Prova in itinere di OTTIMIZZAZIONE

28–4–2005

1. Sia $\{\mathbf{x}_k\}$ una successione convergente a $\mathbf{x}^*$: dare la definizione di velocità di convergenza di tipo lineare, quadratica, superlineare.

2. Dare la definizione di *direzione di discesa*, illustrandone le proprietà, e fornire esempi di particolari direzioni di discesa. Sia $x \in \Re^2$ e $f(x) = 2(x_1-1)^2+x_1(x_2+2)^3$: individuare una direzione di discesa p_0 per il primo passo di un algoritmo di ottimizzazione applicato a $f(x)$, con $x_0^T = (1,0)$; calcolare quindi x_1 (scegliere $\alpha = 1$).

   La direzione di Newton δ_0^{NR} risulta di discesa? Motivare la risposta.

3. Discutere i limiti del metodo di Newton, indicando i possibili rimedi e possibili metodi alternativi.

4. Enunciare e commentare le regole di Wolfe–Powell.

5. Descrivere i metodi quasi–Newton per equazioni non lineari e derivare l'update *Good Broyden* (B_{gb}) motivando i vari passaggi. Dimostrare che, sotto opportune ipotesi, l'update B_{gb} risolve

   $$\min_{\overline{B}\delta_k=\gamma_k} ||\overline{B} - B_k||$$

   dove $B_k \approx J_k$, con $\delta_k = x_{k+1} - x_k$ e $\gamma_k = F(x_{k+1}) - F(x_k)$, $k = 0, 1, \ldots$, $F : \Re^n \to \Re^n$.

6. Enunciare il teorema di Dennis–Morè.

7. Discutere gli aspetti implementativi legati alla risoluzione di $B_k\delta_k = -F_k$ nei metodi quasi–Newton.

8. Dato il sistema non lineare $F(x) = 0$ lo si risolva con un metodo quasi–Newton: la direzione s_k che si ottiene è una direzione di discesa per la cifra di merito $\phi(x) = \frac{1}{2}||F||_2^2$ in x_k? Perchè? In caso di risposta non affermativa, illustrare gli accorgimenti che occorre introdurre nell'implementazione dei metodi di questo tipo.

9. Data la funzione quadratica convessa $f(x_k + s) = g^T s + \frac{1}{2}s^T\nabla^2 f(x_k)s$, determinare il passo di massima discesa $\alpha_k\delta_k$ a partire da x_k.

10. Mostrare che l'update SR1 (simmetrico di rango 1) è invariante allo scaling se l'approssimazione iniziale dell'hessiana è scelta opportunamente. Ovvero, data la trasformazione $\hat{x} = Ax + b$, con $A \in \Re^{n\times n}$ invertibile, $b \in \Re^n$ e $\hat{f}(\hat{x}) = f(x)$, bisogna mostrare che, applicando tale metodo a $\hat{f}(\hat{x})$ con $\hat{x}_0 = Ax_0 + b$ e hessiana iniziale $\hat{B}_0 = A^{-T}B_0A^{-1}$, la successione $\{\hat{x}_k\}$ è tale che $\hat{x}_k = Ax_k + b$.

    (SUGGERIMENTO: occorre prima ricavare la relazione tra i gradienti $\nabla_x f$ e $\nabla_{\hat{x}}\hat{f}$ e quella tra le hessiane.)