Algorisme Subgradient

(Bazaraa Cap. 8.9 p.339)

Considerem ara el problema (P) de $\boxed{\begin{array}{c} Min \;\; f(x) \\ x \in X \end{array}}$ essent X convex i tancat.

Donat $y \in R^n$ la projecció de y sobre X ve donada per trobar el punt sobre X de distància mínima a y.

$$P_X(y) = argmin\{ \; \|y - x\| \mid x \in X \; \}$$

(Generalment es tria $\|\cdot\|$ la euclidiana)

<u>Algorisme subgradient genèric</u> (Uzawa 1958)

0) Sigui $x^0 \in X$; $\begin{array}{cl} \zeta^0 = f(x^0) & \text{(límit superior)} \\ x_I \leftarrow x^0 & \text{(incumbent)} \end{array}$ $\quad k$=0

1) Determinar $v_k \in \partial f(x^k)$

- Si $v_k = 0$ <u>STOP</u> $\Leftrightarrow$ x^k solució de (P)
- Sigui $d_k = -\frac{v_k}{\|v_k\|}$
- Calcular una passa $\lambda_k > 0$ i obtenir nou iterat segons:
  $x^{k+1} = P_X(x^k + \lambda_k d_k)$ (Projecció de la passa)

<u>Si</u> $f(x^{k+1}) < \zeta^k$	llavors	$\zeta^{k+1} = f(x^{k+1})$
		$x_I = x_k$
	altrament	$\zeta^{k+1} = \zeta^k$

El criteri d'aturada que s'implementi en la pràctica tindrà en compte el nº d'iteracions.
La versió usada en l'optimització del dual lagrangià pressuposa que es coneix a priori f^* i llavors el criteri del tipus $\zeta^k \leq f^* + \varepsilon$ és l'utilitzat.
De totes formes cal tenir present el següent resultat: (Teo 8.9 2 Bazaraa)

<u>Teo</u> Pel problema (P) $\begin{array}{c} Min \;\; f(x) \\ x \in X \end{array}$
consideris l'optimització subgradient aplicada a (P) i suposis que les passes λ_k satisfan:

a) $\{\lambda_k\} \rightarrow 0_+$

b) $\sum_{k=0}^{\infty} \lambda_k = \infty$ (sèrie divergent)

Llavors es compleix

1) L'algoritme acaba en nº finit d'iteracions en sol. òptima.
ó

2) La seqüència d'iteracions verifica que

$$\zeta^k \rightarrow f^* \;\; (k \rightarrow \infty)$$

Cal ser cautelòs amb el resultat, la seqüència harmònica $\lambda_k = 1/k$ verifica a), b) però l'algorisme pot presentar un pèssim rendiment.
La passa òptima recomanada ve donada per

$$\lambda^* = (x^* - x^k)^\top d_k = \frac{(x_k - x^*)^\top v^k}{\|v^k\|}$$

ja que $f(x^*) - f(x^k) \geq (x^* - x^k)^\top v^k$; $\lambda^* \geq \frac{f(x^k)-f^*}{\|v^k\|}$

per tant s'acostuma a prendre

$$\lambda_k = \beta_k \cdot \frac{f(x^*) - f^*}{\|v^k\|}, \;\; \beta_k > 0$$

on $\varepsilon_1 < \beta_k \leq 2 - \varepsilon_2$ *per* $\varepsilon_1,\ \varepsilon_2 > 0$

(usant una estimació per f^*).