mathcal{S}$$

La Figura 12 muestra ejemplos de subgradientes de funciones convexas y cóncavas.

En la figura vemos que $f(\bar{\mathbf{x}}) + \boldsymbol{\xi}^\mathsf{T}(\mathbf{x} - \bar{\mathbf{x}})$ corresponde a un hiperplano soporte del epígrafo (o hipógrafo) de f. El subgradiente ξ es la pendiente del hiperplano soporte.

Más precisamente, el hiperplano soporte del epígrafo (o hipógrafo) de f está dado por:

$$\mathbf{p}^\mathsf{T} \left[\begin{pmatrix} \mathbf{x} \\ y \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{x}} \\ f(\bar{\mathbf{x}}) \end{pmatrix} \right] = 0, \quad \text{con } \mathbf{p} = \begin{pmatrix} \boldsymbol{\xi} \\ -1 \end{pmatrix}$$

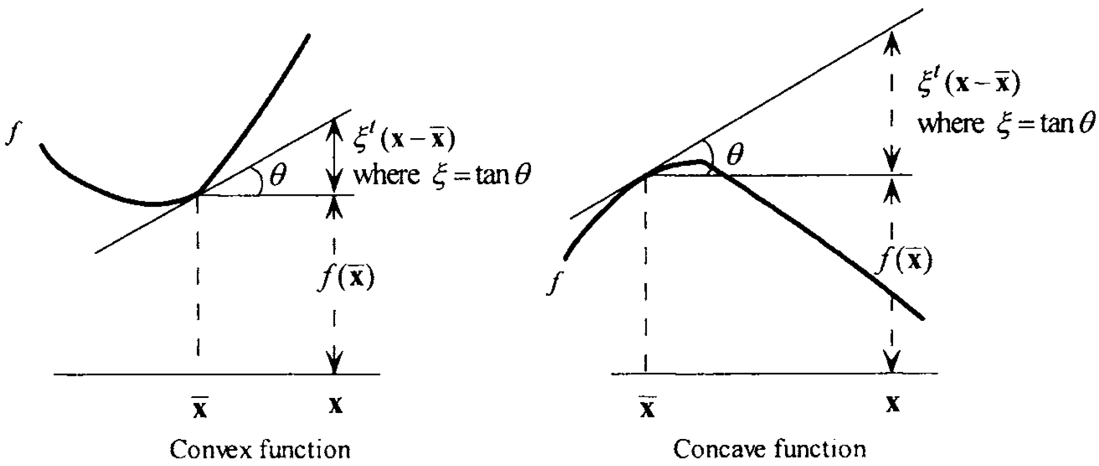

Figura 12: Interpretación geométrica de subgradientes.

Reemplazando tenemos

$$(\boldsymbol{\xi}^{\mathsf{T}} \ \ -1)\begin{pmatrix} \mathbf{x} \\ y \end{pmatrix} - (\boldsymbol{\xi}^{\mathsf{T}} \ \ -1)\begin{pmatrix} \bar{\mathbf{x}} \\ f(\bar{\mathbf{x}}) \end{pmatrix} = 0,$$

de donde obtenemos

$$y = f(\bar{\mathbf{x}}) + \boldsymbol{\xi}^{\mathsf{T}}(\mathbf{x} - \bar{\mathbf{x}})$$

Ejemplo 2 *Sea $f(x) = \text{mín}\{f_1(x), f_2(x)\}$, donde f_1 y f_2 se definen como:*

$$\begin{array}{ll} f_1(x) = 4 - |x|, & x \in \mathbb{R} \\ f_2(x) = 4 - (x-2)^2, & x \in \mathbb{R} \end{array}$$

Debido a que $f_2(x) \geq f_1(x)$ para $1 \leq x \leq 4$, f se puede representar como:

$$f(x) = \begin{cases} 4 - x, & x \in [1,4] \\ 4 - (x-2)^2, & x \notin [1,4] \end{cases}$$

La función cóncava f se muestra en la Figura 13 en líneas negras gruesas. Tenemos que

$$\xi = \begin{cases} -1, & x \in (1,4) \\ -2(x-2), & x < 1, \ x > 4 \end{cases}$$

En $x = 1$ y en $x = 4$ los subgradientes no son únicos ya que existen muchos hiperplanos soporte.

$$\begin{cases} \xi \in [-1, 2], & x = 1 \\ \xi \in [-4, -1], & x = 4 \end{cases}$$

El siguiente teorema muestra que toda función convexa o cóncava tiene al menos un subgradiente en el interior de su dominio.

Teorema 9 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa. Luego para $\bar{\mathbf{x}} \in \text{int}\mathcal{S}$, existe un vector $\boldsymbol{\xi}$ tal que el hiperplano*

$$H = \{(\mathbf{x}, y) : y = f(\bar{\mathbf{x}}) + \boldsymbol{\xi}^{\mathsf{T}}(\mathbf{x} - \bar{\mathbf{x}})\}$$

soporta epi$f$ *en $[\bar{\mathbf{x}}, f(\bar{\mathbf{x}})]$. En particular,*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \boldsymbol{\xi}^{\mathsf{T}}(\mathbf{x} - \bar{\mathbf{x}}), \quad \text{para todo } \mathbf{x} \in \mathcal{S},$$

es decir, $\boldsymbol{\xi}$ es un subgradiente de f en $\bar{\mathbf{x}}$.

Demostración. Ver demostración del Teorema 3.2.5 de Bazaraa. □

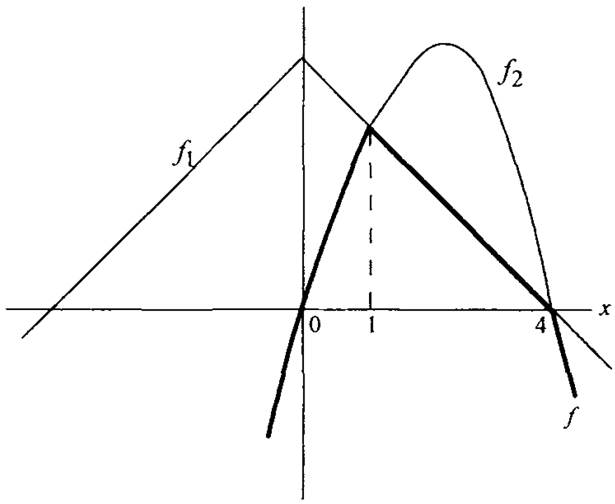

Figura 13: Función del Ejemplo 2.

2.4. Funciones Convexas Diferenciables

Lema 3 *Si f es diferenciable en $\bar{\mathbf{x}}$, luego la derivada direccional existe a lo largo de cualquier vector $\mathbf{v}$, y se tiene*

$$\nabla_{\mathbf{v}} f(\bar{\mathbf{x}}) = \lim_{\lambda \to 0^+} \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{v}) - f(\bar{\mathbf{x}})}{\lambda} = \nabla f(\bar{\mathbf{x}})\mathbf{v}$$

donde $\nabla f(\bar{\mathbf{x}}) \in \mathbb{R}^{1\times n}$ es el gradiente de f en $\bar{\mathbf{x}}$.

Una función convexa diferenciable tiene un único subgradiente $\boldsymbol{\xi}^\mathsf{T} = \nabla f(\mathbf{x})$.

Lema 4 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa. Si f es diferenciable en $\bar{\mathbf{x}} \in \operatorname{int}\mathcal{S}$, luego $\nabla f(\mathbf{x})$ es el único subgradiente de f en $\bar{\mathbf{x}}$.*

Demostración. Ver demostración en Lema 3.3.2 de Bazaraa. □

Teorema 10 (Condición de primer orden) *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo, y sea $f \in C^1$, $f : \mathcal{S} \to \mathbb{R}$. Luego, f es convexa en $\mathcal{S}$ si y solo si para todo $\bar{\mathbf{x}} \in \mathcal{S}$ se cumple:*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall\ \mathbf{x} \in \mathcal{S}$$

Demostración. Primero suponga que f es convexa. Luego para todo $\lambda \in [0, 1]$ tenemos

$$f(\lambda\mathbf{x} + (1-\lambda)\bar{\mathbf{x}}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\bar{\mathbf{x}})$$

Luego, para $0 < \lambda \leq 1$ tenemos

$$\frac{f(\bar{\mathbf{x}} + \lambda(\mathbf{x} - \bar{\mathbf{x}})) - f(\bar{\mathbf{x}})}{\lambda} \leq f(\mathbf{x}) - f(\bar{\mathbf{x}})$$

Tomando el límite cuando $\lambda \to 0^+$ obtenemos

$$\nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) \leq f(\mathbf{x}) - f(\bar{\mathbf{x}})$$

A la inversa, suponga que

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall\ \bar{\mathbf{x}}, \mathbf{x} \in \mathcal{S} \tag{5}$$

Seleccionando $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ y $\lambda \in [0,1]$ con $\bar{\mathbf{x}} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ y alternativamente $\mathbf{x} = \mathbf{x}_1$ o $\mathbf{x} = \mathbf{x}_2$, tenemos:

$$f(\mathbf{x}_1) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x}_1 - \bar{\mathbf{x}}) \tag{6}$$
$$f(\mathbf{x}_2) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x}_2 - \bar{\mathbf{x}}) \tag{7}$$

Multiplicando (6) por λ y (7) por $(1-\lambda)$ y sumando:

$$\lambda f(\mathbf{x}_1) + (1-\lambda)(\mathbf{x}_2) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})\left(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 - \bar{\mathbf{x}}\right)$$

Sustituyendo $\bar{\mathbf{x}} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ obenemos

$$\lambda f(\mathbf{x}_1) + (1-\lambda)(\mathbf{x}_2) \geq f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)$$

lo cual concluye la demostración. □

El enunciado del Teorema 10 se ilustra en la Figura 14. La función afin $f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})$ es una cota inferior de $f(\mathbf{x})$ para cualquier $\bar{x} \in \mathcal{S}$.

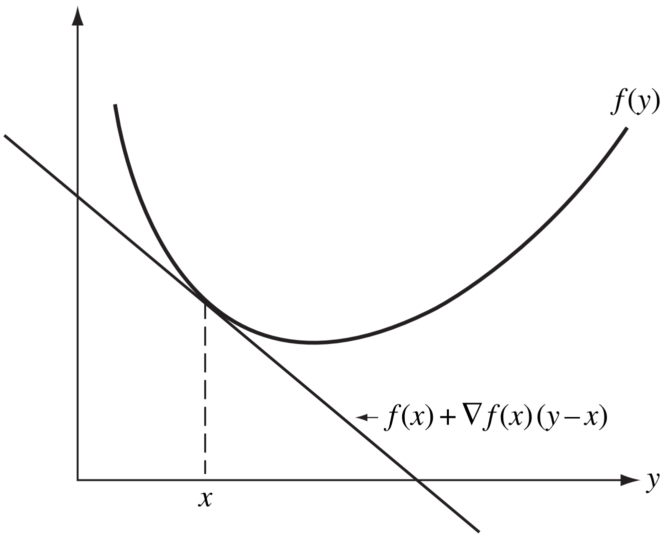

Figura 14: Ilustración del Teorema 10.

Teorema 11 (Condición de segundo orden) *Sea $\mathcal{S} \neq \emptyset$ un conjunto abierto y convexo en $\mathbb{R}^n$, y sea $f \in C^2$, $f : \mathcal{S} \to \mathbb{R}$. Luego, f es convexa en $\mathcal{S}$ si y solo si la matriz hessiana de f es positiva semidefinida (PSD) para todo $\mathbf{x} \in \mathcal{S}$.*

Demostración. Suponga que f es convexa, y sea $\bar{\mathbf{x}} \in \mathcal{S}$. Debemos mostrar que $\mathbf{x}^\mathsf{T}[\nabla^2 f(\bar{\mathbf{x}})]\mathbf{x} \geq 0, \forall\, \mathbf{x} \in \mathbb{R}^n$. $\mathcal{S}$ es abierto, luego para cualquier $\mathbf{x} \in \mathbb{R}^n$, $\bar{\mathbf{x}} + \lambda\mathbf{x} \in \mathcal{S}$ para $|\lambda| \neq 0$ y suficientemente pequeño. Luego, por el Teorema 10 y dado que $f \in C^2$, obtenemos las siguientes dos expresiones:

$$f(\bar{\mathbf{x}} + \lambda\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \lambda\nabla f(\bar{x})\mathbf{x} \tag{8}$$
$$f(\bar{\mathbf{x}} + \lambda\mathbf{x}) = f(\bar{\mathbf{x}}) + \lambda\nabla f(\bar{x})\mathbf{x} + \frac{1}{2}\lambda^2\mathbf{x}^\mathsf{T}[\nabla^2 f(\bar{\mathbf{x}})]\mathbf{x} + o(\lambda^2\|\mathbf{x}\|^2) \tag{9}$$

Restando (9) de (8):

$$\frac{1}{2}\lambda^2\mathbf{x}^\mathsf{T}[\nabla^2 f(\bar{\mathbf{x}})]\mathbf{x} + o(\lambda^2\|\mathbf{x}\|^2) \geq 0$$

Dividiendo por λ^2 y para $\lambda \to 0$, obtenemos:

$$\frac{1}{2}\lambda^2 \mathbf{x}^\mathsf{T}[\nabla^2 f(\bar{\mathbf{x}})]\mathbf{x} \geq 0$$

A la inversa, supongamos ahora que $\nabla^2 f(\bar{\mathbf{x}})$ es positiva semidefinida para todo $\mathbf{x} \in \mathcal{S}$. Considere $\mathbf{x}, \bar{\mathbf{x}} \in \mathcal{S}$. Por el teorema del valor medio:

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^\mathsf{T}[\nabla^2 f(\hat{\mathbf{x}})](\mathbf{x} - \bar{\mathbf{x}}),$$

donde $\hat{\mathbf{x}} = \lambda\bar{\mathbf{x}} + (1 - \lambda)\mathbf{x}$ para algún $\lambda \in (0, 1)$. Notar que $\hat{\mathbf{x}} \in \mathcal{S}$. Luego, por hipótesis, $(\mathbf{x} - \bar{\mathbf{x}})^\mathsf{T}[\nabla^2 f(\hat{\mathbf{x}})](\mathbf{x} - \bar{\mathbf{x}}) \geq 0$, de donde concluimos que

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall\, \mathbf{x}, \bar{\mathbf{x}} \in \mathcal{S}.$$

Luego, se cumple la condición de primer orden del Teorema 10 y f es convexa. □

El Teorema 11 es útil para verificar la convexidad o concavidad de una función dos veces diferenciable. En particular, si la función es cuadrática, luego la matriz hessiana es independiente del punto $\mathbf{x}$. En este caso, verificar la convexidad se reduce a verificar que una matriz constante sea positiva semidefinida.

Teorema 12 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f \in C^2$, $f : \mathcal{S} \to \mathbb{R}$. Si $\nabla^2 f(\mathbf{x})$ es definida positiva $\forall\, \mathbf{x} \in \mathcal{S}$, luego f es estrictamente convexa. Por el contrario, si f es estrictamente convexa, luego $\nabla^2 f(\mathbf{x})$ es definida semipositiva $\forall\, \mathbf{x} \in \mathcal{S}$. Sin embargo, si f es estrictamente convexa y cuadrática, luego $\nabla^2 f(\mathbf{x})$ es definida positiva $\forall\, \mathbf{x} \in \mathcal{S}$.*

Demostración. Ver Teorema 3.3.8 de Bazaraa. □

Ejemplo 3 *Considere la función $f(x) = x^4$. Tenemos que $\nabla^2 f(x) = 12x^2$ es definida positiva para todo $x \neq 0$ pero es positiva semidefinida para $x = 0$. Notar que $f(x) = x^4$ es estrictamente convexa a pesar de que $\nabla^2 f(0) = 0$.*

3. Minimización y Maximización de Funciones Convexas

Los problemas de optimización

$$\min_{\mathbf{x}\in\mathcal{S}} \; f(\mathbf{x})$$

donde f es una función convexa y $\mathcal{S}$ es un conjunto convexo, se conocen como *problemas de optimización convexos.*

El principal resultado para problemas de optimización convexos es que cada mínimo local es también un mínimo global.

Teorema 13 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa en $\mathcal{S}$. Considerar el problema $\min_{\mathbf{x}\in\mathcal{S}} \; f(\mathbf{x})$. Sea $\mathbf{x}^\star \in \mathcal{S}$ una solución óptima local.*

1) *Luego, $\mathbf{x}^\star$ es una solución óptima global.*

2) *Si $\mathbf{x}^\star$ es un mínimo local estricto, o si f es estrictamente convexa, luego $\mathbf{x}^\star$ es la única solución óptima global.*

Demostración. Demostraremos el item (1). Dado que $\mathbf{x}^\star$ es un mínimo local, luego existe un entorno-ε, $B_\varepsilon(\mathbf{x}^\star)$, tal que

$$f(\mathbf{x}) \geq f(\mathbf{x}^\star), \quad \forall\, \mathbf{x} \in \mathcal{S} \cap B_\varepsilon(\mathbf{x}^\star)$$

Por contradicción, suponga que $\mathbf{x}^\star$ no es un óptimo global. Luego, existe $f(\hat{\mathbf{x}}) < f(\mathbf{x}^\star)$ para algún $\hat{\mathbf{x}} \in \mathcal{S}$. Por la convexidad de f la siguiente desigualdad se cumple para todo $\lambda \in [0, 1]$:

$$f(\lambda\hat{\mathbf{x}} + (1-\lambda)\mathbf{x}^\star) \leq \lambda f(\hat{\mathbf{x}}) + (1-\lambda) f(\mathbf{x}^\star) < \lambda f(\mathbf{x}^\star) + (1-\lambda) f(\mathbf{x}^\star) = f(\mathbf{x}^\star)$$

Pero para $\lambda > 0$ suficientemente pequeño, $\lambda\hat{\mathbf{x}} + (1-\lambda)\mathbf{x}^\star \in \mathcal{S} \cap B_\varepsilon(\mathbf{x}^\star)$, por lo que la desigualdad anterior resulta en una contradicción. □

Teorema 14 *Sea $\mathcal{S} \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y sea $f \in C^1$, $f : \mathcal{S} \to \mathbb{R}$ una función convexa en $\mathcal{S}$. Sea $\mathbf{x}^\star \in \mathcal{S}$ tal que para todo $\mathbf{x} \in \mathcal{S}$:*

$$\nabla f(\mathbf{x}^\star)(\mathbf{x} - \mathbf{x}^\star) \geq 0,$$

luego $\mathbf{x}^\star$ es un mínimo global de f en $\mathcal{S}$.

Demostración. Dado que $\mathbf{x} - \mathbf{x}^\star$ es una *dirección factible* en $\mathbf{x}^\star$, la condición $\nabla f(\mathbf{x}^\star)(\mathbf{x} - \mathbf{x}^\star) \geq 0$ es equivalente a una *condición necesaria de primer orden* para un mínimo. Por otro lado, la condición de primer orden para funciones convexas establece que

$$f(\mathbf{x}) \geq f(\mathbf{x}^\star) + \nabla f(\mathbf{x}^\star)(\mathbf{x} - \mathbf{x}^\star) \geq f(\mathbf{x}^\star)$$

lo cual concluye la demostración. □

La maximización de una función convexa en un conjunto convexo suele presentar numerosos máximos locales. Sin embargo, es posible demostrar un resultado importante.

Teorema 15 *Sea $\mathcal{S} \neq \emptyset$ un conjunto compacto y convexo en $\mathbb{R}^n$, y sea $f : \mathcal{S} \to \mathbb{R}$ una función convexa en $\mathcal{S}$. Luego, si f tiene un máximo en $\mathcal{S}$, éste ocurre en un punto extremo de $\mathcal{S}$.*

Demostración. Ver Luenberger p 181, Teorema 3. □