Subgradient algorithm

(Bazaraa's book Chap. 8.9 p.339)

Consider the following problem (P) $\boxed{\begin{array}{c} Min\ \ f(x) \\ x \in X \end{array}}$ where X is closed and convex.

For given $y \in R^n$ the projection of y on X is the point on X with minimum euclidean distance to y.

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

<u>Genèric subgradient algorithm</u> (Uzawa 1958)

0) Let $x^0 \in X$; $\begin{array}{rl} \zeta^0 = f(x^0) & \text{(upper limit)} \\ x_I \leftarrow x^0 & \text{(current bound)} \end{array}$ k=0

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

- If $v_k = 0$ then <u>STOP</u> $\Leftrightarrow$ x^k is a solution of (P)
- Let $d_k = -\frac{v_k}{\|v_k\|}$
- Calculate a step $\lambda_k > 0$ and obtain a new iterate accordingly to: $x^{k+1} = P_X(x^k + \lambda_k d_k)$ (Step's projection)

<u>If</u> $f(x^{k+1}) < \zeta^k$	then	$\zeta^{k+1} = f(x^{k+1})$
		$x_I = x_k$
	otherwise	$\zeta^{k+1} = \zeta^k$

Usually the stopping criterion will take into account a maximum number of iterations. The version used in the optimization of the dual lagrangian function assumes that f^* is known a priori and then $\zeta^k \le f^* + \varepsilon$ is used. Anyway, the following result (Theorem 8.9.2 in Bazaraa's book) must be taken into account:
<u>Teo</u> For problem (P) $Min_{x \in X}\ \ f(x)$ consider the subgradient optimization algorithm applied to problem (P) and assume that step lengths λ_k satisfy:

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

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

then

1) The algorithm ends in a finite number of iterations attaining an optimal solution.
   or

2) The sequence of iterations verify

$$\zeta^k \to f^* \ \ (k \to \infty)$$

The previous result must be taken prudently because, as an instance, with the harmonic sequence $\lambda_k = 1/k$ either a) or b) will be verified but the algorithm presents a very poor performance.
The optimal recommended step is:

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

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

Thus, the following step length is usually adopted:

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

where $\varepsilon_1 < \beta_k \leq 2 - \varepsilon_2$ for $\varepsilon_1, \ \varepsilon_2 > 0$

(notice that an estimation for f^* is required).