

UMA PROPOSTA PARA MELHORAR O DESEMPENHO DO PRECONDICIONADOR SEPARADOR NO MÉTODO DE PONTOS INTERIORES

Cecilia Orellana Castro

IMECC- DMA- Universidade Estadual de Campinas
Cidade Universitária Zeferino Vaz, 13081-970, Campinas -SP, Brasil
o.cecilita@hotmail.com

Aurelio Ribeiro Leite Oliveira

IMECC- DMA - Universidade Estadual de Campinas
Cidade Universitária Zeferino Vaz, 13081-970, Campinas -SP, Brasil
aurelio@ime.unicamp.br

RESUMO

Os métodos de pontos interiores do tipo primal-dual para otimização linear são muito competitivos na resolução de problemas de grande porte, um de seus atrativos é sua complexidade polinomial. Em cada iteração destes métodos deve-se resolver sistemas lineares que tornam-se mal condicionados perto da solução ótima. O preconditionador Separador desenvolveu-se para auxiliar neste problema, ele precisa de uma base que depende fortemente da iteração corrente. Propõe-se uma modificação na escolha da base do preconditionador Separador, o novo critério de reordenamento das colunas da matriz de restrições é amparado num resultado teórico que visa aprimorar o número de condição da matriz preconditionada para obter um melhor desempenho do método dos gradientes conjugados. Experimentos numéricos demonstram bons resultados usando esta nova abordagem.

PALAVRAS CHAVE. Método de Pontos Interiores, Precondicionador Separador, Gradientes Conjugados.

Área Principal: Programação Matemática

ABSTRACT

The primal-dual interior point methods for linear optimization are very competitive in solving large-scale problems, one of its attractions is its polynomial complexity. In each iteration of these methods must solve linear systems that become ill-conditioned close to the optimal solution. The Splitting preconditioner was developed to avoid this problem, it needs a basis that strongly depends on the current iteration. We propose a modification to the choice of the basis, the new criteria for the reordering of restrictions matrix columns is supported in a theoretical result that aims to improve the condition number of the preconditioned matrix and, thus obtaining a better performance of the conjugate gradient method. Numerical experiments have shown good results using this new approach.

KEYWORDS. Interior Point Methods. Splitting preconditioner. Conjugate gradient method.

Main Area: Mathematical Programming

1. Introdução

Nas últimas décadas os métodos de pontos interiores do tipo primal-dual tornaram-se uma poderosa ferramenta para resolver problemas de programação linear de grande porte. Para tanto, é necessário o desenvolvimento de implementações eficientes destes métodos. Além de reduzir o número total de iterações no algoritmo é interessante estudar técnicas para acelerar cada iteração. O objetivo deste trabalho é aprimorar as últimas iterações deste método. Em cada iteração a tarefa mais cara consiste em resolver sistemas lineares para encontrar a direção de busca, trabalha-se com uma abordagem híbrida para resolver estes sistemas usando métodos iterativos.

Nesta abordagem, o método dos gradientes conjugados é preconditionado em duas fases. Na primeira fase, usa-se o preconditionador Fatoração Controlada de Cholesky e na segunda fase, o preconditionador Separador. Na segunda fase, a base do preconditionador Separador é modificada com uma nova reordenação das colunas da matriz de restrições, a justificativa teórica desta proposta é apresentada na Proposição 1 deste trabalho. Foi feita uma implementação desta nova abordagem e uma comparação com a versão atualmente utilizada, veja M. Velazco, et. al.(2011) obtendo resultados competitivos.

2. Fundamentos do método primal-dual de pontos interiores para programação linear

Considere o par primal-dual do problema de programação linear canalizado.

$$(P) \begin{cases} \min & c^T x \\ \text{s. a.} & Ax = b \\ & x + s = u \\ & x, s \geq 0 \end{cases} \quad (D) \begin{cases} \max & b^T y - u^T w \\ \text{s. a.} & A^T y - w \leq c \\ & w \geq 0 \\ & y \in \mathbb{R}^m \end{cases}$$

Onde $x, s, w \in \mathbb{R}^n$ e A uma matriz de tamanho $m \times n$ que será considerada de posto completo.

Se na formulação primal é aplicada a penalização barreira logarítmica nas restrições de não negatividade, tem-se o problema

$$\min c^T x - \mu \sum_{i=1}^n \log x_i - \mu \sum_{i=1}^n \log s_i \quad \text{s. a.} \quad Ax = b, \quad x + s = u, \quad x, s > 0. \quad (1)$$

Dado que o problema em (1) é convexo as condições de otimalidade de primeira ordem são suficientes e necessárias. Considere o lagrangeano ℓ e suas derivadas parciais.

$$\ell(x, s, y, w) = c^T x - \mu \sum_{i=1}^n \log x_i - \mu \sum_{i=1}^n \log s_i + y^T (b - Ax) + w,$$

- (i) $\nabla_y \ell = b - Ax$;
- (ii) $\nabla_w \ell = u - x - s$;
- (iii) $\nabla_x \ell = c - \mu X^{-1} e - A^T y - w$;
- (iv) $\nabla_s \ell = -\mu S^{-1} e - w$.

Onde $X^{-1} = \text{diag}(x_1^{-1}, \dots, x_n^{-1})$, $S = \text{diag}(s_1^{-1}, \dots, s_n^{-1})$, e $e^T = (1, \dots, 1) \in \mathbb{R}^n$

Se $z \in \mathbb{R}^n$ é definida como $z = \mu X^{-1} e$, as condições de otimalidade de (1) são:

- (i) $Ax = b$;
- (ii) $x + s = u \quad x, s > 0$;
- (iii) $A^T y + z - w = c \quad z, w > 0$;
- (iv) $SWe = \mu e$;
- (v) $XZe = \mu e$.

As equações acima são também conhecidas como as *equações da Trajetória Central*.

A direção de busca numa iteração do método primal-dual é obtida do método de Newton aplicado as condições de otimalidade de (1), o que implica resolver o seguinte sistema de equações.

$$\begin{pmatrix} A & 0 & 0 & 0 & 0 \\ I_n & I_n & 0 & 0 & 0 \\ 0 & 0 & A^T & -I_n & I_n \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta s \\ \Delta y \\ \Delta w \\ \Delta z \end{pmatrix} = \begin{pmatrix} r_b \\ r_u \\ r_c \\ \mu e - XZe \\ \mu e - SWe \end{pmatrix}, \quad (2)$$

onde $r_b = b - Ax$, $r_u = u - x - s$ e $r_c = c + w - z - A^T y$.

Embora os primeiros algoritmos de métodos de pontos interiores primal-dual trabalhavam com pontos factíveis, atualmente é comum trabalhar com os algoritmos de pontos interiores inactíveis, onde as sequências

$$r_b^k = b - Ax^k, \quad r_u^k = u - x^k - s^k, \quad r_c^k = c + w^k - z^k - A^T y^k, \quad \mu^k = \frac{(x^k)^T z^k + (s^k)^T w^k}{2n},$$

decrecem até atingir uma tolerância predeterminada.

3. Algoritmo primal-dual do método de Pontos Interiores

A maioria dos algoritmos primal-dual de pontos interiores inactíveis encontrados na literatura estão baseados na seguinte estrutura

Algoritmo 1: Algoritmo Primal-Dual

Entrada: Dado $(x_0, s_0, y_0, w_0, z_0)$ sendo $x_0 > 0$, $z_0 > 0$, $w_0 > 0$ e $s_0 > 0$ e Máx = 100.

Para $k \leftarrow 1$ **até** Máx **fazer**

- 1) Calcule $\mu^k = \frac{(x^k)^T z^k + (s^k)^T w^k}{2n}$ e escolha $\sigma^k \in [0, 1]$;
- 2) Resolva o sistema de equações (2);
- 3) Calcule o comprimento de passo $\alpha^k = \min \{1, \tau^k \rho_x^k, \tau^k \rho_z^k, \tau^k \rho_s^k, \tau^k \rho_w^k\}$,
 onde $\tau^k \in (0, 1)$, $\rho_x^k = \frac{-1}{\min\left(\frac{\Delta x_i^k}{x_i^k}\right)}$, $\rho_z^k = \frac{-1}{\min\left(\frac{\Delta z_i^k}{z_i^k}\right)}$, $\rho_s^k = \frac{-1}{\min\left(\frac{\Delta s_i^k}{s_i^k}\right)}$ e

$$\rho_w^k = \frac{-1}{\min\left(\frac{\Delta w_i^k}{w_i^k}\right)}.$$
- 4) Calcule o novo ponto

$$\begin{aligned} (x^{k+1}, s^{k+1}, y^{k+1}, w^{k+1}, z^{k+1}) = \\ (x^k, s^k, y^k, w^k, z^k) + \alpha^k (\Delta x^k, \Delta s^k, \Delta y^k, \Delta w^k, \Delta z^k) \end{aligned}$$

O critério de convergência usado é baseado nas condições de otimalidade:

$$\frac{\|b - Ax\|}{\|b\| + 1} \leq \epsilon, \quad \frac{\|u - x - s\|}{\|u\| + 1} \leq \epsilon, \quad \frac{\|c - A^T y - z + w\|}{\|c\| + 1} \leq \epsilon,$$

$$\frac{|c^T x - b^T y + u^T w|}{|c^T x| + |b^T y - b^T w| + 1} \leq \epsilon, \quad \epsilon = 10^{-8}.$$

O sistema de equações do passo (2) no Algoritmo 1 é reduzido a duas formulações muito usadas no cálculo da direção de busca. A primeira delas conhecida como Sistema Aumentado é um sistema de equações com matriz simétrica e indefinida de tamanho $n + m$.

$$\begin{pmatrix} -D & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} r \\ h \end{pmatrix}, \quad (3)$$

onde $D = X^{-1}Z + S^{-1}W$, $r = r_c - X^{-1}(\sigma\mu e - XZe) + S^{-1}(\sigma\mu e - SWe) - S^{-1}Wr_u$ e $h = r_b$. A segunda formulação é chamada Sistema de Equações Normais

$$AD^{-1}A^T \Delta y = h + AD^{-1}r. \quad (4)$$

O nome, deve-se à semelhança com equações normais para o problema dos quadrados mínimos no caso em que $h = 0$ e matriz de coeficientes é $D^{-1/2}A^T$. A matriz do sistema (4) é simétrica e definida positiva de tamanho m .

As variantes do algoritmo de pontos interiores mais usadas são os métodos seguidores de caminho e o algoritmo de Mehrotra Predictor-Corretor. Apresenta-se na próxima seção uma descrição do segundo método, pois este é usado na implementação do código PCx, Wright(1997).

4. Método Predictor-Corretor

A principal característica deste algoritmo é que a direção de busca consiste de dois componentes:

- Uma direção afim-escala ou preditora, que consiste em um passo na direção de Newton puro com $\sigma = 0$.
- Uma direção corretora e de centragem, que tenta compensar a não-linearidade da direção preditora com uma escolha adaptativa do parâmetro de centragem σ

O algoritmo de Mehrotra gera seqüências de iterações infactíveis. Suponha que o ponto $(x^k, s^k, y^k, w^k, z^k)$ tenha as componentes x^k, s^k, w^k, z^k não negativas e a medida de dualidade seja μ^k , o cálculo da direção afim-escala é obtido do seguinte sistema de equações:

$$\begin{pmatrix} A & 0 & 0 & 0 & 0 \\ I_n & I_n & 0 & 0 & 0 \\ 0 & 0 & A^T & -I_n & I_n \\ Z^k & 0 & 0 & X^k & 0 \\ 0 & W^k & 0 & 0 & S^k \end{pmatrix} \begin{pmatrix} \Delta x^{af} \\ \Delta s^{af} \\ \Delta y^{af} \\ \Delta w^{af} \\ \Delta z^{af} \end{pmatrix} = \begin{pmatrix} r_b^k \\ r_u^k \\ r_c^k \\ -X^k Z^k e \\ -S^k W^k e \end{pmatrix}. \quad (5)$$

O comprimento de passo nesta direção é dado por:

$$\alpha_x^{af} = \operatorname{argmax}\{\alpha \in [0, 1] : x^k + \alpha \Delta x^{af} > 0\};$$

$$\alpha_s^{af} = \operatorname{argmax}\{\alpha \in [0, 1] : s^k + \alpha \Delta s^{af} > 0\};$$

$$\alpha_z^{af} = \operatorname{argmax}\{\alpha \in [0, 1] : z^k + \alpha \Delta z^{af} > 0\};$$

$$\alpha_w^{af} = \operatorname{argmax}\{\alpha \in [0, 1] : w^k + \alpha \Delta w^{af} > 0\};$$

Para medir a eficiência do passo na direção preditora é calculado o gap de dualidade afim

$$\mu_{af} = \frac{(x + \alpha_x^{af} \Delta x^{af})^T (z + \alpha_z^{af} \Delta z^{af}) + (s + \alpha_s^{af} \Delta s^{af})^T (w + \alpha_w^{af} \Delta w^{af})}{2n}.$$

O cálculo do parâmetro de centragem σ , considera duas situações:

- Se $\mu_{af} \ll \mu^k$, então a direção preditora é uma boa direção de busca, pois permite uma significativa redução do gap μ sem deixar o ortante positivo $(x, s, z, w) > 0$. Assim não será necessária uma centragem muito grande, isto é, $\sigma \approx 0$.
- Se $\mu_{af} \approx \mu^k$, a direção preditora fez um pequeno progresso na diminuição de μ , logo é preciso fazer uma centragem maior (mais perto da trajetória central) com o objetivo de que o ponto esteja numa melhor posição para atingir um maior decréscimo de μ na seguinte iteração.

Mehrotra sugere uma heurística bem sucedida

$$\sigma = \left(\frac{\mu_{af}}{\mu^k} \right)^3.$$

Se $\mu_{af} \ll \mu^k$, então $\sigma \approx 0$ e se $\mu_{af} \approx \mu^k$, então $\sigma \approx 1$.

Suponha que $\alpha^{af} = 1$ seja um passo otimista na direção preditora, então

$$\begin{aligned} r_b^{k+1} &= b - Ax^{k+1} = b - A(x^k + \Delta x^{af}) = r_b^k - A\Delta x^{af} = 0; \\ r_u^{k+1} &= u - x^{k+1} - s^{k+1} = u - (x^k + \Delta x^{af}) - (s^k + \Delta s^{af}) = r_u^k - (\Delta x^{af} + \Delta s^{af}) = 0; \\ r_c^{k+1} &= A^T y^{k+1} + z^{k+1} - w^{k+1} - c \\ &= A^T (y^k + \Delta y^{af}) + (z^k + \Delta z^{af}) - (w^k + \Delta w^{af}) = r_c^k + A^T \Delta y^{af} + \Delta z^{af} - \Delta w^{af} = 0. \end{aligned}$$

Isto é, a factibilidade é atingida num só passo, porém as folgas complementares

$$\begin{aligned} (x_i + \Delta x_i^{af})(z_i + \Delta z_i^{af}) &= x_i z_i + x_i \Delta z_i^{af} + z_i \Delta x_i^{af} + \Delta x_i^{af} \Delta z_i^{af} = \Delta x_i^{af} \Delta z_i^{af} \\ e \quad (s_i + \Delta s_i^{af})(w_i + \Delta w_i^{af}) &= s_i w_i + s_i \Delta w_i^{af} + w_i \Delta s_i^{af} + \Delta s_i^{af} \Delta w_i^{af} = \Delta s_i^{af} \Delta w_i^{af} \end{aligned}$$

não são nulas pois o passo de Newton puro é apenas uma aproximação linear de primeira ordem.

A componente corretora e de centragem, tenta compensar esse desvio. Parece natural criar uma direção que faça com que $\Delta x_i^{af} \Delta z_i^{af} = 0$ e $\Delta s_i^{af} \Delta w_i^{af} = 0$, porém, existe o risco de perder a positividade de (x, s, z, w) antes de atingir a otimalidade. Assim, a ideia é fazer uma centragem usando σ , este foi definido de tal jeito que depende da posição da iteração atual no ortante positivo. Logo, a direção corretora e de centragem visa obter

$$\Delta x_i^{af} \Delta z_i^{af} = \sigma \mu \quad e \quad \Delta s_i^{af} \Delta w_i^{af} = \sigma \mu.$$

Esta direção é obtida pelo seguinte sistema de equações:

$$\begin{pmatrix} A & 0 & 0 & 0 & 0 \\ I_n & I_n & 0 & 0 & 0 \\ 0 & 0 & A^T & -I_n & I_n \\ Z^k & 0 & 0 & X^k & 0 \\ 0 & W^k & 0 & 0 & S^k \end{pmatrix} \begin{pmatrix} \Delta x^{cc} \\ \Delta s^{cc} \\ \Delta y^{cc} \\ \Delta w^{cc} \\ \Delta z^{cc} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sigma \mu e - \Delta X^{af} \Delta Z^{af} e \\ \sigma \mu e - \Delta S^{af} \Delta W^{af} e \end{pmatrix}. \quad (6)$$

Finalmente, veja Wright(1997), a direção de busca na variante de Mehrotra é obtida pela soma

$$(\Delta x^k \Delta s^k \Delta y^k \Delta w^k \Delta z^k)^T = (\Delta x^{af} \Delta s^{af} \Delta y^{af} \Delta w^{af} \Delta z^{af})^T + (\Delta x^{cc} \Delta s^{cc} \Delta y^{cc} \Delta w^{cc} \Delta z^{cc})^T.$$

Assim, encontrar a direção de busca usando o algoritmo de Mehrotra implica resolver dois sistemas lineares em cada iteração.

Em problemas de grande porte é comum o uso de métodos iterativos, porém eles são muito sensíveis ao número de condição da matriz do sistema linear.

A partir da seguinte notação da matriz das equações normais

$$AD^{-1}A^T = \sum_{i=1}^n d_i^{-1} A^i (A^i)^T = \sum_{i=1}^n \left(\frac{z_i}{x_i} + \frac{w_i}{s_i} \right)^{-1} A^i (A^i)^T,$$

observa-se que uma coluna densa A^i implica que a matriz $AD^{-1}A^T$ perca esparsidade e outra razão para evitar este inconveniente é o uso dos métodos iterativos.

Por outro lado, os autovalores λ de $AD^{-1}A^T$ satisfazem a desigualdade

$$\lambda \leq \sum_{i=1}^n d_i^{-1} \|A^i\|,$$

isto é, os autovalores estão limitados superiormente por quantidades que podem se tornar muito grandes nas últimas iterações do método de pontos interiores.

Mais precisamente, em Gondzio (2012) foi demonstrado que se todas as iterações estão na seguinte vizinhança de pontos infactíveis

$$N_\beta = \left\{ (x, s, y, w, z) \in \mathcal{F} \mid \beta\mu \leq x_j z_j, s_j w_j \leq \frac{1}{\beta}\mu \right\},$$

onde

$$\mathcal{F} = \left\{ (x, s, y, w, z) \mid \begin{array}{l} \|b - Ax\| \leq \beta_b \mu \|b - Ax^0\|, \\ \|u - x - s\| \leq \beta_u \mu \|u - x^0 - s^0\|, \\ \|c - A^T y - z + w\| \leq \beta_c \mu \|c - A^T y^0 - z^0 + w^0\| \end{array} \right\},$$

com $\beta, \beta_b, \beta_u, \beta_c \in (0, 1)$, então

$$\kappa(AD^{-1}A^T) \leq \kappa^2(A) \mathcal{O}(\mu_k^{-2}).$$

Para melhorar o condicionamento de uma matriz são usados os preconditionadores.

5. Precondicionamento

Precondicionar o sistema linear $Kx = b$, significa transformá-lo em outro sistema linear com melhor número de condição e cuja solução seja facilmente relacionada com x . O preconditionamento pode ser aplicado pela esquerda, direita e por ambos os lados:

$$\begin{aligned} P^{-1}Ky &= P^{-1}b, & \text{com } x &= y, \\ KP^{-1}y &= b, & \text{com } x &= P^{-1}y \text{ e} \\ P^{-1}KP^{-T}z &= P^{-1}b, & \text{com } x &= P^{-T}z, \end{aligned}$$

respectivamente. Observe que as matrizes preconditionadas por esquerda $P^{-1}K$ e direita KP^{-1} são matrizes equivalentes pois $P(P^{-1}K)P^{-1} = KP^{-1}$, logo têm o mesmo espectro.

Um preconditionador que é aplicado por ambos os lados é o preconditionador Fatoração Controlada de Cholesky (FCC), Campos (1995). Trata-se de uma versão da Fatoração Incompleta de Cholesky que ignora os preenchimentos durante o processo. Este preconditionador aplicado à matriz das Equações Normais tem produzido bons resultados nas primeiras iterações do método, já nas últimas iterações torna-se menos eficiente devido ao mal condicionamento da matriz $AD^{-1}A^T$. Mais tarde, Oliveira (1997) propõe o preconditionador Separador para auxiliar no mal condicionamento dos sistemas de equações que aparecem nas últimas iterações do método de pontos interiores em problemas de otimização linear. Este preconditionador será estudado na subseção 5.1. Além disso, uma abordagem híbrida usando estes dois preconditionadores foi proposta por Bocanegra et. al. (2007), posteriormente foram apresentadas heurísticas para aprimorar esta abordagem, veja Velazco et. al. (2011).

5.1. Preconditionador Separador

O preconditionador Separador é um preconditionador que pode ser obtido tanto da matriz do Sistema Aumentado quanto da matriz das Equações Normais. Dado que A é uma matriz de posto completo, A contém uma submatriz B de tamanho $m \times m$ com colunas linearmente independentes.

Suponha que depois de uma permutação de colunas, $A = [B, N]$, $D = \begin{pmatrix} D_B & 0 \\ 0 & D_N \end{pmatrix}$, $\Delta x = [\Delta x_B, \Delta x_N]^T$, $r = [r_B, r_N]^T$ e $r_y = h$, o sistema Aumentado (3) pode ser escrito da seguinte maneira:

$$\begin{pmatrix} -D_B & 0 & B^T \\ 0 & -D_N & N^T \\ B & N & 0 \end{pmatrix} \begin{pmatrix} \Delta x_B \\ \Delta x_N \\ \Delta y \end{pmatrix} = \begin{pmatrix} r_B \\ r_N \\ r_y \end{pmatrix}. \quad (7)$$

O preconditionador Separador para a matriz do sistema Aumentado é dado pela seguinte matriz:

$$P^{-1} = \begin{pmatrix} D_B^{-1/2} & 0 & D_B^{1/2}B^{-1} \\ 0 & D_N^{-1/2} & 0 \\ D_B^{-1/2} & 0 & 0 \end{pmatrix}.$$

Por outro lado, o preconditionador Separador para a matriz das Equações Normais é dado por:

$$P^{-1} = D_B^{1/2}B^{-1}.$$

Para obter a direção de busca do método primal-dual de pontos interiores usando o preconditionador Separador, os seguintes sistemas lineares devem ser resolvidos:

$$\begin{aligned} (I + WW^T)\tilde{\Delta}y &= D_B^{1/2}B^{-1}(h + AD^{-1}r); \\ \Delta y &= B^{-T}D_B^{1/2}\tilde{\Delta}y; \\ \Delta x_B &= D_B^{-1}(B^T\Delta y - r_B); \\ \Delta x_N &= D_N^{-1}(N^T\Delta y - r_N). \end{aligned}$$

Onde $W = D_B^{1/2}B^{-1}ND_N^{-1/2}$. Lembrando que $\Delta x = [\Delta x_B, \Delta x_N]^T$, segue-se

$$\begin{aligned} \Delta s &= r_u - \Delta x; \\ \Delta w &= S^{-1}(\sigma\mu e - SWe - W\Delta s); \\ \Delta z &= X^{-1}(\sigma\mu e - XZe - Z\Delta x). \end{aligned}$$

A matriz B é obtida usando a fatoração LU da matriz A depois de uma cuidadosa implementação que pode ser cara por gerar muito preenchimento, isto acontece pois o conjunto de colunas linearmente independentes é desconhecido antes da fatoração LU. Por esta razão é necessário um bom critério de ordenamento da base que melhore o condicionamento da matriz preconditionada pelo preconditionador Separador.

6. Análise Espectral da matriz preconditionada pelo Precondicionador Separador para uma nova base B

O desempenho do método dos gradientes conjugados depende do espectro da matriz. Nesta seção, estuda-se a matriz preconditionada pelo preconditionador Separador

$$I + WW^T = I + (D_B^{1/2} B^{-1} N D_N^{-1/2})(D_N^{-1/2} N^T B^{-T} D_B^{1/2}). \quad (8)$$

Observa-se em (8) que a matriz $I + WW^T$ muda a cada iteração pois a matriz D depende da iteração atual. Além disso, nas últimas iterações, espera-se que $\mu = \frac{x^T z + s^T w}{2n} \approx 0$ para atingir a otimalidade, isto acontece se e somente se $x_i z_i \approx 0$ e $s_i w_i \approx 0$ o que implica que existem índices $i \in \{1, 2, \dots, n\}$ tais que $d_i \approx 0$ e $d_j \approx \infty$. A quantidade de índices tais que $d_i \approx 0$ não necessariamente é m , mesmo com problemas não degenerados, como acontece no método Simplex. Assim, uma boa escolha de índices $\{i : d_i \in D_B, A^i \in B\}$ determinará o bom condicionamento de $I + WW^T$.

Os autores Oliveira e Sorensen, veja Oliveira(1997) sugerem a escolha das primeiras m colunas linearmente independentes A^j da matriz A ordenadas pelos índices j que correspondem às colunas de AD^{-1} com as maiores norma-1. Posteriormente M. Velazco, et. al.(2011) fizeram testes para escolha da base B usando uma ordenação das colunas de A considerando as colunas com maiores norma-1 das matrizes $AD^{-1/2}$ e $AD^{-3/2}$ sem melhoras significativas. Porém a heurística baseada na escolha da base B usando as colunas de A considerando a maior norma-2 das colunas da matriz AD^{-1} foi bem sucedida e atualmente esta ordenação é usada.

Neste trabalho, depois de estudar o espectro da matriz preconditionada pelo preconditionador Separador, apresenta-se um resultado que motiva uma nova reordenação da base B para melhorar o condicionamento da matriz $I + WW^T$. Assim, pode-se acelerar o cálculo da direção de busca nas últimas iterações do método primal-dual de pontos interiores.

Suponha que M seja $I + WW^T$ e (λ, v) seja um autopar de M ; isto é,

$$v + WW^T v = \lambda v.$$

Multiplicando pela esquerda por v^T , tem-se $\|v\|^2 + \|W^T v\|^2 = |\lambda| \|v\|^2$. Daí

$$|\lambda| = 1 + \frac{\|W^T v\|^2}{\|v\|^2} \geq 1. \quad (9)$$

Assim, os autovalores de M estão limitados inferiormente por 1.

Dado que a matriz M é simétrica e definida positiva, pela equação (9), tem-se $\kappa(M) = \frac{\lambda_{max}}{\lambda_{min}} \leq \lambda_{max}$. Por outro lado, da desigualdade $\|W^T v\| \leq \|W^T\| \|v\|$, obtem-se

$$|\lambda| = 1 + \frac{\|W^T v\|^2}{\|v\|^2} \leq 1 + \|W^T\|^2.$$

Portanto, para limitar o número de condição de M é necessário limitar $\|W^T\|$.

A norma de Frobenius de W é usada para trabalhar na minimização de $\|W^T\|$ colocando condições sobre as colunas ou linhas de A :

$$\|W^T\|_F^2 = \sum_{j=1}^m \|W^j\|_2^2, \quad (10)$$

onde W^j é a j -ésima coluna de W .

Proposição 1. *Considere a matriz $W = D_B^{1/2} B^{-1} N D_N^{-1/2}$, se B for composta pelas primeiras m colunas A^j linearmente independentes de A tais que j corresponde às maiores $\|A^j d_j^{-1/2}\|_2$, então $\|W\|_F$ é minimizada.*

Demonstração. Da desigualdade $\|W\|_F \leq \|D_B^{1/2} B^{-1}\|_F \|N D_N^{-1/2}\|_F$, observa-se que é preciso minimizar $\|D_B^{1/2} B^{-1}\|_F$ e $\|N D_N^{-1/2}\|_F$.

Para minimizar a primeira norma $\|D_B^{1/2} B^{-1}\|_F$ a partir da seguinte desigualdade matricial $\|B D_B^{-1/2}\| \geq \frac{1}{\|D_B^{1/2} B^{-1}\|}$, pode-se concluir que

$$\max \|B D_B^{-1/2}\|_F \quad \text{implica} \quad \max \frac{1}{\|D_B^{1/2} B^{-1}\|_F} \quad \text{se, e somente se,} \quad \min \|D_B^{1/2} B^{-1}\|_F.$$

Portanto, maximizar $\|B D_B^{-1/2}\|_F$ implica minimizar $\|D_B^{1/2} B^{-1}\|_F$. Observe que

$$A D^{-1/2} = \begin{bmatrix} B D_B^{-1/2} & , & N D_N^{-1/2} \end{bmatrix}, \quad (11)$$

pela definição da norma de Frobenius dada em (10), para maximizar $\|B D_B^{-1/2}\|_F$ se deve escolher as colunas de $A D^{-1/2}$ com as maiores norma-2. Assim, B deve ser composta por colunas linearmente independentes A^j de A tais que j corresponde às maiores $\|A^j d_j^{-1/2}\|_2$.

Por outro lado, para minimizar a segunda norma $\|N D_N^{-1/2}\|_F$, a partir de (11) é imediato que se deve escolher as colunas de $A D^{-1/2}$ com menores norma-2.

Da equação (11), também é possível notar que as condições de maximizar $\|B D_B^{-1/2}\|_F$ e minimizar $\|N D_N^{-1/2}\|_F$ são equivalentes. \square

A nova reordenação de base B usando a proposição anterior foi implementada e comparada com a reordenação atual baseada no trabalho de Velazco et.al. (2011).

A tarefa de determinar uma boa base é muito importante, pois o preconditionador Separador tem a interessante propriedade de utilizar a matriz B em algumas iterações além da iteração na que foi calculada, em consequência o preconditionador é barato para se calcular em algumas iterações do método de pontos interiores.

A mudança da base B é determinada pelo número de iterações do método dos gradientes conjugados usados para resolver o sistema linear preconditionado.

7. Experimentos Numéricos

Os experimentos numéricos foram realizados utilizando a versão modificada do PCx, Czyzyk et.al.(1999), nesta versão, o método direto usado para a solução dos sistemas lineares foi substituído por um método iterativo, o método dos gradientes conjugados preconditionado. O preconditionamento dos sistemas de equações usa a abordagem híbrida proposta por Bocanegra et.al.(2007) e troca de fase proposta por Velazco (2011).

A base B do preconditionador Separador muda quando $8 * n_g \geq m$, onde n_g denota o número de iterações do método dos gradientes conjugados preconditionado numa iteração do método de pontos interiores.

Os testes realizados comparam os resultados das abordagens $PCxm$ e $PCxc$, sendo $PCxm$ a abordagem com ordenação da base B proposta pela heurística de M. Velazco et. al. (2011), e $PCxc$ a abordagem com a ordenação da base B usando nossa proposição apresentada na seção anterior.

Todos os testes foram realizados em um processador Intel i7 com 16 Gb de memória, em ambiente Linux usando os compiladores gcc e gfortran. Os problemas utilizados são de domínio público extraídos da biblioteca NETLIB e QAP.

Tabela 1: Problemas Testes

Prob	Tamanho	
	Linha	Coluna
25fv47	825	1571
bnl1	643	1175
chr22b	5587	5335
chr25a	8149	7825
cre-a	3516	4067
cre-b	9648	72447
cre-c	3068	3678
cre-d	8926	69980
els19	4350	9937
ex01	246	1379
ex02	238	1378
ex05	833	6980
ex09	1846	16422
ganges	1309	1681
ken13	28632	42659
ken18	105127	154699
maros	840	1443
nesm	662	2923
rou20	7359	33840
scr15	2234	4635
scr20	5079	12180
ste36a	27686	109653
stocfor2	2157	2031
stocfor3	16675	15695

Na Tabela 1 se apresenta o número de linhas e de colunas dos problemas testados. Para avaliar a eficiência da ordenação proposta comparamos o número total de iterações do método de pontos interiores IPM, o tempo em que cada problema é resolvido e o número total de iterações do método de gradientes conjugados preconditionado PCG em todas as iterações do método de pontos interiores.

Tabela 2: Número de iterações do IPM, Tempo e número de iterações do PCG

Prob	Iterações IPM		Tempo		Iterações PCG	
	PCxm	PCxc	PCxm	PCxc	PCxm	PCxc
25fv47	29	26	1,80	1,35	5122	2951
bnl1	40	40	0,75	0,76	2814	2634
chr22b	29	29	19,33	17,94	938	909
chr25a	29	29	42,94	40,03	2785	2964
cre-a	27	27	7,67	7,65	176	187
cre-b	43	43	43,40	42,29	108	166
cre-c	27	27	5,83	5,02	151	155
cre-d	42	42	28,41	27,33	79	133
els19	31	31	44,24	35,49	3243	3212
ex01	28	28	0,41	0,34	1448	1036
ex02	46	37	0,95	0,67	6431	3717
ex05	39	39	5,82	4,92	2332	2290
ex09	45	52	52,14	54,95	11862	15541
ganges	18	18	0,63	0,63	326	383
ken13	29	29	93,80	92,35	33	34
ken18	41	41	1040,20	1011,89	409	504
maros	40	25	2,31	1,10	13813	5086
nesm	31	31	1,57	1,28	4943	3729
rou20	24	24	757,49	420,04	1470	1889
scr15	24	24	7,66	6,61	2013	1768
scr20	21	21	60,08	55,45	1534	2369
ste36a	37	37	14078,34	5523,68	14128	13793
stocfor2	21	21	1,13	1,17	467	477
stocfor3	32	32	87,90	87,51	5110	5154

Observa-se na Tabela 2 que as abordagens comparadas se comportam de forma parecida nas colunas correspondentes às iterações do IPM pois o nosso objetivo foi acelerar o tempo do cálculo da direção de busca em cada iteração do IPM. Apesar disso ocorrem melhoras em três problemas, 25fv47, ex02 e maros. O problema ex09 incrementou o número de iterações do IPM.

Com respeito ao tempo, o desempenho do PCxc foi melhor na maior parte dos problemas testados, as reduções de tempo consideráveis foram marcadas em negrito. Destacamos os problemas els19, ken18, rou20 e ste36a. Esta melhora aconteceu porque a base B com reordenação PCxc exige menos iterações do método PCG em algumas iterações do IPM o que implica que a nova reordenação da base melhora o condicionamento que o preconditionador Separador oferece às equações normais. Outra razão muito importante é que no total de iterações do IPM correspondentes ao preconditionador Separador na abordagem usando o PCxc foram necessários menos cálculos da matriz B quando comparados com o critério de reordenação de PCxm.

A propriedade de manter a matriz B da iteração k nas seguintes iterações é uma estratégia barata, observe que

$$\text{se } P_k^{-1} = D_B^{1/2} B^{-1} \text{ então } P_{k+1}^{-1} = \hat{D}_B^{1/2} B^{-1}.$$

Porém, a matriz preconditionada $I + \hat{W}\hat{W}^T$ da iteração $k + 1$ não tem mais o melhor

conjunto de colunas e portanto seu desempenho não é tão bom quando comparado com iteração k .

8. Conclusões

O desempenho do preconditionador Separador com a nova reordenação das colunas da matriz de restrições baseada na Proposição 1 obteve resultados superiores à abordagem baseada na heurística de reordenação proposta por M. Velazco(2011), particularmente nos problemas de grande porte há diferenças notáveis.

Novas estratégias assim como um critério de troca de base junto com técnicas que permitam a reutilização da base B de maneira mais eficiente serão investigadas devendo produzir resultados ainda melhores.

Finalmente, ressaltamos que reduzir tempo em não criar uma nova matriz B implica mais iterações no método PCG, por esta razão existem problemas tais como **ken18** com mais iterações no método PCG, porém resolvido em menos tempo.

Agradecimentos

Este trabalho contou com o apoio financeiro da FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Referências

- Bocanegra, S., Campos, F.F., Oliveira, A.R.L.** (2007) Using a hybrid preconditioner for solving large-scale linear systems arising from interior point methods, *Computational Optimization and Applications* 36(1-2): 149-164. *Special issue on "Linear Algebra Issues arising in Interior Point Methods"*
- Campos, F.F.** (1995) Analysis of conjugate Gradients-type methods for solving linear equations. *PhD Thesis, Oxford University Computing Laboratory, Oxford*
- Czyzyk, J., Mehrotra, S., Wagner, M. e Wright, S. J.** (1999), PCx an interior point code for linear programming, *Optimization Methods and Software* 11-2(1-4): 397-430.
- Gondzio, J.**(2012) Matrix-Free Interior Point Method, *Computational Optimization and Applications*, Volume 51, Issue 2, pp 457-480.
- Oliveira, A. R. L.** (1997) A new class of preconditioners for large-scale linear systems from interior point methods for linear programming. *PhD Thesis, TR97-11, Department of Computational and Applied Mathematics, Rice University, Houston TX*
- Velazco, M., Oliveira, A.R.L., Campos, F.F.** (2011) Heuristics for implementation of a hybrid preconditioner for interior-point methods, *Pesqui. Oper. vol.31 no.3 Rio de Janeiro*
- Wright, S. J.** *Primal-Dual Interior-Point Methods* , SIAM Publications, SIAM, Philadelphia, PA, USA,1997.