Profa. Cristiane - 18/03/2006
Aula 5 - Métodos de Otimização Irrestrita (continuação)

3.3 - Algoritmo com Convergência Global

(-) Para impedir passos que se aproximem muito rapidamente de zero iremos requisitar que $\|d_k\| \geq \sigma \|f(x_k)\|$, $\forall\ k \in \mathcal{N}$, onde $\sigma > 0$ é uma constante.

(-) Para impedir passos grandes com pouco decréscimo na busca linear, pediremos que λ_k verifique $f(x_k + \lambda_k d_k) < f(x_k) + \alpha \nabla^t f(x_k) \lambda_k d_k$, $\forall\ k \in \mathcal{N}$, onde $\alpha \in (0, 1)$ é uma constante.

Condição de Armijo (ou decréscimo suficiente): $\alpha \nabla f(x_k) \lambda_k d_k < 0$. **Interpretação geométrica**

(-) Para impedir que as direções sejam quase ortogonais a $\nabla f(x_k)$ iremos impor que dada uma constante $\sigma \in (0, 1)$, $\nabla f(x_k) d_k \leq -\sigma \|\nabla f(x_k)\|\ \|d_k\|$, $\forall\ k \in \mathcal{N}$.

Se β é o ângulo entre $\nabla f(x_k)$ e d_k, então $\cos\beta = \dfrac{\nabla f(x_k) d_k}{\|\nabla f(x_k)\|\ \|d_k\|}$.

Vamos definir uma algortimo para minimizar funções sem restrições que seja o mais geral possível e que incorpore essas condições. Sejam $\sigma > 0$, α e $\theta \in (0, 1)$ constantes dadas. Se $x_k \in \Re^n$ é tal que $\nabla f(x_k) \neq 0$, os passos para determinar x_{k+1} são:

ALG. 2 *Passo 1:* Escolher $d_k \in \Re^n$ tal que: (i) $\|d_k\| > \sigma\ \|\nabla f(x_k)\|$;

(ii) $\nabla^t f(x_k) d_k \leq -\sigma \|f(x_k)\|\ \|d_k\|$.

Passo 2: (Busca linear) (i) $\lambda = 1$;

(ii) Se $f(x_k + \lambda d_k) < f(x_k) + \alpha\ \lambda\ \nabla f(x_k) d_k$ ir para (iv);

(iii) Escolher $\overline{\lambda} \in [0.1\lambda, 0.9\lambda]$. Fazer $\lambda = \overline{\lambda}$ e ir para (ii);

(iv) Fazer $\lambda_k = \lambda$ e $x_{k+1} = x_k + \lambda_k d_k$.

Teorema (Convergência Global do ALG. 2): O ALG.2 pára com algum valor de k tal que $\nabla f(x_k) = 0$, ou gera uma sequência infinita $\{x_k\}$ tal que qualquer ponto de acumulação dela é um ponto estacionário de f.

Ordem de Convergêcia: $e_{k+1} = \|x_{k+1} - x^*\|$ e $e_k = \|x_k - x^*\|$. É desejável que a partir de algum índice k_0, seja verdade que $e_{k+1} \leq r\ e_k$, $\forall r \in (0, 1)$.

Definição 1: Se $e_{k+1} \leq r\ e_k$ se verifica para algum $r \in (0, 1)$, diremos que $\{x_k\}$ converge com ordem linear e taxa não superior a r.

Definição 2: Se $\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = 0$ diremos que a sequência $\{x_k\}$ converge com ordem super linear.

Definição 3: Se $e_{k+1} \leq a\ (e_k)^p$, onde $a > 0$ e $p > 1$, diremos que a sequência $\{x_k\}$ converge para x^* com ordem não inferior a p. Se $p = 2$, diremos que a convergência é quadrática.