PENALITY METHOD

Prendiamo un problema vincolato come segue:

$$\begin{cases} \min f(x) \\ g_i(x) \leq 0 \qquad \forall i = 1 \dots m \end{cases} \qquad (P)$$

Dove quindi definiamo $X = \{x \in \mathbb{R}^n : g_i(x) \leq 0\}$

Penality function

Definiamo la funzione di penalità $p(x)$:

$$p(x) = \sum_{i=1}^{m} \max\left(0, g_i(x)\right)^2$$

Consideriamo ora il problema senza vincoli

$$\begin{cases} \min f(x) + \frac{1}{\varepsilon} p(x) = p_\varepsilon(x) \\ x \in \mathbb{R}^n \end{cases} \qquad (P_\varepsilon)$$

Da notare che per sua definizione $p_\varepsilon(x) = f(x)$ $\forall x \in X$ perché in tal caso $p(x) = 0$. Altrimenti $\forall x \notin X$ abbiamo $p_\varepsilon(x) > f(x)$

PROPOSIZIONI

1) SE f e g_i sono continue e differenziabili, ALLORA $p_\varepsilon(x)$ è continua e differenziabile e il suo gradiente:

$$\nabla p_\varepsilon(x) = \nabla f(x) + \frac{2}{\varepsilon} \sum_{i=1}^{m} \max\{0, g_i(x)\} \nabla g_i(x)$$

2) SE f e g_i sono **convesse** ALLORA p_ε è convessa

3) Ogni problema P_ε è un rilassamento di P, ovvero che $v(P_\varepsilon) \leq v(P)$ $\forall \varepsilon > 0$.

4) SE x^* risolve P_ε e $x^* \in X$ ALLORA x^*_ε è ottimo per P

5) SE $0 < \varepsilon_2 < \varepsilon_1$ ALLORA $v(P_{\varepsilon_1}) \leq v(P_{\varepsilon_2})$

PASSI DEL METODO

1) Setta $\varepsilon_0 > 0$, $\tau \in (0, 1)$, $K = 0$

2) Trovare la soluzione ottima x_K per P_{ε_K}

3) SE $x_K \in X$ ALLORA **STOP**

ALTRIMENTI $\varepsilon_{K+1} = \tau \varepsilon_K$, $K{+}{+}$ e ricomincia da 1.

TEOREMA

1) SE f è coerciva, ALLORA la sequenza $\{x_K\}$ è limitata e ogni punto stazionario è una soluzione ottima per P

2) SE $\{x_K\}$ converge a x^*, ALLORA x^* è ottimo per P

3) SE $\{x_K\}$ converge a x^* e i gradienti dei vincoli attivi in x^* sono linearmente indipendenti, ALLORA x^* è una soluzione ottima per P e la sequenza dei vettori $\{\lambda_K\}$ definita:

$$\lambda_K^i = \frac{2}{\varepsilon_K} \max\{0, g_i(x_K)\}$$

converge al vettore λ^* che è un moltiplicatore KKT associato a x^*.

EXACT PENALITY METHOD

Consideriamo ora un problema CONVESSO, e definiamo ugualmente a prima P, $\tilde{f}(x)$, $\tilde{P}_\varepsilon$.

Nota: per questi problemi non abbiamo bisogno di usare una sequenza $\varepsilon_k \to 0$ per trovare l'ottimo di P, esiste infatti un preciso ε che minimizza $\tilde{P}_\varepsilon$ che coincide con il minimo di P.

PROPOSIZIONE

Supponendo x^* ottimo per P e λ^* ottimo per il KKT vettore dei moltiplicatori associato a x^* ALLORA esiste un set di soluzioni ottime per P e $\tilde{P}_\varepsilon$ uguale per entrambi purché $\varepsilon \in \left(0, \frac{1}{\|\lambda^*\|_\infty}\right)$

PASSI DEL METODO

1) Setta $\varepsilon_0 > 0$, $\tau \in (0, 1)$, $k = 0$
2) Trovare la soluzione ottima x_k per $\tilde{P}_{\varepsilon_k}$
3) SE $x_k \in X$ ALLORA **STOP**

   ALTRIMENTI $\varepsilon_{k+1} = \tau \varepsilon_k$, $k{+}{+}$ e ricomincia da 1.

TEOREMA

L'exac penality method termina dopo un numero finito di iterazioni con la soluzione ottima di P.

BARRIER METHODS

Questo metodo genera una sequenza di punti ammissibili che approssima l'ottimo per P.

$$\begin{cases} \min f(x) \\ g_i(x) \leq 0 \quad \forall i = 1 \dots m \end{cases} : P$$

Considerando il primale P, sotto le seguenti assunzioni vale la STRONG DUALITY:

1) f e g_i convesse, continue e differenziabili 2 volte
2) $\exists\, x^*$ soluzione ottima
3) $\exists\, \bar{x} : g_i(\bar{x}) < 0$

LOGARITHMIC BARRIER

In questo metodo consideriamo l'intorno di X: $int(X)$ come set ammissibile, e approssimiamo di conseguenza P:

$$\begin{cases} \min f(x) - \varepsilon \sum_{i=1}^{m} \log(-g_i(x)) = \psi_\varepsilon(x) \\ x \in int(X) \end{cases}$$

Dove definiamo $B(x) = -\sum_{i=1}^{m} \log(-g_i(x))$ la funzione barriera logaritmica:

$$\begin{cases} \min f(x) - \varepsilon B(x) \\ x \in int(X) \end{cases} : P_{log}$$

Nota: per x che tende alla frontiera di X, $\Psi_\varepsilon(x) \to +\infty$

PROPRIETÀ DI $B(x)$

1) dominio di $B = \text{int}(X)$

2) B è convessa

3) B è derivabile con:

a) $$\nabla B(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)} \nabla g_i(x)$$

b) $$\nabla^2 B(x) = -\sum_{i=1}^{m} \frac{1}{g_i^2(x)} \nabla g_i(x) \nabla g_i^T(x) + \sum_{i=1}^{m} \frac{1}{-g_i^2(x)} \nabla^2 g_i(x)$$

Se x_ε^* è l'ottimo di P allora $\nabla f(x_\varepsilon^*) + \varepsilon \nabla B(x_\varepsilon^*) = \emptyset$.

Definito il Lagrangiano di P:

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

Possiamo definire la sua soluzione ottima λ_ε^*:

$$\lambda_\varepsilon^* = \left(g_1(x_\varepsilon^*), \dots, g_m(x_\varepsilon^*) \right) > \emptyset$$

Da cui $\nabla L(x_\varepsilon^*, \lambda_\varepsilon^*) = \emptyset$.

Ma dato che abbiamo assunto P convesso e che la strong duality valga, abbiamo che

$$v(P) = \max_{\lambda \in \mathbb{R}_+^m} \left[\inf_{x \in \mathbb{R}^n} L(x, \lambda) \right]$$

Da cui consegue che

$$v(P) \geq \min_{x} L(x, \lambda) = L(x^*_\varepsilon, \lambda^*_\varepsilon)$$

$$f(x^*_\varepsilon) \geq v(P) \geq L(x^*_\varepsilon, \lambda^*_\varepsilon) = f(x^*_\varepsilon) + \underbrace{m\,\varepsilon}_{\text{gap ottimale.}}$$

KKT

Il problema KKT viene approssimato

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0 \\ -\lambda_i\, g_i(x) = 0 \quad \forall i = 1 \dots m \\ \lambda > 0 \\ g(x) \leq 0 \end{cases} \Longrightarrow \begin{cases} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0 \\ -\lambda_i\, g_i(x) = \varepsilon \quad \forall i = 1 \dots m \\ \lambda > 0 \\ g(x) \leq 0 \end{cases}$$

PASSI LOGARITHMIC BARRIER METHOD

0) Setta $\delta > 0$, $\tau \in (0, 1)$, $\varepsilon_1 > 0$, $x_0 \in int(X)$, $k = 1$

1) Trova la soluzione ottima x_k di P_{log}, con x_{k-1} come punto di partenza

3) SE $m\,\varepsilon_k < \delta$ ALLORA STOP

ALTRIMENTI

a) $\varepsilon_{k+1} = \tau\, \varepsilon_k$

b) k++

c) ricomincia da 1.

COME SCEGLIERE $x_0 \in int(X)$

Per farlo possiamo considerare il problema ausiliario:

$$\begin{cases} \min\limits_{x,\, s} \; s \\ g_i(x) \leq s \quad \forall i = 1 \dots m \end{cases}$$

1) Prendiamo un $\bar{x} \in \mathbb{R}^n$ e troviamo $\bar{s} > \max\limits_{i=1\dots m} \{ g_i(\bar{x}) \}$

2) Cerchiamo (x^*, s^*) ottimo dell'ausiliario

3) SE $s^* < 0$ ALLORA $x^* \in int(X)$

ALTRIMENTI $int(X) = \emptyset$.