Profa. Cristiane - 13/05/2006
Aula 12 - Métodos de Otimização Restrita (continuação)

4.3 - Algoritmos de descida - Método do Gradiente Projetado

Motivado pelo método da descida de maior passo. A idéia é projetar a direção $-\nabla f(\cdot)$ sobre a superfície de trabalho, para definir a direção do movimento.
Considere problemas da forma:

$$\begin{array}{lll} (P1) & Min & f(x) \\ & s.a & a_i(x) = b_i \quad i \in I_1 \\ & & a_i(x) \leq b_i \quad i \in I_2 \end{array}$$

1^o) Uma solução factível, se existente, pode ser obtida via fase I de um procedimento de programação linear.
2^o) Partindo de um ponto factível x, existem q restrições ativas $a_i x = b_i$ e outras inativas $a_i x < b_i$. O conjunto $W(x)$ será o conjunto das restrições ativas.
3^o) A partir do ponto x, procura-se por uma direção factível d satisfazendo $\nabla f(x)^t d < 0$, de maneira que um crescimento na direção d provoque um decrescimento na função f.
4^o) Inicialmente, consideram-se direções d tais que $a_i d = 0$, $i \in W(x)$, de tal forma que todas as restrições de trabalho permaneçam ativas. Portanto, o vetor direção d está no subespaço tangente M, definido pelas restrições do conjunto de trabalho. O vetor direção usável é obtido pela projeção de $-\nabla f(x)$ nesse subespaço.

4.3.1 - Cálculo da projeção

Seja A_q a matriz composta pelas restrições ativas. Assumindo regularidade, A_q é uma matriz qxn de posto $q < n$.

O subespaço tangente M no qual d deve ficar é o subespaço de vetores que satisfazem $A_q d = 0$. Portanto, o subespaço N composto pelos vetores na forma $A_q^t \lambda$ ($\lambda \in \Re^q$) é ortogonal a M.

Na verdade, qualquer vetor pode ser escrito como a soma de vetores de cada um desses dois subespaços complementares: em particular, $-\nabla f(x_k)$ pode ser escrito como

$$-\nabla f(x_k) = d_k + A_q^t \lambda_k, \quad \text{onde } d_k \in M \text{ e } \lambda_k \in \Re^q. \text{ Com a restrição } \ A_q d_k = 0, \text{ pode-se escrever}$$

$$A_q d_k = -A_q \nabla f(x_k) - A_q A_q^t \lambda_k = 0 \quad \text{que leva a} \quad \lambda_k = -(A_q A_q^t)^{-1} A_q \nabla f(x_k) \text{ e portanto}$$

$$d_k = -\left[I - A_q^t (A_q A_q^t)^{-1} A_q\right] \nabla f(x_k) = -P_k \nabla f(x_k)$$

4.3.2 - Escolha do passo

Quando α começa a sair do zero, $x+\alpha d$ continua factível (incialmente) e o valor de f decresce. Deve-se encontrar o tamanho do segmento de reta factível a partir de x, e minimizar-se f sobre esse segmento.

Se o mínimo ocorre no extremo, então uma nova restrição tornou-se ativa, e deve ser acrescentada ao conjunto de trabalho. Se a projeção de $-\nabla f(x_k)$ for zero, tem-se $\nabla f(x_k) + A_q^t \lambda_k = 0$ e o ponto x_k satisfaz as condições necessárias para um mínimo sobre a superfície M.

Se além disso, as componentes de λ_k correspondentes às restrições ativas forem todas não-negativas, as condições KT para o problema original estão satisfeitas em x_k, e o processo termina. Nesse caso, λ_k encontrado pela projeção é praticamente o multiplicador de Lagrange do problema original (a menos dos multiplicadores nulos, associados às restrições ativas).

Se, entretanto, ao menos um dos componentes de λ_k é negativo, é possível (relaxando-se a desigualdade correspondente) mover-se numa direção e melhorar a solução.

4.3.3 - O Algoritmo do Método do Gradiente Projetado

ALG. 7: Seja $x_k \in \Re^n$ factível dado:

Passo 1: Encontre M (subespaço das restrições ativas) e construa A_q e $W(x)$.
Passo 2: Calcule $P = I - A_q^t (A_q A_q^t)^{-1} A_q$ e $d = -P \nabla f(x)$
Passo 3: Se $d \neq 0$, encontre α_1 e α_2 tais que:

$$\alpha_1 = \max\{\alpha : x + \alpha d \text{ é factível}\} \quad \text{e} \quad \alpha_2 = \min\{f(x+\alpha d) : 0 \leq \alpha \leq \alpha_1\}$$

Faça $x \leftarrow x + \alpha_2 d$ e volte para o *Passo 1.*
Passo 4: Se $d = 0$, encontre $\lambda_k = -A_q^t (A_q A_q^t)^{-1} A_q \nabla f(x_k)$

a) Se $\lambda_j \geq 0$ para todo j correspondente as restrições ativas, pare; x satisfaz as condições de KT.

b) Caso contrário, retire da matriz A_q a linha correspondente à desiguldade de λ mais negativo (e exclua a restrição correspondente de $W(x)$) e volte para o *Passo 2.*