Método de Lagrangiano aumentado (ALGENCAN)

$\rho_0 > 0$, $\tau \in [0,1)$, $x^0 \in \mathbb{R}^n$,
$\bar{\lambda}_i^0 \in [\lambda_{min}, \lambda_{máx}]$, $\forall i$
$\bar{\mu}_j^0 \in [0, \mu_{máx}]$, $\forall j$; $\{\varepsilon_k\} \downarrow 0$

Parar?

sim → PARE

NÃO

Resolver o subproblema

$$\min_x f(x) + \frac{\rho_k}{2}\left[\left\|h(x) + \frac{\bar{\lambda}^k}{\rho_k}\right\|^2 + \left\|\left(g(x) + \frac{\bar{\mu}}{\rho_k}\right)_+\right\|^2\right]$$

obtendo o iterando x^k. O ponto x^k deve satisfazer $\|\nabla_x L_{\rho_k}(x^k, \bar{\lambda}^k, \bar{\mu}^k)\|_\infty \leq \varepsilon_k$

Atualizar multiplicadores projetados:

$$\bar{\lambda}^{k+1} = \max\{\lambda_{min}, \min\{\lambda_{máx}, \rho_k h(x^k) + \bar{\lambda}^k\}\}$$

$$\bar{\mu}^{k+1} = \max\{0, \min\{\mu_{máx}, \rho_k g(x^k) + \bar{\mu}^k\}\}$$

Atualize ρ segundo o critério (**)

$k \leftarrow k+1$

(**) Se $K=0$ (primeira iteração) ou

$$\max\{\|h(x^K)\|_\infty, \|V^K\|_\infty\} \leq \tau \max\{\|h(x^{K-1})\|_\infty, \|V^{K-1}\|_\infty\}$$

então $\rho_{K+1} = \rho_K$ (não aumentamos ρ). Caso contrário, $\rho_{K+1} = 10\,\rho_K$.

- $\tau \in [0,1)$: parâmetro.
- $V^K = \min\{\bar{\mu}^K, -g(x^K)\}$.

Critério de parada: paramos quando x^k é ~ KKT do problema original

$$P: \min_x f(x) \text{ s.a. } h(x)=0,\ g(x)\leq 0.$$

Parâmetro: $\varepsilon_{opt} > 0$: precisão de parada.

$$\left\| \nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) \right\| \leq \varepsilon_{opt}$$

$$\mu_j^k \geq 0,\ \forall j.$$

$$\left| \min\{\mu_j^k, -g_j(x^k)\} \right| \leq \varepsilon_{opt},\ \forall j$$

$$\|h(x^k)\|_\infty \leq \varepsilon_{opt},\quad \|g(x^k)_+\|_\infty \leq \varepsilon_{opt}$$

- Padrão: $\varepsilon_{opt} = 10^{-8}$.
- $\lambda^k = \rho_k h(x^k) + \bar{\lambda}^k$, $\mu^k = (\rho_k g(x^k) + \bar{\mu}^k)_+$.

Convergência

$P: \min_x f(x)$ s.a. $h(x)=0$, $g(x)\leq 0$.

Objetivo: mostrar que ALGENCAN atinge/converge à pontos KKT de P.

Obs: vamos desconsiderar o critério de parada. Assim, o método gera uma infinita.

Seja $\{x^k\}$ a sequência gerada pelo método

Queremos mostrar que qualquer ponto de acumulação x^* de $\{x^k\}$ é KKT.

1) "o método cumpre a complementaridade"

Lema: Seja x^* um ponto de acumulação de $\{x^k\}$. Se $g_j(x^*)<0$ então

$\mu_j^{k+1}=(\rho_k g_j(x^k)+\bar{\mu}_j^k)_+=0$, $\forall k \gg 1$.

Em particular, $\mu_j^{k+1} g_j(x^k) = 0$, $\forall k \gg 1$.

Prova:

CASO 1: $\rho_k \to \infty$.

No método, $\bar{\mu}^k \in [0, \mu_{máx}]$, $\forall k$, e assim $\{\bar{\mu}^k\}$ é limitada. Como $g_j(x^*) < 0$, temos $g_j(x^k) \leq \frac{g_j(x^*)}{2} < 0$, $\forall k \gg 1$.

Daí

$$\rho_k g_j(x^k) + \bar{\mu}_j^k < \underbrace{\rho_k}_{\to\infty} \underbrace{\frac{g_j(x^*)}{2}}_{<0} + \underbrace{\bar{\mu}_j^k}_{<\infty} \longrightarrow -\infty.$$

Assim, $\mu_j^{k+1} = \max\{0, \rho_k g_j(x^k) + \bar{\mu}_j^k\} = 0$, $\forall k \gg 1$.

CASO 2: $\{\rho_k\}$ é limitada.

Neste caso, o controle de viabilidade deu certo $\forall k \gg 1$. Em particular, $\|V^{k-1}\|_\infty \leq \tau \|V^k\|_\infty$. Como $\tau < 1$, Temos $V^k \to 0$. Como

$V_j^k = \min\{\bar{\mu}_j^k, -g_j(x^k)\}$ e

$g_j(x^k) \lesssim \frac{g_j(x^*)}{2} < 0 \quad \forall k \gg 1$, temos

$\bar{\mu}_j^k \to 0$. Assim, $\rho_k g_j(x^k) + \bar{\mu}_j^k \leq$

$$\underbrace{\rho_k}_{>0} \underbrace{\frac{g_j(x^*)}{2}}_{<0} + \underbrace{\bar{\mu}_j^k}_{\to 0} < 0, \quad \forall k \gg 1.$$ Novamente,

$$\mu_j^{k+1} = \max\{0, \rho_k g_j(x^k) + \bar{\mu}_j^k\} = 0, \quad \forall k \gg 1 \quad \blacksquare$$

---

(2) "<u>o método atinge KKT se x^* for viável</u>"

<u>Teorema:</u> Seja x^* ponto de acumulação de $\{x^k\}$. Suponha que x^* seja <u>regular</u> e <u>viável</u>. Então x^* é ponto KKT de P.

Prova: No método, temos

$$\|\nabla L_{\rho_k}(x^k, \bar{\lambda}^k, \bar{\mu}^k)\|_\infty \leq \varepsilon_k \downarrow 0.$$

Isto é,

$$(1) \quad \Big\| \nabla f(x^k) + \sum_{i=1}^{m} \overbrace{(\rho_k h_i(x^k) + \bar{\lambda}_i)}^{\lambda_i^{k+1}} \nabla h_i(x^k) + \sum_{j=1}^{p} \underbrace{(\rho_k g_j(x^k) + \bar{\mu}_j^k)_+}_{\mu_j^{k+1}} \nabla g_j(x^k) \Big\| \leq \varepsilon_k \to 0$$

Definimos

$$\delta_{k+1} = \|(\lambda^{k+1}, \mu^{k+1})\|_\infty, \quad \forall k.$$

Vamos mostrar que $\{\delta_k\}$ é uma sequência limitada.

Suponhamos por contradição que exista subsequência de $\{\delta_k\}$ ilimitada. Digamos que $\lim_{k\to\infty} \delta_k = \infty$. Dividindo (1) por δ_{k+1}, obtemos

$$(2) \quad \underbrace{\frac{\nabla f(x^k)}{\delta_{k+1}}}_{\to 0} + \sum_{i=1}^{m} \frac{\lambda_i^{k+1}}{\delta_{k+1}} \nabla h_i(x^k) + \sum_{j=1}^{p} \frac{\mu_j^{k+1}}{\delta_{k+1}} \nabla g_j(x^k) \leq \frac{\varepsilon_k}{\delta_{k+1}} \to 0$$

Pela definição de δ_{k+1}, temos

$$\delta_{k+1} = \|(\lambda^{k+1}, \mu^{k+1})\|_\infty$$
$$= \max\{|\lambda_1^{k+1}|, \ldots, |\lambda_m^{k+1}|, |\mu_1^{k+1}|, \ldots, |\mu_p^{k+1}|\},$$

e logo

$$\left\{\frac{\lambda_i^{k+1}}{\delta_{k+1}}\right\} \longrightarrow 1 \text{ ou } -1 \quad \text{para algum } i$$

ou

$$\left\{\frac{\mu_j^{k+1}}{\delta_{k+1}}\right\} \longrightarrow 1 \quad \text{para algum } j.$$

Passando o limite em (2) obtemos

$$\sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla g_j(x^*) = 0,$$

onde pelo menos um λ_i^* ou μ_j^* é $\neq 0$.
Pelo lema anterior, $\mu_j^{k+1} = 0$, $\forall k \gg 1$ sempre que $g_j(x^*) < 0$. Assim,

$$\sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j : g_j(x^*)=0} \mu_j^* \nabla g_j(x^*) = 0,$$

onde nem todos λ_i^*'s e μ_j^*'s são nulos.
Isso contraria o fato de x^* ser regular.
Assim, concluímos que $\{\delta_k\}$ é limitada.
Desta forma, $\{\delta_{k+1} = \|(\lambda^{k+1}, \mu^{k+1})\|_\infty\}$ admite ponto de acumulação, digamos
$\lambda_i^* = \lim \lambda_i^{k+1}$, $\forall i$, $\mu_j^* = \lim \mu_j^{k+1}$, $\forall j$.
Pelo Lema anterior, $\mu_j^* = 0$ se $g_j(x^*) < 0$.
Assim, passando o limite em (1), obtemos

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

Finalmente, note que $\mu_j^* = \lim \mu_j^{k+1} \geq 0$ $\forall j$, e $\mu_j^* g_j(x^*) = 0$, $\forall j$. Ou seja, x^* é ponto KKT de P ▩

③ "<u>o método atinge um ponto estacionário da inviabilidade de P</u>"

P: $\min f(x)$ s.a. $h(x)=0$, $g(x)\leq 0$

Problema da inviabilidade:

$$\min_x \underbrace{\|h(x)\|^2 + \|g(x)_+\|^2}_{\phi(x)}$$

x^* é "estacionário da inviabilidade" se o gradiente de $\phi(x)$ em x^* é nulo.

($\Leftarrow$ x^* é KKT do prob. de inviabilidade)

"Mesmo que x^* não seja viável, ser estacionário da inviabilidade dá expectativa de que x^* é menos inviável possível, pois é candidato a minimizador de $\phi(x) = \|h(x)\|^2 + \|g(x)_+\|^2$"

Isso é relevante, em particular, quando P é inviável.

→ "o método fornece x^* estacionário da inviabilidade":

<u>Teorema</u>: Seja x^* ponto de acumulação de $\{x^k\}$ (viável ou não). Então

$$\nabla \phi(x^*) = 0$$

(é estacionário/KKT do problema de inviabilidade).

<u>Prova</u>: consultar o livro referência ("Otimização prática usando L.A.", de Martínez) ou "slide" na página da disciplina. ∎