Multiplicadores de Lagrange y dualidad

Problemas con solo restricciones de igualdad

minimize $f(x)$

subject to $h_i(x) = 0, \qquad i = 1, \ldots, m$

- Sea x* un mínimo local y regular ($\nabla h_i(x^*)$: son linealmente independientes), entonces existen $\lambda_1^*, \ldots, \lambda_m^*$ tales que:

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0$$

- Interpretación y ejemplos.

Problemas con restricciones de igualdad y desigualdad

minimize $f(x)$

subject to $h(x) = 0, \qquad g(x) \leq 0$

donde $f:\Re^n \to \Re, h:\Re^n \to \Re^m, g:\Re^n \to \Re^r$

son funciones continuas diferenciables.

Asumimos que el dominio $D = \bigcap_{i=1}^{m} \text{dominio}(f_i) \bigcap \bigcap_{i=1}^{r} \text{dominio}(g_i)$ es no vacío

Construimos el Lagrangiano $L: \Re^n x \Re^m x \Re^r \to \Re$

asociado al problema anterior como:

$$L(x,\lambda,\nu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{i=1}^{r} \nu_i g_i(x)$$

λ_i : multiplicadores de Lagrange asociados con la restricción de igualdad i

ν_i : multiplicadores de Lagrange asociados con la restricción de desigualdad i

La función dual de Lagrange

Es una función $d : \Re^m \times \Re^r \rightarrow \Re$

$$d(\lambda,\nu) = \inf_x L(x,\lambda,\nu) = \inf_x \left(f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{i=1}^{r} \nu_i g_i(x) \right)$$

como $d(\lambda,\nu)$es el mínimo puntual de una familia de funciones lineales de (λ,ν) es una función cóncava aún cuando el problema no sea convexo.

Cota inferior del óptimo

Para todo $(\lambda, \nu \geq 0)$, *sea* p^* el valor óptimo del problema de optimización entonces,

$d(\lambda, \nu) \leq p^*$

Si $\bar{x}$ es un puto factible del problema de optimización, entonces:

$$\left(\sum_{i=1}^{m} \lambda_i h_i(\bar{x}) + \sum_{i=1}^{r} \nu_i g_i(\bar{x})\right) \leq 0$$

por lo tanto

$L(\bar{x}, \lambda, \nu) \leq f(\bar{x})$

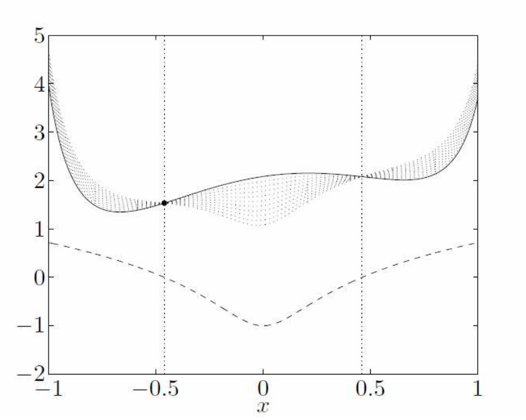
por lo tanto:

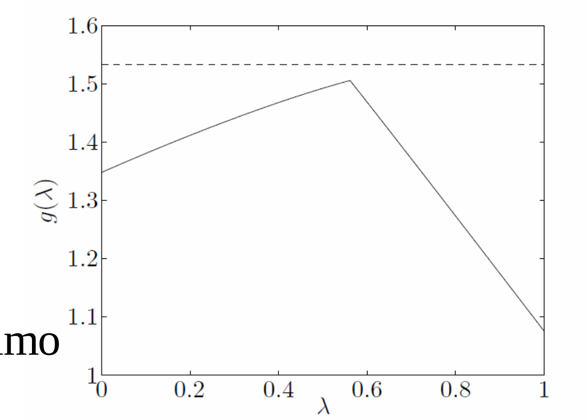
$d(\lambda, \nu) \leq f(\bar{x})$, como vale para cualquier punto factible vale para el óptimo

$d(\lambda, \nu) \leq p^*$

$\mathrm{Dominio}(g) = \{\lambda, \nu : d(\lambda, \nu) > -\infty\}$

λ, ν son factibles del dual si $\nu \geq 0, \lambda, \nu \in Dom(d)$





El Lagrangiano y la cota inferior se pueden interpretar como una linealización de la función indicatriz de las restricciones

El problema dual de Lagrange

- ¿Cuál es la mejor cota inferior que puede ser obtenida a través de la función dual de Lagrange?

$$\text{maximizar} \quad d(\lambda, \nu)$$

$$sujeto \quad \nu \geq 0$$

a los valores óptimos (λ^*, ν^*) se los llama óptimos del dual o multiplicadores de Lagrange óptimos.

Este problema es siempre un problema de optimización convexa

Dualidad débil

- Esta propiedad se deduce de la cota inferior vista antes

Si llamamos d^* al óptimo del problema dual

$d^* \leq p^*$ (dualidad débil)

habitualmente se denomina a p^*-d^* como la

diferencia (gap) de dualidad.

Esta diferencia es siempre positiva.

Dualidad fuerte. Teorema de Slater

- La pregunta es cuando ¿d* será igual a p*?
- Esto en general no es cierto pero si el problema es convexo con restricciones convexas, habitualmente (pero no siempre) se tiene dualidad fuerte.
- Lo que establece el teorema de Slater es una condición adicional a la convexidad para que se verifique la dualidad fuerte. Hay otros resultados.

Slater. Dualidad fuerte

- Se considera el siguiente problema convexo (f_0,f_i convexas, h_i funciones lineales)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

- La condición de Slater para que en un problema convexo sea válida la dualidad fuerte es que

Exista x un punto del dominio D tal que verifique:

$$f_i(x) < 0 \;\; i = 1, \ldots, m \quad \text{y } h_i(x) = 0$$

Ejemplos de dualidad

$$f(x) = 1/2(x_1^2 + x_2^2)$$

$$s.t.$$

$$g(x) = x_1 - 1 \leq 0,$$

$$x \in X = R^2$$

Optimización discreta:

$$f(x) = -x$$

$$g(x) = x - 1/2 \leq 0$$

$$x \in X = \{0,1\}$$

Optimización sobre un simplex:

$$\min \ f(x)$$

$$\sum_{i=1}^{n} x_i = r$$

$$x_i \geq 0$$

Interpretación de la dualidad y los multiplicadores de Lagrange

- Interpretación económica.

Condiciones de optimalidad

- Complementariedad de la relajación
- Cuando existe dualidad fuerte:

$$\begin{aligned} f_0(x^\star) &= g(\lambda^\star, \nu^\star) \\ &= \inf_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^\star f_i(x) + \sum_{i=1}^{p} \nu_i^\star h_i(x) \right) \\ &\leq f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star h_i(x^\star) \\ &\leq f_0(x^\star). \end{aligned}$$

$$f_0(x^\star) = f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star h_i(x^\star)$$

Complementariedad de la relajación

- Lo cual implica: $\sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) = 0$
- Como cada término es no-positivo, se tiene la propiedad:

$$\lambda_i^\star f_i(x^\star) = 0, \quad i = 1, \ldots, m.$$

- Lo que es equivalente a:

$$\lambda_i^\star > 0 \implies f_i(x^\star) = 0,$$
$$f_i(x^\star) < 0 \implies \lambda_i^\star = 0.$$

- Si el multiplicador de Lagrange es 0, la restricción puede estar activa (no es biyectivo)

$$\lambda_i^* = 0 \not\to f_i(x^*) < 0 \qquad f_i(x^*) = 0 \not\to \lambda_i^* > 0$$

Condiciones Karush-Kuhn-Tucker (KKT) para el caso no convexo.

- Asumiendo que hay dualidad fuerte y que las f0,fi,hi son diferenciables, como $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$ el gradiente debe ser 0 en x*

$$\nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(x^\star) = 0.$$

- Por lo tanto juntando todo lo visto se tienen las siguientes condiciones necesarias de optimalidad

$$\begin{array}{rcll}
f_i(x^\star) & \leq & 0, & i = 1, \ldots, m \\
h_i(x^\star) & = & 0, & i = 1, \ldots, p \\
\lambda_i^\star & \geq & 0, & i = 1, \ldots, m \\
\lambda_i^\star f_i(x^\star) & = & 0, & i = 1, \ldots, m \\
\nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(x^\star) & = & 0, &
\end{array}$$

Condiciones KKT (caso convexo)

- En este caso también son suficientes para que sean el óptimo del primal y el dual sin diferencia de dualidad.

$$
\begin{array}{rcll}
f_i(\tilde{x}) & \leq & 0, & i = 1, \ldots, m \\
h_i(\tilde{x}) & = & 0, & i = 1, \ldots, p \\
\tilde{\lambda}_i & \geq & 0, & i = 1, \ldots, m \\
\tilde{\lambda}_i f_i(\tilde{x}) & = & 0, & i = 1, \ldots, m \\
\nabla f_0(\tilde{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{x}) & = & 0, &
\end{array}
$$

- Ejemplos

Análisis de sensibilidad

Consideramos el problema:

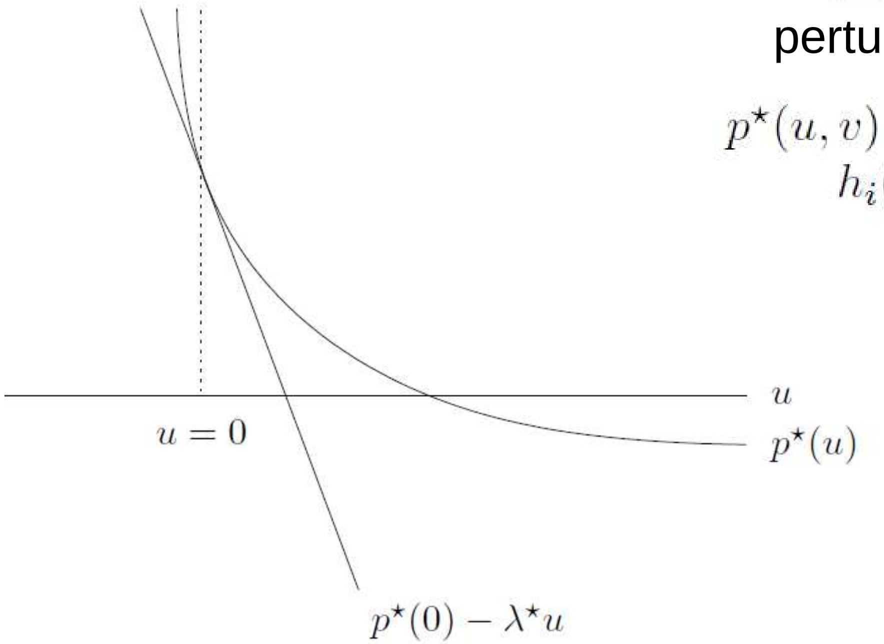
$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le u_i, \quad i = 1, \ldots, m \\ & h_i(x) = v_i, \quad i = 1, \ldots, p, \end{array}$$

Si $p^\star(u,v)$ es la solución óptima del problema perturbado, es decir:

$$p^\star(u,v) = \inf\{f_0(x) \mid \exists x \in \mathcal{D},\ f_i(x) \le u_i,\ i = 1, \ldots, m, \\ h_i(x) = v_i,\ i = 1, \ldots, p\}.$$

Si $p^\star(u,v)$ es diferenciable en (0,0) y existe dualidad fuerte, entonces las variables duales óptimas $\lambda^\star, \nu^\star$ cumplen que

$$\lambda_i^\star = -\frac{\partial p^\star(0,0)}{\partial u_i}, \qquad \nu_i^\star = -\frac{\partial p^\star(0,0)}{\partial v_i}$$



Idea general de los métodos de barrera.

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b, \end{array}$$

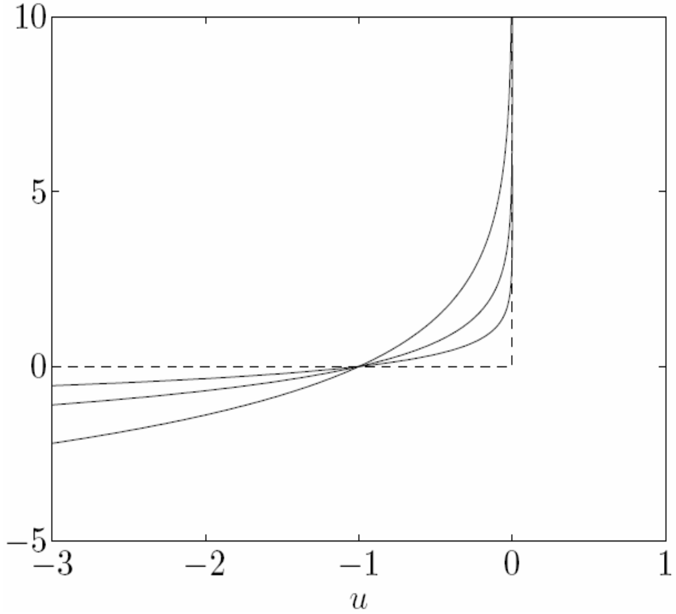
- Aproximar las restricciones por una función barrera (una aproximación de la indicatriz)

$$\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{i=1}^{m} I_-(f_i(x)) \\ \text{subject to} & Ax = b, \end{array}$$

$$I_-(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$$

Idea general de los métodos de barrera.

$$\begin{array}{ll} \text{minimize} & f_0(x) - (1/t)\sum_{i=1}^{m} \log(-f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$



- A medida que t se hace más grande la aproximación mejora.