Tema 4

Introducción a la programación no lineal

4.1. Matrices semidefinidas positivas y definidas positivas

Definición 4.1. *Sea $\boldsymbol{A}$ una matriz simétrica $n \times n$.*

- $\boldsymbol{A}$ *es* ***semidefinida positiva*** *si* $\boldsymbol{x}^t \boldsymbol{A} \boldsymbol{x} \geq 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n$.
- $\boldsymbol{A}$ *es* ***definida positiva*** *si* $\boldsymbol{x}^t \boldsymbol{A} \boldsymbol{x} > 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$.

Lema 4.1. *Sea* $\boldsymbol{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. *Se verifican los siguientes resultados:*

1) $\boldsymbol{A}$ *es semidefinida positiva si y solo si* $a \geq 0$, $c \geq 0$ *y* $ac - b^2 \geq 0$.

2) $\boldsymbol{A}$ *es definida positiva si y solo si* $a > 0$, $c > 0$ *y* $ac - b^2 > 0$.

Teorema 4.1. *Sea $\boldsymbol{A}$ una matriz simétrica. Entonces $\boldsymbol{A}$ es definida positiva si y solo si es semidefinida positiva y no singular.*

Teorema 4.2. *Sea $\boldsymbol{A}$ una matriz simétrica con autovalores reales. Se verifican los siguientes resultados:*

1) $\boldsymbol{A}$ *es semidefinida positiva si y solo si todos sus autovalores son no negativos.*

2) $\boldsymbol{A}$ *es definida positiva si y solo si todos sus autovalores son positivos.*

4.2. Funciones convexas y cóncavas. Generalizaciones

4.2.1. Convexidad y concavidad en un conjunto

Definición 4.2. *Sean $S \subseteq \mathbb{R}^n$ un conjunto convexo no vacío y $f : S \to \mathbb{R}$.*

- f *es* ***convexa*** *en* S *si* $f(\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2}) \leq \lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2}) \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S, \quad \forall \lambda \in (0,1)$.
- f *es* ***estrictamente convexa*** *en* S *si* $f(\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2}) < \lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2}) \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ *tales que* $\boldsymbol{x_1} \neq \boldsymbol{x_2}, \quad \forall \lambda \in (0,1)$.
- f *es* ***cuasiconvexa*** *en* S *si* $f(\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2}) \leq \text{máx}\{f(\boldsymbol{x_1}), f(\boldsymbol{x_2})\} \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S, \quad \forall \lambda \in (0,1)$.
- f *es* ***estrictamente cuasiconvexa*** *en* S *si* $f(\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2}) < \text{máx}\{f(\boldsymbol{x_1}), f(\boldsymbol{x_2})\} \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ *tales que* $f(\boldsymbol{x_1}) \neq f(\boldsymbol{x_2}), \quad \forall \lambda \in (0,1)$.
- f *es* ***fuertemente cuasiconvexa*** *en* S *si* $f(\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2}) < \text{máx}\{f(\boldsymbol{x_1}), f(\boldsymbol{x_2})\} \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ *tales que* $\boldsymbol{x_1} \neq \boldsymbol{x_2}, \quad \forall \lambda \in (0,1)$.

Definición 4.3. *Sean $S \subseteq \mathbb{R}^n$ un conjunto abierto no vacío y $f : S \to \mathbb{R}$ diferenciable en S.*

- f *es* ***pseudoconvexa*** *en* S *si* $f(\boldsymbol{x_2}) \geq f(\boldsymbol{x_1}) \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ *tales que* $\boldsymbol{\nabla} f(\boldsymbol{x_1})^t(\boldsymbol{x_2} - \boldsymbol{x_1}) \geq 0$.
- f *es* ***estrictamente pseudoconvexa*** *en* S *si* $f(\boldsymbol{x_2}) > f(\boldsymbol{x_1}) \ \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ *tales que* $\boldsymbol{x_1} \neq \boldsymbol{x_2}$ *y* $\boldsymbol{\nabla} f(\boldsymbol{x_1})^t(\boldsymbol{x_2} - \boldsymbol{x_1}) \geq 0$.

Definición 4.4. *Sean $S \subseteq \mathbb{R}^n$ un conjunto convexo no vacío y $f : S \to \mathbb{R}$.*

- f *es* ***cóncava*** *en* S *si* $-f$ *es convexa en* S.
- f *es* ***estrictamente cóncava*** *en* S *si* $-f$ *es estrictamente convexa en* S.
- f *es* ***cuasicóncava*** *en* S *si* $-f$ *es cuasiconvexa en* S.
- f *es* ***estrictamente cuasicóncava*** *en* S *si* $-f$ *es estrictamente cuasiconvexa en* S.
- f *es* ***fuertemente cuasicóncava*** *en* S *si* $-f$ *es fuertemente cuasiconvexa en* S.

Definición 4.5. *Sean $S \subseteq I\!R^n$ un conjunto abierto no vacío y $f : S \to I\!R$ diferenciable en S.*

- *f es **pseudocóncava** en S si $-f$ es pseudoconvexa en S.*
- *f es **estrictamente pseudocóncava** en S si $-f$ es estrictamente pseudoconvexa en S.*

Los conceptos y resultados que se exponen de ahora en adelante están referidos a funciones convexas y sus generalizaciones, aunque también podrían enunciarse de forma análoga para funciones cóncavas y sus generalizaciones (véanse las definiciones 4.4 y 4.5).

4.2.2. Convexidad en un punto

Definición 4.6. *Sean $S \subseteq I\!R^n$ un conjunto convexo, $\overline{\boldsymbol{x}} \in S$ y $f : S \to I\!R$.*

- *f es **convexa** en $\overline{\boldsymbol{x}}$ si $f(\lambda\overline{\boldsymbol{x}} + (1-\lambda)\boldsymbol{x}) \leq \lambda f(\overline{\boldsymbol{x}}) + (1-\lambda)f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S, \quad \forall \lambda \in (0,1)$.*
- *f es **estrictamente convexa** en $\overline{\boldsymbol{x}}$ si $f(\lambda\overline{\boldsymbol{x}} + (1-\lambda)\boldsymbol{x}) < \lambda f(\overline{\boldsymbol{x}}) + (1-\lambda)f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S$ tal que $\boldsymbol{x} \neq \overline{\boldsymbol{x}}, \quad \forall \lambda \in (0,1)$.*
- *f es **cuasiconvexa** en $\overline{\boldsymbol{x}}$ si $f(\lambda\overline{\boldsymbol{x}} + (1-\lambda)\boldsymbol{x}) \leq \text{máx}\{f(\overline{\boldsymbol{x}}), f(\boldsymbol{x})\} \quad \forall \boldsymbol{x} \in S, \quad \forall \lambda \in (0,1)$.*
- *f es **estrictamente cuasiconvexa** en $\overline{\boldsymbol{x}}$ si $f(\lambda\overline{\boldsymbol{x}} + (1-\lambda)\boldsymbol{x}) < \text{máx}\{f(\overline{\boldsymbol{x}}), f(\boldsymbol{x})\} \quad \forall \boldsymbol{x} \in S$ tal que $f(\boldsymbol{x}) \neq f(\overline{\boldsymbol{x}}), \quad \forall \lambda \in (0,1)$.*
- *f es **fuertemente cuasiconvexa** en $\overline{\boldsymbol{x}}$ si $f(\lambda\overline{\boldsymbol{x}} + (1-\lambda)\boldsymbol{x}) < \text{máx}\{f(\overline{\boldsymbol{x}}), f(\boldsymbol{x})\} \quad \forall \boldsymbol{x} \in S$ tal que $\boldsymbol{x} \neq \overline{\boldsymbol{x}}, \quad \forall \lambda \in (0,1)$.*

Definición 4.7. *Sean $S \subseteq I\!R^n$, $\overline{\boldsymbol{x}} \in \mathring{S}$ y $f : S \to I\!R$ diferenciable en $\overline{\boldsymbol{x}}$.*

- *f es **pseudoconvexa** en $\overline{\boldsymbol{x}}$ si $f(\boldsymbol{x}) \geq f(\overline{\boldsymbol{x}}) \quad \forall \boldsymbol{x} \in S$ tal que $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})^t(\boldsymbol{x} - \overline{\boldsymbol{x}}) \geq 0$.*
- *f es **estrictamente pseudoconvexa** en $\overline{\boldsymbol{x}}$ si $f(\boldsymbol{x}) > f(\overline{\boldsymbol{x}}) \quad \forall \boldsymbol{x} \in S$ tal que $\boldsymbol{x} \neq \overline{\boldsymbol{x}}$ y $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})^t(\boldsymbol{x} - \overline{\boldsymbol{x}}) \geq 0$.*

4.2.3. Relaciones entre los distintos tipos de convexidad

- Convexidad estricta $\Rightarrow$ convexidad
- Convexidad $\Rightarrow$ cuasiconvexidad

- Convexidad $\Rightarrow$ cuasiconvexidad estricta
- Convexidad estricta $\Rightarrow$ cuasiconvexidad fuerte
- Cuasiconvexidad estricta + semicontinuidad inferior $\Rightarrow$ cuasiconvexidad
- Cuasiconvexidad fuerte $\Rightarrow$ cuasiconvexidad
- Cuasiconvexidad fuerte $\Rightarrow$ cuasiconvexidad estricta
- Convexidad + diferenciabilidad $\Rightarrow$ pseudoconvexidad
- Convexidad estricta + diferenciabilidad $\Rightarrow$ pseudoconvexidad estricta
- Pseudoconvexidad $\Rightarrow$ cuasiconvexidad
- Pseudoconvexidad $\Rightarrow$ cuasiconvexidad estricta
- Pseudoconvexidad estricta $\Rightarrow$ cuasiconvexidad fuerte
- Pseudoconvexidad estricta $\Rightarrow$ pseudoconvexidad

4.2.4. Propiedades de las funciones convexas, estrictamente convexas y cuasiconvexas

Proposición 4.1. *Sean $S \subseteq \mathbb{R}^n$ un conjunto convexo no vacío y $f : S \to \mathbb{R}$. Entonces f es convexa en S si y solo si $f(\sum_{i=1}^{m} \lambda_i \boldsymbol{x_i}) \leq \sum_{i=1}^{m} \lambda_i f(\boldsymbol{x_i}) \quad \forall m \in \mathbb{N}, \forall \boldsymbol{x_1}, \ldots, \boldsymbol{x_m} \in S, \forall \lambda_1, \ldots, \lambda_m \geq 0$ tales que $\sum_{i=1}^{m} \lambda_i = 1$.*

Lema 4.2. *Sean $S \subseteq \mathbb{R}^n$ un conjunto convexo no vacío y $f_1, \ldots, f_k : S \to \mathbb{R}$ funciones convexas en S, donde $k \in \mathbb{N}$. Se verifican los siguientes resultados:*

1) $\sum_{i=1}^{k} \alpha_i f_i$ *es una función convexa en* $S \quad \forall \alpha_1, \ldots, \alpha_k > 0$.

2) $\max\{f_1, \ldots, f_k\}$ *es una función convexa en* S.

Teorema 4.3. *Sean $S \subseteq \mathbb{R}^n$ un conjunto convexo no vacío y $f : S \to \mathbb{R}$ una función convexa en S. Entonces f es continua en $\mathring{S}$.*

Teorema 4.4. *Sea $S \subseteq I\!R^n$ un conjunto no vacío, convexo y abierto, y sea $f : S \to I\!R$ una función diferenciable en S. Entonces se verifican los siguientes resultados:*

1) f es convexa en S si y solo si $(\boldsymbol{\nabla} f(\boldsymbol{x_2}) - \boldsymbol{\nabla} f(\boldsymbol{x_1}))^t(\boldsymbol{x_2} - \boldsymbol{x_1}) \geq 0 \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$.

2) f es estrictamente convexa en S si y solo si $(\boldsymbol{\nabla} f(\boldsymbol{x_2}) - \boldsymbol{\nabla} f(\boldsymbol{x_1}))^t(\boldsymbol{x_2} - \boldsymbol{x_1}) > 0 \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ tales que $\boldsymbol{x_1} \neq \boldsymbol{x_2}$.

3) f es cuasiconvexa en S si y solo si $\boldsymbol{\nabla} f(\boldsymbol{x_2})^t(\boldsymbol{x_2} - \boldsymbol{x_1}) \geq 0 \quad \forall \boldsymbol{x_1}, \boldsymbol{x_2} \in S$ tales que $f(\boldsymbol{x_2}) \geq f(\boldsymbol{x_1})$.

Teorema 4.5. *Sea $S \subseteq I\!R^n$ un conjunto no vacío, convexo y abierto, y sea $f : S \to I\!R$ una función dos veces diferenciable en S. Se verifican los siguientes resultados:*

1) f es convexa en S si y solo si $\boldsymbol{H}(\overline{\boldsymbol{x}})$ es semidefinida positiva $\forall\, \overline{\boldsymbol{x}} \in S$.

2) Si f es estrictamente convexa en S, entonces $\boldsymbol{H}(\overline{\boldsymbol{x}})$ es semidefinida positiva $\forall\, \overline{\boldsymbol{x}} \in S$.

3) Si $\boldsymbol{H}(\overline{\boldsymbol{x}})$ es definida positiva $\forall\, \overline{\boldsymbol{x}} \in S$, entonces f es estrictamente convexa en S.

4.3. Condiciones de optimalidad en problemas de maximización genéricos

Considérese el siguiente problema de programación matemática:

$$\begin{array}{ll} \text{Maximizar} & f(\boldsymbol{x}) \\ \text{sujeto a:} & \boldsymbol{x} \in S, \end{array} \tag{P}$$

donde $S \subseteq I\!R^n$ es un conjunto no vacío y $f : I\!R^n \to I\!R$.

Definición 4.8. *Sea $\overline{\boldsymbol{x}} \in S$.*

- *$\overline{\boldsymbol{x}}$ es un **máximo global** de (P) si $f(\overline{\boldsymbol{x}}) \geq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S$.*
- *$\overline{\boldsymbol{x}}$ es un **máximo local** de (P) si $\exists\, \varepsilon > 0$ tal que $f(\overline{\boldsymbol{x}}) \geq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S \cap B(\overline{\boldsymbol{x}}, \varepsilon)$.*
- *$\overline{\boldsymbol{x}}$ es un **máximo local estricto** de (P) si $\exists\, \varepsilon > 0$ tal que $f(\overline{\boldsymbol{x}}) > f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S \cap B(\overline{\boldsymbol{x}}, \varepsilon)$ tal que $\boldsymbol{x} \neq \overline{\boldsymbol{x}}$.*

Se observa que si un punto es un máximo global o un máximo local estricto, también es un máximo local.

Teorema 4.6. *Sean S convexo y f convexa y diferenciable en $\mathbb{R}^n$. Si $\overline{\boldsymbol{x}}$ es un máximo local de (P), entonces $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})^t(\boldsymbol{x}-\overline{\boldsymbol{x}}) \leq 0 \quad \forall \boldsymbol{x} \in S$.*

Teorema 4.7. *Sea S un politopo. Si f es convexa en $\mathbb{R}^n$ o f es cuasiconvexa y continua en S, entonces existe algún punto extremo de S que es un máximo global de (P).*

4.4. Condiciones de optimalidad en problemas de minimización genéricos

Considérese el siguiente problema de programación matemática:

$$\begin{array}{ll} \text{Minimizar} & f(\boldsymbol{x}) \\ \text{sujeto a:} & \boldsymbol{x} \in S, \end{array} \tag{P}$$

donde $S \subseteq \mathbb{R}^n$ es un conjunto no vacío y $f : \mathbb{R}^n \to \mathbb{R}$.

Definición 4.9. *Sea $\overline{\boldsymbol{x}} \in S$.*

- *$\overline{\boldsymbol{x}}$ es un **mínimo global** de (P) si $f(\overline{\boldsymbol{x}}) \leq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S$.*
- *$\overline{\boldsymbol{x}}$ es un **mínimo local** de (P) si $\exists\, \varepsilon > 0$ tal que $f(\overline{\boldsymbol{x}}) \leq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S \cap B(\overline{\boldsymbol{x}}, \varepsilon)$.*
- *$\overline{\boldsymbol{x}}$ es un **mínimo local estricto** de (P) si $\exists\, \varepsilon > 0$ tal que $f(\overline{\boldsymbol{x}}) < f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S \cap B(\overline{\boldsymbol{x}}, \varepsilon)$ tal que $\boldsymbol{x} \neq \overline{\boldsymbol{x}}$.*

Se observa que si un punto es un mínimo global o un mínimo local estricto, también es un mínimo local.

Teorema 4.8. *Sean S convexo, $\overline{\boldsymbol{x}} \in S$ y f convexa y diferenciable en $\mathbb{R}^n$. Se verifican los siguientes resultados:*

1) $\overline{\boldsymbol{x}}$ es un mínimo global de (P) si y solo si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})^t(\boldsymbol{x}-\overline{\boldsymbol{x}}) \geq 0 \quad \forall \boldsymbol{x} \in S$.

2) Si S es abierto, entonces $\overline{\boldsymbol{x}}$ es un mínimo global de (P) si y solo si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$.

Teorema 4.9. *Sean S convexo y $\overline{\boldsymbol{x}}$ un mínimo local de (P). Se verifican los siguientes resultados:*

1) Si f es estrictamente cuasiconvexa en S, entonces $\overline{\boldsymbol{x}}$ es un mínimo global de (P).

2) Si f es fuertemente cuasiconvexa en S, entonces $\overline{\boldsymbol{x}}$ es el único mínimo global de (P).

Teorema 4.10. *Sean S convexo y $\overline{\boldsymbol{x}}$ un mínimo local estricto de (P). Si f es convexa en S, entonces $\overline{\boldsymbol{x}}$ es el único mínimo global de (P).*

Definición 4.10. *Sea $\overline{\boldsymbol{x}} \in \overline{S}$. Un vector $\boldsymbol{d} \in I\!R^n \setminus \{\boldsymbol{0}\}$ es una **dirección factible** de S en $\overline{\boldsymbol{x}}$ si $\exists\, \delta > 0$ tal que $\overline{\boldsymbol{x}} + \lambda \boldsymbol{d} \in S \quad \forall \lambda \in (0, \delta)$.*

*El **cono de direcciones factibles** de S en $\overline{\boldsymbol{x}}$ es el conjunto D de todas las direcciones factibles de S en $\overline{\boldsymbol{x}}$, esto es, $D = \{\boldsymbol{d} \in I\!R^n \setminus \{\boldsymbol{0}\} \mid \exists\, \delta_{\boldsymbol{d}} > 0$ tal que $\overline{\boldsymbol{x}} + \lambda \boldsymbol{d} \in S \quad \forall \lambda \in (0, \delta_{\boldsymbol{d}})\}$.*

Definición 4.11. *Sea $\overline{\boldsymbol{x}} \in \overline{S}$. Un vector $\boldsymbol{d} \in I\!R^n$ es una **dirección de descenso** de f en $\overline{\boldsymbol{x}}$ si $\exists\, \delta > 0$ tal que $f(\overline{\boldsymbol{x}} + \lambda \boldsymbol{d}) < f(\overline{\boldsymbol{x}}) \quad \forall \lambda \in (0, \delta)$.*

*El **cono de direcciones de descenso** de f en $\overline{\boldsymbol{x}}$ es el conjunto F de todas las direcciones de descenso de f en $\overline{\boldsymbol{x}}$, esto es, $F = \{\boldsymbol{d} \in I\!R^n \mid \exists\, \delta_{\boldsymbol{d}} > 0$ tal que $f(\overline{\boldsymbol{x}} + \lambda \boldsymbol{d}) < f(\overline{\boldsymbol{x}}) \quad \forall \lambda \in (0, \delta_{\boldsymbol{d}})\}$.*

Teorema 4.11. *Sean $\overline{\boldsymbol{x}} \in S$ y f diferenciable en $\overline{\boldsymbol{x}}$. Si $\exists\, \boldsymbol{d} \in I\!R^n$ tal que $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})^t \boldsymbol{d} < 0$, entonces $\boldsymbol{d}$ es una dirección de descenso de f en $\overline{\boldsymbol{x}}$.*

Dado un punto $\overline{\boldsymbol{x}} \in S$ tal que f es diferenciable en $\overline{\boldsymbol{x}}$, se define el conjunto

$$F_0 = \{\boldsymbol{d} \in I\!R^n \mid \boldsymbol{\nabla} f(\overline{\boldsymbol{x}})^t \boldsymbol{d} < 0\}.$$

Teorema 4.12. *Sean $\overline{\boldsymbol{x}} \in S$ y f diferenciable en $\overline{\boldsymbol{x}}$. Se verifican los siguientes resultados:*

1) Si $\overline{\boldsymbol{x}}$ es un mínimo local de (P), entonces $F_0 \cap D = \emptyset$.

2) Si $F_0 \cap D = \emptyset$, f es pseudoconvexa en $\overline{\boldsymbol{x}}$ y $\exists\, \varepsilon > 0$ tal que $\boldsymbol{x} - \overline{\boldsymbol{x}} \in D \quad \forall \boldsymbol{x} \in S \cap B(\overline{\boldsymbol{x}}, \varepsilon)$ con $\boldsymbol{x} \neq \overline{\boldsymbol{x}}$, entonces $\overline{\boldsymbol{x}}$ es un mínimo local de (P).

Es obvio que si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$, entonces $F_0 \cap D = \emptyset$.

4.5. Condiciones de optimalidad en problemas de minimización sin restricciones

Considérese el siguiente problema de programación matemática:

$$\begin{array}{ll} \text{Minimizar} & f(\boldsymbol{x}) \\ \text{sujeto a:} & \boldsymbol{x} \in I\!R^n, \end{array} \qquad (P)$$

donde $f : I\!R^n \to I\!R$.

Teorema 4.13. *Sean $\overline{\boldsymbol{x}} \in I\!R^n$ y f diferenciable en $\overline{\boldsymbol{x}}$. Si $\overline{\boldsymbol{x}}$ es un mínimo local de (P), entonces $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$.*

Teorema 4.14. *Sean $\overline{\boldsymbol{x}} \in I\!R^n$ y f dos veces diferenciable en $\overline{\boldsymbol{x}}$. Si $\overline{\boldsymbol{x}}$ es un mínimo local de (P), entonces $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$ y $\boldsymbol{H}(\overline{\boldsymbol{x}})$ es semidefinida positiva.*

Teorema 4.15. *Sean $\overline{\boldsymbol{x}} \in I\!R^n$ y f dos veces diferenciable en $\overline{\boldsymbol{x}}$. Si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$ y $\boldsymbol{H}(\overline{\boldsymbol{x}})$ es definida positiva, entonces $\overline{\boldsymbol{x}}$ es un mínimo local estricto de (P).*

Teorema 4.16. *Sean $\overline{\boldsymbol{x}} \in I\!R^n$ y f pseudoconvexa en $\overline{\boldsymbol{x}}$. Entonces $\overline{\boldsymbol{x}}$ es un mínimo global de (P) si y solo si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$.*

4.6. Condiciones de optimalidad de Fritz John y de Karush-Kuhn-Tucker en problemas de minimización con restricciones de desigualdad

Considérese el siguiente problema de programación matemática:

$$\begin{array}{ll} \text{Minimizar} & f(\boldsymbol{x}) \\ \text{sujeto a:} & \boldsymbol{x} \in S, \end{array} \qquad (P)$$

donde $S = \{\boldsymbol{x} \in X \mid g_i(\boldsymbol{x}) \leq 0 \quad \forall i \in \{1, \ldots, m\}\}$, $X \subseteq I\!R^n$ es un conjunto abierto no vacío, $m \in I\!N$, $g_i : I\!R^n \to I\!R \quad \forall i \in \{1, \ldots, m\}$ y $f : I\!R^n \to I\!R$.

Dado $\overline{\boldsymbol{x}} \in S$, se denota por I al conjunto de índices $i \in \{1, \ldots, m\}$ de las restricciones **activas** en $\overline{\boldsymbol{x}}$, es decir, $I = \{i \in \{1, \ldots, m\} \mid g_i(\overline{\boldsymbol{x}}) = 0\}$. Si las funciones $\{g_i\}_{i \in I}$ son diferenciables en $\overline{\boldsymbol{x}}$, se define el conjunto $G_0 = \{\boldsymbol{d} \in I\!R^n \mid \boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})^t \boldsymbol{d} < 0 \quad \forall i \in I\}$.

Teorema 4.17. *Sean* $\overline{\boldsymbol{x}} \in S$*,* f *y* $\{g_i\}_{i\in I}$ *diferenciables en* $\overline{\boldsymbol{x}}$*, y* $\{g_i\}_{i\in\{1,\dots,m\}\setminus I}$ *continuas en* $\overline{\boldsymbol{x}}$*. Se verifican los siguientes resultados:*

1) *Si* $\overline{\boldsymbol{x}}$ *es un mínimo local de* (P)*, entonces* $F_0 \cap G_0 = \emptyset$*.*

2) *Si* $F_0 \cap G_0 = \emptyset$*,* f *es pseudoconvexa en* $\overline{\boldsymbol{x}}$ *y* $\exists\, \varepsilon > 0$ *tal que* $\{g_i\}_{i\in I}$ *son estrictamente pseudoconvexas en* $B(\overline{\boldsymbol{x}}, \varepsilon)$*, entonces* $\overline{\boldsymbol{x}}$ *es un mínimo local de* (P)*.*

Es obvio que si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})= \mathbf{0}$ o $\exists\, i \in I$ con $\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})= \mathbf{0}$, entonces $F_0 \cap G_0 = \emptyset$.

Teorema 4.18 (Condiciones necesarias de Fritz John)**.** *Sean* $\overline{\boldsymbol{x}} \in S, f$ *y* $\{g_i\}_{i\in I}$ *diferenciables en* $\overline{\boldsymbol{x}}$*, y* $\{g_i\}_{i\in\{1,\dots,m\}\setminus I}$ *continuas en* $\overline{\boldsymbol{x}}$*. Si* $\overline{\boldsymbol{x}}$ *es un mínimo local de* (P)*, se verifican los siguientes resultados:*

1) *Existen escalares* $u_0, \{u_i\}_{i\in I}$ *no todos nulos tales que*

$$u_0 \boldsymbol{\nabla} f(\overline{\boldsymbol{x}})+ \sum_{i\in I} u_i \boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})= \mathbf{0},$$
$$u_0 \geq 0,\ u_i \geq 0 \quad \forall i \in I$$

2) *Si las funciones* $\{g_i\}_{i\in\{1,\dots,m\}\setminus I}$ *son diferenciables en* $\overline{\boldsymbol{x}}$*, entonces existen escalares* u_0*,* $\{u_i\}_{i\in\{1,\dots,m\}}$ *no todos nulos tales que*

$$u_0 \boldsymbol{\nabla} f(\overline{\boldsymbol{x}})+ \sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})= \mathbf{0},$$
$$u_i g_i(\overline{\boldsymbol{x}}) = 0 \quad \forall i \in \{1, \dots, m\},$$
$$u_0 \geq 0,\ u_i \geq 0 \quad \forall i \in \{1, \dots, m\}$$

Es obvio que si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})= \mathbf{0}$ o $\exists\, i \in I$ con $\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})= \mathbf{0}$, entonces $\overline{\boldsymbol{x}}$ satisface las condiciones necesarias de Fritz John.

Un punto $\overline{\boldsymbol{x}} \in S$ que satisface dichas condiciones se denomina **punto de Fritz John** o, abreviadamente, **punto FJ**. Los escalares $u_0, \{u_i\}_{i\in\{1,\dots,m\}}$ se llaman **multiplicadores de Lagrange**.

Se observa que todo mínimo global de (P) es un punto FJ.

Teorema 4.19 (Condiciones suficientes de Fritz John)**.** *Sean* $\overline{\boldsymbol{x}}$ *un punto FJ,* f *pseudoconvexa en* $\overline{\boldsymbol{x}}$ *y* $\{g_i\}_{i\in I}$ *estrictamente pseudoconvexas y cuasiconvexas en* $\overline{\boldsymbol{x}}$*. Entonces* $\overline{\boldsymbol{x}}$ *es un mínimo global de* (P)*.*

Teorema 4.20 (Condiciones necesarias de Karush-Kuhn-Tucker)**.** *Sean* $\overline{\boldsymbol{x}} \in S$*,* f *y* $\{g_i\}_{i\in I}$ *diferenciables en* $\overline{\boldsymbol{x}}$*,* $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ *continuas en* $\overline{\boldsymbol{x}}$ *y* $\{\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})\}_{i\in I}$ *linealmente independientes. Si* $\overline{\boldsymbol{x}}$ *es un mínimo local de* (P)*, se verifican los siguientes resultados:*

1) Existen escalares $\{u_i\}_{i\in I}$ *tales que*

$$\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})+\sum_{i\in I} u_i \boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})=\boldsymbol{0},$$
$$u_i \geq 0 \quad \forall i \in I$$

2) Si las funciones $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ *son diferenciables en* $\overline{\boldsymbol{x}}$*, entonces existen escalares* $\{u_i\}_{i\in\{1,\ldots,m\}}$ *tales que*

$$\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})+\sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})=\boldsymbol{0},$$
$$u_i g_i(\overline{\boldsymbol{x}}) = 0 \quad \forall i \in \{1,\ldots,m\},$$
$$u_i \geq 0 \quad \forall i \in \{1,\ldots,m\}$$

Es obvio que si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})=\boldsymbol{0}$, $\overline{\boldsymbol{x}}$ satisface las condiciones necesarias de Karush-Kuhn-Tucker.

Un punto $\overline{\boldsymbol{x}}\in S$ que satisface dichas condiciones se denomina **punto de Karush-Kuhn-Tucker** o, abreviadamente, **punto KKT**. Los escalares $\{u_i\}_{i\in\{1,\ldots,m\}}$ se llaman **multiplicadores de Lagrange**.

Se observa que todo punto KKT es un punto FJ, pero no todo mínimo global de (P) ha de ser un punto KKT. Por otra parte, se verifica que un punto $\overline{\boldsymbol{x}} \in S$ es un punto KKT si y solo si $-\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})$ pertenece al cono generado por los vectores $\{\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})\}_{i\in I}$.

Teorema 4.21 (Condiciones suficientes de Karush-Kuhn-Tucker)**.** *Sean* $\overline{\boldsymbol{x}}$ *un punto KKT,* f *pseudoconvexa en* $\overline{\boldsymbol{x}}$ *y* $\{g_i\}_{i\in I}$ *cuasiconvexas en* $\overline{\boldsymbol{x}}$*. Entonces* $\overline{\boldsymbol{x}}$ *es un mínimo global de* (P)*.*

4.7. Condiciones de optimalidad de Fritz John y de Karush-Kuhn-Tucker en problemas de minimización con restricciones de desigualdad e igualdad

Considérese el siguiente problema de programación matemática:

$$\begin{array}{ll} \text{Minimizar} & f(\boldsymbol{x}) \\ \text{sujeto a:} & \boldsymbol{x} \in S, \end{array} \qquad (P)$$

donde $S = \{\boldsymbol{x} \in X \mid g_i(\boldsymbol{x}) \leq 0 \quad \forall i \in \{1,\ldots,m\},\ h_i(\boldsymbol{x}) = 0 \quad \forall i \in \{1,\ldots,l\}\}$, $X \subseteq I\!R^n$ es un conjunto abierto no vacío, $m \in I\!N \cup \{0\}$, $l \in I\!N$, $g_i : I\!R^n \to I\!R \quad \forall i \in \{1,\ldots,m\}$, $h_i : I\!R^n \to I\!R$ $\forall i \in \{1,\ldots,l\}$ y $f : I\!R^n \to I\!R$.

Dado $\overline{\boldsymbol{x}} \in S$ tal que las funciones $\{h_i\}_{i\in\{1,\ldots,l\}}$ son diferenciables en $\overline{\boldsymbol{x}}$, se define el conjunto $H_0 = \{\boldsymbol{d} \in I\!R^n \mid \boldsymbol{\nabla} h_i(\overline{\boldsymbol{x}})^t\boldsymbol{d} = 0 \quad \forall i \in \{1,\ldots,l\}\}$.

Teorema 4.22. *Sean $\overline{\boldsymbol{x}} \in S$, f y $\{g_i\}_{i\in I}$ diferenciables en $\overline{\boldsymbol{x}}$, $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ continuas en $\overline{\boldsymbol{x}}$, $\{h_i\}_{i\in\{1,\ldots,l\}}$ continuamente diferenciables en $\overline{\boldsymbol{x}}$ y $\{\boldsymbol{\nabla} h_i(\overline{\boldsymbol{x}})\}_{i\in\{1,\ldots,l\}}$ linealmente independientes. Se verifican los siguientes resultados:*

1) Si $\overline{\boldsymbol{x}}$ es un mínimo local de (P), entonces $F_0 \cap G_0 \cap H_0 = \emptyset$.

2) Si $F_0 \cap G_0 \cap H_0 = \emptyset$, f es pseudoconvexa en $\overline{\boldsymbol{x}}$, $\exists\, \varepsilon > 0$ tal que $\{g_i\}_{i\in I}$ son estrictamente pseudoconvexas en $B(\overline{\boldsymbol{x}},\varepsilon)$ y $\{h_i\}_{i\in\{1,\ldots,l\}}$ son afines, entonces $\overline{\boldsymbol{x}}$ es un mínimo local de (P).

Es obvio que si $\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})= \boldsymbol{0}$ o $\exists\, i \in I$ con $\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})= \boldsymbol{0}$, entonces $F_0 \cap G_0 \cap H_0 = \emptyset$.

Teorema 4.23 (Condiciones necesarias de Fritz John)**.** *Sean $\overline{\boldsymbol{x}} \in S, f$ y $\{g_i\}_{i\in I}$ diferenciables en $\overline{\boldsymbol{x}}$, $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ continuas en $\overline{\boldsymbol{x}}$ y $\{h_i\}_{i\in\{1,\ldots,l\}}$ continuamente diferenciables en $\overline{\boldsymbol{x}}$. Si $\overline{\boldsymbol{x}}$ es un mínimo local de (P), se verifican los siguientes resultados:*

1) Existen escalares $u_0, \{u_i\}_{i\in I}, \{v_i\}_{i\in\{1,\ldots,l\}}$ no todos nulos tales que

$$u_0\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})+\sum_{i\in I} u_i\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})+\sum_{i=1}^{l} v_i\boldsymbol{\nabla} h_i(\overline{\boldsymbol{x}})= \boldsymbol{0},$$

$$u_0 \geq 0,\ u_i \geq 0 \quad \forall i \in I$$

2) Si las funciones $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ son diferenciables en $\overline{\boldsymbol{x}}$, entonces existen escalares u_0, $\{u_i\}_{i\in\{1,\ldots,m\}}, \{v_i\}_{i\in\{1,\ldots,l\}}$ no todos nulos tales que

$$u_0\boldsymbol{\nabla} f(\overline{\boldsymbol{x}})+\sum_{i=1}^{m} u_i\boldsymbol{\nabla} g_i(\overline{\boldsymbol{x}})+\sum_{i=1}^{l} v_i\boldsymbol{\nabla} h_i(\overline{\boldsymbol{x}})= \boldsymbol{0},$$

$$u_i g_i(\overline{\boldsymbol{x}}) = 0 \quad \forall i \in \{1,\ldots,m\},$$

$$u_0 \geq 0,\ u_i \geq 0 \quad \forall i \in \{1,\ldots,m\}$$

Es obvio que si $\nabla f(\overline{\boldsymbol{x}})= \mathbf{0}$ o $\exists\, i \in I$ con $\nabla g_i(\overline{\boldsymbol{x}})= \mathbf{0}$ o $\exists\, i \in \{1,\ldots,l\}$ con $\nabla h_i(\overline{\boldsymbol{x}})= \mathbf{0}$, entonces $\overline{\boldsymbol{x}}$ satisface las condiciones necesarias de Fritz John.

Un punto $\overline{\boldsymbol{x}} \in S$ que satisface dichas condiciones se denomina **punto de Fritz John** o, abreviadamente, **punto FJ**. Los escalares $u_0, \{u_i\}_{i\in\{1,\ldots,m\}}, \{v_i\}_{i\in\{1,\ldots,l\}}$ se llaman **multiplicadores de Lagrange**.

Se observa que todo mínimo global de (P) es un punto FJ.

Teorema 4.24 (Condiciones necesarias de Karush-Kuhn-Tucker)**.** *Sean $\overline{\boldsymbol{x}} \in S$, f y $\{g_i\}_{i\in I}$ diferenciables en $\overline{\boldsymbol{x}}$, $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ continuas en $\overline{\boldsymbol{x}}$, $\{h_i\}_{i\in\{1,\ldots,l\}}$ continuamente diferenciables en $\overline{\boldsymbol{x}}$ y $\{\{\nabla g_i(\overline{\boldsymbol{x}})\}_{i\in I}, \{\nabla h_i(\overline{\boldsymbol{x}})\}_{i\in\{1,\ldots,l\}}\}$ linealmente independientes. Si $\overline{\boldsymbol{x}}$ es un mínimo local de (P), se verifican los siguientes resultados:*

1) Existen escalares $\{u_i\}_{i\in I}, \{v_i\}_{i\in\{1,\ldots,l\}}$ tales que

$$\nabla f(\overline{\boldsymbol{x}})+\sum_{i\in I} u_i \nabla g_i(\overline{\boldsymbol{x}})+\sum_{i=1}^{l} v_i \nabla h_i(\overline{\boldsymbol{x}})= \mathbf{0},$$
$$u_i \geq 0 \quad \forall i \in I$$

2) Si las funciones $\{g_i\}_{i\in\{1,\ldots,m\}\setminus I}$ son diferenciables en $\overline{\boldsymbol{x}}$, entonces existen escalares $\{u_i\}_{i\in\{1,\ldots,m\}}, \{v_i\}_{i\in\{1,\ldots,l\}}$ tales que

$$\nabla f(\overline{\boldsymbol{x}})+\sum_{i=1}^{m} u_i \nabla g_i(\overline{\boldsymbol{x}})+\sum_{i=1}^{l} v_i \nabla h_i(\overline{\boldsymbol{x}})= \mathbf{0},$$
$$u_i g_i(\overline{\boldsymbol{x}}) = 0 \quad \forall i \in \{1,\ldots,m\},$$
$$u_i \geq 0 \quad \forall i \in \{1,\ldots,m\}$$

Es obvio que si $\nabla f(\overline{\boldsymbol{x}})= \mathbf{0}$, $\overline{\boldsymbol{x}}$ satisface las condiciones necesarias de Karush-Kuhn-Tucker.

Un punto $\overline{\boldsymbol{x}} \in S$ que satisface dichas condiciones se denomina **punto de Karush-Kuhn-Tucker** o, abreviadamente, **punto KKT**. Los escalares $\{u_i\}_{i\in\{1,\ldots,m\}}, \{v_i\}_{i\in\{1,\ldots,l\}}$ se llaman **multiplicadores de Lagrange**.

Se observa que todo punto KKT es un punto FJ, pero no todo mínimo global de (P) ha de ser un punto KKT.

Teorema 4.25 (Condiciones suficientes de Karush-Kuhn-Tucker)**.** *Sean $\overline{\boldsymbol{x}}$ un punto KKT, f pseudoconvexa en $\overline{\boldsymbol{x}}$, $\{g_i\}_{i\in I}$ cuasiconvexas en $\overline{\boldsymbol{x}}$, $\{h_i\}_{i\in\{1,\ldots,l\}\, con\ v_i>0}$ cuasiconvexas en $\overline{\boldsymbol{x}}$ y $\{h_i\}_{i\in\{1,\ldots,l\}\, con\ v_i<0}$ cuasicóncavas en $\overline{\boldsymbol{x}}$. Entonces $\overline{\boldsymbol{x}}$ es un mínimo global de (P).*