# Optimización con Restricciones
# Fundamentos y Marco Teórico

Dr. Ing. Claudio Risso (crisso@fing.edu.uy)

**Instituto de Computación (FING - UDELAR)**

Curso Optimización Continua y Aplicaciones (OCA)
(Octubre 2023)

# Programación Matemática

El problema de referencia para esta parte del curso será el de programación matemática:

$$(P)\begin{cases} \min f(x) \\ g_i(x) \leq 0, \quad 1 \leq i \leq p, \\ h_j(x) = 0, \quad 1 \leq j \leq q, \\ x \in X \end{cases}$$

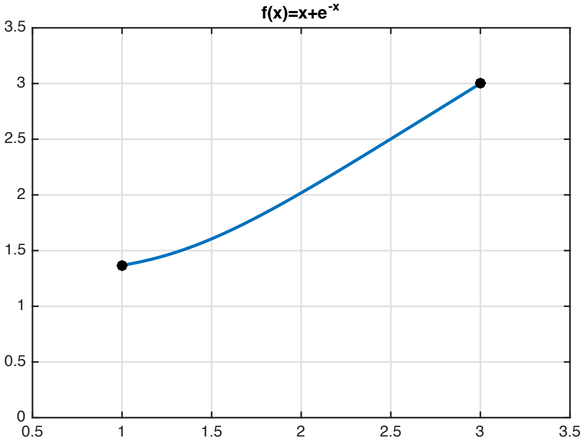
El problema $(P)$ tiene $n$ variables y $m = p + q$ restricciones.

Como norma, asumiremos que $X = \mathbb{R}^n$ si no se especifica.

Observar que incluso siendo todas las funciones diferenciables, ya no podemos asumir simplemente que un óptimo local del problema debe cumplir $\nabla f(x) = \vec{0}$.

# Programación Matemática

Observar que incluso siendo todas las funciones diferenciables, ya no podemos asumir simplemente que un óptimo local del problema debe cumplir $\nabla f(x) = \vec{0}$.



# Programación Matemática

El problema de referencia para esta parte del curso será el de programación matemática:

$$(P)\begin{cases} \min f(x) \\ g_i(x) \leq 0, \quad 1 \leq i \leq p, \\ h_j(x) = 0, \quad 1 \leq j \leq q, \\ x \in X \end{cases}$$

El problema $(P)$ tiene $n$ variables y $m = p + q$ restricciones.

Como norma, asumiremos que $X = \mathbb{R}^n$ si no se especifica.

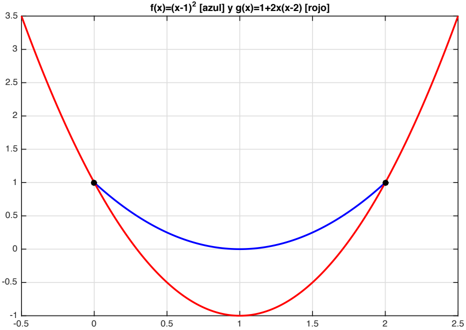
Observar que incluso siendo todas las funciones diferenciables, ya no podemos asumir simplemente que un óptimo local del problema debe cumplir $\nabla f(x) = \vec{0}$.

**¿Qué propiedades teóricas debemos usar para encontrar óptimos en problemas con restricciones?**

# Optimización y Relajaciones

## Definición

*Dados dos problemas* $(G)\begin{cases}\min f(x)\\ x\in F\end{cases}$ *y* $(G_L)\begin{cases}\min f_L(x)\\ x\in F_L\end{cases}$, *se dice que* $(G_L)$ *es una relajación de* $(G)$, *sii se cumplen las siguientes dos condiciones: i)* $F\subseteq F_L$ *y ii)* $f_L(x)\leq f(x), \forall x\in F$.




$$(P)\begin{cases}\min\,(x-1)^2\\ 0\leq x\leq 2\end{cases}$$

$$(P_L)\begin{cases}\min 1+2x(x-2)\\ -\frac{1}{2}\leq x\leq\frac{5}{2}\end{cases}$$

$(P_L)$ es una relajación de $(P)$

# Optimización y Relajaciones

**Propiedad**

*Sea* $(G_L)$ *una relajación de* $(G)$, $\bar{x}_L$ *y* $\bar{x}$ *óptimos de* $(G_L)$ *y* $(G)$ *respectivamente, y se cumple que* $\bar{x}_L \in F$, *entonces, también se cumple que:* $f_L(\bar{x}_L) \leq f(\bar{x}) \leq f(\bar{x}_L)$.

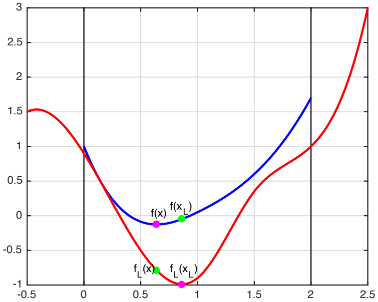
**PRUEBA:** Observar que $\bar{x} \in F_L$ (ya que $x \in F$ y $F \subseteq F_L$). Como $\bar{x}_L$ es el óptimo de $(G_L)$, se cumple $f_L(\bar{x}_L) \leq f_L(\bar{x})$, ya que en $(G_L)$, $\bar{x}$ es simplemente un punto más. Además, $f_L(\bar{x}) \leq f(\bar{x})$, porque $x \in F$ y $f_L(\cdot)$ es la función de la relajación. Eso concluye que $f_L(\bar{x}_L) \leq f(\bar{x})$, la primera parte de la desigualdad.

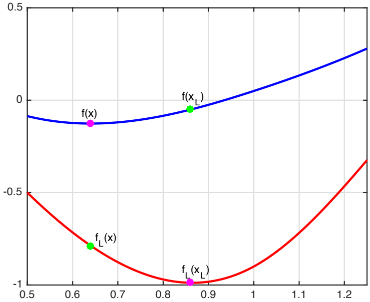
La segunda parte $(f(\bar{x}) \leq f(\bar{x}_L))$ se explica porque $\bar{x}$ es el óptimo de $(G)$, $\bar{x}_L \in F$ (por hipótesis) y por tanto, ese punto es uno más en el problema del cual $\bar{x}$ es óptimo.

# Optimización y Relajaciones

## Propiedad

*Sea* $(G_L)$ *una relajación de* $(G)$, $\bar{x}_L$ *y* $\bar{x}$ *óptimos de* $(G_L)$ *y* $(G)$ *respectivamente, y se cumple que* $\bar{x}_L \in F$, *entonces, también se cumple que:* $f_L(\bar{x}_L) \leq f(\bar{x}) \leq f(\bar{x}_L)$.



# Optimización y Relajaciones

**Propiedad**

*Sea $(G_L)$ una relajación de $(G)$, $\bar{x}_L$ y $\bar{x}$ óptimos de $(G_L)$ y $(G)$ respectivamente, y se cumple que $\bar{x}_L \in F$, entonces, también se cumple que: $f_L(\bar{x}_L) \leq f(\bar{x}) \leq f(\bar{x}_L)$.*

**Corolario (Óptimo mediante Relajación)**

*Sea $(G_L)$ una relajación de $(G)$, y $\bar{x}_L$ la solución óptima de $(G_L)$. Si se cumple que $\bar{x}_L \in F$ y $f_L(\bar{x}_L) = f(\bar{x}_L)$, entonces $\bar{x}_L$ también es solución óptima de $(G)$.*

Esta última propiedad es uno de los pilares para generalizar las condiciones teóricas a los problemas continuos con restricciones, e incluso las técnicas exactas en la optimización combinatoria.

# Ejemplo (Relajación Entera)

Si las variables de un problema de optimización están definidas sobre el conjunto $X = \mathbb{Z}^n$, llamamos Relajación Entera al problema equivalente en el que $X_L = \mathbb{R}^n$.

Es una relajación porque los reales son más que los enteros, y cuando coinciden, la función objetivo vale lo mismo.

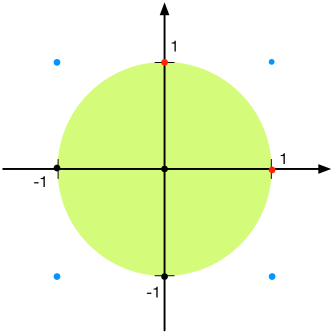
$$(P)\begin{cases} \min x^2 + 4y^2 - 12x - 8y + 10 \\ x, y \in \mathbb{Z} \end{cases}$$

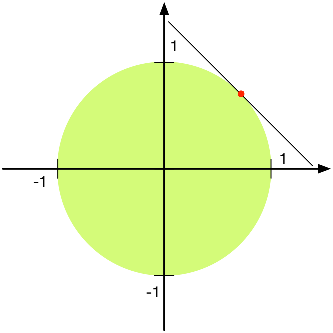
Ya lo resolvimos con $(P_L)$ en $\mathbb{R}^2$ y la solución era (6,1), entonces ésa también es la solución para $(P)$, que está definido en $\mathbb{Z}^2$.

**El argumento no es general porque el óptimo en $(P_L)$ puede no ser entero.**

# Ejemplo (Relajación Entera)

**El argumento no es general porque el óptimo en $(P_L)$ puede no ser entero.**




Lo habíamos visto en el problema $(P) \begin{cases} \max x + y \\ x^2 + y^2 \leq 1 \end{cases}$

# Ejemplo (Relajación Entera)

**El argumento no es general porque el óptimo en $(P_L)$ puede no ser entero.**

Lo que sí vale siempre es que el óptimo de $(P_L)$ es una cota inferior del de $(P)$.

Además, el mismo punto $\bar{x}$ solución de $(P_L)$ tiene valor en sí mismo. Si por ejemplo, $\bar{x}_2 = 3.2$, podemos generar dos sub-problemas (proceso conocido como branching), uno en el que se agrega la restricción $x_2 \leq 3$ y otro en el que $x_2 \geq 4$.

Variantes de estas ideas se usan en los algoritmos *branch-and-boud*, *branch-and-cut*, etc., que son los Algoritmos Exactos por excelencia en la programación entera.

**Esos algoritmos están fuera del alcance de este curso.**

# Relajación Lagrangeana

Es la reducción fundamental en la programación continua.

$$(P)\begin{cases} \min f(x) \\ g_i(x) \leq 0, \quad 1 \leq i \leq p, \\ h_j(x) = 0, \quad 1 \leq j \leq q \end{cases}$$

Dado un problema $(P)$ y un conjunto de constantes $\lambda_i \geq 0$ $(1 \leq i \leq p)$ y $\mu_j$ $(1 \leq j \leq q)$, construye una relajación $(P_{\lambda,\mu})$ sin restricciones.

$$(P_{\lambda,\mu})\left\{ \ \min f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x) \right.$$

Observar que $F \subseteq F_{\lambda,\mu}$, porque $(P_{\lambda,\mu})$ no tiene restricciones. Además, como en la región factible $F$ de $(P)$ se verifica $g_i(x) \leq 0$ y $h_j(x) = 0$, $\boldsymbol{f_{\lambda,\mu}(x)} = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x) \leq f(x)$, ya que $\lambda_i g_i(x) \leq 0$ y $\mu_j h_j(x) = 0$.

# Relajación Lagrangeana

Es la reducción fundamental en la programación continua.

$$(P)\begin{cases} \min f(x) \\ g_i(x) \leq 0, \quad 1 \leq i \leq p, \\ h_j(x) = 0, \quad 1 \leq j \leq q \end{cases}$$

Dado un problema $(P)$ y un conjunto de constantes $\lambda_i \geq 0$ $(1 \leq i \leq p)$ y $\mu_j$ $(1 \leq j \leq q)$, construye una relajación $(P_{\lambda,\mu})$ sin restricciones.

$$(P_{\lambda,\mu})\left\{ \min f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x) \right.$$

No es necesario relajar todas las restricciones para tener una Relajación Lagrangeana. Pero si lo hacemos, tendremos un problema donde sí vale que en el óptimo se debe anular el gradiente (cuando las funciones son diferenciables).

# Relajación Lagrangeana

Es la reducción fundamental en la programación continua.

$$(P)\begin{cases} \min f(x) \\ g_i(x) \leq 0, \quad 1 \leq i \leq p, \\ h_j(x) = 0, \quad 1 \leq j \leq q \end{cases}$$

Dado un problema $(P)$ y un conjunto de constantes $\lambda_i \geq 0$ $(1 \leq i \leq p)$ y $\mu_j$ $(1 \leq j \leq q)$, construye una relajación $(P_{\lambda,\mu})$ sin restricciones.

$$(P_{\lambda,\mu})\left\{\ \min f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x)\right.$$

**Los $\lambda$'s y $\mu$'s no son variables, son parámetros del problema relajado. Constantes, a efectos de derivar o usar propiedades.**

Con eso en mente, se puede usar alguna de las técnicas de optimización sin restricciones vistas previamente en el curso.

# Relajación Lagrangeana

**Teorema (Fundamental de la Relajación Lagrangeana)**

*Sea $\bar{x}$ una solución de $(P_{\lambda,\mu})$ que cumple:*

1. $g_i(\bar{x}) \leq 0 \, (1 \leq i \leq p)$ *y* $h_j(\bar{x}) = 0 \, (1 \leq j \leq q)$, *[$\bar{x}$ factible en $(P)$]*
2. $\lambda_i \geq 0 \, (1 \leq i \leq p)$, *[$(P_{\lambda,\mu})$ es una relajación de $(P)$]*
3. $\lambda_i \cdot g_i(\bar{x}) = 0 \, (1 \leq i \leq p)$, *[$f_{\lambda,\mu}(\bar{x}) = f(\bar{x})$]*

*Entonces $\bar{x}$ también es solución óptima de $(P)$.*

**PRUEBA:** Se cumple porque se dan las condiciones del corolario anterior (ver Óptimo mediante Relajación), que aparecen sombreadas al lado de las hipótesis.

Si tuviéramos propiedades (o algoritmos) para encontrar el óptimo global del problema $(P_{\lambda,\mu})$, el TMFRL da ecuaciones adicionales que nos asegurarían que ese óptimo también vale para $(P)$.

# Suficiencia para un Óptimo (Karush-Kuhn-Tucker)

## Teorema (Las condiciones KKT son suficientes)

*En un problema $(P)$ donde las funciones $f$ y $g_i$ son convexas, las $h_j$ afines, y además todas son diferenciables, es suficiente encontrar $\bar{x}$, $\bar{\lambda}=\{\lambda_i\}$ y $\bar{\mu}=\{\mu_j\}$ que cumplan las condiciones KKT siguientes:*

1. $\nabla f_{\bar{\lambda},\bar{\mu}}(\bar{x}) = \nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = \vec{0}$
2. $g_i(\bar{x}) \leq 0 \, (1 \leq i \leq p)$ *y* $h_j(\bar{x}) = 0$
3. $\lambda_i \geq 0 \, (1 \leq i \leq p)$
4. $\lambda_i \cdot g_i(\bar{x}) = 0 \, (1 \leq i \leq p)$ *[condición de complementaridad]*

*para tener un óptimo global $\bar{x}$ del problema.*

**PRUEBA:** $f_{\lambda,\mu}(x)$ siempre es diferenciable (combinación lineal de diferenciables). Cuando $q = 0$ (i.e. $\{h_j\} = \emptyset$), sabemos que $f(x) + \sum_{i=1}^{p} \lambda_i g_i(x)$ es convexa porque $\lambda_i \geq 0$ (tercera condición).

# Suficiencia para un Óptimo (Karush-Kuhn-Tucker)

## Teorema (Las condiciones KKT son suficientes)

*En un problema $(P)$ donde las funciones $f$ y $g_i$ son convexas, las $h_j$ afines, y además todas son diferenciables, es suficiente encontrar $\bar{x}$, $\bar{\lambda}=\{\lambda_i\}$ y $\bar{\mu}=\{\mu_j\}$ que cumplan las condiciones KKT siguientes:*

1. $\nabla f_{\bar{\lambda},\bar{\mu}}(\bar{x}) = \nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = \vec{0}$
2. $g_i(\bar{x}) \leq 0 \, (1 \leq i \leq p)$ *y* $h_j(\bar{x}) = 0$
3. $\lambda_i \geq 0 \, (1 \leq i \leq p)$
4. $\lambda_i \cdot g_i(\bar{x}) = 0 \, (1 \leq i \leq p)$ *[condición de complementaridad]*

*para tener un óptimo global $\bar{x}$ del problema.*

**PRUEBA:** Luego, $\nabla f_{\bar{\lambda}}(\bar{x}) = \vec{0}$ (condición 1) es suficiente para un óptimo global del problema $(P_{\lambda})$, y el resto de las condiciones son las del TMFRL.

# Suficiencia para un Óptimo (Karush-Kuhn-Tucker)

## Teorema (Las condiciones KKT son suficientes)

*En un problema $(P)$ donde las funciones $f$ y $g_i$ son convexas, las $h_j$ afines, y además todas son diferenciables, es suficiente encontrar $\bar{x}$, $\bar{\lambda}=\{\lambda_i\}$ y $\bar{\mu}=\{\mu_j\}$ que cumplan las condiciones KKT siguientes:*

1. $\nabla f_{\bar{\lambda},\bar{\mu}}(\bar{x}) = \nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = \vec{0}$
2. $g_i(\bar{x}) \leq 0 \, (1 \leq i \leq p)$ *y* $h_j(\bar{x}) = 0$
3. $\lambda_i \geq 0 \, (1 \leq i \leq p)$
4. $\lambda_i \cdot g_i(\bar{x}) = 0 \, (1 \leq i \leq p)$ *[condición de complementaridad]*

*para tener un óptimo global $\bar{x}$ del problema.*

**PRUEBA:** Las funciones afines son las que preservan lineas y paralelismos (e.g. las funciones lineales son afines). Mantienen la convexidad, incluso al ser multiplicadas por negativos.

# Suficiencia para un Óptimo (Karush-Kuhn-Tucker)

## Teorema (Las condiciones KKT son suficientes)

*En un problema $(P)$ donde las funciones $f$ y $g_i$ son convexas, las $h_j$ afines, y además todas son diferenciables, es suficiente encontrar $\bar{x}$, $\bar{\lambda}=\{\lambda_i\}$ y $\bar{\mu}=\{\mu_j\}$ que cumplan las condiciones KKT siguientes:*

1. $\nabla f_{\bar{\lambda},\bar{\mu}}(\bar{x}) = \nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = \vec{0}$
2. $g_i(\bar{x}) \leq 0\,(1 \leq i \leq p)$ *y* $h_j(\bar{x}) = 0$
3. $\lambda_i \geq 0\,(1 \leq i \leq p)$
4. $\lambda_i \cdot g_i(\bar{x}) = 0\,(1 \leq i \leq p)$ *[condición de complementaridad]*

*para tener un óptimo global $\bar{x}$ del problema.*

**PRUEBA:** Si $q > 0$, pueden haber $\mu_j$'s$< 0$ y aun así sostenerse la convexidad de $f_{\lambda,\mu}(x)$ por la afinidad en las $h_j$, lo que completa la suficiencia de las condiciones.

# Ejemplo KKT

Sea $(P)\begin{cases} \max x+y \\ x^2+y^2 \leq 1 \end{cases}$, equivalente a $(P')\begin{cases} \min -(x+y) \\ x^2+y^2-1 \leq 0 \end{cases}$.

Luego, $f_\lambda(x,y) = -x-y+\lambda(x^2+y^2-1)$ y resolvemos el sistema:

i) $-1+2\lambda x = 0$ y $-1+2\lambda y = 0$,

ii) $x^2+y^2-1 \leq 0$ (equivalente a $x^2+y^2 \leq 1$),

iii) $\lambda \geq 0$

iv) $\lambda(x^2+y^2-1) = 0$.

Si $\lambda = 0$ no puede cumplirse i). Entonces $\lambda > 0$ (por iii), $x^2+y^2 = 1$ (por iv y cumple ii) y $x = y = \frac{1}{2\lambda}$ (por i).
Substituyendo $x$ y $y$ llegamos a $(\frac{1}{2\lambda})^2 + (\frac{1}{2\lambda})^2 = 1 = \frac{1}{2\lambda^2}$, de donde $\lambda = \frac{1}{\sqrt{2}}$, $\bar{x} = \bar{y} = \frac{1}{\sqrt{2}}$ es el óptimo y $f(\bar{x},\bar{y})$ vale $-\sqrt{2}$ en $(P')$.
El óptimo en $(P)$ se da en el mismo punto, pero vale $\sqrt{2}$.

# Ejemplo KKT

Sea $(P)\begin{cases} \max x+y \\ x^2+y^2 \leq 1 \end{cases}$, equivalente a $(P')\begin{cases} \min -(x+y) \\ x^2+y^2-1 \leq 0 \end{cases}$.

Luego, $f_\lambda(x,y) = -x-y+\lambda(x^2+y^2-1)$ y resolvemos el sistema:

i) $-1+2\lambda x = 0$ y $-1+2\lambda y = 0$,

ii) $x^2+y^2-1 \leq 0$ (equivalente a $x^2+y^2 \leq 1$),

iii) $\lambda \geq 0$

iv) $\lambda(x^2+y^2-1) = 0$.

**¿Qué hubiera pasado si no hubiésemos encontrado solución al sistema? ¿Podemos decir que el problema no tiene óptimo?**

**En este caso tiene óptimo por Weierstrass. En general ¿estamos seguros que KKT se cumple? ¿podemos asegurar que no hay óptimo si sus ecuaciones no tienen solución?**

# Suficiencia para un Óptimo (Karush-Kuhn-Tucker)

Las condiciones de KKT son en general para clasificar óptimos locales. En el caso convexo, se clasifican ambos tipos simultáneamente y no hacemos la diferencia.

Pero las condiciones KKT son suficientes para clasificar óptimos locales incluso bajo hipótesis más generales, como que las funciones sean Invexas: $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable es invexa sii para todo $x, y \in \mathbb{R}^n$ existe $\eta(x, y)$ tal que $f(y) - f(x) \geq \nabla^T f(x) \cdot \eta(x, y)$. Cuando $\eta(x, y) = y - x$ es la convexidad.

# Suficiencia para un Óptimo (Karush-Kuhn-Tucker)

Las condiciones de KKT son en general para clasificar óptimos locales. En el caso convexo, se clasifican ambos tipos simultáneamente y no hacemos la diferencia.

Pero las condiciones KKT son suficientes para clasificar óptimos locales incluso bajo hipótesis más generales, como que las funciones sean Invexas. **No las vemos en este curso.**

En la prueba del TM previo se exigió que las funciones $f$ y $g_i$ sean convexas, y en caso de existir las $h_j$, que ellas sean afines. En esas hipótesis se cumple la Invexidad

Otra forma general de garantizar la suficiencia de un óptimo local la constituyen las *Second Order Sufficient Conditions* (SOSC), que generalizan la idea de que la Hessiana sea semidefinida positiva. **Tampoco las vemos en el curso.**

# Otro Caso (optimización con igualdades)

Cuando no hay desigualdades (i.e. $\{g_i\}=\emptyset$), el problema queda:

$$(P)\begin{cases} \min f(x) \\ h_j(x) = 0, \quad 1 \leq j \leq q \end{cases},$$

y las condiciones KKT son las condiciones de Lagrange:

$$\begin{cases} \nabla f_{\bar{\mu}}(\bar{x}) = \nabla f(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = \vec{0} \\ h_j(\bar{x}) = 0, \ 1 \leq j \leq q \end{cases}$$

Pensemos en un caso simple en $\mathbb{R}^2$, con una única restricción:

$$(P)\begin{cases} f(x,y) = x^2 - 2x + 4y^2 \\ xy = 4 \ \text{(i.e. } g(x,y) = xy - 4 = 0) \end{cases}.$$

# Otro Caso (optimización con igualdades)

Pensemos en un caso simple en $\mathbb{R}^2$, con una única restricción:

$$(P)\begin{cases} f(x,y) = x^2 - 2x + 4y^2 \\ xy = 4 \ \text{(i.e. } g(x,y) = xy - 4 = 0) \end{cases}.$$

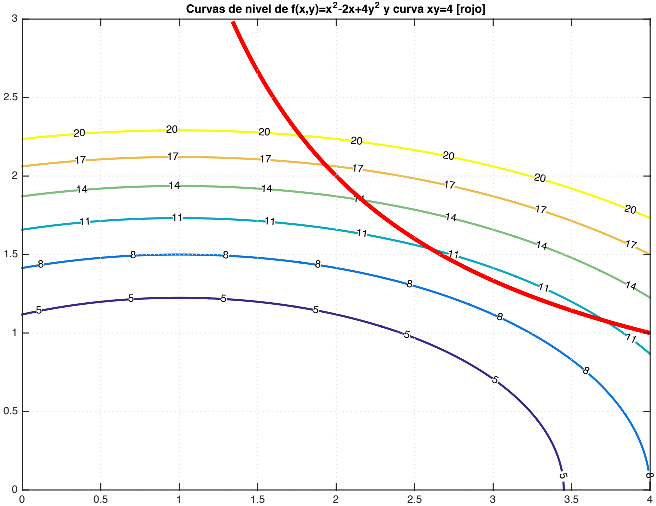
Las ecuaciones quedan: $\nabla f_\mu(x,y) = \begin{bmatrix} 2x-2 \\ 8y \end{bmatrix} + \mu \begin{bmatrix} y \\ x \end{bmatrix} = \vec{0}$,

con $xy = 4$. Luego $y = \frac{4}{x}$, $\frac{32}{x} + \mu x = 0 \Rightarrow \mu = -\frac{32}{x^2}$. Finalmente, $2x - 2 - \frac{32}{x^2} \cdot \frac{4}{x} = 0$, así que $x^3(x-1) - 64 = 0$.

La raíz está entre 3 y 4. Podemos usar Newton, y conseguimos la secuencia $x_0 = 3$, $x_1 = \frac{253}{81}$, $x_2 = \frac{2717}{872}$, $x_3 = \frac{6431}{2064}$. La diferencia entre los últimos es $3.1e^{-5}$, así que $x \approx 3.1158$, $y \approx 1.2838$, $\mu \approx -3.2962$ y el óptimo vale $f(\bar{x}, \bar{y}) \approx 10.069$.

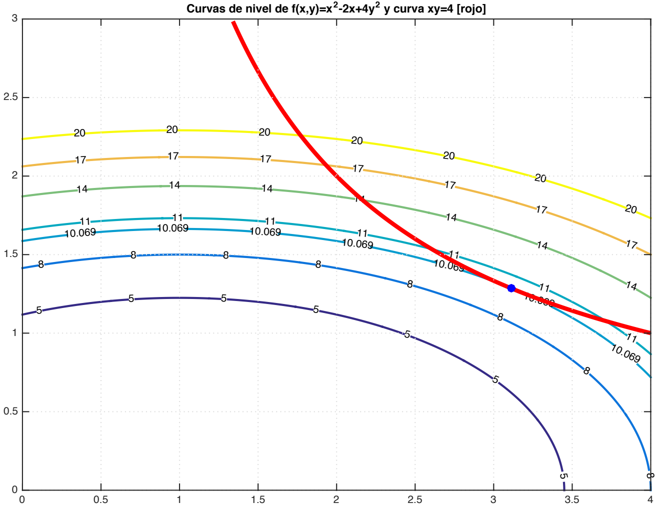
# Otro Caso (optimización con igualdades)

¿Cómo podríamos reconocer la solución intuitivamente?

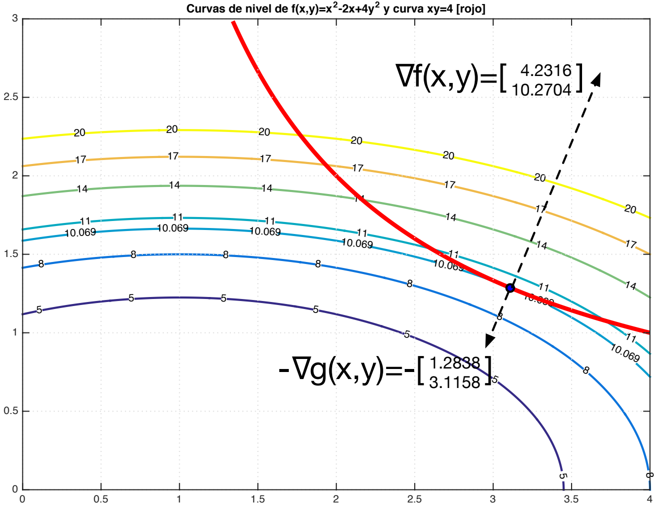


# Otro Caso (optimización con igualdades)

¿Cómo podríamos reconocer la solución intuitivamente?



# Otro Caso (optimización con igualdades)




Buscando que las normales a $xy=4$ y $f(x,y)$ sean colineales.

# Necesidad de las condiciones KKT

Que las condiciones KKT sean necesarias se sustenta en la geometría de la región factible, más que en argumentos del cálculo.

Pensemos en el caso con desigualdades (el otro es Lagrange).

Si el óptimo se alcanza en el interior de la región factible, el problema es *sin restricciones* luego $\bar{\lambda} = \vec{0}$ y $\nabla f(\bar{x}) = \vec{0}$, entonces estamos como en la primera parte del curso.

Estudiamos el caso en que $\bar{x}$ está en el frontera.

**Definición (Cono Normal en un $\bar{x}$ de la frontera)**

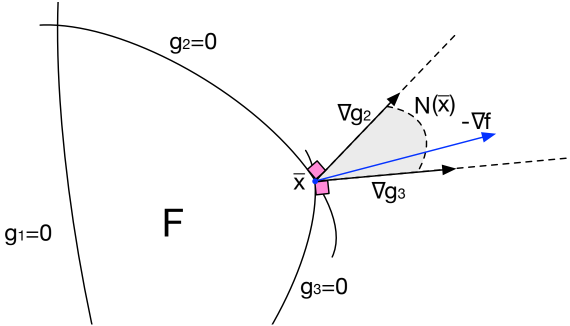
*Dado $(P)$ y $\bar{x} \in F$ (su reg. fact.) tal que $I_a \subseteq \{1, \ldots, p\}$ es el conjunto donde $g_i(\bar{x}) = 0$ sii $i \in I_a$ (i.e. restricciones activas). Llamamos $\mathcal{N}(\bar{x}) = \{v \in \mathbb{R}^n : v = \sum_{i \in I_a} \alpha_i \nabla g_i(\bar{x}), \text{ con } \alpha_i \geq 0\}$*

Observar que $g_i(\bar{x}) < 0$ si $i \notin I_a$ porque $\bar{x}$ es factible en $(P)$.

# Necesidad de las condiciones KKT

## Definición (Cono Normal en un $\bar{x}$ de la frontera)

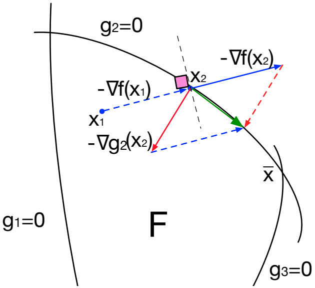
*Dado $(P)$ y $\bar{x} \in F$ (su reg. fact.) tal que $I_a \subseteq \{1, \ldots, p\}$ es el conjunto donde $g_i(\bar{x}) = 0$ sii $i \in I_a$ (i.e. restricciones activas). Llamamos $\mathcal{N}(\bar{x}) = \{v \in \mathbb{R}^n : v = \sum_{i \in I_a} \alpha_i \nabla g_i(\bar{x}),\ con\ \alpha_i \geq 0\}$*



# Necesidad de las condiciones KKT

Si pensamos en $-\nabla f(x)$ como una fuera que actúa sobre una partícula $x \in F$, ella tiende a moverse en esa dirección hasta alcanzar una pared/restricción $i$. Llegado ese punto, la pared aplica una fuerza de reacción hacia el interior, normal a la pared, esto es, colineal a $-\nabla g_i(x)$.

Partiendo de $x_1$ nos moveríamos hasta $x_2$, para recibir una reacción $-\lambda_2 \nabla g_2(x_2)$ de la pared, que combinada con $-\nabla f(x_2)$, nos conduciría en la dirección verde. Al llegar a $\bar{x}$ ya no habría una dirección de decrecimiento interior a $F$, por la fuerza normal adicional producto de $g_3(\bar{x})$.



# Necesidad de las condiciones KKT

**Definición (Cono Normal en un $\bar{x}$ de la frontera)**

*Dado $(P)$ y $\bar{x} \in F$ (su reg. fact.) tal que $I_a \subseteq \{1, \ldots, p\}$ es el conjunto donde $g_i(\bar{x}) = 0$ sii $i \in I_a$ (i.e. restricciones activas). Llamamos $\mathcal{N}(\bar{x}) = \{v \in \mathbb{R}^n : v = \sum_{i \in I_a} \alpha_i \nabla g_i(\bar{x}),\ \text{con } \alpha_i \geq 0\}$*

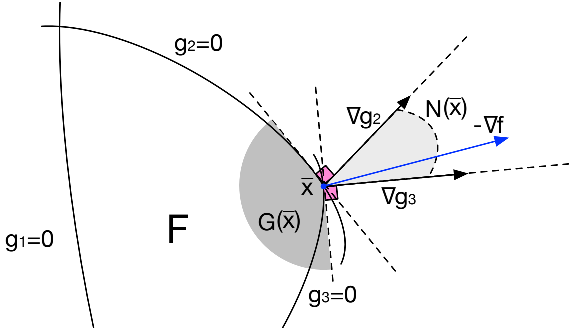
Alcanzado el punto $\bar{x}$, existe una combinación lineal de fuerzas que se anula: $-\nabla f(\bar{x}) - \lambda_2 \nabla g_2(\bar{x}) - \lambda_3 \nabla g_3(\bar{x}) = \vec{0}$, es decir, $-\nabla f(\bar{x}) = \lambda_2 \nabla g_2(\bar{x}) + \lambda_3 \nabla g_3(\bar{x})$, con $\lambda_2, \lambda_3 > 0$.

La condición $\nabla f_{\bar{\lambda}}(\bar{x}) = \vec{0}$ con $\bar{\lambda} \geq 0$ equivale a que la dirección de máximo decrecimiento de $f$ en $\bar{x}$ esté en $\mathcal{N}(\bar{x})$.

Se cumple que el cono $\mathcal{G}(\bar{x}) = \{v \in \mathbb{R}^n : \nabla^T g_i(\bar{x}) \cdot v \leq 0, \forall i \in I_a\}$ (el Cono Dual) es el ortogonal a $\mathcal{N}(\bar{x})$, esto es $\mathcal{G}(\bar{x}) = \mathcal{N}^{\perp}(\bar{x})$.

# Necesidad de las condiciones KKT

Se cumple que el cono $\mathcal{G}(\bar{x}) = \{v \in \mathbb{R}^n : \nabla^T g_i(\bar{x}) \cdot v \leq 0, \forall i \in I_a\}$ (el Cono Dual) es el ortogonal a $\mathcal{N}(\bar{x})$, esto es $\mathcal{G}(\bar{x}) = \mathcal{N}^{\perp}(\bar{x})$.




Para que KKT sea necesario, basta que la tangente a cualquier trayectoria originada en $\bar{x}$ esté contenida en $\mathcal{G}(\bar{x})$: que las direcciones para decrecer sean dentro del cono normal, fuera de $F$.

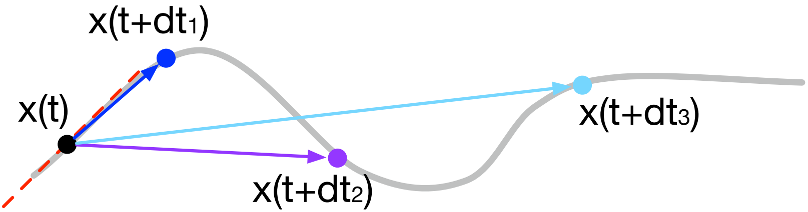
# Repaso sobre curvas en $\mathbb{R}^n$

- Llamaremos *curva* a toda función $x : \mathbb{R} \to \mathbb{R}^n$ diferenciable.
- Sin pérdida de generalidad, asumimos que $x : [0, 1) \to \mathbb{R}^n$.
- En el curso hemos hecho uso de curvas, concretamente, si en un método de descenso tomamos una dirección $\vec{d_n}$ desde un punto $x_n$, la función de descenso $h(t) = f(x_n + \vec{d_n} \cdot t)$ es en realidad $h(t) = f(x(t))$ con $x(t) = x_n + \vec{d_n} \cdot t$.
- En ese caso es claro que la dirección tangente es única e igual a $\vec{d_n}$, y las rectas han sido suficientes para modelar e implementar los métodos de descenso.
- Sin embargo, no son exhaustivas a la hora de capturar algunas propiedades, como la condición de ser un óptimo local.

# Repaso sobre curvas en $\mathbb{R}^n$

**Definición (Tangente a una curva $x(t)$ en un punto $\bar{x}$)**

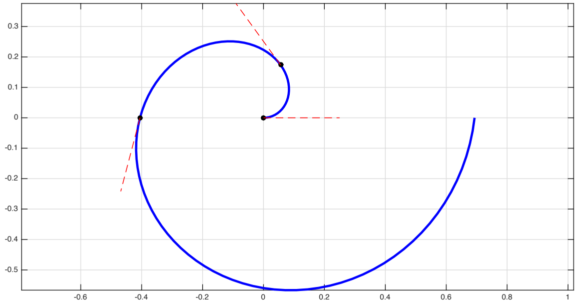
*Dada una curva $x : \mathbb{R} \to \mathbb{R}^n$ (diferenciable) y un punto $\bar{x} = x(\bar{t})$, llamamos tangente a la curva $x$ en $\bar{x}$ al $\lim_{dt \to 0} \frac{x(\bar{t} + dt) - x(\bar{t})}{dt}$.*




Al derivar $h(t) = f(x_n + \vec{d_n} \cdot t)$ en la búsqueda lineal, aplicábamos la regla de la cadena: $h'(t) = \nabla^T f(x(t)) \cdot x'(t)$, seguida del hecho que $x'(t) = \vec{d_n}$. La regla vale incluso si $x(t)$ no es una recta.

# Repaso sobre curvas en $\mathbb{R}^n$

Consideremos la curva $x(t) = \log(1+t)\begin{bmatrix}\cos(2\pi t)\\ \sin(2\pi t)\end{bmatrix}$, $t \geq 0$, de tangente: $\dot{x}(t) = \frac{1}{1+t}\begin{bmatrix}\cos(2\pi t)\\ \sin(2\pi t)\end{bmatrix} + 2\pi\log(1+t)\begin{bmatrix}-\sin(2\pi t)\\ \cos(2\pi t)\end{bmatrix}$.




En la figura se han representado los puntos correspondientes a $t = 0$, $t = 0.2$ y $t = 0.5$, y sus tangentes divididas por $4||\dot{x}(t)||$.

# Repaso sobre curvas en $\mathbb{R}^n$

Y para la $x(t)$ anterior, consideremos las funciones $f, g : \mathbb{R}^2 \to \mathbb{R}$ definidas como:

$$g(x) = \begin{cases} -1 & \text{existe } \; t \backslash x(t) = x \\ 0 & \text{otro caso} \end{cases} \quad , \quad f(x) = ||x||_2 \cdot g(x) \; .$$

Observar que $x_0 = (0, 0)$ no es un mínimo local, ya que en todo entorno existen puntos de la curva con valores negativos.

Pero sí aparenta serlo, cuando nos limitamos a las rectas que parten de $x_0$, ya que sobre ellas no hay decrecimiento.

El extender la condición para poder incluir a $x(t)$, verificamos que $h(t) = f(x(t))$ cumple $h'(0) = -1 < 0$.

**La anterior es la razón por la cual necesitamos considerar todas las curvas admisibles para clasificar un óptimo.**

# Necesidad de las condiciones KKT

**Definición (Cono Tangente en un punto $\bar{x}$ de la frontera)**

*Dado $(P)$ de región factible $F$ no-trivial y $\bar{x}$ en su borde, llamamos $\mathcal{T}(\bar{x}) = \{\vec{d} \in \mathbb{R}^n : \vec{d}$ tangente a una curva admisible desde $\bar{x}\}$ al Cono Tangente en $\bar{x}$. Decimos que el punto $\bar{x}$ es Regular en $(P)$ cuando $\mathcal{T}(\bar{x}) = \mathcal{G}(\bar{x})$.*

**Lema (Farkas-Minkowski)**

*Sea $A \in \mathbb{R}^{m\times n}$ no singular cualquiera, con $n > m$ y $b \in \mathbb{R}^m$. Se cumple que $\exists x \geq 0$ $(x \in \mathbb{R}^n)$, tal que $Ax = b$ si y sólo si, $b^T u \geq 0$ para todo $u \in \mathbb{R}^m$ tal que $A^T u \geq 0$.*

Lo probaremos como parte del práctico, con argumentos puramente algebraicos en el problema de Programación Lineal.

# Necesidad de las condiciones KKT

## Teorema (Las condiciones KKT son necesarias)

*En un problema $(P)$ de región factible no-trivial, donde las funciones $f$ y $g_i$ son diferenciables, si existe un óptimo local de $(P)$ en $\bar{x}$, punto regular en la frontera de $(P)$, las condiciones KKT deben cumplirse.*

**PRUEBA:** Sabemos que $g_i(\bar{x}) \leq 0$ porque $\bar{x} \in F$. Hay que probar que existe $\bar{\lambda} = \{\lambda_i\}$ que cumple KKT. Si $i \notin I_a$ (no activa), tomo $\lambda_i = 0$ y nos concentramos en las activas.

Cualquier curva admisible puede representarse como $x : [0, 1) \to \mathbb{R}^n$ diferenciable donde: $x(t) \in F, t \in [0, 1)$ y $x(0) = \bar{x}$. Como $\bar{x}$ es óptimo local, debe verificarse que $h'(0) \geq 0$, donde $h(t) = f(x(t))$, esto es: la función $f$ no puede decrecer sobre curvas admisibles.

# Necesidad de las condiciones KKT

**Teorema (Las condiciones KKT son necesarias)**

*En un problema $(P)$ de región factible no-trivial, donde las funciones $f$ y $g_i$ son diferenciables, si existe un óptimo local de $(P)$ en $\bar{x}$, punto regular en la frontera de $(P)$, las condiciones KKT deben cumplirse.*

**PRUEBA:** Luego, $h'(0) = \nabla^T f(x(0)) \cdot x'(0) = \nabla^T f(\bar{x}) \cdot \vec{d} \geq 0$, para todo $\vec{d} \in \mathcal{T}(\bar{x}) = \mathcal{G}(\bar{x})$ porque $\bar{x}$ es regular.
Luego $\nabla^T f(\bar{x}) \cdot \vec{d} \geq 0$ para todo $\vec{d}$ tal que $\nabla^T g_i(\bar{x}) \cdot \vec{d} \leq 0, i \in I_a$.
Tomo $b = \nabla f(\bar{x})$ y $A = [\cdots | -\nabla g_i(\bar{x}) | \cdots]$ y se cumple que: $b^T \cdot \vec{d} \geq 0$ para todo $\vec{d}$ tal que $A^T \cdot \vec{d} \geq 0$.
Por Farkas-Minkowski, existe $\bar{\alpha} \geq 0$ ($\bar{\alpha} = \{\alpha_i\}$) tal que $A \cdot \bar{\alpha} = b$, que equivale a: $\nabla f(\bar{x}) = -\sum_{i \in I_a} \alpha_i \nabla g_i(\bar{x})$.

# Necesidad de las condiciones KKT

**Teorema (Las condiciones KKT son necesarias)**

*En un problema $(P)$ de región factible no-trivial, donde las funciones $f$ y $g_i$ son diferenciables, si existe un óptimo local de $(P)$ en $\bar{x}$, punto regular en la frontera de $(P)$, las condiciones KKT deben cumplirse.*

**PRUEBA:** Completando $\bar{\lambda}$ (era $\lambda_i = 0$ si $i \notin I_a$) con $\lambda_i = \alpha_i$ cuando $i \in I_a$, concluimos que $\nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) = \vec{0}$ con $\lambda_i \geq 0, 1 \leq i \leq p$.

Eso prueba las ecuaciones 1 a 3 de KKT. La complementaridad (ecuación 4) se cumple por construcción, ya que $\lambda_i = 0$ si $i \notin I_a$ (i.e. si $g_i(\bar{x}) < 0$), y si $\lambda_i > 0$ es porque $i \in I_a$ y $g_i(\bar{x}) = 0$.

# Necesidad de las condiciones KKT

**Teorema (Las condiciones KKT son necesarias)**

*En un problema $(P)$ de región factible no-trivial, donde las funciones $f$ y $g_i$ son diferenciables, si existe un óptimo local de $(P)$ en $\bar{x}$, punto regular en la frontera de $(P)$, las condiciones KKT deben cumplirse.*

Así como la Invexidad y las Second Order Sufficient Conditions garantizaban suficiencia de las condiciones KKT, hay condiciones de regularidad (llamadas *constraint qualifications*) para las cuales las condiciones KKT son necesarias.

Tampoco las veremos en este curso porque las regiones factibles convexas no-triviales las cumplen.

# Necesidad y suficiencia de las condiciones KKT (convexidad)

**Teorema (Las condiciones KKT son necesarias en el caso convexo)**

*En un problema $(P)$ donde las $f$ y $g_i$ son convexas, las $h_j$ afines, $F$ es no-trivial, y las funciones son diferenciables, las condiciones KKT son necesarias y suficientes para la existencia de óptimos.*

La suficiencia ya la vimos. La necesidad se debe a que los dominios convexos no triviales cumplen las condiciones de regularidad, y el dominio $F$ de $(P)$ es la intersección de convexos, ya que las restricciones $g_i(x) \leq 0$ definen conjuntos convexos ($g_i$'s son convexas), las $h_j(x) = 0$ también (son afines) y la región factible es entonces intersección de conexos.

Que el conjunto sea no-trivial equivale a que exista $\hat{x}$ tal que $g_i(\hat{x}) < 0$ con $i = 1, \ldots, p$, **y debe cumplirse para garantizar que las hipótesis del teorema se cumplen.**

# Necesidad y suficiencia de las condiciones KKT (convexidad)

**Teorema (Las condiciones KKT son necesarias en el caso convexo)**

*En un problema $(P)$ donde las $f$ y $g_i$ son convexas, las $h_j$ afines, $F$ es no-trivial, y las funciones son diferenciables, las condiciones KKT son necesarias y suficientes para la existencia de óptimos.*

Pensar en $(P)\begin{cases}\min y\\ (x+1)^2+y^2 \leq 1\\ (x-1)^2+y^2 \leq 1\end{cases}$. La función objetivo y las restricciones son diferenciables y convexas. Existe un único punto factible $(0,0)$ que por lo tanto es óptimo, y sin embargo, no hay solución a las condiciones KKT para el problema $(P)$.

Verificar que $\nabla f_{\lambda}(0,0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} -2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (no se cumple KKT) y que el punto factible es único.

# Otro ejemplo de Kuhn-Tucker

Tomemos ahora como referencia a $(M) = \begin{cases} \min f(x) \\ x \geq 0 \end{cases}, x \in \mathbb{R}^n$,

que equivale a $(M') = \begin{cases} \min f(x) \\ g_i(x) = -x_i \leq 0 \end{cases}$.

Usemos KKT tomando $f_\lambda(x) = f(x) - \sum_{i=1}^{n} \lambda_i x_i$ y llegamos a:

$$\begin{cases} \nabla f(x) - \lambda = \vec{0} \\ x_i \geq 0, \lambda_i \geq 0 \\ \lambda_i x_i = 0 \end{cases} \quad \text{que equivale a} \begin{cases} \nabla^T f(x) \cdot x = 0 \\ \nabla f(x) \geq 0, x \geq 0 \end{cases}.$$

La primera igualdad $x_1 \frac{\partial f(x)}{\partial x_1} + \cdots + x_n \frac{\partial f(x)}{\partial x_n} = 0$ reemplaza a las anteriores porque: $\lambda_i = \frac{\partial f(x)}{\partial x_i}$, y el hecho que la suma sea 0 implica que cada producto lo es, ya que los operandos son no-negativos.

OBS: Cuando $f : \mathbb{R} \to \mathbb{R}$, la condición anterior (caso convexo) equivale a que exista $\bar{x} > 0$ con $\frac{df(\bar{x})}{dx} = 0$, o que $\frac{df(0)}{dx} > 0$.

# Otro ejemplo de Kuhn-Tucker

Usamos las condiciones $\begin{cases} \nabla^T f(x) \cdot x = 0 \\ \nabla f(x) \geq 0, x \geq 0 \end{cases}$ para resolver:

$$(M) \begin{cases} \min\, (x+2)^2 + 2(y+1)^2 \\ x, y \geq 0 \end{cases}$$

$\nabla f(x,y) = \begin{bmatrix} 2(x+2) \\ 4(y+1) \end{bmatrix}$, de donde: i) $2x(x+2) + 4y(y+1) = 0$, ii) $2(x+2) \geq 0$, iii) $4(y+1) \geq 0$ y iv) $x, y \geq 0$. Si se cumple iv) se deben cumplir ii) y iii), así que tenemos que resolver:

$$\begin{cases} x(x+2) + 2y(y+1) = 0 \\ x, y \geq 0 \end{cases},$$

cuya única solución es $(\bar{x}, \bar{y}) = (0, 0)$. Así, el valor óptimo es 6.

**Resolver con *quadprog* (Matlab) o *pqpnonneg/qp* (Octave).**

# Programación Lineal y su Forma Canónica

**Definición**

*Diremos que un problema de Programación Matemática es de Programación Lineal, cuando las funciones $f$, $g_i$, $h_j$ son lineales.*

Observar que en los problemas de Programación Lineal (LP), las funciones son diferenciales y afines (las $g_i$ son convexas), por tanto, las condiciones de KKT son necesarias y suficientes. Diremos que un LP está en su *Forma Canónica*, cuando:

$$(LP) \begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases} , A \in \mathbb{R}^{m \times n},$$

Si hubiera alguna restricción lineal de igualdad $h^T x = d$, puedo escribirla con dos ecuaciones $h^T x \geq d$ y $h^T x \leq d$, siendo la última equivalente a $[-h]^T x \geq [-d]$.

# Programación Lineal y su Forma Canónica

Diremos que un LP está en su *Forma Canónica*, cuando:

$$(LP)\begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases}, A \in \mathbb{R}^{m\times n},$$

Si hubiera alguna restricción lineal de igualdad $h^T x = d$, puedo escribirla con dos ecuaciones $h^T x \geq d$ y $h^T x \leq d$, siendo la última equivalente a $[-h]^T x \geq [-d]$.

Si la restricción de signo de una variable fuera $x_i \leq 0$, tomo el cambio de variables $x_i' = -x_i \geq 0$, y si $x_i$ no tuviera restricción de signo, uso en cambio dos variables positivas y tomo $x_i = x_i' - x_i''$ agregando una ecuación que siempre tiene solución.

**Puedo pensar entonces que cualquier (LP) puede escribirse en forma canónica.**

# Condiciones Teóricas en la Programación Lineal

Para (LP) en su forma canónica, relajo usando $\lambda \in \mathbb{R}^m$, $\lambda_j \geq 0$:

$$(LP_\lambda) \begin{cases} \min b^T\lambda + (c - A^T\lambda)^T x \\ x \geq 0 \end{cases}$$

Hallamos el óptimo de $(LP_\lambda)$ con $f_\lambda(x) = b^T\lambda + (c - A^T\lambda)^T x$ y las condiciones KKT como en "Otro ejemplo de Kuhn-Tucker".
Como $\nabla f_\lambda(x) = c - A^T\lambda$, llegamos a: $(c - A^T\lambda)^T x = 0$, con $A^T\lambda \leq c$ y $x \geq 0$. Para que se cumpla el TMFRL y tengamos óptimo en (LP) hay que agregar: $\lambda \geq 0$, $Ax \geq b$ y $(Ax - b)^T\lambda = 0$.
Las condiciones KKT para (LP) en su forma canónica son:

$$\begin{cases} Ax \geq b, & A^T\lambda \leq c, \\ x \geq 0, & \lambda \geq 0, \\ (Ax-b)^T\lambda=0, & (c-A^T\lambda)^T x=0 \end{cases}$$

# Condiciones Teóricas en la Programación Lineal

Para (LP) en su forma canónica, relajo usando $\lambda \in \mathbb{R}^m$, $\lambda_j \geq 0$:

$$(LP_\lambda) \begin{cases} \min b^T\lambda + (c - A^T\lambda)^T x \\ x \geq 0 \end{cases}$$

Las condiciones KKT para (LP) en su forma canónica son:

$$\begin{cases} Ax \geq b, & A^T\lambda \leq c, \\ x \geq 0, & \lambda \geq 0, \\ (Ax-b)^T\lambda=0, & (c-A^T\lambda)^T x=0 \end{cases}$$

Las restricciones remarcadas aseguran que a $\lambda$ dado, $x = x_\lambda$ es óptimo del problema $(LP_\lambda)$.

# Condiciones Teóricas en la Programación Lineal

Para (LP) en su forma canónica, relajo usando $\lambda \in \mathbb{R}^m$, $\lambda_j \geq 0$:

$$(LP_\lambda) \begin{cases} \min b^T\lambda + (c - A^T\lambda)^T x \\ x \geq 0 \end{cases}$$

Las condiciones KKT para (LP) en su forma canónica son:

$$\begin{cases} Ax \geq b, & A^T\lambda \leq c, \\ x \geq 0, & \lambda \geq 0, \\ (Ax-b)^T\lambda=0, & (c-A^T\lambda)^T x=0 \end{cases}$$

Las restricciones remarcadas aseguran que a $\lambda$ dado, $x = x_\lambda$ es óptimo del problema $(LP_\lambda)$.

Mientras que éstas aseguran que $x = x_\lambda$ también es óptimo del problema $(LP)$, ya que son las del TFRL.

# Ejemplo de Programación Lineal

Considere el problema: $\begin{cases} \min x+y \\ 2x+y \geq 2, \\ x+2y \geq 3, \\ x,y \geq 0 \end{cases}$ ¿Es $(\frac{1}{3},\frac{4}{3})$ su solución?

El punto cumple $Az=b$ con $z=(\frac{1}{3},\frac{4}{3})^T>0$, así que es óptimo si y sólo si existen $\lambda_1,\lambda_2 \geq 0$ tales que: $c=A^T\lambda = A\lambda$ porque $A$ es simétrica en este ejemplo.

Resolvemos $\begin{cases} 2\lambda_1+\lambda_2 &= 1 \\ \lambda_1+2\lambda_2 &= 1 \end{cases}$ y llegamos a que $\lambda_1=\lambda_2=\frac{1}{3}>0$ es la única solución del sistema, y por tanto es el único punto óptimo. El valor óptimo del problema es $\frac{5}{3}$.

**Resolver usando *linprog* (Matlab) o *glpk* (Octave).**

# Dualidad

Para lidiar con las restricciones, la estrategia usada hasta ahora ha pasado por relajarlas, resolver $(P_\lambda)$ con $\lambda \geq 0$ para un $(P)$ dado.

Ejemplo: $(P) \begin{cases} \min 2x^2 + 5y^2 + 2xy - 12x - 8y + 10 \\ x + y \leq 3 \end{cases}$

Así, $f_\lambda(x,y) = 2x^2 + 5y^2 + 2xy + (\lambda - 12)x + (\lambda - 8)y + 10 - 3\lambda$ y sabemos que para $\lambda = 0$, el óptimo de $(P_\lambda)$ es el de $f$ (ver BFGS), por lo que $f_0 = -8.2222$ y se alcanza en $(\frac{26}{9}, \frac{2}{9})$.

Podemos verificar que si $\lambda = 1$, el óptimo es $f_1(\frac{8}{3}, \frac{1}{6}) = -8.25 \leq f(\frac{8}{3}, \frac{1}{6}) = -8.0833$ y para $f_{0.5}(\frac{25}{9}, \frac{7}{36}) = -8.2014 \leq f(\frac{25}{9}, \frac{7}{36}) = -8.1875$. Si $\lambda \geq 0$, una solución $\bar{z}_\lambda$ de $(P_\lambda)$ establece una cota inferior de $(P)$. Cuando además $\bar{z}_\lambda$ es factible en $(P)$, $f(\bar{z}_\lambda)$ es una cota superior al óptimo de ambos problemas.

**¿Cuál es el máximo entre las relajaciones?**

# Dualidad

## Definición

*Dado un problema $(P)$ de programación matemática y una relajación $(P_{\lambda,\mu})$, llamaremos Dual de esa relajación a $(D)$ $\{\max \Phi(\lambda,\mu) | \lambda \geq 0\}$, donde $\Phi(\lambda,\mu)$ es el mínimo de $(P_{\lambda,\mu})$.*

Ejemplo: $(P)\begin{cases} \min 2x^2 + 5y^2 + 2xy - 12x - 8y + 10 \\ x + y \leq 3 \end{cases}$

En este caso, podemos verificar que $\bar{z}_\lambda = [\frac{26}{9} - \frac{2\lambda}{9}, \frac{2}{9} - \frac{\lambda}{18}]$, y por tanto: $\Phi(\lambda) = -\frac{5}{36}\lambda^2 + \frac{\lambda}{9} - \frac{74}{9}$. El problema dual es:

$$(D)\begin{cases} \max_\lambda -\frac{5}{36}\lambda^2 + \frac{\lambda}{9} - \frac{74}{9} \\ \lambda \geq 0 \end{cases}$$

El problema $(D)$ tiene una sola variable y sabemos que el óptimo no es en $\lambda$=0, así que $\Phi'(\bar{\lambda}) = -\frac{5}{18}\bar{\lambda} + \frac{1}{9} = 0$, $\bar{\lambda} = \frac{2}{5}$ y $\bar{z}_{\bar{\lambda}} = (\frac{14}{5}, \frac{1}{5})$.

# Dualidad

Ejemplo: $(P)\begin{cases} \min 2x^2+5y^2+2xy-12x-8y+10 \\ x+y\leq 3 \end{cases}$

En este caso, podemos verificar que $\bar{z}_\lambda = [\frac{26}{9}-\frac{2\lambda}{9}, \frac{2}{9}-\frac{\lambda}{18}]$, y por tanto: $\Phi(\lambda) = -\frac{5}{36}\lambda^2+\frac{\lambda}{9}-\frac{74}{9}$. El problema dual es:

$$(D)\begin{cases} \max_\lambda -\frac{5}{36}\lambda^2+\frac{\lambda}{9}-\frac{74}{9} \\ \lambda\geq 0 \end{cases}$$

El problema $(D)$ tiene una sola variable y sabemos que el óptimo no es en $\lambda{=}0$, así que $\Phi'(\bar{\lambda}){=}-\frac{5}{18}\bar{\lambda}{+}\frac{1}{9}=0$, $\bar{\lambda}{=}\frac{2}{5}$ y $\bar{z}_{\bar{\lambda}}{=}(\frac{14}{5},\frac{1}{5})$. Verificar que $\bar{\lambda}$ y $\bar{z}_{\bar{\lambda}}$ anteriores cumplen KKT para $(P)$. En primer lugar, debe ser $x+y=3$ (porque $\lambda>0$) (eso se cumple). Además, $\nabla f_\lambda(x,y) = \begin{bmatrix} 4x+2y+\lambda-12 \\ 2x+10y+\lambda-8 \end{bmatrix}$ se anula en el punto, luego $\bar{\lambda}$ y $\bar{z}_{\bar{\lambda}}$ cumplen KKT, el óptimo de $(P)$ es -$\frac{41}{5}$ y se alcanza en $(\frac{14}{5},\frac{1}{5})$.

# Dualidad (observaciones)

Aun sin conocer el óptimo de $(P)$ en el problema anterior, podíamos asegurar que era mayor o igual a -8.25 y -8.2014, ya que los $\bar{x}$'s asociados eran primal-factibles: $\frac{24}{9} + \frac{1}{6} \approx 2.8333 < 3$, y $\frac{25}{9} + \frac{7}{36} \approx 2.9722 < 3$. En particular, el segundo valor termina estando a menos de 0.02% de error.

Lo anterior no se cumple cuando $\lambda = 0$, porque en ese caso, el punto optimal $(\frac{26}{9}, \frac{2}{9})$ no está en la región factible de $(P)$, ya que $\frac{26}{9} + \frac{2}{9} \approx 3.1111 > 3$.

De hecho, también podíamos adelantar que el óptimo de $f(x, y)$ (sin restricciones) debía ser menor o igual que el de $(P)$, porque al no tener restricciones es una relajación.

**Veremos que la función $\Phi(\lambda)$ es cóncava.**

# Dualidad (Propiedades)

## Propiedades

*Dado un problema $(P)$ y su dual $(D)$, si $x$ y $\lambda$ son factibles en $(P)$ y $(D)$ respectivamente, se cumple:*

- $\Phi(\lambda) \leq f(x)$
- *Si* $\Phi(\lambda) = f(x)$ *entonces* $x$ *y* $\lambda$ *son óptimos*
- *Si* $f$ *y* $g_i$ *son convexas,* $h_j$ *afines,* $\exists x_0, g_i(x_0) < 0$ *(no-trivial), y existe* $\bar{x}$ *óptimo de* $(P)$*, entonces existe* $\bar{\lambda}$ *óptimo de* $(D)$*, tal que* $\Phi(\bar{\lambda}) = f(\bar{x})$.

Como consecuencia de la primera propiedad, podemos asegurar que si $(P)$ no es acotado inferiormente entonces $(D)$ no es factible.

Análogamente, si $(D)$ no es acotado superiormente, el problema $(P)$ es no-factible (el conjunto factible es vacío).

# Dualidad (Propiedades)

Como consecuencia de la primera propiedad, podemos asegurar que si $(P)$ no es acotado inferiormente entonces $(D)$ no es factible.

Análogamente, si $(D)$ no es acotado superiormente, el problema $(P)$ es no-factible (el conjunto factible es vacío).

La segunda propiedad se conoce como *Dualidad Débil* y la tercera como *Dualidad Fuerte*.

Las relajaciones y la dualidad permiten determinar que en un paso de un método se encontró el óptimo, o estimar cotas del error en ese paso, de disponer de puntos factibles en ambos problemas.

En algunos problemas se conoce el dual en forma teórica, y las propiedades anteriores son de extrema utilidad.

La Programación Lineal y la Cuadrática caen en el caso anterior.

# Dualidad (Propiedades)

## Propiedad

*Dado un problema $(P)$ de programación matemática, la función $\Phi(\lambda, \mu)$ de su dual $(D)$ es cóncava.*

**PRUEBA:** $\Phi(\lambda, \mu) = \min_x\{f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x)\}$.
Para un $x$ fijo cualquiera, $f_{\lambda,\mu}(x)$ es lineal en $\lambda$ y $\mu$, entonces es afin (cóncava y convexa).

Así como el máximo de funciones convexas es convexo, el mínimo de cóncavas es cóncava.

$\Phi(\lambda, \mu)$ es el mínimo de un conjunto de (infinitas) funciones cóncavas, y por tanto también es cóncava.

# Dualidad (Caso Lineal)

Consideramos la Forma Canónica: $(LP) \begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases}$, $A \in \mathbb{R}^{m \times n}$,

con $(LP_\lambda) \begin{cases} \min b^T\lambda + (c - A^T\lambda)^T x \\ x \geq 0 \end{cases}$, donde $\lambda \geq 0$.

Observar que si existiera $1 \leq j \leq m$ tal que $(c - A^T\lambda)_j < 0$, el problema $(LP_\lambda)$ no estaría bien definido, ya que basta tomar $x_j \to \infty$ (y $x_i = 0, i \neq j$) para ver que no es acotado inferiormente. Además, cuando $(LP_\lambda)$ sí está bien definido, la solución es $b^T\lambda$ y se alcanza tomando $x_j = 0$ para todo $(c - A^T\lambda)_j > 0$, de donde $(c - A^T\lambda)^T x = 0$ (condición de complementariedad).

Por tanto, el dual de (LP) es: $(D) \begin{cases} \max b^T\lambda \\ A^T\lambda \leq c \\ \lambda \geq 0 \end{cases}$

# Dualidad (Caso Lineal)

Consideramos la Forma Canónica: $(LP)\begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases}$, $A \in \mathbb{R}^{m\times n}$,

con $(LP_\lambda)\begin{cases} \min b^T\lambda + (c - A^T\lambda)^T x \\ x \geq 0 \end{cases}$, donde $\lambda \geq 0$.

Observar que si existiera $1 \leq j \leq m$ tal que $(c - A^T\lambda)_j < 0$, el problema $(LP_\lambda)$ no estaría bien definido, ya que basta tomar $x_j \to \infty$ (y $x_i = 0, i \neq j$) para ver que no es acotado inferiormente. Además, cuando $(LP_\lambda)$ sí está bien definido, la solución es $b^T\lambda$ y se alcanza tomando $x_j = 0$ para todo $(c - A^T\lambda)_j > 0$, de donde $(c - A^T\lambda)^T x = 0$ (condición de complementariedad).

**Observar además que si calculamos el dual del dual, volvemos a obtener el problema primal.**

# Dualidad (Caso Lineal)

La dualidad nos da una nueva interpretación de las condiciones KKT en este caso. Siendo $(P)\begin{cases}\min c^T x\\ Ax \geq b\\ x \geq 0\end{cases}$ y $(D)\begin{cases}\max b^T \lambda\\ A^T\lambda \leq c\\ \lambda \geq 0\end{cases}$,

cuando $(P)$ y $(D)$ son factibles, $(P)$ es acotado inferiormente.

Un problema lineal factible y acotado tiene óptimo (ver Simplex).

Como las funciones son convexas, se cumple la Dualidad Fuerte, y existe $\bar{\lambda} \in F_{(D)}$ tal que $c^T\bar{x} = b^T\bar{\lambda}$ (ambos son óptimos).

Se cumple $c^T\bar{x} - \bar{\lambda}^T A\bar{x} = (c - A^T\bar{\lambda})^T\bar{x} = 0$ (página anterior).

Luego, $c^T\bar{x} = \bar{\lambda}^T A\bar{x} = \bar{x}^T A^T\bar{\lambda} = b^T\bar{\lambda} \Rightarrow (A\bar{x} - b)^T\bar{\lambda} = 0$.

Así que KKT resulta de combinar que:

# Dualidad (Caso Lineal)

La dualidad nos da una nueva interpretación de las condiciones KKT en este caso. Siendo $(P)\begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases}$ y $(D)\begin{cases} \max b^T \lambda \\ A^T\lambda \leq c \\ \lambda \geq 0 \end{cases}$,

cuando $(P)$ y $(D)$ son factibles, $(P)$ es acotado inferiormente.

Un problema lineal factible y acotado tiene óptimo (ver Simplex).

Como las funciones son convexas, se cumple la Dualidad Fuerte, y existe $\bar{\lambda} \in F_{(D)}$ tal que $c^T\bar{x} = b^T\bar{\lambda}$ (ambos son óptimos).

Se cumple $c^T\bar{x} - \bar{\lambda}^T A\bar{x} = (c - A^T\bar{\lambda})^T\bar{x} = 0$ (página anterior).

Luego, $c^T\bar{x} = \bar{\lambda}^T A\bar{x} = \bar{x}^T A^T\bar{\lambda} = b^T\bar{\lambda} \Rightarrow (A\bar{x} - b)^T\bar{\lambda} = 0$.

Así que KKT resulta de combinar que:

**El punto $x$ sea primal-factible**

# Dualidad (Caso Lineal)

La dualidad nos da una nueva interpretación de las condiciones KKT en este caso. Siendo $(P)\begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases}$ y $(D)\begin{cases} \max b^T \lambda \\ A^T \lambda \leq c \\ \lambda \geq 0 \end{cases}$,

cuando $(P)$ y $(D)$ son factibles, $(P)$ es acotado inferiormente.

Un problema lineal factible y acotado tiene óptimo (ver Simplex).

Como las funciones son convexas, se cumple la Dualidad Fuerte, y existe $\bar{\lambda} \in F_{(D)}$ tal que $c^T \bar{x} = b^T \bar{\lambda}$ (ambos son óptimos).

Se cumple $c^T \bar{x} - \bar{\lambda}^T A\bar{x} = (c - A^T \bar{\lambda})^T \bar{x} = 0$ (página anterior).

Luego, $c^T \bar{x} = \bar{\lambda}^T A\bar{x} = \bar{x}^T A^T \bar{\lambda} = b^T \bar{\lambda} \Rightarrow (A\bar{x} - b)^T \bar{\lambda} = 0$.

Así que KKT resulta de combinar que:

**El punto $\lambda$ sea dual-factible**

# Dualidad (Caso Lineal)

La dualidad nos da una nueva interpretación de las condiciones KKT en este caso. Siendo $(P)\begin{cases}\min c^T x\\ Ax \geq b\\ x \geq 0\end{cases}$ y $(D)\begin{cases}\max b^T\lambda\\ A^T\lambda \leq c\\ \lambda \geq 0\end{cases}$,

cuando $(P)$ y $(D)$ son factibles, $(P)$ es acotado inferiormente.

Un problema lineal factible y acotado tiene óptimo (ver Simplex).

Como las funciones son convexas, se cumple la Dualidad Fuerte, y existe $\bar{\lambda} \in F_{(D)}$ tal que $c^T\bar{x} = b^T\bar{\lambda}$ (ambos son óptimos).

Se cumple $c^T\bar{x} - \bar{\lambda}^T A\bar{x} = (c - A^T\bar{\lambda})^T\bar{x} = 0$ (página anterior).

Luego, $c^T\bar{x} = \bar{\lambda}^T A\bar{x} = \bar{x}^T A^T\bar{\lambda} = b^T\bar{\lambda} \Rightarrow (A\bar{x} - b)^T\bar{\lambda} = 0$.

Así que KKT resulta de combinar que:

**Que se dé la Dualidad-Fuerte, que equivale a las complementariedades cruzadas: $(Ax - b)^T\lambda = 0$, y $(c - A^T\lambda)^T x = 0$.**

# Dualidad (Interpretación Económica)

Considere el problema de diseñar una dieta óptima para satisfacer las necesidades diarias de dos nutrientes $A$ y $B$, que ascienden a 13 y 8 unidades respectivamente.

Hay cinco alimentos posibles, cuyos costos unitarios y aportes de cada nutriente (por unidad consumida) son los siguientes:

| Alimento | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Costo | 11 | 7 | 9 | 12 | 10 |
| Aporte 1 | 1 | 0 | 2 | 1 | 1 |
| Aporte 2 | 0 | 1 | 1 | 2 | 1 |

**¿Cuál es problema** $(P)$ **y su dual** $(D)$**?**

# Dualidad (Interpretación Económica)

| **Alimento** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|
| Costo | 11 | 7 | 9 | 12 | 10 |
| Aporte 1 | 1 | 0 | 2 | 1 | 1 |
| Aporte 2 | 0 | 1 | 1 | 2 | 1 |

$$(P)\begin{cases} \min 11x_1 + 7x_2 + 9x_3 + 12x_4 + 10x_5 \\ x_1 + 2x_3 + x_4 + x_5 \geq 13, \\ x_2 + x_3 + 2x_4 + x_5 \geq 8, \\ x_i \geq 0. \end{cases}$$

$$(D)\begin{cases} \max 13\lambda_1 + 8\lambda_2 \\ \lambda_1 \leq 11, \\ \lambda_2 \leq 7, \\ 2\lambda_1 + \lambda_2 \leq 9, \\ \lambda_1 + 2\lambda_2 \leq 12, \\ \lambda_1 + \lambda_2 \leq 10, \\ \lambda_i \geq 0. \end{cases}$$

# Dualidad (Interpretación Económica)

$$(P)\begin{cases} \min 11x_1 + 7x_2 + 9x_3 + 12x_4 + 10x_5 \\ x_1 + 2x_3 + x_4 + x_5 \geq 13, \\ x_2 + x_3 + 2x_4 + x_5 \geq 8, \\ x_i \geq 0. \end{cases} \qquad (D)\begin{cases} \max 13\lambda_1 + 8\lambda_2 \\ \lambda_1 \leq 11, \\ \lambda_2 \leq 7, \\ 2\lambda_1 + \lambda_2 \leq 9, \\ \lambda_1 + 2\lambda_2 \leq 12, \\ \lambda_1 + \lambda_2 \leq 10, \\ \lambda_i \geq 0. \end{cases}$$

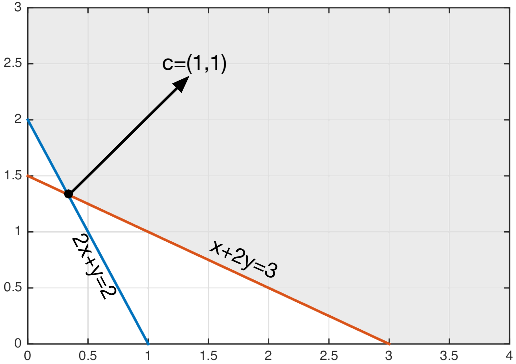
El problema $(D)$ es el de un fabricante de comprimidos para los nutrientes $A$ y $B$, que quiere determinar el precio de venta de cada comprimido para maximizar sus ingresos, siendo económicamente competitivo con la oferta de alimentos alternativa.

# Análisis de Sensibilidad

Observar que en un problema $(LP)$ canónico no-trivial, el óptimo no puede darse en el interior de la región factible porque $\nabla f(x) = c \neq \vec{0}$.

Entonces, el óptimo se alcanza en la frontera, y de hecho, lo esperable es que sea solamente en un vértice.

$$\begin{cases} \min x + y \\ 2x + y \geq 2, \\ x + 2y \geq 3, \\ x, y \geq 0 \end{cases}$$



# Análisis de Sensibilidad

Observar que en un problema $(LP)$ canónico no-trivial, el óptimo no puede darse en el interior de la región factible porque $\nabla f(x) = c \neq \vec{0}$.

Entonces, el óptimo se alcanza en la frontera, y de hecho, lo esperable es que sea solamente en un vértice.

Y podemos entonces perturbar ligeramente los datos del problema sin cambiar de vértice, ni en el problema ni en su dual.

Para una perturbación $\epsilon \in \mathbb{R}^m$, la dualidad fuerte indica que los óptimos de los problemas $P(\epsilon)$ y $D(\epsilon)$ coinciden:

$$(P(\epsilon)) \begin{cases} \min c^T x \\ Ax \geq b + \epsilon \\ x \geq 0 \end{cases} \text{ y } (D(\epsilon)) \begin{cases} \max (b+\epsilon)^T \lambda \\ A^T \lambda \leq c \\ \lambda \geq 0 \end{cases}$$

El análisis sobre cómo cambia $\Psi(\epsilon) = \min\{c^T x | Ax \geq b + \epsilon, x \geq 0\}$ es más simple en $D(\epsilon)$, si conocemos su óptimo $\bar{\lambda}$.

# Análisis de Sensibilidad

Y podemos entonces perturbar ligeramente los datos del problema sin cambiar de vértice, ni en el problema ni en su dual.

Para una perturbación $\epsilon \in \mathbb{R}^m$, la dualidad fuerte indica que los óptimos de los problemas $P(\epsilon)$ y $D(\epsilon)$ coinciden:

$$(P(\epsilon)) \begin{cases} \min c^T x \\ Ax \geq b + \epsilon \\ x \geq 0 \end{cases} \text{ y } (D(\epsilon)) \begin{cases} \max\, (b+\epsilon)^T \lambda \\ A^T \lambda \leq c \\ \lambda \geq 0 \end{cases}$$

El análisis sobre cómo cambia $\Psi(\epsilon) = \min\{c^T x | Ax \geq b + \epsilon, x \geq 0\}$ es más simple en $D(\epsilon)$, si conocemos su óptimo $\bar{\lambda}$.
Por la complementariedad, si $(Ax - b)_j > 0$, debe ser $\lambda_j = 0$, y una perturbación en esa componente no cambia el resultado: $\frac{\partial \Psi(0)}{\partial \epsilon_j} = 0$.

Si $\bar{\lambda}_j > 0$, se cumple $\Delta\Psi(\delta\epsilon_j) = \bar{\lambda}_j \delta\epsilon_j$. Así, $\frac{\partial \Psi(0)}{\partial \epsilon_j} = \bar{\lambda}_j$.

# Análisis de Sensibilidad

## Propiedad

*Dado* $\Psi(\epsilon,\rho)=\min\{f(x)|g_i(x)\leq\epsilon_i, h_j(x)=\rho_j\}$*, se cumple que:* $\frac{\partial\Psi(0,0)}{\partial\epsilon_i}=-\bar{\lambda}_i$ *y* $\frac{\partial\Psi(0,0)}{\partial\rho_j}=-\bar{\mu}_j$*, donde* $(\bar{\lambda},\bar{\mu})$ *componen el óptimo del dual de* $(P)=\min\{f(x)|g_i(x)\leq 0, h_j(x)=0\}$.

El $-\bar{\lambda}_i$ es porque ahora la restricción es $\leq$ (era $\geq$ en la otra).

Ejemplo: $(P(\epsilon))\begin{cases}\min 2x^2+5y^2+2xy-12x-8y+10\\ x+y\leq 3+\epsilon\end{cases}$

Observar que $\nabla f_\lambda$ no depende de $\epsilon$, así que el punto óptimo en $(P_\lambda(\epsilon))$ no cambia, sigue siendo: $\bar{z}_\lambda=[\frac{26}{9}-\frac{2\lambda}{9},\frac{2}{9}-\frac{\lambda}{18}]$.

Sí cambia $\Phi(\lambda)=-\frac{5}{36}\lambda^2+\frac{\lambda}{9}-\frac{74}{9}-\epsilon\lambda$, de donde $\bar{\lambda}=\frac{2}{5}-\frac{18\epsilon}{5}$, $\bar{z}_{\bar{\lambda}}=(\frac{14}{5}+\frac{4\epsilon}{5},\frac{1}{5}+\frac{\epsilon}{5})$ y $\Psi(\epsilon)=-\frac{41}{5}-\frac{2\epsilon}{5}+\frac{9\epsilon^2}{5}$.

# Análisis de Sensibilidad

El $-\bar{\lambda}_i$ es porque ahora la restricción es $\leq$ (era $\geq$ en la otra).

Ejemplo: $(P(\epsilon))\begin{cases} \min 2x^2+5y^2+2xy-12x-8y+10 \\ x+y\leq 3+\epsilon \end{cases}$

Observar que $\nabla f_\lambda$ no depende de $\epsilon$, así que el punto óptimo en $(P_\lambda(\epsilon))$ no cambia, sigue siendo: $\bar{z}_\lambda=[\frac{26}{9}-\frac{2\lambda}{9},\frac{2}{9}-\frac{\lambda}{18}]$.

Sí cambia $\Phi(\lambda)=-\frac{5}{36}\lambda^2+\frac{\lambda}{9}-\frac{74}{9}-\epsilon\lambda$, de donde $\bar{\lambda}=\frac{2}{5}-\frac{18\epsilon}{5}$, $\bar{z}_{\bar{\lambda}}=(\frac{14}{5}+\frac{4\epsilon}{5},\frac{1}{5}+\frac{\epsilon}{5})$ y $\Psi(\epsilon)=-\frac{41}{5}-\frac{2\epsilon}{5}+\frac{9\epsilon^2}{5}$.

Se verifica entonces que $\frac{\partial\Psi(0)}{\partial\epsilon}=-\frac{2}{5}=-\bar{\lambda}$.

Para verificar experimentalmente que $\frac{\partial\Psi(0)}{\partial\rho}=-\bar{\mu}$ (i.e. con una igualdad), observar que como $\bar{\lambda}>0$ en este caso, se cumple que $(P(\epsilon))$ es equivalente al problema con $x+y=3+\epsilon$.

Los valores $\frac{\partial\Psi(0,0)}{\partial\epsilon_i}$ y $\frac{\partial\Psi(0,0)}{\partial\rho_j}$ son los *costos marginales asociados a esas restricciones* y coinciden con los valores duales óptimos.

# Sensibilidad (prueba con igualdades)

## Propiedad

*Dado* $\Psi(\rho) = \min\{f(x)|h_j(x) = \rho_j\}$, *se cumple que* $\frac{\partial \Psi(0)}{\partial \rho_j} = -\bar{\mu}_j$, *donde* $\bar{\mu}=\{\mu_j\}$ *es el óptimo del dual de* $(P)=\min\{f(x)|h_j(x)=0\}$.

Consideramos $x(\rho)$, óptimo de $\Psi(\rho)$. Se cumple que $x(0) = \bar{x}$, óptimo de $(P)$ con multiplicadores $\bar{\mu}$. Se cumple que $\Psi(\rho)=f(x(\rho))=f(x(\rho))+\sum_{j=1}^{q} \mu_j h_j(x(\rho))-\sum_{j=1}^{q} \mu_j h_j(x(\rho))$ y que $\Psi(\rho)=f_{\bar{\mu}}(x(\rho)) - \sum_{j=1}^{q} \mu_j \rho_j$, en el segundo término se ha usado que $h_j(x(\rho)) = \rho_j$.

Luego, $\nabla\Psi(\rho) = \nabla^T f_{\bar{\mu}}(x(\rho)) \cdot x'(\rho) - \bar{\mu}$, y como $\bar{x}$ es el óptimo de $(P)$, $\nabla^T f_{\bar{\mu}}(x(0)) = \nabla^T f_{\bar{\mu}}(\bar{x}) = 0$ y $\nabla\Psi(0) = -\bar{\mu}$.

Para las desigualdades, pensar que si $\epsilon$ es chico, las restricciones activas se mantienen y es equivalente a tener igualdades.

# Análisis de Sensibilidad (Interpretación Económica)

Sabiendo que los precios óptimos del fabricante de comprimidos son $\lambda_1 = 2$ y $\lambda_2 = 5$.

| **Alimento** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|
| **Costo** | 11 | 7 | 9 | 12 | 10 |
| **Aporte 1** | 1 | 0 | 2 | 1 | 1 |
| **Aporte 2** | 0 | 1 | 1 | 2 | 1 |

$$(D) \begin{cases} \max 13\lambda_1 + 8\lambda_2 \\ \lambda_1 \leq 11, \\ \lambda_2 \leq 7, \\ 2\lambda_1 + \lambda_2 \leq 9, \\ \lambda_1 + 2\lambda_2 \leq 12, \\ \lambda_1 + \lambda_2 \leq 10, \\ \lambda_i \geq 0. \end{cases}$$

¿Cuáles son los alimentos que conspiran contra sus ingresos?

# Análisis de Sensibilidad (Interpretación Económica)

Sabiendo que los precios óptimos del fabricante de comprimidos son $\lambda_1 = 2$ y $\lambda_2 = 5$.

| Alimento | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Costo | 11 | 7 | 9 | 12 | 10 |
| Aporte 1 | 1 | 0 | 2 | 1 | 1 |
| Aporte 2 | 0 | 1 | 1 | 2 | 1 |

$$(D) \begin{cases} \max 13\lambda_1 + 8\lambda_2 & \\ \lambda_1 \leq 11, & (x_1) \\ \lambda_2 \leq 7, & (x_2) \\ 2\lambda_1 + \lambda_2 \leq 9, & (x_3) \\ \lambda_1 + 2\lambda_2 \leq 12, & (x_4) \\ \lambda_1 + \lambda_2 \leq 10, & (x_5) \\ \lambda_i \geq 0. & \end{cases}$$

¿Cuáles son los alimentos que conspiran contra sus ingresos?

Son los alimentos 3 y 4. En los demás las restricciones no están activas y el beneficio marginal por cambiar sus parámetros es 0.

# Análisis de Sensibilidad (Interpretación Económica)

Sabiendo que los precios óptimos del fabricante de comprimidos son $\lambda_1 = 2$ y $\lambda_2 = 5$.

| Alimento | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Costo | 11 | 7 | 9 | 12 | 10 |
| Aporte 1 | 1 | 0 | 2 | 1 | 1 |
| Aporte 2 | 0 | 1 | 1 | 2 | 1 |

$$(D)\begin{cases} \max 13\lambda_1 + 8\lambda_2 & \\ \lambda_1 \leq 11, & (x_1) \\ \lambda_2 \leq 7, & (x_2) \\ 2\lambda_1 + \lambda_2 \leq 9, & (x_3) \\ \lambda_1 + 2\lambda_2 \leq 12, & (x_4) \\ \lambda_1 + \lambda_2 \leq 10, & (x_5) \\ \lambda_i \geq 0. & \end{cases}$$

¿Cuáles son los alimentos que conspiran contra sus ingresos?

Son los alimentos 3 y 4. En los demás las restricciones no están activas y el beneficio marginal por cambiar sus parámetros es 0.

¿Cuál es la dieta óptima que siguen sus potenciales clientes?

# Análisis de Sensibilidad (Interpretación Económica)

Sabiendo que los precios óptimos del fabricante de comprimidos son $\lambda_1 = 2$ y $\lambda_2 = 5$.

| Alimento | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Costo | 11 | 7 | 9 | 12 | 10 |
| Aporte 1 | 1 | 0 | 2 | 1 | 1 |
| Aporte 2 | 0 | 1 | 1 | 2 | 1 |

$$(D)\begin{cases} \max 13\lambda_1 + 8\lambda_2 & \\ \lambda_1 \leq 11, & (x_1) \\ \lambda_2 \leq 7, & (x_2) \\ 2\lambda_1 + \lambda_2 \leq 9, & (x_3) \\ \lambda_1 + 2\lambda_2 \leq 12, & (x_4) \\ \lambda_1 + \lambda_2 \leq 10, & (x_5) \\ \lambda_i \geq 0. & \end{cases}$$

¿Cuáles son los alimentos que conspiran contra sus ingresos?

Son los alimentos 3 y 4. En los demás las restricciones no están activas y el beneficio marginal por cambiar sus parámetros es 0.

¿Cuál es la dieta óptima que siguen sus potenciales clientes?

Como $\lambda_1, \lambda_2 > 0 \Rightarrow$ las restricciones primales están activas. La solución de $2x_3 + x_4 = 13$, $x_3 + 2x_4 = 8$ es $x_3 = 6$ y $x_4 = 1$.