Note implementative: backtracking basato su interpolazione quadratica

Determinati $\mathbf{x}_k$ e una direzione di discesa $\mathbf{p}_k$, si deve calcolare il valore $\alpha_k > 0$ tale che $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$ soddisfi la regola di Armijo (o le regole di Wolfe Powell). Per determinare tale valore, si può procedere mediante la costruzione del polinomio interpolatore

$$p(\alpha) = \phi_0 + b\alpha + c\alpha^2$$

che deve passare per i punti $(0, \phi_0),\ (\alpha_p, \phi_p),\ (\alpha_c, \phi_c)$, con $\phi(\alpha) = ||\mathbf{F}(\mathbf{x}_k + \alpha \mathbf{p}_k)||_2^2$ (dovendo risolvere $\mathbf{F}(\mathbf{x}) = 0$)[1], $\phi_0 = \phi(0),\ \phi_c = \phi(\alpha_c),\ \phi_p = \phi(\alpha_p)$, dove α_c e α_p sono rispettivamente il valore corrente e precedente di α.
I coefficienti $b,\ c$ sono determinati imponendo il passaggio per i punti $(\alpha_p, \phi_p),\ (\alpha_c, \phi_c)$:

$$\begin{cases} \phi_p = \phi_0 + b\alpha_p + c\alpha_p^2 \\ \phi_c = \phi_0 + b\alpha_c + c\alpha_c^2 \end{cases} \tag{1}$$

Si calcola poi il punto di minimo della parabola $\overline{\alpha} = -b/(2 * c)$ (con $c > 0$) e si pone

$$\alpha_t = \overline{\alpha}$$

dove α_t è un valore tentativo. Se tale valore soddisfa la condizione di Armijo (o le regole WP) allora $\alpha_k = \alpha_t$, altrimenti $\alpha_p = \alpha_c,\ \alpha_c = \alpha_t$ e si itera la procedura, ricalcolando i coefficienti $b,\ c$ ed il nuovo valore $\overline{\alpha}$ finchè α_t soddisfa la regola di Armijo.

N.B.: ricordarsi di porre un limite massimo (es. 15/20) al numero di iterazioni in cui si calcola $\overline{\alpha}$ e di inserire le regole di salvaguardia ($0.1\alpha_c = \alpha_{min} \leq \alpha_t \leq \alpha_{max} = 0.5\alpha_c$).

Calcolo dei coefficienti

A partire dalla (1) si ottiene:

$$b = \frac{1}{\alpha_c - \alpha_p}\left(\frac{\alpha_c}{\alpha_p}(\phi_p - \phi_0) + \frac{\alpha_p}{\alpha_c}(\phi_0 - \phi_c)\right)$$

$$c = \frac{1}{\alpha_c - \alpha_p}\left(\frac{\phi_0 - \phi_p}{\alpha_p} + \frac{\phi_c - \phi_0}{\alpha_c}\right)$$

Affinchè si riesca a trovare il minimo (che approssima il valore desiderato) la parabola deve avere derivata seconda positiva: quindi occorre verificare che $c > 0$, altrimenti si sceglie $\alpha_t = \alpha_{max}$ (v. esempio in Fig. 1).

[1] Nel caso di un problema $\min_{\Re^n} f(\mathbf{x})$, si definisce $\phi(\alpha) = f(\mathbf{x_k} + \alpha \mathbf{p_k})$.

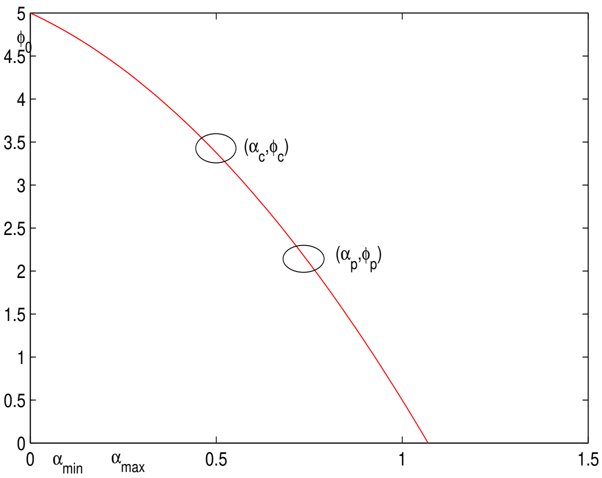

Figure 1: Caso in cui la parabola interpolante ha la concavità rivolta verso il basso.