

Solução Iterativa dos Sistemas Lineares do Método de Pontos Interiores

C.T.L.S. GHIDINI^{1*}, A.R.L. OLIVEIRA² e M. SILVA³,

Recebido em 28 novembro, 2013 / Aceito em 27 agosto, 2014

RESUMO. Nesse trabalho, consideramos o método preditor-corretor, que é uma das variantes mais importante do método de pontos interiores devido à sua eficiência e convergência rápida. No método preditor-corretor, é preciso resolver dois sistemas lineares a cada iteração para determinar a direção preditora-corretora. A resolução desses sistemas é o passo que requer mais tempo de processamento, devendo assim ser realizada de maneira eficiente. Para obter a solução dos sistemas lineares do método preditor-corretor consideramos dois métodos iterativos de Krylov: MINRES e método dos gradientes conjugados. Para que estes métodos convirjam mais rapidamente um pré-condicionador especialmente desenvolvido para os sistemas lineares oriundos dos métodos de pontos interiores é usado. Experimentos computacionais em um conjunto variado de problemas de programação linear foram realizados com o intuito de analisar a eficiência e robustez dos métodos de solução dos sistemas.

Palavras-chave: métodos de pontos interiores, sistemas lineares, métodos iterativos.

1 INTRODUÇÃO

O trabalho de Karmarkar em 1984 [6] revolucionou a área de programação linear por apresentar um algoritmo com complexidade polinomial e excelente desempenho quando aplicado em problemas lineares de grande porte. A partir daí, foi desenvolvido uma nova linha de pesquisa: os métodos de pontos interiores. Esses métodos buscam a solução ótima de um problema de programação linear, percorrendo o interior da região de factibilidade, determinada pelas restrições do problema. Eles podem ser classificados em primal, dual e primal-dual ou ainda por métodos afim escala, método de redução de potencial e métodos de trajetória central.

Neste trabalho, estamos considerando o método preditor-corretor, que é um método do tipo primal-dual de trajetória central. O passo mais importante do método preditor-corretor consiste

*Autor correspondente: Carla Ghidini

¹FCA – Faculdade de Ciências Aplicadas, UNICAMP – Universidade Estadual de Campinas, 13484-350 Limeira, SP, Brasil. E-mail: carla.ghidini@fca.unicamp.br

²Departamento de Matemática Aplicada, IMECC – Instituto de Matemática, Estatística e Computação Científica, UNICAMP – Universidade Estadual de Campinas, 13083-859 Campinas, SP, Brasil. E-mail: aurelio@ime.unicamp.br

³Departamento de Matemática Aplicada, IMECC – Instituto de Matemática, Estatística e Computação Científica, UNICAMP – Universidade Estadual de Campinas, 13083-859 Campinas, SP, Brasil. E-mail: marilene@ime.unicamp.br

em encontrar as direções preditora (direção afim escala) e corretora. Para isso, alguns sistemas de equações lineares devem ser resolvidos a cada iteração, o que requer grande esforço computacional.

Uma forma de diminuir a complexidade computacional consiste em reduzir esses sistemas a sistemas de equações normais ou sistemas aumentados e depois aplicar algum método iterativo para resolvê-los.

Para resolver os sistemas de equações normais e aumentado foi feita uma implementação cuidadosa do método MINRES com reortogonalização [16]. Esta implementação foi comparada com uma implementação do método dos gradientes conjugados (GC) já existente [15]. Ambos métodos utilizam o subespaço de Krylov como subespaço de busca para encontrar uma solução aproximada do sistema, o que é uma estratégia vantajosa, uma vez que esse subespaço pode ser muito menor que o subespaço de busca completo.

Para que os métodos iterativos tenham um melhor desempenho é realizado o pré-condicionamento da matriz de restrições dos sistemas utilizando o pré-condicionador separador. Além disso, com o intuito de resolver um número maior de problemas e de forma mais eficiente é considerada uma abordagem que utiliza, inicialmente, o GC e depois de um certo número de iterações, se o critério pré-estabelecido para troca for satisfeito, o MINRES passa a ser utilizado. Mais detalhes sobre essa técnica são fornecidos na Seção 5.

2 MÉTODO DE PONTOS INTERIORES – PRIMAL DUAL

Considere o problema de programação linear na forma padrão:

$$\begin{aligned} \text{Min} \quad & c^t x \\ \text{s.a.} \quad & Ax = b \\ & x \geq 0, \end{aligned} \quad (2.1)$$

em que $A \in \mathbb{R}^{m \times n}$, $\text{posto}(A) = m$, $x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$. Associado ao problema (2.1), denominado de problema primal, temos o seguinte problema dual:

$$\begin{aligned} \text{Max} \quad & b^t y \\ \text{s.a.} \quad & A^t y + z = c \\ & z \geq 0, \end{aligned} \quad (2.2)$$

em que $y \in \mathbb{R}^m$ é o vetor de variáveis duais livres e $z \in \mathbb{R}^n$ de variáveis de folga.

As condições de otimalidade de primeira ordem (Karush-Kuhn-Tucker) de (2.1) e (2.2) são dadas por [18]:

$$\left\{ \begin{array}{l} Ax - b = 0 \\ A^t y + z - c = 0 \\ XZe = 0 \\ (x, z) \geq 0 \end{array} \right. \quad (2.3)$$

sendo $X = \text{diag}(x)$, $Z = \text{diag}(z)$ e $e \in \mathbb{R}^n$, tal que $e = (1, 1, \dots, 1)^t$.

Os métodos de pontos interiores do tipo primal-dual consistem em aplicar o método de Newton às condições de otimalidade (2.3), partindo de um ponto interior e mantendo interior a cada iteração. A direção afim escala (direção de Newton) é obtida resolvendo o seguinte sistema:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}y \\ \hat{d}z \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix}$$

em que $r_p = b - Ax$, $r_d = c - A^t y - z$, $r_a = -XZe$. Eliminando a variável $\hat{d}z$, obtemos o sistema aumentado:

$$\begin{bmatrix} -D^{-1} & A^t \\ A & 0 \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}y \end{bmatrix} = \begin{bmatrix} r_d - X^{-1}r_a \\ r_p \end{bmatrix} \quad (2.4)$$

em que $D = XZ^{-1}$. A forma mais usada para resolver o sistema (2.4) consiste em reduzi-lo, por meio da eliminação da variável $\hat{d}x$, a um sistema de equações normais $ADA^t \hat{d}y = AD(r_d - X^{-1}r_a) + r_p$:

$$(ADA^t) \hat{d}y = AD(r_d - X^{-1}r_a) + r_p. \quad (2.5)$$

Em geral, não é possível realizar um passo completo ao longo da direção de Newton, visto que $(x, z) \geq 0$ deve ser satisfeita. Dessa forma, o novo ponto é dado por:

$$(x^{k+1}, y^{k+1}, z^{k+1}) = (x^k, y^k, z^k) + \hat{\alpha}(\hat{d}x^k, \hat{d}y^k, \hat{d}z^k),$$

em que $\hat{\alpha} = (\hat{\alpha}_p, \hat{\alpha}_d, \hat{\alpha}_z)^t$. Os passos $\hat{\alpha}_p$ e $\hat{\alpha}_d$ são os passos primal e dual que preservam a não negatividade das variáveis x e z , respectivamente. Esses valores são determinados da seguinte forma:

$$\hat{\alpha}_p = \min \left\{ 1, \tau \min_i \left(-\frac{x_i^k}{\hat{d}x_i^k} \mid \hat{d}x_i^k < 0 \right) \right\}, \quad (2.6)$$

$$\hat{\alpha}_d = \min \left\{ 1, \tau \min_i \left(-\frac{z_i^k}{\hat{d}z_i^k} \mid \hat{d}z_i^k < 0 \right) \right\}, \quad (2.7)$$

em que $\tau \in (0, 1)$.

2.1 Método Preditor-Corretor

O método preditor-corretor desenvolvido por Mehrotra [9] consiste em utilizar uma direção composta por três componentes: direção afim escala, direção de centragem e direção de correção não linear [12]. Fazendo a correção não linear e introduzindo a perturbação para centragem, temos o sistema:

$$\begin{aligned} A\bar{d}x &= 0 \\ A^t \bar{d}y + \bar{d}z &= 0 \\ Z\bar{d}x + X\bar{d}z &= \mu e - \hat{D}_x \hat{D}_z e, \end{aligned}$$

em que $\mu = \left(\frac{\hat{y}}{\hat{y}}\right)^p \left(\frac{\hat{y}}{\hat{y}}\right)$, $\gamma = (x^k)^t z^k$ e $\hat{y} = \left(x^k + \hat{\alpha}_p \hat{d}x\right)^t \left(z^k + \hat{\alpha}_d \hat{d}z\right)$.

Mehrotra em [9] sugere $p = 2$ ou $p = 3$. Além disso, ao tomar $\hat{\alpha}_p = \hat{\alpha}_d = 1$, obtemos um ponto primal e dual factível. Assim, $r_p = r_d = 0$ e r_a na restrição de complementariedade é dado por $r_a = \hat{D}_x \hat{D}_z e$, em que $\hat{D}_x = \text{diag}(\hat{d}_x)$ e $\hat{D}_z = \text{diag}(\hat{d}_z)$.

A direção preditora-corretora é dada por: $d = \hat{d} + \bar{d}$ e determinada resolvendo o sistema:

$$\begin{aligned} A d x &= r_p \\ A^t d y + d z &= r_d \\ Z d x + X d z &= r_s, \end{aligned}$$

em que $r_s = r_a + \mu e - \hat{D}_x \hat{D}_z e$.

2.2 Pré-condicionador Separador

Para uma convergência mais rápida dos métodos iterativos o pré-condicionamento da matriz de restrições do sistema linear se faz necessário. O pré-condicionador separador [14] é específico para os sistemas lineares oriundos dos métodos de pontos interiores. Quando aplicado ao método dos gradientes conjugados evita o cálculo da matriz de equações normais e sua fatoração. Apresenta resultados muito bons próximos a uma solução do problema, mas não é eficiente nas primeiras iterações. Este pré-condicionador é calculado como segue: Considere a matriz de equações normais em (2.5). Tome $A = [B \ N]P$, em que $P \in \mathbb{R}^{n \times n}$ é uma matriz de permutação tal que $B \in \mathbb{R}^{m \times m}$ é não singular, então

$$\begin{aligned} ADA^t &= [B \ N]PDP^t \begin{bmatrix} B^t \\ N^t \end{bmatrix} \\ &= [B \ N] \begin{bmatrix} D_B & 0 \\ 0 & D_N \end{bmatrix} \begin{bmatrix} B^t \\ N^t \end{bmatrix} \\ &= BD_B B^t + ND_N N^t. \end{aligned}$$

O pré-condicionador é dado por: $D_B^{-\frac{1}{2}} B^{-1}$ e a matriz pré-condicionada M é como segue:

$$M = D_B^{-\frac{1}{2}} B^{-1} (ADA^t) B^{-t} D_B^{-\frac{1}{2}} = I_m + GG^t,$$

em que

$$G = D_B^{-\frac{1}{2}} B^{-1} N D_N^{\frac{1}{2}}.$$

Note que próximo a uma solução pelo menos $n - m$ elementos em D^{-1} são grandes. Dessa forma, uma escolha adequada das colunas de B faz com que os elementos em $D_B^{-1/2}$ e D_N sejam muito pequenos nesta situação. Neste caso, G aproxima-se da matriz nula, M se aproxima da matriz identidade e ambos, o menor e o maior autovalor de M , se aproximam do valor 1, assim como o número de condição $\kappa_2(M)$ ⁴.

⁴O número de condição de uma matriz M é definido por: $\kappa_2(M) = \|M\|_2 \|M^{-1}\|_2$.

3 MÉTODOS DO SUBESPAÇO DE KRYLOV

Considere o sistema linear:

$$Ax = b, \quad (3.1)$$

em que b é um vetor do \mathbb{R}^n e A é uma matriz $m \times n$. Sejam \mathcal{K} e \mathcal{L} dois subespaços do \mathbb{R}^n de dimensão n . Uma técnica de projeção consiste em encontrar em \mathcal{K} uma solução aproximada \hat{x} para (3.1), de forma que o resíduo $b - Ax$ seja ortogonal à \mathcal{L} , em que \mathcal{L} denota o espaço de restrições.

Métodos do subespaço de Krylov são baseados em processos de projeção em subespaços de Krylov. O subespaço de Krylov de ordem m associado à A e b é o subespaço gerado pelos vetores da sequência de Krylov: $\mathcal{K}^m(A, b) = \text{span}\{b, Ab, A^2b, \dots, A^{m-1}b\}$. Nos métodos de Krylov, \mathcal{K}^m é dado por: $\mathcal{K}^m(A, r^0) = \text{span}\{r^0, Ar^0, A^2r^0, \dots, A^{m-1}r^0\}$, em que $r^0 = b - Ax^0$ é o resíduo da solução inicial x^0 .

As diferentes versões dos métodos do subespaço de Krylov surgem a partir de diferentes escolhas do subespaço \mathcal{L}^m e das formas pelas quais o sistema é pré-condicionado. Os métodos GC e MINRES foram originados escolhendo $\mathcal{L}^m = A\mathcal{K}^m$.

3.1 Método de Arnoldi

O método de Arnoldi é um método de projeção ortogonal [10] utilizado para construir uma base ortonormal para o subespaço de Krylov $\mathcal{K}^m(A, b)$. Este procedimento começa com $v_1 = \frac{r^0}{\|r^0\|_2}$. Em seguida, calcula-se Av_1 e a partir dele constrói-se um vetor ortogonal a v_1 . Normaliza-se este vetor e obtém-se v_2 . Dessa forma, se v_1, \dots, v_j for uma base ortogonal para $\mathcal{K}^j(A, r^0)$, então para encontrar v_{j+1} basta calcular $t = Av_j$ e ortonormalizá-lo com respeito a v_1, \dots, v_j . Isto produz um método para a criação de uma base ortonormal para $\mathcal{K}^{j+1}(A, r_0)$. Os vetores v_1, \dots, v_{j+1} formam uma base para $\mathcal{K}^{j+1}(A, r_0)$, a menos que t seja igual a zero. A ortogonalização gera a relação expressa em termos de v_j : $AV_{m-1} = V_m H_{m,m-1}$, em que V_m denota a matriz com colunas v_1 até v_m . A matriz $H_{m,m-1}$ é uma matriz Hessenberg superior de ordem $m \times m - 1$, em que os elementos $h_{i,j}$ são definidos pelo método de Arnoldi.

A princípio o processo de ortogonalização pode ser realizado de diferentes formas, mas normalmente a abordagem mais usada é o processo de Gram-Schmidt [4].

3.2 Método de Lanczos

Assim como o método de Arnoldi, o método de Lanczos [8] é utilizado para gerar uma base ortonormal para um subespaço de Krylov associado à A e b . Porém, em Lanczos A deve ser uma matriz simétrica. Note que se A é simétrica, então $H_{m-1,m-1} = V_{m-1}^t A V_{m-1}$ é uma matriz tridiagonal. Assim, durante o processo de ortogonalização de A , vários termos h_{ij} se anulam, o que reduz H a uma matriz tridiagonal simétrica e faz com que o custo computacional associado à ortogonalização e ao armazenamento sejam reduzidos, em comparação ao método de Arnoldi.

O processo de Lanczos aproxima uma matriz simétrica A em uma matriz simétrica tridiagonal $T_{k+1,k}$ ⁵ com uma linha adicional na parte inferior.

Se definirmos T_k como sendo as primeiras k linhas de $T_{k+1,k}$, então T_k é uma matriz quadrada e simétrica. Dessa forma:

$$T_{k+1,k} = \begin{bmatrix} T_k \\ \beta_{k+1}e_k^t \end{bmatrix} \quad \text{e} \quad T_k = \begin{bmatrix} T_{k-1} & \beta_k e_{k-1} \\ \beta_k e_{k-1}^t & \alpha_k \end{bmatrix},$$

em que e_k^t representa o vetor canônico. O processo Lanczos determina o vetor v_{k+1} tomando como vetor inicial $v_0 = 0$ e $\beta_1 v_1 = b$, em que β_1 serve para normalizar v_1 : $\beta_{k+1} v_{k+1} = p_k - \alpha_k v_k - \beta_k v_{k-1}$, em que $\alpha_k = v_k^t p_k$ e $p_k = Av_k$. O valor β_{k+1} é usado para normalizar v_{k+1} . Na forma matricial,

$$AV_k = V_{k+1}T_{k+1,k}, \quad \text{em que } V_k = [v_1 | \dots | v_k]. \tag{3.2}$$

Em aritmética exata, as colunas de V_k são vetores colunas ortonormais. O processo termina quando $\beta_{k+1} = 0$ ($k \leq n$), então obtemos $AV_k = V_k T_k$.

Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica e $v_1 \in \mathbb{R}^n$ um vetor unitário. Então, a partir do processo de Lanczos, obtemos a relação garantida pelo Teorema 9.11 em [4]: $AV_k = V_k T_k + r_k e_k^t$.

3.3 Método dos Gradientes Conjugados

O método dos Gradientes Conjugados (GC) resolve sistemas de equações lineares, cuja matriz dos coeficientes é simétrica e definida positiva. Este método é baseado em Lanczos, a partir do qual é obtida uma base ortogonal de vetores para o subespaço de Krylov $\mathcal{K}^k(A, r_0)$. Além disso, o método GC é um método fundamentado na abordagem de Ritz-Galerkin. Na condição de Ritz-Galerkin, o novo resíduo $b - Ax_{k+1}$ deve ser ortogonal ao subespaço gerado por v_1, \dots, v_k , ou seja, $V_k^t(b - Ax_k) = 0$.

Sejam $b = \beta_1 v_1$ e e_1 o primeiro vetor canônico unitário em \mathbb{R}^k , segue que

$$V_k^t b = V_k^t (\beta_1 v_1) = \beta_1 = \beta_1 e_1,$$

visto que os vetores v_1, \dots, v_k formam uma base ortonormal para o subespaço de Krylov $\mathcal{K}^k(A, b)$. Procuramos por uma solução aproximada x_k para o sistema $Ax = b$ tal que $x_k \in \mathcal{K}^k(A, b)$. Dessa forma, podemos escrever x_k como uma combinação dos vetores da base de

$${}^5 T_{k+1,k} = \begin{bmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_k & \\ & & & \beta_k & \alpha_k & \\ & & & & \beta_{k+1} & \end{bmatrix}, \quad \text{em que } \alpha_j = h_{ij}, \beta_j = h_{j-1,j}.$$

mínimos de (3.3) obtida via fatoração QR é definida pelo seguinte sistema triangular: $R_k y_k = t_k$, em que $\|r_k\|_2 = \|b - Ax_k\|_2 = \|\phi_k\|_2^2$. Cada $Q_{i,i+1}$ de Q_k é definida em termos de c_i e s_i , isto é,

$$Q_{i,i+1} = \begin{bmatrix} I_{i-1} & & & \\ & c_i & s_i & \\ & s_i & -c_i & \\ & & & I_{k-i} \end{bmatrix}.$$

A partir das Rotações de Givens [4, Seção 5.1.8] podemos construir a matriz Q_k , tal que $Q_k T_{k+1,k}$ seja uma matriz triangular superior. Como a matriz $T_{k+1,k}$ é tridiagonal é preciso eliminar a linha abaixo da diagonal principal, assim dados t_{ii} e $t_{i+1,i}$, elementos da matriz $T_{k+1,k}$, obtemos os fatores c_i e s_i : (*Givens*($t_{ii}, t_{i+1,i}$) \rightarrow (c_i, s_i, t_{ii})). O MINRES determina a solução aproximada $x_k \in \mathcal{K}^k(A, b)$ do problema original $Ax = b$ da seguinte forma: $x_k = V_k y_k$. Pela relação $R_k y_k = t_k$ temos que $y_k = R_k^{-1} t_k$. Logo,

$$\begin{aligned} x_k &= V_k y_k = V_k R_k^{-1} t_k \\ &= D_k \begin{bmatrix} t_{k-1} \\ \tau_k \end{bmatrix} = \begin{bmatrix} D_{k-1} & d_k \end{bmatrix} \begin{bmatrix} t_{k-1} \\ \tau_k \end{bmatrix} \\ &= x_{k-1} + \tau_k d_k, \end{aligned} \tag{3.4}$$

em que $V_k R_k^{-1} = D_k$, ou melhor, $V_k = D_k R_k$ e

$$d_k = \frac{v_k - \delta_k^{(2)} d_{k-1} - \varepsilon_k^{(1)} d_{k-2}}{\gamma_k^{(2)}}. \tag{3.5}$$

4 IMPLEMENTAÇÕES

Considere o sistema de equações lineares simétrico $Ax = b$ e um pré-condicionador para este sistema $M = L^t L$. Então o sistema pré-condicionado é equivalente a ao sistema $(L^{-1} A L^{-t})(L^t x) = L^{-1} b$, em que $L^{-1} A L^{-t}$ é uma matriz simétrica e definida positiva. Dado o sistema linear $Ax = b$, obtemos MINRESP (MINRES pré-condicionado) aplicando MINRES ao sistema equivalente $\hat{A} \hat{x} = \hat{b}$, em que $\hat{A} = L^{-1} A L^{-1}$, $\hat{x} = L^{-1} x$ e $\hat{b} = L^{-1} b$. Seja $\hat{x} \in \mathcal{K}(\hat{A}, \hat{b})$, $\hat{v}_0 = 0$ e $\hat{\beta}_1 \hat{v}_1 = \hat{b}$. Definimos $\hat{z}_k = \hat{\beta}_k L \hat{v}_k$ e $\hat{q}_k = \hat{\beta}_k L^{-1} \hat{v}_k$ tal que $M \hat{q}_k = \hat{z}_k$. Então,

$$\hat{\beta}_k = \left\| \hat{\beta}_k \hat{v}_k \right\|_2 = \left\| L^{-1} \hat{z}_k \right\|_2 = ((\hat{z}_k^t (L^{-1})^t) (L^{-1} \hat{z}_k))^{\frac{1}{2}}.$$

Visto que $M^{-1} = (L^t)^{-1} L^{-1}$ e $M \hat{q}_k = \hat{z}_k$ temos $\hat{\beta}_k = (\hat{z}_k^t M^{-1} \hat{z}_k)^{\frac{1}{2}} = \sqrt{\hat{q}_k^t \hat{z}_k}$. Dessa forma, obtemos o processo de Lanczos pré-condicionado:

$$\begin{aligned} \hat{p}_k &= \hat{A} \hat{v}_k = \frac{1}{\hat{\beta}} L^{-1} A \hat{q}_k, & \hat{\alpha}_k &= \hat{v}_k^t \hat{p}_k = \frac{1}{\hat{\beta}^2} \hat{q}_k^t A \hat{q}_k & \text{e} \\ \hat{\beta}_{k+1} \hat{v}_{k+1} &= L^{-1} A L^{-1} \hat{v}_k - \hat{\alpha}_k \hat{v}_k - \hat{\beta}_k \hat{v}_{k-1}. \end{aligned}$$

Temos que q_k é a solução do sistema $M\hat{q}_k = \hat{z}_k$. Assim, multiplicando a equação $\hat{\beta}_{k+1}\hat{v}_{k+1}$ por L , obtemos:

$$\hat{z}_{k+1} = \hat{\beta}_{k+1}L\hat{v}_{k+1} = \frac{1}{\hat{\beta}_k}A\hat{q}_k - \frac{\hat{\alpha}_k}{\hat{\beta}_k}\hat{z}_k - \frac{\hat{\beta}_k}{\hat{\beta}_{k-1}}\hat{z}_{k-1}.$$

Dados o pré-condicionador M e *maxit* (número máximo de iterações), obtemos a partir do Algoritmo 1 uma solução aproximada x para o sistema $Ax = b$.

Algoritmo 4.1 MINRES PRÉ-CONDICIONADO

função MINRESP(A,b,M,maxit)

$$z_0 = 0; z_1 = b; \hat{c}_0 = -1;$$

Resolva $Mq_1 = z_1$.

$$\hat{\beta}_1 = \sqrt{b^t q_1}; \hat{\delta}_1^{(1)} = 0; \hat{s}_0 = 0; k = 1; d_0 = d_{-1} = x_0 = 0;$$

enquanto Nenhuma condição de parada é verdadeira **faça**

$$p_k = Aq_k; \hat{\alpha} = \frac{1}{\hat{\beta}_k}q_k^t p_k; z_{k+1} = \frac{1}{\hat{\beta}_k}p_k - \frac{\hat{\alpha}_k}{\hat{\beta}_k}z_k - \frac{\hat{\beta}_k}{\hat{\beta}_{k-1}}z_{k-1};$$

Resolva $Mq_{k+1} = z_{k+1}$.

$$\hat{\beta}_{k+1} = \sqrt{q_{k+1}^t z_{k+1}}; \hat{\delta}_k^{(2)} = c_{k-1}\hat{\delta}_k^{(1)} + \hat{s}_{k-1}\hat{\alpha}_k;$$

$$\hat{\gamma}_k^{(1)} = \hat{s}_{k-1}\hat{\delta}_k^{(1)} - \hat{c}_{k-1}\hat{\alpha}_k; \hat{\varepsilon}_{k+1}^{(1)} = \hat{s}_{k-1}\hat{\beta}_{k+1}; \hat{\delta}_{k+1}^{(1)} = -\hat{c}_{k-1}\hat{\beta}_{k+1};$$

$$\mathbf{Givens}(\hat{\gamma}_k^{(1)}, \hat{\beta}_{k+1}) \rightarrow (\hat{c}_k, \hat{s}_k, \hat{\gamma}_k^{(2)});$$

$$\hat{t}_k = \hat{c}_k\hat{\phi}_{k-1}; \hat{\phi}_k = \hat{s}_k\hat{\phi}_{k-1}; \hat{\psi}_{k-1} = \hat{\phi}_{k-1}\sqrt{(\hat{\gamma}_k^{(1)})^2 + (\hat{\delta}_{k+1}^{(1)})^2};$$

se $\hat{\gamma}_k^{(2)} \neq 0$ **então**

$$d_k = \frac{1}{\hat{\gamma}_k^{(2)}}\left(\frac{1}{\hat{\beta}_k}q_k - \hat{\delta}_k^{(2)}d_{k-1} - \hat{\varepsilon}_k^{(1)}d_{k-2}\right); x_k = x_{k-1} + \hat{t}_k d_k;$$

fim se $k \leftarrow k + 1$

$$\mathbf{fim enquanto} \hat{\phi} = \hat{\phi}_k; \quad x = x_k; \hat{\psi} = \hat{\phi}_k\sqrt{(\hat{\gamma}_{k+1}^{(1)})^2 + (\hat{\delta}_{k+2}^{(1)})^2};$$

fim função

O método MINRESP é construído utilizando o processo de Lanczos pré-condicionado. A cada etapa de Lanczos pré-condicionado, resolvemos o subproblema de quadrados mínimos equivalente ao apresentado em (3.3) associado ao sistema $\hat{A}\hat{x} = \hat{b}$. Para atualizar a solução do problema original basta multiplicar as equações equivalentes à (3.4) e (3.5), resultadas do MINRES aplicado ao sistema pré-condicionado $\hat{A}\hat{x} = \hat{b}$, por L^{-1} :

$$d_k = L^{-1}\hat{d}_k = \frac{1}{\hat{\gamma}_k^{(2)}}\left(\frac{1}{\hat{\beta}_k}q_k - \hat{\delta}_k^{(2)}d_{k-1} - \hat{\varepsilon}_k^{(1)}d_{k-2}\right) \quad \text{e} \quad x_k = L^{-1}\hat{x}_k = x_{k-1} + \hat{t}_k d_k.$$

Um pré-condicionador adequado para GC deve ser simétrico e definido positivo. A ideia do método do GC pré-condicionado (GCP) consiste em aplicar o método GC, como apresentado na Seção 3.3, para resolver um sistema pré-condicionado. Considere o sistema simétrico e definido

positivo $Ax = b$. Pré-condicionando esse sistema obtemos o sistema equivalente: $\hat{A}\hat{x} = \hat{b}$, em que L é uma matriz simétrica e definida positiva. Definimos o pré-condicionador simétrico e definido positivo M por: $M = L^tL$. Seja $r_k = b - Ax_k$ o resíduo do sistema original e z_k o resíduo do sistema pré-condicionado tal que $Mz_k = r_k$. Obtemos o processo iterativo do GCP aplicando o GC para resolver o sistema equivalente, isto é,

$$\begin{aligned}\alpha_k &= \frac{\hat{r}_{k-1}^t \hat{r}_{k-1}}{\hat{q}_k^t \hat{A} \hat{q}_k}; \hat{x}_k = \hat{x}_{k-1} + \alpha_k \hat{q}_k; \hat{r}_k = \hat{r}_{k-1} - \alpha_k \hat{A} \hat{q}_k; \\ \lambda_{k+1} &= \frac{\hat{r}_k^t \hat{r}_k}{\hat{r}_{k-1}^t \hat{r}_{k-1}}; \hat{q}_{k+1} = \hat{r}_k + \lambda_{k+1} \hat{q}_k.\end{aligned}\quad (4.1)$$

Algoritmo 4.2 Gradientes Conjugados Pré-condicionado

função GCP(A,b,maxit)

$\beta_1 = \|b\|_2; r_0 = b; q_1 = r_0; x_0 = 0; k = 1;$

enquanto Nenhuma condição de parada é verdadeira **faça**

Resolva $Mz_k = r_k$.

se $\eta_k \leq 0$ **então**

$x_k = 0;$

Pare;

$\triangleright q_k$ é um vetor nulo.

fim se

$\alpha_k = \frac{r_{k-1}^t z_{k-1}}{q_k^t A q_k}; x_k = x_{k-1} + \alpha_k q_k; r_k = r_{k-1} - \alpha_k A q_k;$

$\lambda_{k+1} = \frac{r_k^t z_k}{r_{k-1}^t z_{k-1}}; q_{k+1} = z_k + \lambda_{k+1} q_k;$

$k \leftarrow k + 1$

fim enquanto. $x = x_k;$

fim função

Visto que a matriz ADA^t do sistema de equações normais (2.5) é simétrica e definida positiva, podemos utilizar o GCP para determinar as direções de Newton nos métodos de pontos interiores. Esta versão do método do GCP fornece uma solução aproximada do sistema original diretamente sem ter que calculá-la a partir da solução aproximada do sistema de pré-condicionado. Os resultados teóricos e experimentais apresentados em [2] fornecem evidências que muitas vezes, em sistemas simétricos e definido positivo, o MINRES pode parar muito mais cedo que o GC. Em alguns casos, o MINRES pode convergir mais rapidamente que o GC. Porém, geralmente o GC converge mais rapidamente do que o MINRES considerando tanto $\|x^* - x_k\|_A$ ⁶ quanto $\|x^* - x_k\|_2$. Nos trabalhos [13] e [5], são feitas comparações entre os métodos MINRES, CG e também outros métodos de resolução de sistemas. A seguir, apresentamos o Algoritmo 3 para a nova abordagem proposta, a qual utiliza os métodos GCP e MINRESP juntos.

⁶A norma-A é definida como:

$$\|w\|_A = \sqrt{w^t A w}.$$

Algoritmo 4.3 Método Híbrido

```

função HIBRIDO(A,b,maxit)
  para  $i \leftarrow 1$  até  $maxit$  faça
    GCP(A, b,  $maxit$ )  $\rightarrow (x)$ ;
    se  $i > it$  então
      MINRESP(A, b,  $maxit$ )  $\rightarrow (x)$ ;
    fim se
  fim para
fim função

```

Para resolver um sistema $Ax = b$, o Algoritmo 3 utiliza, inicialmente, o GCP. Se este método não convergir até um certo número de iterações (it) pré-determinado, então o MINRESP passa a ser utilizado para resolver o sistema. Nos experimentos, consideramos $it =$ número de linhas da matriz dos coeficientes.

5 EXPERIMENTOS COMPUTACIONAIS

Para comparar o desempenho do método preditor-corretor ao resolver problemas de programação linear usando os métodos iterativos GC e MINRES, ambos pré-condicionados pelo pré-condicionador separador e uma nova abordagem híbrida, a qual utiliza, inicialmente, o CGP e em seguida o MINRESP, foram realizados experimentos computacionais em um Intel Core i5, 4 GB RAM, 2,3 GHZ, 225 GB HD, com sistema operacional Linux. O código foi implementado em linguagem C e os 63 problemas de programação linear testados, pertencentes as coleções Netlib [3], Qaplib [1], Kennington [7], STOCHLP [17] e MISC [11], que estão no formato MPS, foram lidos usando *callable library* do CPLEX.

O pré-condicionador separador requer o cálculo da fatoração LU da matriz do sistema. Uma nova fatoração LU é calculada de uma iteração para a outra sempre que os métodos GCP e MINRESP necessitam muitas iterações para atingir a convergência. Em [14] é proposto calcular uma segunda fatoração LU(refatoração) quando a primeira fatoração é muito densa a fim de melhorar o desempenho. Esta técnica foi considerada nos experimentos. As Tabelas 1 e 2 comparam o tempo total e número de iterações, respectivamente, do método preditor-corretor ao resolver os problemas de programação linear usando os métodos iterativos GCP e MINRESP. Foram resolvidos 38 problemas pelo GCP e/ou MINRESP. Nas colunas Refact. a refatoração está sendo realizada.

Os primeiros experimentos foram feitos considerando apenas o sistema aumentado e depois o sistema de equações normais. No primeiro caso, o método GCP não convergiu para nenhum problema enquanto o MINRESP mostrou-se mais robusto e apresentou bons resultados resolvendo 25 problemas, sendo que 2 deles foram resolvidos somente nesse caso. Considerando o sistema de equações normais os dois métodos apresentaram resultados satisfatórios e equilibrados. Quando a refatoração é feita, o GCP resolveu 3 problemas a mais que o MINRESP e em menos tempo em 50% dos problemas, porém a diferença no tempo entre os métodos não

Tabela 1: Tempo total do método de pontos interiores em segundos.

Problema	Sistema Aumentado				Equações Normais			
	Refact.		Fact.		Refact.		Fact.	
	GCP	MINRESP	GCP	MINRESP	GCP	MINRESP	GCP	MINRESP
afiro	*	0.00	*	0.00	0.00	0.00	0.00	0.00
agg	*	*	*	*	*	*	1.20	*
fit1d	*	0.49	*	0.36	*	0.28	*	0.06
fit2d	*	14.78	*	*	*	15.49	*	2.48
ganges	*	*	*	*	1.52	*	*	*
gen4	*	*	*	16.32	*	*	*	*
kb2	*	0.04	*	*	0.02	*	0.01	*
ken-07	*	0.86	*	0.88	0.61	0.75	0.60	0.27
nug05	*	*	*	0.1	0.08	0.09	0.70	0.03
nug06	*	*	*	*	0.42	0.52	*	*
nug07	*	*	*	*	4.66	*	5.19	4.66
nug30	*	*	*	*	*	5.57	*	*
pds-02	*	7.8	*	7.82	4.68	*	4.69	2.1
qap8	*	*	*	*	*	*	18.73	*
recipe	*	0.03	*	0.03	0.02	0.02	0.02	0.00
sc50a	*	0.01	*	0.00	0.00	0.00	0.00	0.00
sc50b	*	0.00	*	0.00	0.00	0.00	0.00	0.00
sc105	*	0.03	*	0.01	0.01	0.00	0.01	0.01
sc205	*	*	*	*	*	*	0.21	*
scagr7	*	0.03	*	0.02	0.01	0.01	0.01	0.01
scsd1	*	0.10	*	0.08	*	0.06	*	0.02
scsd6	*	0.24	*	0.23	0.14	0.16	0.14	0.05
scsd8	*	*	*	*	0.54	*	0.52	*
sctap1	*	0.33	*	0.32	0.25	0.26	0.23	0.08
sctap2	*	1.14	*	1.14	2.25	0.89	2.25	0.29
sctap3	*	21.72	*	*	66.2	*	*	*
share2b	*	0.31	*	0.21	0.11	0.12	*	*
shell	*	3.12	*	3.18	0.57	0.94	0.53	*
ship04l	*	0.36	*	0.31	0.24	0.27	0.18	0.08
ship04s	*	*	*	*	*	*	*	0.21
ship08l	*	*	*	*	*	*	*	0.37
ship08s	*	0.36	*	0.34	0.25	0.28	0.22	0.08
ship12l	*	1.17	*	1.22	0.8	0.87	0.71	0.26
ship12s	*	0.48	*	0.46	0.33	0.37	0.3	0.12
standgub	*	1.35	*	1.44	0.59	1.57	0.57	0.24
standmps	*	1.93	*	2.52	0.79	1.7	1.40	0.48
stocfor1	*	0.03	*	0.05	0.04	0.36	0.05	*
truss	*	*	*	*	12.45	*	*	*

Tabela 2: Total de iterações do método de pontos interiores.

Problema	Sistema Aumentado				Equações Normais			
	Refact.		Fact.		Refact.		Fact.	
	GCP	MINRESP	GCP	MINRESP	GCP	MINRESP	GCP	MINRESP
agg	*	*	*	*	*	*	36	*
fit1d	*	31	*	26	*	24	*	24
fit2d	*	95	*	*	*	121	*	88
ganges	*	*	*	*	20	*	*	*
gen4	*	*	*	43	*	*	*	*
kb2	*	22	*	*	11	*	11	*
nug05	*	*	*	8	8	8	8	8
nug06	*	*	*	*	9	9	*	*
nug07	*	*	*	*	14	*	12	14
nug30	*	*	*	*	*	29	*	*
pds-02	*	29	*	29	29	*	29	29
qap8	*	*	*	*	*	*	10	*
sc205	*	*	*	*	*	*	18	*
scagr7	*	16	*	16	10	16	16	16
scsd1	*	11	*	11	*	11	*	11
scsd8	*	*	*	*	11	*	11	*
sctap2	*	17	*	17	24	17	24	17
sctap3	*	32	*	*	91	*	*	*
share2b	*	29	*	19	19	19	*	*
shell	*	36	*	26	23	24	23	*
ship04s	*	*	*	*	*	*	*	20
ship08l	*	*	*	*	*	*	*	16
standgub	*	25	*	24	22	23	22	22
standmps	*	28	*	30	30	29	32	28
stocfor1	*	13	*	17	15	54	16	*
truss	*	*	*	*	18	*	*	*

foi muito significativa. Com relação ao número de iterações o GCP fez menos iterações em 9 problemas contra 6 problemas do MINRESP.

Considerando o caso em que a refatoração não é feita, observa-se um melhor desempenho do MINRESP, que foi mais rápido em todos os problemas resolvidos por ambos métodos. A diferença nos tempos foi bem significativa, chegando a ser 23 vezes menor em um dos problemas. Com relação ao número de iterações os resultados foram muito próximos.

Na Tabela 2, foram apresentados os resultados somente dos problemas em que houveram diferenças no número total de iterações. Com os experimentos já realizados, observamos que aproximadamente 34% dos problemas testados foram resolvidos somente por um dos métodos MINRESP ou GCP. Isso motivou a realização de mais um teste considerando o Algoritmo 3.

De acordo com os resultados obtidos, temos que o método MINRESP sem refatorar e usando o sistema de equações normais apresentou o melhor desempenho com relação ao tempo e como o GCP não funciona bem com o sistema aumentado, decidimos utilizar o método híbrido com sistema de equações normais e sem refatoração. A Tabela 3 compara os resultados desse novo teste.

Tabela 3: Comparação entre GCP, MINRESP e Híbrido.

Problema	GCP		MINRESP		Híbrido	
	Tempo(s)	It.	Tempo(s)	It.	Tempo(s)	It.
adlittle	*	*	*	*	0.01	13
afiro	0.00	9	0.00	9	0.00	9
agg	1.20	36	*	*	0.39	23
agg3	*	*	*	*	2.17	25
bandm	*	*	*	*	2.90	47
beaconfd	*	*	*	*	1.93	64
chr22b	*	*	*	*	30.64	30
chr25a	*	*	*	*	60.26	31
cre-a	*	*	*	*	28.48	33
czprob	*	*	*	*	0.59	33
els19	*	*	*	*	810.02	31
finnis	*	*	*	*	1.21	26
fit1d	*	*	0.06	24	0.29	25
fit1p	*	*	*	*	1.11	23
fit2d	*	*	2.48	88	11.18	57
ganges	*	*	*	*	1.37	20
israel	*	*	*	*	0.48	22
kb2	0.01	11	*	*	0	11
ken11	*	*	*	*	15.91	23
ken13	*	*	*	*	68.33	25
ken-07	0.60	16	0.27	16	0.54	16
lotfi	*	*	*	*	3.17	101
nug05	0.70	8	0.03	8	0.07	8
nug06	*	*	*	*	0.43	9
nug07	5.19	12	4.66	14	2.72	12
nug08	*	*	*	*	10.45	10
pds-02	4.69	29	2.10	29	4.06	29
pds-06	*	*	*	*	57.87	38
pds-10	*	*	*	*	188.27	54
qap8	18.73	10	*	*	9.79	10
recipe	0.02	11	0.00	11	0.03	11

Tabela 3: (continuação)

Problema	GCP		MINRESP		Híbrido	
	Tempo(s)	It.	Tempo(s)	It.	Tempo(s)	It.
sc50a	0.00	10	0.00	10	0.00	10
sc50b	0.00	9	0.00	9	0.00	9
sc105	0.01	10	0.01	10	0.01	10
sc205	0.21	18	*	*	0.04	11
scagr7	0.01	10	0.01	16	0.01	16
scr15	*	*	*	*	72.19	23
scr20	*	*	*	*	1219.38	24
scsd1	*	*	0.02	11	0.07	11
scsd6	0.14	11	0.05	11	0.15	11
scsd8	0.52	11	*	*	0.54	11
scsd8-2b-64	*	*	*	*	17.83	12
scsd8-2c-64	*	*	*	*	20.77	11
scsd8-2r-432	*	*	*	*	48.28	18
sctap1	0.23	19	0.08	19	0.2	19
sctap2	2.25	24	0.29	17	0.69	17
sctap3	*	*	*	*	1.24	20
share1b	*	*	*	*	0.15	26
share2b	*	*	*	*	0.05	12
shell	0.53	23	*	*	0.29	23
ship04l	0.18	15	0.08	15	0.23	15
ship04s	*	*	0.21	20	0.14	16
ship08l	*	*	0.37	16	0.6	18
ship08s	0.22	16	0.08	16	0.24	16
ship12l	0.71	18	0.26	18	0.76	18
ship12s	0.30	17	0.12	17	0.32	17
standgub	0.57	22	0.24	22	0.54	22
standmps	1.4000	32	0.48	28	0.73	30
stocfor1	0.0500	16	*	*	0.02	13
stocfor2	*	*	*	*	7.91	32
truss	*	*	*	*	11.52	18

O método híbrido resolveu 61 dos 63 problemas testados, o que significa que 25 problemas passaram a ser solucionados com a nova abordagem. Além disso, este método foi mais rápido em 62% dos problemas, o MINRESP em 27% e o GCP em aproximadamente 2% dos problemas. O número de iterações foi menor em 59% dos problemas para o híbrido, 6% para o MINRESP e em torno de 2% para o GCP. Considerando os problemas resolvidos apenas pelo GCP e híbrido,

vemos que somente um deles foi resolvido em menor tempo pelo GCP. Com relação ao número de iterações, o valor foi menor ou igual para o híbrido em todos os problemas. Comparando os resultados do MINRESP com o híbrido vemos que somente 2 problemas foram resolvidos mais rápido pelo híbrido e apenas 3 problemas tiveram o número de iterações menor.

6 CONCLUSÕES

Neste trabalho, para a resolução dos sistemas lineares a cada iteração do método preditor-corretor consideramos tanto o sistema aumentado quanto o sistema de equações normais. O método MINRESP com reortogonalização foi implementado e comparado com o método dos gradientes conjugados, ambos métodos foram pré-condicionados pelo pré-condicionador separador. Experimentos computacionais foram realizados com o objetivo de determinar qual método é mais eficiente e robusto. Os resultados mostraram que o MINRESP é mais robusto que o GCP quando somente o sistema aumentado é considerado. Para o sistema de equações normais com e sem refatoração os resultados foram semelhantes para o número de iterações, porém o tempo foi menor ou igual em todos os problemas para o MINRESP sem refatorar. Um método híbrido também foi considerado, o qual resolveu 40% a mais de problemas (inclusive problemas de dimensões maiores) além de ser mais rápido e fazer menos iterações que os métodos GCP e MINRESP para o sistema de equações normais sem refatoração. Dessa forma, concluímos que este método é o mais robusto e eficiente e será ainda mais investigado e utilizado em trabalhos futuros.

AGRADECIMENTOS

Ao CNPq e a FAPESP pelo apoio financeiro.

ABSTRACT. In this paper, we consider the predictor-corrector method, which is one of the most important variants of the interior point methods due to its efficiency and fast convergence. In the predictor-corrector method it is necessary to solve two linear systems at each iteration to determine the search direction. The solution of such systems is the step that requires more processing time and should therefore be carried out efficiently. Two iterative Krylov methods are considered to solve these linear systems of: the MINRES and the conjugate gradient method. A preconditioner specially developed for the linear systems arising from the interior point methods is used so that the iterative methods converge faster. Computational experiments in a some sets of linear programming problems are performed in order to analyse the efficiency and robustness of the linear systems solution methods.

Keywords: interior point methods, linear systems, iterative methods.

REFERÊNCIAS

- [1] R.S. Burkard, S. Karisch & F. Rendl. Qaplib – a quadratic assignment problem library. *European Journal of Operations Research*, **55** (1991), 115–119.

- [2] D.C. Fong & M. A. Saunders. CG versus MINRES: An empirical comparison. *SQU Journal for Science*, **17**(1) (2012), 44–62.
- [3] D.M. Gay. Electronic mail distribution of linear programming test problems. *Mathematical Programming Society Committee on Algorithms COAL Newsletter*, (1985), 10–12.
- [4] G.H. Golub & C.F. Van Loan. *Matrix Computations*. 3^a ed. Baltimore, Md., London: The Johns Hopkins University Press, (1996).
- [5] I.B. Kalambi. A Comparison of three Iterative Methods for the Solution of Linear Equations. *Journal of Applied Sciences and Environmental Management*, **12**(4) (2009), 53–55.
- [6] N. Karmarkar. *A new polynomial-time algorithm for linear programming*. *Combinato-rica*, **4** (1984), 373–395.
- [7] J.L. Kennington, S. Niemi, W.J. Carolan, J.E. Hill & S.J. Wichmann. An empirical evaluation of the korbx algorithms for military airlift applications. *Oper. Res*, **38** (1990), 240–248.
- [8] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bureau Standards*, **45**(4) (1950), 255–282.
- [9] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, **2**(4) (1992), 575–601.
- [10] C.D. Meyer. *Matrix analysis and applied linear algebra*. SIAM Publications, SIAM, (2000).
- [11] Miscellaneous lp models. *Technical report, Hungarian Academy of Sciences OR Lab.* url, www.sztaki.hu/~meszaros/publicftp/lptestset/misc.
- [12] R.D.C. Monteiro, I. Adler & M.G.C. Resende. A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension. *Mathematics of Operations Research*, **15** (1990), 191–214.
- [13] J. Noreen. A comparison of direct and indirect solvers for linear system of equations. *International Journal of Emerging Science*, **2**(2) (2012), 310–321.
- [14] A.R.L. Oliveira & D.C. Sorensen. A new class of preconditioners for large-scale linear systems from interior point methods for linear programming. *Linear Algebra and Its Applications*, **394** (2005), 1–24.
- [15] Y. Saad. *Iterative Methods for Sparse Linear Systems*, SIAM, (2003).
- [16] D.C. Sorensen. Implicit application of polynomial filters in a k -step Arnoldi method. *SIAM Journal on Matrix Analysis and Applications*, **13** (1992), 357–385.
- [17] Stochastic lp test sets. *Technical report, Hungarian Academy of Sciences OR Lab.* url, www.sztaki.hu/~meszaros/publicftp/lptestset/stochlp.
- [18] S.J. Wright. *Primal-dual interior-point methods*. SIAM Publications, SIAM, Philadelphia, PA, USA, (1997).