TEMA Tend. Mat. Apl. Comput., **13**, No. 3 (2012), 219-231.
doi: 10.5540/tema.2012.013.03.0219

Método de Pontos Interiores Barreira Logarítmica Preditor-Corretor Especializado para o Problema de Regressão pela Norma Lp[1]

D.R. CANTANE[2], Departamento de Bioestatística, Instituto de Biociências, IBB, UNESP - Univ. Estadual Paulista, Distrito de Rubião Júnior, s/n, 18618-970 Botucatu, SP, Brasil.

E.G. CONTHARTEZE [3], A.R.L. OLIVEIRA[4], Instituto de Matemática, Estatística e Computação Científica, IMECC, UNICAMP - Universidade Estadual de Campinas, Rua Sérgio Buarque de Holanda, 651, 13083-859 Cidade Universitária "Zeferino Vaz", Distr. Barão Geraldo, Campinas, SP, Brasil.

Resumo. Os métodos de pontos interiores barreira logarítmica e preditor-corretor são aplicados ao problema de regressão pela norma L_p com algumas particularidades com o objetivo de obter uma implementação eficiente. O problema de regressão tem inúmeras aplicações em diversas áreas. A norma-2 é muito popular, entre outros motivos, por permitir uma solução direta. Por sua vez, a norma-1 permite reduzir o efeito de pontos discrepantes enquanto que a norma-∞ garante proteção contra o pior caso. A norma$-p$ permite pensar estas características de diferentes formas, adaptando o método ao problema a ser resolvido. A implementação do método de pontos interiores desenvolvida é comparada com métodos existentes.

Palavras-chave. Métodos de pontos interiores, problema de regressão, norma L_p.

1. Introdução

O método IRLS *iteratively reweighted least-squares* [5] foi por muito tempo a única alternativa prática para a resolução do problema de regressão pela norma L_p. Mais recentemente este método foi aperfeiçoado, no que diz respeito a robustez, por meio da inclusão de uma busca linear [4]. No mesmo trabalho, foi também proposto um novo método que apresenta características similares aos métodos de pontos interiores. Este método apresentou resultados computacionais superiores ao IRLS.

Ambos métodos apresentados em [4] têm uma importante desvantagem: a busca linear é computacionalmente cara. Os métodos de pontos interiores aplicados a

[1]Agradecimentos à FAPESP, CNPq e CAPES; Trabalho apresentado no Congresso de Matemática Aplicada e Computacional - CMAC 2011.
[2]dcantane@ibb.unesp.br.
[3]contharteze@yahoo.com.br.
[4]aurelio@ime.unicamp.br

Recebido em 21 Outubro 2011; Aceito em 07 Setembro 2012.

este problema obtém resultados computacionais superiores, repetindo o desempenho obtido na minimização pelas normas L_1 e L_∞ em [8] e [7], respectivamente.

Em [1] e [6] são descritos os métodos de pontos interiores barreira logarítmica, barreira logarítmica preditor corretor, primal dual barreira logarítmica e primal dual barreira logarítmica preditor corretor para resolver problemas de regressão pela norma L_p, com $1 < p < 2$. Esses métodos foram implementados em MATLAB e os resultados computacionais comprovam que são mais eficientes que o GNCS em [1] e [6], um método de Newton globalizado que usa as condições de folgas complementares.

O objetivo desse trabalho é aperfeiçoar e implementar métodos de pontos interiores para resolver de maneira eficiente problemas de otimização não-linear

$$\min_{x \in \Re^n} \quad \|Ax - b\|_p^p, \tag{1.1}$$

onde $A \in \Re^{m \times n}$ é uma matriz de posto completo, $b \in \Re^m$, $m > n$ e $1 < p < \infty$. O sistema linear $Ax = b$ pode ser inconsistente e a matriz A é uma matriz de Vandermonde no caso particular do problema de regressão polinomial.

Apresentamos neste trabalho os métodos barreira logarítmica e barreira logarítmica preditor corretor, desenvolvidos [1] e [6], com modificações nos pontos iniciais, no cálculo dos resíduos e direções, com a intenção de reduzir o número de operações e fazer com que estes métodos sejam computacionalmente mais eficientes. Os novos métodos implementados são comparados com os métodos descritos em [1] e [6].

2. O problema de regressão pela norma L_p

O problema de regressão (1.1), quando $p = 1$ e $p = \infty$, pode ser formulado por programação linear e os métodos de pontos interiores aplicados a esses problemas permitem a exploração da estrutura matricial do problema de forma muito eficiente. A norma-2 permite solução direta. A norma L_p, com $1 < p < \infty$ é estritamente convexa. O problema (1.1) possui formulação convexa e portanto é garantida a existência de mínimo global. O problema de regressão pode ser reescrito como

$$\begin{array}{ll} \min & \|r\|_p^p \\ \text{s.a} & Ax + r = b. \end{array} \tag{2.1}$$

Definindo $r = u - v$, com $u \geq 0$ e $v \geq 0$

$$\begin{array}{ll} \min \phi(u, v) = & \displaystyle\sum_{i=1}^{m} (u_i + v_i)^p \\ \text{s.a} & Ax + u - v - b = 0 \\ & (u, v) \geq 0. \end{array} \tag{2.2}$$

Observe que minimizar $\|r\|_p^p = \displaystyle\sum_{i=1}^{m} |u_i - v_i|^p$ é equivalente a minimizar $\phi(u, v) = \displaystyle\sum_{i=1}^{m} (u_i + v_i)^p$ desde que se verifiquem as restrições $u_i v_i = 0$, $i = 1, ..., m$. Temos que, dado r_i sempre podemos escrevê-lo como $u_i - v_i$ de forma que $u_i v_i = 0$, por exemplo, se $r_i < 0$, basta fazer $u_i = 0$ e $v_i = -r_i$, por outro lado, se $r_i \geq 0$, podemos fazer $u_i = r_i$ e $v_i = 0$, assim teremos $\|r\|_p^p = \sum_{i=1}^{m} |u_i - v_i|^p = \sum_{i=1}^{m} (u_i + v_i)^p$.

3. Modificação nos métodos dos pontos interiores aplicados ao problema de regressão pela norma L_p

Os métodos de pontos interiores desenvolvidos em [1]: barreira logarítimica e barreira logarítimica preditor corretor foram modificados. As modificações foram feitas no ponto inicial, inserindo o parâmetro σ nos vetores u e v e no cálculo dos resíduos R_1 e R_2, que não eram considerados anteriormente. Essas modifições foram realizadas no intuito de reduzir o número de operações necessárias por iteração. Além disso, os métodos também foram implementados na linguagem C, e não somente em Matlab, como em [1], com o objetivo de obter maior eficiência computacional.

3.1. Método barreira logarítmica

O método descrito a seguir foi desenvolvido em [1]. Nesse caso, as alterações foram realizadas no ponto inicial. Considerando o problema (2.2), a função objetivo é denotada em termos de u e v por $\phi(u,v) = \sum_{i=1}^{n} (u_i + v_i)^p$, o gradiente $\nabla\phi(u,v)$ é denotado por $g = \begin{bmatrix} g_u \\ g_v \end{bmatrix}$, onde $g_{u_i} = g_{v_i} = p(u_i + v_i)^{p-1}$ e $\nabla^2\phi = \begin{bmatrix} \nabla g_u \\ \nabla g_v \end{bmatrix}$, onde $G = \nabla g_{u_{ij}} = \nabla g_{v_{ij}} = \begin{cases} \dfrac{p(p-1)}{(u_i+v_i)^{2-p}}, & \text{se } i = j, \\ 0, & \text{se } i \neq j \end{cases}$ é uma matriz diagonal.

Utilizando o Método Barreira Logarítmica [3], temos

$$\begin{aligned} \min \ & \sum_{i=1}^{n} (u_i + v_i)^p - \mu \sum_{i=1}^{n} \ln(u_i) - \mu \sum_{i=1}^{n} \ln(v_i) \\ \text{sa} \ & Ax + u - v - b = 0, \end{aligned}$$

onde $\mu > 0$ é o parâmetro barreira ($\mu \to 0$). A Lagrangeana é dada por

$$L = \sum_{i=1}^{n} (u_i + v_i)^p - \mu \sum_{i=1}^{n} \ln(u_i) - \mu \sum_{i=1}^{n} \ln(v_i) + y^t(Ax + u - v - b),$$

onde y é o multiplicador de Lagrange.

Aplicando as condições de otimalidade [9], obtemos

$$\underbrace{\nabla L}_{J(x,y,u,v)} = \begin{bmatrix} A^t y \\ Ax + u - v - b \\ g - \mu U^{-1} e + y \\ g - \mu V^{-1} e - y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{3.1}$$

onde U e V são matrizes diagonais cujos elementos não nulos são u e v, respectivamente e $e = (1, 1, \ldots, 1)^t$. O sistema (3.1) pode ser reescrito como

$$\underbrace{\nabla L}_{J(x,y,u,v)} = \begin{bmatrix} A^t y \\ Ax + u - v - b \\ U(g+y) \\ V(g-y) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu e \\ \mu e \end{bmatrix}. \tag{3.2}$$

Usando o método de Newton [9], chegamos a

$$\begin{bmatrix} 0 & A^t & 0 & 0 \\ A & 0 & I & -I \\ 0 & U & diag(g+y) + UG & UG \\ 0 & -V & VG & diag(g-y) + VG \end{bmatrix} \begin{bmatrix} dx \\ dy \\ du \\ dv \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}, \tag{3.3}$$

onde

$$\begin{aligned} r_1 &= -A^t y, \\ r_2 &= -Ax - u + v + b, \\ r_3 &= -U(g+y) + \mu e, \\ r_4 &= -V(g-y) + \mu e. \end{aligned}$$

Resolvendo o sistema, obtemos as direções dx, dy, du e dv. Por meio de eliminação de variáveis, o sistema se reduz a

$$\begin{aligned} r &= r_1 + A^t (D_2)^{-1} \left\{ D_1 r_2 + (D_u + UG) \, r_4 - (D_v + VG) \, r_3 \right\}, \\ dx &= \left(A^t (D_2)^{-1} D_1 A \right)^{-1} r, \\ dy &= (D_2)^{-1} \left\{ D_1 \left(A dx - r_2 \right) - (D_u + UG) \, r_4 + (D_v + VG) \, r_3 \right\}, \\ dv &= (D_1)^{-1} \left[D_u \left(r_4 + V dy \right) - VG r_3 + VUG dy \right], \\ du &= (D_u)^{-1} (r_3 - U dy - UG dv), \end{aligned}$$

onde

$$\begin{aligned} D_u &= \left[diag(g+y) + UG \right], \\ D_v &= \left[diag(g-y) + VG \right], \\ D_1 &= \left[D_u diag(g-y) + diag(g+y) VG \right], \\ D_2 &= \left[V \left(D_u + UG \right) + U \left(D_v + VG \right) \right]. \end{aligned}$$

Calculadas as direções, o tamanho do passo α é definido de forma a manter as variáveis u e v estritamente positivas

$$\alpha = \min \left\{ \tau \left(\min_{du_i < 0} - \frac{u_i}{du_i} \right), \tau \left(\min_{dv_i < 0} - \frac{v_i}{dv_i} \right), 1 \right\}, \text{ onde } \tau = 0,99995.$$

Conhecendo as direções e o tamanho do passo, a atualização das variáveis e do parâmetro barreira são dados por:

$$x^{k+1} = x^k + \alpha^k dx^k, y^{k+1} = y^k + \alpha^k dy^k,$$

$$u^{k+1} = u^k + \alpha^k du^k, v^{k+1} = v^k + \alpha^k dv^k, \mu^{k+1} = \frac{\mu^k}{\beta}, \text{ onde } \beta > 1.$$

Os critérios de convergência são baseados nas condições de otimalidade (3.1) e na diferença dos valores de N atual e da iteração anterior

$$N^k = \frac{\left\|\nabla L^k\right\|}{(1+\|x^k\|+\|u^k\|+\|v^k\|+\|y^k\|)(2m)} \leq \epsilon, \ \left|N^{k+1}-N^k\right| \leq \epsilon_1.$$

O ponto inicial é calculado baseado nas idéias desenvolvidas em [2], com modificações nos vetores u e v

$$x^0 = (A^tA)^{-1}A^tb, \quad r^0 = b - Ax^0, \quad y^0 = \frac{\kappa r^0}{\|r^0\|_\infty},$$

$$\begin{cases} r_i^0+\sigma, & \text{se } r_i^0 \geq 0, \\ \sigma, & \text{caso contrário,} \end{cases} \quad \begin{cases} \sigma, & \text{se } r_i^0 \geq 0, \\ \sigma - r_i^0, & \text{caso contrário,} \end{cases} \quad \text{onde } 0 < \sigma \leq 1.$$

Em [2] u e v não estão definidos, no entanto, esta escolha foi proposta em Oliveira, Nascimento e Lyra [8] para satisfazer a relação $u^0+v^0 = |\sigma r^0|$. Estabelecemos $\kappa = 0,975$.

3.2. Método Preditor-Corretor

Este método, desenvolvido também em [6], é apresentado com modificações no cálculo dos resíduos R_1 e R_2, além do ponto inicial. Considere o sistema de equações (3.3). No método preditor-corretor, inicialmente considera-se a direção afim, onde o parâmetro barreira $\mu = 0$. Assim, precisa-se resolver o sistema

$$\begin{bmatrix} 0 & A^t & 0 & 0 \\ A & 0 & I & -I \\ 0 & U & diag(g+y)+UG & UG \\ 0 & -V & VG & diag(g-y)+VG \end{bmatrix} \begin{bmatrix} \bar{dx} \\ \bar{dy} \\ \bar{du} \\ \bar{dv} \end{bmatrix} = \begin{bmatrix} \bar{r_1} \\ \bar{r_2} \\ \bar{r_3} \\ \bar{r_4} \end{bmatrix}, \tag{3.4}$$

onde

$$\begin{aligned} \bar{r_1} &= -A^t y, \\ \bar{r_2} &= -Ax - u + v + b, \\ \bar{r_3} &= -U(g+y), \\ \bar{r_4} &= -V(g-y). \end{aligned}$$

Resolvendo o sistema (3.4), obtém-se as direções do passo preditor

$$\begin{aligned} \bar{dx} &= (A^tDA)^{-1}\bar{r}, \\ \bar{dy} &= D[\bar{r_2} - A\bar{dx} - D_u^{-1}\bar{r_3} + (D_v - VUG^2D_u^{-1})^{-1}(I + D_u^{-1}UG)(\bar{r_4} - D_u^{-1}VG\bar{r_3})], \\ \bar{dv} &= (D_v - VUG^2D_u^{-1})^{-1}[\bar{r_4} + V\bar{dy} + D_u^{-1}(-VG\bar{r_3} + VUG\bar{dy})], \\ \bar{du} &= D_u^{-1}(\bar{r_3} - U\bar{dy} - UG\bar{dv}), \end{aligned}$$

onde

$$\begin{aligned} D_u &= [diag(g+y) + UG], \\ D_v &= [diag(g-y) + VG], \\ D &= [-D_u^{-1}U - (D_v - VUG^2D_u^{-1})^{-1}V(I + D_u^{-1}UG)^2]^{-1}, \\ \bar{r} &= -\left\{\bar{r_1} - A^tD[\bar{r_2} - D_u^{-1}\bar{r_3} + (D_v - VUG^2D_u^{-1})^{-1}(I + D_u^{-1}UG)(\bar{r_4} - D_u^{-1}VG\bar{r_3})]\right\}. \end{aligned}$$

Para calcular os novos resíduos, deve-se substituir x por $x + dx$, y por $y + dy$, u por $u + du$ e v por $v + dv$ no sistema de equações (3.2) e compará-lo ao sistema de equações (3.3). As diferenças encontradas, R_1 e R_2 dados abaixo, são os resíduos. O primeiro sistema de equações do sistema (3.2) é dado por

$$A^t(y + dy) = 0 \Rightarrow A^t dy = -A^t y.$$

Comparando com o sistema (3.3), vemos que não existe resíduo. O segundo sistema de equações do sistema (3.2) é dado por

$$A(x + dx) + u + du - v - dv - b = 0 \Rightarrow Adx + du - dv = b - Ax - u + v.$$

Comparando com o sistema (3.3), vemos que nesse sistema também não existe resíduo. O terceiro sistema de equações do sistema (3.2) é dado por

$$(U + dU)(\tilde{g} + y + dy) = \mu e,$$

onde $\tilde{g}_i = p(u_i + du_i + v_i + dv_i)^{p-1}$, assim $\tilde{g} = \tilde{g}_i$, $i = 1, 2, ..., m$. Entretanto, $p(u_i + du_i + v_i + dv_i)^{p-1}$ pode resultar em um número imaginário, para corrigir esse problema devemos fazer $\tilde{g}_i = p(u_i + \beta du_i + v_i + \beta dv_i)^{p-1}$, onde β é da forma

$$\beta = \min \left\{ \tau \left(\min_{(d_{u_i} + d_{v_i}) < 0} -\frac{(u_i + v_i)}{(d_{u_i} + d_{v_i})} \right), 1 \right\}, \text{ onde } \tau = 0{,}99995. \tag{3.5}$$

Com base nos testes computacionais, $\tilde{g}_i$ não afeta a convergência deste método. Nesse sistema de equações, temos o resíduo

$R_1 = ((U + dU)(\tilde{g} + y + dy) - \mu e) - (Udy + diag(g)du + diag(y)du + UGdu + UGdv + Ug + Uy - \mu e) \Rightarrow R_1 = (U + dU)\tilde{g} + dUdy - dUg - UGdu - UGdv - Ug.$

O quarto sistema de equações do sistema (3.2) é dado por

$$(V + dV)(\tilde{g} - y - dy) = \mu e,$$

onde $\tilde{g} = p(u + \beta du + v + \beta dv)^{p-1}$ e β é calculado como em (3.5). Nesse sistema de equações, temos o resíduo

$$R_2 = ((V + dV)(\tilde{g} - y - dy) - \mu e) - (-Vdy + VGdu + diag(g - y)dv + VGdv + Vg - Vy - \mu e)$$

$$\Rightarrow R_2 = (V + dV)\tilde{g} - dVdy - VGdu - dVg - VGdv - Vg.$$

No passo corretor, pode-se impor um valor para o parâmetro barreira μ e resolver o sistema

$$\begin{bmatrix} 0 & A^t & 0 & 0 \\ A & 0 & I & -I \\ 0 & U & diag(g + y) + UG & UG \\ 0 & -V & VG & diag(g - y) + VG \end{bmatrix} \begin{bmatrix} \hat{dx} \\ \hat{dy} \\ \hat{du} \\ \hat{dv} \end{bmatrix} = \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \\ \hat{r}_3 \\ \hat{r}_4 \end{bmatrix},$$

onde

$$
\begin{aligned}
\hat{r}_1 &= -A^t y,\\
\hat{r}_2 &= -Ax - u + v + b,\\
\hat{r}_3 &= -U(g+y) + \mu e - R_1,\\
\hat{r}_4 &= -V(g-y) + \mu e - R_2,\\
R_1 &= (U + d\bar{U})\tilde{g} + d\bar{U}d\bar{y} - d\bar{U}g - UGd\bar{u} - UGd\bar{v} - Ug,\\
R_2 &= (V + d\bar{V})\tilde{g} - d\bar{V}dy - VGd\bar{u} - d\bar{V}g - VGd\bar{v} - Vg.\\
\tilde{g} &= p(u + \beta du + v + \beta dv)^{p-1},\\
\beta &= \min\left\{\tau\left(\min_{(\bar{d_{u_i}}+\bar{d_{v_i}})<0} -\frac{(u_i+v_i)}{(\bar{d_{u_i}}+\bar{d_{v_i}})}\right), 1\right\}, \text{ onde } \tau = 0.9995.
\end{aligned}
$$

Obtém-se as direções corretoras

$$
\begin{aligned}
\hat{dx} &= \left(A^t(D_2)^{-1}D_1A\right)^{-1}\hat{r},\\
\hat{dy} &= (D_2)^{-1}\left\{D_1\left(A\hat{dx} - \hat{r_2}\right) - (D_u + UG)\,\hat{r_4} + (D_v + VG)\,\hat{r_3}\right\},\\
\hat{dv} &= (D_1)^{-1}\left[D_u\left(\hat{r_4} + V\hat{dy}\right) - VG\hat{r_3} + VUG\hat{dy}\right],\\
\hat{du} &= (D_u)^{-1}(\hat{r_3} - U\hat{dy} - UG\hat{dv}),
\end{aligned}
$$

onde

$$
\begin{aligned}
D_u &= [diag(g+y) + UG],\\
D_v &= [diag(g-y) + VG],\\
D_1 &= [D_u diag(g-y) + diag(g+y)VG],\\
D_2 &= [V(D_u + UG) + U(D_v + VG)],\\
\hat{r} &= \hat{r_1} + A^t(D_2)^{-1}\left\{D_1\hat{r_2} + (D_u + UG)\,\hat{r_4} - (D_v + VG)\,\hat{r_3}\right\}.
\end{aligned}
$$

O tamanho do passo, a atualização das variáveis, o parâmetro barreira, o critério de convergência e o ponto inicial são calculados como no método barreira logarítmica.

4. Experimentos computacionais

A seguir, encontram-se os resultados computacionais em Matlab comparando essas modificações com os resultados obtidos em [1] e [6]. Os testes computacionais foram feitos no Matlab R2009a, sistema computacional Linux, memória 4GB, processador Intel(R) Core(TM)2. Para todos os problemas foi suposto que modelo de regressão adotado é a aproximação dos pontos por um polinômio de grau um (uma reta).

Utiliza-se as seguintes notações nas Tabelas a seguir: BL - Método Barreira Logarítmica; BL (antigo) - Método Barreira Logarítmica existente [1]; BLPC - Método Barreira Logarítmica Preditor-Corretor; BLPC (antigo) - Método Barreira Logarítmica Preditor-Corretor existente [1].

4.1. Problema de grande porte

Esse problema, descrito em [1] e [6], compreende valores de juros diários ao longo de 40 anos, totalizando 10958 valores observados, os dados foram normalizados no intervalo $[0,1]$ e assim, temos que a matriz de Vandermonde possui $m = 10958$ linhas. Utiliza-se $\mu = 0,001$, $\tau = 0,99995$, $\beta = 10$, $\epsilon = 10^{-10}$, $\epsilon_1 = 10^{-8}$, onde ϵ e ϵ_1 são do critério de convergência, e $\sigma = 10^{-1}$.

Pode-se observar nas Tabelas 1 e 2 que o número de iterações foi reduzido. Na Tabela 1 o tempo computacional também foi reduzido, porém na Tabela 2 o tempo computacional piorou, pois a nova forma de encontrar os resíduos no método barreira logarítmica preditor-corretor exige mais cálculos que a forma antiga.

4.2. Função cosseno

Esse problema contém 20001 pontos. A segunda coluna da matriz A é formada por elementos no intervalo $[0, 2\pi]$, o vetor b consiste no cosseno desses valores. Utiliza-se $\mu = 0,001$, $\tau = 0,99995$, $\beta = 10$, $\epsilon = 10^{-10}$, $\epsilon_1 = 10^{-8}$, onde ϵ e ϵ_1 são usados para a verificação do critério de convergência, e $\sigma = 10^{-10}$.

Analisando as Tabelas 3 e 4 percebe-se uma redução no número de iterações e no tempo computacional na maioria dos casos, o valor da função objetivo não se alterou.

4.3. Função logaritmo

Esse problema contém 15 mil pontos, os valores da segunda coluna de A são obtidos dividindo-se o intervalo $[1,4]$ em 15 mil. O vetor b é o valor do logaritmo neperiano do valor correspondente à segunda coluna de A. Obtém-se a matriz $A_{15000\times 2}$ e o vetor b. Utiliza-se $\mu = 0,001$, $\tau = 0,99995$, $\beta = 10$, $\epsilon = 10^{-10}$, $\epsilon_1 = 10^{-8}$, onde ϵ e ϵ_1 são usados para a verificação do critério de convergência, e $\sigma = 10^{-2}$.

Na Tabela 5 observa-se que para $p = 1,1$ o número de iterações aumentou, mas o valor da função objetivo reduziu. Nas Tabelas 5 e 6 o número de iterações diminuiu em aproximadamente 28% dos casos, no entanto, o tempo computacional foi reduzido na maior parte deles.

4.4. Função seno

Os valores da segunda coluna de A são obtidos dividindo-se o intervalo $[-2,2]$ em 40 mil e o vetor b é o valor do seno hiperbólico do valor correspondente na segunda coluna de A. Utiliza-se $\mu = 0,001$, $\tau = 0,99995$, $\beta = 10$, $\epsilon = 10^{-10}$, $\epsilon_1 = 10^{-8}$, onde ϵ e ϵ_1 são do critério de convergência, e $\sigma = 10^{-1}$.

Na Tabela 7 pode-se observar que o número de iterações e o tempo computacional foram reduzidos à medida que o valor de p aumenta. Na Tabela 8 o número de iterações e o tempo computacional diminuiram na maioria dos casos e a função objetivo foi ligeiramente maior em poucos casos.

Com um critério de parada mais rígido, os dois métodos tendem a obter um número de iterações semelhantes, o que beneficia a nova abordagem.

5. Conclusões

Nesse trabalho, os métodos de pontos interiores já existentes são aperfeiçoados. Modifica-se a forma de calcular o ponto inicial, as direções e os resíduos dos métodos. Como consequência, o número de iterações e o valor da função objetivo foram reduzidos na maioria dos casos. Testes computacionais foram realizados para comprovar que pode-se escolher σ, que é usado no cálculo do ponto inicial, de forma a diminuir ainda mais o valor da função objetivo.

Com as modificações no cálculo dos resíduos, o tempo por iteração aumenta ligeiramente, pois a nova forma de calculá-los é mais trabalhosa que a forma antiga, porém consegue-se diminuir o número de iterações no método barreira logarítmica preditor-corretor.

O problema (2.2) possui mínimo global, assim a solução ótima é única. Ao alterar o ponto inicial, o valor da função objetivo é menor e, consequentemente, mais próximo da solução ótima global.

A nova formulação obtida através das modificações para os métodos de pontos interiores traz melhoras tanto nos testes computacionais em MATLAB quanto na implementação eficiente na linguagem C. Na implementação eficiente, aproveitamos a estrutura matricial dos métodos para economizar memória, isso é possível porque muitos valores deixam de ser armazenados. Com isso, a complexidade computacional é reduzida e podemos resolver problemas de maior porte de forma eficiente. Em particular, não é necessário armazenar nenhuma matriz durante a solução de um problema.

Pode-se concluir que as modificações realizadas nos métodos BL e BLPC possuem melhor desempenho computacional em comparação aos resultados obtidos em [1] e [6], porém se o problema convergir em poucas iterações e o objetivo for diminuir o tempo computacional, o BLPC(antigo) é mais indicado, pois a forma de calcular os resíduos é mais simples.

O método BLPC é mais indicado quando temos problemas de grande porte, pois tem uma convergência mais rápida e reduz o número de iterações, entretanto, quando os problemas convergem em poucas iterações, o uso do método BL é mais indicado, pois os cálculos envolvidos são mais simples e conseguimos assim uma redução do tempo computacional.

Abstract. The logaritmic barrier and predictor-corrector interior point methods are applied to the L_p norm fitting problem exploiting the matrix structure in order to obtain an efficient implementation. The fitting problem has numerous applications in various areas. The 2-norm is very popular, among other reasons, for allowing a direct solution. The 1-norm allows the reduction of the effect of outliers while the ∞-norm provides protection against the worst case. The p-norm allows to think these characteristics in different ways adapting the method to the problem to be solved. The interior point method implementation to be developed is compared with existing methods.

Keywords. Interior point methods, fitting problems, L_p norm.

Referências

[1] D.R. Cantane, "Métodos de Pontos Interiores Aplicados ao Problema de Regressão pela Norma L_p", dissertação de mestrado ICMC-USP, São Carlos, 2004.

[2] T.F. Coleman, Y. Li, A globally and quadratically convergent affine scaling method for linear l_1 problems, *Math. Programming*, **56** (1992), 189-222.

[3] A.S. El-Bakry, R.A. Tapia, T. Tsuchiya, Y. Zhang, On the formulation and the theory of the Newton interior-point method for nonlinear programming, *Journal of Optimization Theory and Applications*, **89** (1996), 507-541.

[4] Y. Li, A globally convergent method for l_p problems, *SIAM J. Optimization*, **3** (1993), 609-629.

[5] G. Merle, H. Späcth, Computational experience with discrete l_p approximation, *Computing*, **12** (1974), 315-321.

[6] A.R.L. Oliveira, D.R. Cantane, Métodos de pontos interiores aplicados ao problema de regressão pela norma Lp, *TEMA. Tendências em Matemática Aplicada e Computacional*, **5** 2004, p. 281-291.

[7] A.R.L. Oliveira, C. Lyra, Interior point methods for the polynomial L_∞ fitting problems, *Internacional Transactions in Operational Research*, **11** (2004), 309-322.

[8] A.R.L. Oliveira, M.A. Nascimento, C. Lyra, Efficient implementation and benchmark of interior point methods for the polynomial L_1 fitting problems, *Statistics & Data Analysis*, **35** (2000), 119-135.

[9] S.J. Wright, "Primal-Dual Interior-Point Methods", SIAM Publications, SIAM Philadelphia, PA, USA, 1996.

Tabela 1: Problema de Juros - Barreira Logarítmica.

	BL			BL(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	25	0,2035	2,4550e+04	4	0,2581	2,6040e+04
1,2	19	0,2000	2,8274e+04	24	0,2855	2,8274e+04
1,3	13	0,1445	3,2781e+04	19	0,2320	3,2781e+04
1,4	10	0,1159	3,8243e+04	16	0,2006	3,8243e+04
1,5	9	0,0982	4,4881e+04	14	0,1838	4,4881e+04
1,6	8	0,0973	5,2952e+04	13	0,1539	5,2952e+04
1,7	8	0,0955	6,2775e+04	12	0,1577	6,2775e+04
1,8	8	0,0950	7,4734e+04	11	0,1554	7,4734e+04
1,9	7	0,0814	8,9311e+04	9	0,1335	8,9311e+04

Tabela 2: Problema de Juros - Barreira Logarítmica Preditor-Corretor.

	BLPC			BLPC(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	4	0,2311	2,4705e+04	6	0,1325	2,6040e+04
1,2	7	0,3770	2,8274e+04	26	0,3702	2,8274e+04
1,3	7	0,3729	3,2781e+04	19	0,2819	3,2781e+04
1,4	6	0,3233	3,8243e+04	14	0,2136	3,8243e+04
1,5	5	0,2617	4,4881e+04	12	0,1832	4,4881e+04
1,6	5	0,2925	5,2952e+04	10	0,1607	5,2952e+04
1,7	4	0,2269	6,2775e+04	9	0,1495	6,2775e+04
1,8	4	0,2276	7,4734e+04	9	0,1584	7,4734e+04
1,9	4	0,1922	8,9311e+04	9	0,1514	8,9311e+04

Tabela 3: Função Cosseno - Barreira Logarítmica.

	BL			BL(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	2	0,0549	1,2355e+004	11	0,2111	1,2355e+004
1,2	3	0,0699	1,2011e+004	12	0,2202	1,2011e+004
1,3	3	0,0714	1,1693e+004	8	0,1721	1,1693e+004
1,4	3	0,0769	1,1399e+004	6	0,1389	1,1399e+004
1,5	4	0,0092	1,1125e+004	6	0,1381	1,1125e+004
1,6	5	0,0999	1,0869e+004	6	0,1419	1,0869e+004
1,7	6	0,1163	1,0630e+004	6	0,1395	1,0630e+004
1,8	6	0,1132	1,0406e+004	5	0,1214	1,0406e+004
1,9	6	0,1343	1,0195e+004	6	0,1380	1,0195e+004

Tabela 4: Função Cosseno - Barreira Logarítmica Preditor-Corretor.

	BLPC			BLPC(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	2	0,1044	1,2355e+004	2	0,1020	1,2355e+004
1,2	4	0,1780	1,2011e+004	13	0,3024	1,2011e+004
1,3	3	0,1438	1,693e+004	10	0,2450	1,1693e+004
1,4	3	0,1436	1,1399e+004	8	0,2114	1,1399e+004
1,5	3	0,1332	1,1125e+004	7	0,1960	1,1125e+004
1,6	4	0,1748	1,0869e+004	6	0,1687	1,0869e+004
1,7	4	0,1720	1,0630e+004	6	0,1748	1,0630e+004
1,8	5	0,2152	1,0406e+004	6	0,1702	1,0406e+004
1,9	6	0,2422	1,0195e+004	6	0,1748	1,0195e+004

Tabela 5: Função Logaritmo - Barreira Logarítmica.

	BL			BL(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	15	0,1655	607,9389	2	0,0510	635,1600
1,2	2	0,0382	483,1952	3	0,0612	485,3428
1,3	2	0,0381	372,2146	2	0,0511	372,4476
1,4	2	0,0383	287,5839	2	0,0507	287,4594
1,5	2	0,0389	222,8549	2	0,0501	222,6275
1,6	2	0,0381	173,2170	2	0,0507	173,0235
1,7	3	0,0491	135,0870	2	0,0522	134,9931
1,8	2	0,0378	105,6851	2	0,0510	105,9290
1,9	2	0,0377	83,0445	6	0,0929	83,9296

Tabela 6: Função Logaritmo - Barreira Logarítmica Preditor-Corretor.

	BLPC			BLPC(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	1	0,0558	632,6171	1	0,0586	635,4659
1,2	2	0,0731	484,8184	3	0,0773	486,5972
1,3	3	0,0970	372,9440	3	0,0771	372,5131
1,4	2	0,0740	287,5702	3	0,0784	287,5263
1,5	3	0,0895	222,7532	2	0,0653	222,6095
1,6	2	0,0734	173,1270	2	0,0652	172,9877
1,7	2	0,0732	135,0390	2	0,0654	134,9850
1,8	2	0,0751	105,7751	2	0,0653	105,9557
1,9	2	0,0731	83,3012	4	0,0909	83,2252

Tabela 7: Seno Hiperbólico - Barreira Logarítmica.

	BL			BL(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	5	0,2037	7,4691e+003	1	0,1233	7,4721e+003
1,2	5	0,2098	6,5534e+003	1	0,1201	6,5561e+003
1,3	5	0,2068	5,7622e+003	1	0,1203	5,7673e+003
1,4	2	0,1151	5,0782e+003	5	0,2502	5,0635e+003
1,5	2	0,1159	4,4674e+003	4	0,2149	4,4664e+003
1,6	3	0,1465	3,9556e+003	7	0,3151	3,9568e+003
1,7	3	0,1453	3,5238e+003	5	0,2478	3,5253e+003
1,8	5	0,2089	3,1472e+003	6	0,2799	3,1471e+003
1,9	7	0,2629	2,8141e+003	8	0,3382	2,8141e+003

Tabela 8: Seno Hiperbólico - Barreira Logarítmica Preditor-Corretor.

	BLPC			BLPC(antigo)		
p	it	tempo	Fo	it	tempo	Fo
1,1	2	0,2289	7,4725e+003	1	0,1640	7,4721e+003
1,2	2	0,2202	6,5565e+003	1	0,1571	6,5561e+003
1,3	2	0,2156	5,7681e+003	17	0,7589	5,6138e+003
1,4	2	0,2178	5,0873e+003	3	0,2321	5,0691e+003
1,5	2	0,2056	4,4953e+003	10	0,5064	4,4339e+003
1,6	2	0,2239	3,9709e+003	4	0,2705	3,9632e+003
1,7	4	0,3671	3,6159e+003	10	0,4992	3,5244e+003
1,8	6	0,5069	3,1489e+003	6	0,5109	3,1489e+003
1,9	6	0,5073	2,8141e+003	6	0,5032	2,8141e+003