$$\begin{aligned} \min \quad & f_0(x) \\ \text{sujeto a:} \quad & f_i(x) \leq 0, \ i = 1, \dots, m \\ & Ax = b \end{aligned}$$

donde: $f_0, f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ son funciones convexas y dos veces continuamente diferenciables y $A \in \mathbb{R}^{p \times n}$ con $\operatorname{rank} A = p < n$.

Se asume lo siguiente:

+) El problema es solvable, esto es, existe x^* punto óptimo y el valor óptimo es $p^* = f_0(x^*)$.

+) El problema es estrictamente factible*, esto es, existe $x \in D$ tal que $Ax = b$ y $f_i(x) < 0 \ \forall i = 1, \dots, m$

donde: D es el dominio del problema de optimización: $D = \bigcap_{i=0}^{m} \operatorname{dom} f_i$.

condiciones de optimalidad de Karush-Kuhn-Tucker (KKT) del problema anterior:

$$\left.\begin{aligned} & Ax^* = b, \ f_i(x^*) \leq 0 \ \forall i = 1, \dots, m \\ & \lambda^* \succeq 0 \\ & \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + A^T \nu^* = 0 \\ & \lambda_i^* f_i(x^*) = 0 \ \ \forall i = 1, \dots, m \end{aligned}\right\}$$

por las suposiciones anteriores se tiene dualidad fuerte: $\underbrace{f(x^*) - g(\lambda^*, \nu^*)}_{\text{duality gap}}$ es cero

<u>Método de barrera logarítmica</u> (es uno de los métodos por puntos interiores)

Ideas:) Reducir la duality gap entre puntos primal-dual factibles

•) Transformar el problema
$$\begin{aligned} \min \quad & f_0(x) \\ \text{sujeto a:} \quad & f_i(x) \leq 0, \ i = 1, \dots, m \\ & Ax = b \end{aligned}$$
a uno sin desigualdades para usar (bajo ciertas condiciones) el método de Newton de la nota 4.5.2 (clase pasada)

<u>Comentarios</u>:

La transformación se realiza:

$$\begin{aligned} \min \quad & f_0(x) \\ \text{sujeto a:} \quad & f_i(x) \leq 0, \ i = 1, \dots, m \\ & Ax = b \end{aligned} \iff \begin{aligned} \min \quad & f_0(x) + \frac{1}{t}\Phi(x) \\ \text{sujeto a:} \quad & Ax = b \end{aligned} \iff \begin{aligned} \min \quad & t f_0(x) + \Phi(x) \\ \text{sujeto a:} \quad & Ax = b \end{aligned}$$

donde: $\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x))$. La 1ª equivalencia: $\min f_0(x) + \frac{1}{t}\phi(x)$ sujeto a: $Ax=b$ ayuda a la

interpretación y la 2ª equivalencia ayuda a la notación matemática.

La 1era observación para el método de barrera es que debe penalizarse la función objetivo

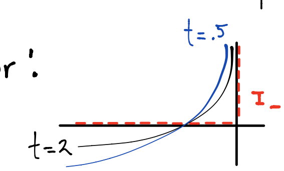
del problema: $\begin{array}{ll} \min & f_0(x) \\ \text{sujeto a:} & f_i(x) \leq 0,\ i=1,\dots,m \\ & Ax=b \end{array}$ si $f_i(x) > 0$ para alguna $i=1,\dots,m$. Para esto se utiliza

$I_-(u) = \begin{cases} 0 & \text{si } u \leq 0 \\ \infty & \text{e.o.c.} \end{cases}$: $\min\limits_{x} f_0(x) + \sum_{i=1}^{m} I_-(f_i(x))$, $Ax=b$ pero I_- no es diferenciable

(no podemos usar método de Newton) por lo que se utiliza $\hat{I}_-(u) = -\frac{1}{t}\log(-u)$ con $t > 0$

la cual es convexa y diferenciable, toma valor ∞ para valores positivos y para

valores de t más grandes la aproximación de $\hat{I}_-$ a I_- es mejor:



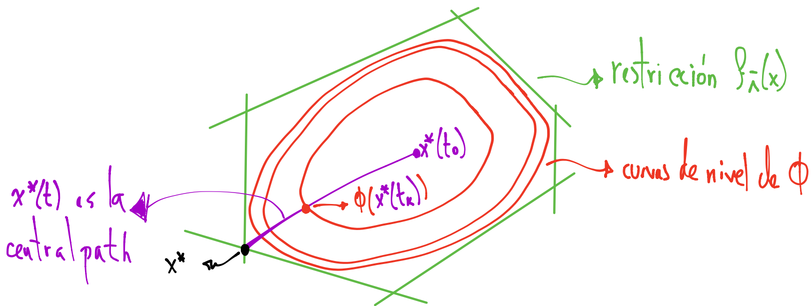
) El método de barrera logarítmica también se le llama path following method pues se resuelven las

condiciones KKT modificadas del problema: $\min\limits_{x} tf(x) + \phi(x)$, $Ax=b$ con $\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x))$ dadas por:

$$0 = t\nabla f_0(x^*(t)) + \nabla\phi(x^*(t)) + A^T\hat{\nu}$$

$$Ax^*(t) = b \quad , \quad f_i(x^*(t)) < 0 \quad i=1,\dots m \text{ (factibilidad estricta)} \qquad \left[\hat{\nu} = \frac{\nu^*}{t}\right]$$

La condición $0 = t\nabla f_0(x^*(t)) + \nabla\phi(x^*(t)) + A^T\hat{\nu}$ se satisface para los central points $x^*(t)$:



Se verifica que las condiciones anteriores se escriben como:

$$A x^*(t) = b, \ f_i(x^*(t)) \leq 0 \ \forall i = 1, \dots, m$$
$$\lambda^*(t) \geq 0$$
$$\nabla f_0(x^*(t)) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*(t)) + A^T \nu^*(t) = 0$$
$$-\lambda_i^*(t) f_i(x^*(t)) = \frac{1}{t} \ \forall i = 1, \dots, m$$

definiendo $\lambda_i^*(t) = \dfrac{1}{t f_i(x^*(t))}$, $\nu^*(t) = \dfrac{\hat{\nu}}{t}$.

Por esto, se resuelve de forma aproximada las condiciones de KKT.

·) El método de barrera logarítmica no requiere trabajar explícitamente con variables duales pues

$$f_0(x^*(t)) - g(\lambda^*(t), \nu^*(t)) = \frac{m}{t} \quad y \quad f_0(x^*(t)) - p^* \leq \frac{m}{t}$$

·) El método consiste en resolver una secuencia de problemas de minimización con restricciones lineales de igualdad en los que se utiliza como punto inicial la solución obtenida en el problema anterior: se calcula $x^*(t)$ para una secuencia de valores crecientes de t, hasta que $t \geq \frac{m}{\varepsilon}$ (que garantiza una solución ε-subóptima del problema original).

Este método es el método de barrera o path-following:

Dado x punto estrictamente factible ($x \in D$ tal que $Ax = b$ y $f_i(x) < 0 \ \forall i = 1, \dots, m$) y $t := t^{(0)} > 0$, $\mu > 1$, tolerancia $\varepsilon > 0$.

Repetir:

1) Centering step:

Calcular $x^*(t)$ al resolver: $\min_{x} \; t f_0(x) + \phi(x)$ iniciando con x.
sujeto a: $Ax = b$

2) Actualizar $x := x^*(t)$.

3) Criterio de paro: salir si $\frac{m}{t} < \varepsilon$.

4) Incrementar t, $t := \mu t$.

Comentarios:

+) El centering step calcula un punto central y este paso también se llama outer iteration. Al final de este paso se tienen puntos primal, dual factibles.

+) El centering step se realiza con métodos de optimización con restricciones lineales por ejemplo el método de Newton se utiliza y así, cada iteración del método de Newton se le llama inner iterations. En cada paso de las inner iterations se tienen puntos primales factibles.

+) La elección del valor de μ involucra un trade-off en el número de iteraciones inner y outer requeridas:

Si μ se elige cercana a 1 entonces el resultado de cada centering step es un buen punto inicial del siguiente y el número de iteraciones inner del método de Newton es pequeño (pues cada iteración inner sigue la central path) pero se requerirán un mayor número de iteraciones outer que inner (pues la duality gap se reduce por una cantidad pequeña).

Si μ es grande se tiene la situación contraria a lo anteriormente descrito:

μ es el factor por el que se reduce la duality gap entre cada iteración outer

+) Si (λ, ν) es un punto dual factible con duality gap $\eta = f_0(x^{(0)}) - g(\lambda, \nu)$ entonces $t^{(0)}$ puede elegirse como $\frac{m}{\eta}$ pues al final del primer centering step se tendrá una duality gap de $\eta = \frac{m}{t^{(0)}}$. Así, se elige $t^{(0)}$ de modo que $\frac{m}{t^{(0)}}$ sea del mismo orden que $f_0(x^{(0)}) - p^*$ o μ veces esta cantidad.

+) $\alpha = 0.01$, $\beta = 0.5$, $\frac{\lambda^2(x)}{2} \leq 10^{-5}$, criterio de paro $\leq 10^{-6}$

$\hookrightarrow \frac{m}{t}$