

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Experimentos numéricos com um sistema linear alternativo para o pré-condicionador separador aplicado a métodos de pontos interiores

Fábio Rodrigues Silva¹Aurelio Ribeiro Leite de Oliveira²

Instituto de Matemática, Estatística e Computação Científica, UNICAMP, Campinas, SP

Marta Ines Velazco Fontova³

Faculdade Campo Limpo Paulista – FACCAMP, Campo Limpo Paulista, SP

Resumo. Em trabalhos anteriores, uma abordagem híbrida de pré-condicionamento dos sistemas lineares oriundos da aplicação da proposta de Mehrotra, para resolver problemas de programação linear, foi desenvolvida e aprimorada. Foram considerados o pré-condicionador da Fatoração Controlada de Cholesky, com preenchimento adaptativo em função da memória disponível no computador, e o pré-condicionador separador, especializado para as iterações finais dos métodos de pontos interiores. Um dos blocos da matriz pré-condicionada, que é diagonal por blocos, foi reduzido a um sistema definido positivo de dimensão igual à diferença entre o número de colunas e linhas da matriz de restrições. Desenvolvemos duas abordagens para resolver o sistema resultante usando o método de Gradientes Conjugados pré-condicionado. Comparamos as três abordagens e aferimos o desempenho das versões utilizando quatro métricas. Os resultados apontam que as versões desenvolvidas neste trabalho têm melhor comportamento computacional e desempenho superior.

Palavras-chave. Programação Linear, Métodos de Pontos Interiores, Pré-condicionamento, Sistema Linear Alternativo, Gradientes Conjugados.

1 Introdução

O método de pontos interiores desenvolvido por Mehrotra [1] resolve o problema de programação linear calculando as direções de busca pela solução de dois sistemas lineares que compartilham a mesma matriz, além de usar uma heurística para estimar o comprimento de passo. Nos trabalhos [2–4], foram estudados métodos iterativos para solução destes sistemas lineares propondo uma abordagem híbrida de pré-condicionamento, desenvolvendo e aprimorando uma heurística para promover a troca de pré-condicionadores e melhorando a estratégia usada para obter uma matriz não singular a partir da matriz de restrições do problema de programação linear, necessária para implementação de um dos

¹ra153871@ime.unicamp.br

²aurelio@ime.unicamp.br

³marta.velazco@gmail.com

pré-condicionadores. A matriz que foi pré-condicionada é a matriz do sistema de equações normais, para ambas as abordagens de pré-condicionamento.

Neste trabalho, apresentamos duas abordagens para aplicação do pré-condicionador separador no sistema aumentado. A matriz pré-condicionada é diagonal por blocos e um destes blocos pode ser reduzido a um sistema simétrico definido positivo de dimensão igual à diferença entre o número de colunas e linhas da matriz de restrições do problema.

2 Métodos de pontos interiores tipo primal-dual

Considere um problema de programação linear na forma padrão e seu respectivo dual, também na forma padrão, dados abaixo

$$\begin{array}{ll} \min & c^\top x \\ \text{s. a} & \mathbf{A}x = b \\ & x \geq 0, \end{array} \quad \begin{array}{ll} \max & b^\top y \\ \text{s. a} & \mathbf{A}^\top y + s = c \\ & s \geq 0, \end{array} \quad (1)$$

onde $x, s, c \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{M}_{m \times n}(\mathbb{R})$, $b, y \in \mathbb{R}^m$, $n > m$ e $\text{posto}(\mathbf{A}) = m$. O método de pontos interiores proposto por Mehrotra pode ser visto como uma aplicação do método de Newton nas condições de Otimalidade de Karush-Kunh-Tucker (KKT) perturbadas, para o par de problemas (1) [5]. Uma modificação no procedimento de cálculo das direções de busca é considerada para que sejam admitidos pontos iniciais inactíveis. Considera-se um *parâmetro de centragem* $\sigma \in [0, 1]$ e a *medida de dualidade* $\mu := \frac{1}{n} \sum_{i=1}^n x_i s_i$, que fornece o valor médio dos pares de produtos $x_i s_i$. Definem-se os resíduos do sistema linear pondo $r_p = b - \mathbf{A}x$, $r_d = c - \mathbf{A}^\top y - s$ e $r_c = \sigma \mu e - \mathbf{X}Se$, e configura-se a direção de busca como sendo a solução para o seguinte sistema linear

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^\top & \mathbf{I} \\ \mathbf{S} & \mathbf{0} & \mathbf{X} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_c \end{bmatrix}. \quad (2)$$

Agora, partindo de um ponto inicial (x^0, y^0, s^0) com $(x^0, s^0) > 0$, são obtidas as direções $(\Delta x^{\text{aff}}, \Delta y^{\text{aff}}, \Delta s^{\text{aff}})$ resolvendo (2) com $\sigma = 0$. Em seguida, calculam-se os comprimentos de passo para as variáveis primais por $\alpha_{\text{aff}}^{\text{pri}} = \arg \max\{\alpha \in [0, 1]; x + \alpha \Delta x^{\text{aff}} \geq 0\}$ e para as variáveis duais por $\alpha_{\text{aff}}^{\text{dual}} = \arg \max\{\alpha \in [0, 1]; s + \alpha \Delta s^{\text{aff}} \geq 0\}$. A medida da eficiência da direção afim-escala é feita usando $\mu_{\text{aff}} = (1/n)(x + \alpha_{\text{aff}}^{\text{pri}} \Delta x^{\text{aff}})^\top (s + \alpha_{\text{aff}}^{\text{dual}} \Delta s^{\text{aff}})$. Em [1] sugere-se a heurística $\sigma = \left(\frac{\mu_{\text{aff}}}{\mu}\right)^3$, para escolher σ próximo de zero se $\mu_{\text{aff}} \ll \mu$ e próximo de 1 caso contrário. O passo combinado centragem-corretor $(\Delta x^{\text{cc}}, \Delta y^{\text{cc}}, \Delta s^{\text{cc}})$ é obtido resolvendo (2) com $r_p = 0$, $r_d = 0$ e $r_c = \sigma \mu e - \Delta \mathbf{X}^{\text{aff}} \Delta \mathbf{S}^{\text{aff}} e$. A direção final de busca configura-se como $(\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)}) = (\Delta x^{\text{aff}}, \Delta y^{\text{aff}}, \Delta s^{\text{aff}}) + (\Delta x^{\text{cc}}, \Delta y^{\text{cc}}, \Delta s^{\text{cc}})$. Novos comprimentos de passo são calculados como $\alpha_{\text{max}}^{\text{pri}} = \arg \max\{\alpha \geq 0; x^{(k)} + \alpha \Delta x^{(k)} \geq 0\}$, $\alpha_{\text{max}}^{\text{dual}} = \arg \max\{\alpha \geq 0; s^{(k)} + \alpha \Delta s^{(k)} \geq 0\}$ e, por fim, $\alpha_{(k)}^{\text{pri}} = \min\{0.99995 \alpha_{\text{max}}^{\text{pri}}, 1\}$ e $\alpha_{(k)}^{\text{dual}} = \min\{0.99995 \alpha_{\text{max}}^{\text{dual}}, 1\}$. Finalmente, os novos pontos primais e duais são calculados, pondo $x^{(k+1)} = x^{(k)} + \alpha_{(k)}^{\text{pri}} \Delta x^{(k)}$ e $(y^{(k+1)}, s^{(k+1)}) = (y^{(k)}, s^{(k)}) + \alpha_{(k)}^{\text{dual}} (\Delta y^{(k)}, \Delta s^{(k)})$.

Podemos eliminar a variável Δs do sistema (2) escrevendo

$$\begin{bmatrix} -\mathbf{D} & \mathbf{A}^\top \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} r_d - \mathbf{X}^{-1}r_c \\ r_p \end{bmatrix}, \quad (3)$$

onde $\mathbf{D} = \mathbf{X}^{-1}\mathbf{S}$ e Δs pode ser recuperado pondo $\Delta s = \mathbf{X}^{-1}(r_c - \mathbf{S}\Delta x)$. O sistema (3) recebe o nome de *sistema aumentado*. Eliminando Δx do sistema (3) escrevemos

$$\mathbf{Z}\Delta y = \mathbf{A}\mathbf{X}e - \sigma\mu\mathbf{A}e - \mathbf{A}\mathbf{D}^{-1}r_c - r_b, \quad (4)$$

onde $\mathbf{Z} = \mathbf{A}\mathbf{D}^{-1}\mathbf{A}^\top$ é o *Complemento de Schur*. O sistema (4) recebe o nome de *equações normais*. Recuperamos Δx fazendo $\Delta x = \mathbf{D}^{-1}\mathbf{A}^\top\Delta y + \mathbf{D}^{-1}r_c - \mathbf{S}^{-1}r_p$.

3 Estratégia híbrida de pré-condicionamento

A estratégia híbrida de pré-condicionamento considerada neste trabalho foi proposta em [3]. Os autores consideram que o processo iterativo de Mehrotra ocorre em duas fases, onde diferentes pré-condicionadores são usados em cada fase. Para as primeiras iterações do método de Mehrotra (fase 1), a estratégia é usar o pré-condicionador da Fatoração Controlada de Cholesky (FCC). Trata-se de uma fatoração incompleta de Cholesky que é construída minimizando a norma de Frobenius da diferença entre o fator completo de Cholesky e o fator incompleto. Tal minimização foi estabelecida envolvendo um termo η que controla o preenchimento do fator incompleto de Cholesky.

Neste mesmo trabalho, uma heurística que estabelece a mudança de pré-condicionadores foi proposta, sendo melhorada em [4]. Inicialmente, um valor η_0 é estabelecido em função do número de elementos não nulos da matriz do sistema de equações normais. A partir daí, no decorrer do processo iterativo de Mehrotra, preenchimento adicional é permitido no fator incompleto de Cholesky por acréscimos sucessivos ao valor η_0 até que se ultrapasse o valor $\bar{\eta}$, dado como parâmetro de entrada na solução de um problema. Quando isso acontece, o pré-condicionador separador proposto em [2] é utilizado em lugar do pré-condicionador da FCC (fase 2).

O pré-condicionador separador foi construído baseado na inversa do sistema aumentado, porém focado em evitar a formação do complemento de Schur na expressão de quaisquer dos blocos da matriz pré-condicionada. Obteve-se, assim, a matriz pré-condicionada

$$\begin{bmatrix} -\mathbf{I}_n + \mathbf{D}^{-\frac{1}{2}}\mathbf{A}^\top\mathbf{J}^\top + \mathbf{J}\mathbf{A}\mathbf{D}^{-\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D}_B \end{bmatrix} \quad (5)$$

e a matriz pré-condicionadora $\mathbf{M}^{-1} = \mathbf{N}^{-\top} = \begin{bmatrix} \mathbf{D}^{-\frac{1}{2}} & \mathbf{J} \\ \mathbf{H} & \mathbf{0} \end{bmatrix}$, onde $\mathbf{J} = \mathbf{H}^\top\mathbf{D}_B^{\frac{1}{2}}\mathbf{B}^{-1}$, $\mathbf{H}\mathbf{P}^\top = [\mathbf{I}_m \ \mathbf{0}]$, $\mathbf{A}\mathbf{P}^\top = [\mathbf{B} \ \mathbf{N}]$, \mathbf{B} é uma matriz não singular e \mathbf{P} é uma matriz de permutação. Temos que $\mathbf{P}\mathbf{D}\mathbf{P}^\top = \begin{bmatrix} \mathbf{D}_B & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_N \end{bmatrix}$, com $\mathbf{D}_B \in \mathbb{M}_m(\mathbb{R})$ e $\mathbf{D}_N \in \mathbb{M}_{n-m}(\mathbb{R})$ matrizes diagonais. O sistema linear cuja matriz é a do bloco superior esquerdo de (5) pode

ser reduzido a resolver um sistema linear definido positivo envolvendo uma das seguintes matrizes:

$$\mathbf{I}_m + \mathbf{D}_B^{\frac{1}{2}} \mathbf{B}^{-1} \mathbf{N} \mathbf{D}_N^{-1} \mathbf{N}^T \mathbf{B}^{-T} \mathbf{D}_B^{\frac{1}{2}} \quad (6a)$$

$$\mathbf{I}_{n-m} + \mathbf{D}_N^{-\frac{1}{2}} \mathbf{N}^T \mathbf{B}^{-T} \mathbf{D}_B \mathbf{B}^{-1} \mathbf{N} \mathbf{D}_N^{-\frac{1}{2}}. \quad (6b)$$

Atualmente, a estratégia híbrida (a qual nos referiremos como PCx_MOD) resolve a matriz pré-condicionada (6a) na fase 2, pois esta tem uma relação que fornece um pré-condicionador para o sistema de equações normais.

4 Abordagens propostas

Neste trabalho, os experimentos numéricos foram orientados pelo uso da matriz (6b) em vez de (6a). Como não foi obtida uma relação que fornecesse um pré-condicionador para o sistema de equações normais a partir desta matriz, as estratégias foram repensadas para o sistema aumentado na fase 2.

Inicialmente, construímos um novo pré-condicionador, análogo àquele fornecido em [2], tomando como base a matriz (6b) e multiplicando-a à esquerda e à direita por $\mathbf{D}_N^{\frac{1}{2}}$, obtendo a matriz

$$\mathbf{D}_N + \mathbf{N}^T \mathbf{B}^{-T} \mathbf{D}_B \mathbf{B}^{-1} \mathbf{N}. \quad (7)$$

Tal abordagem (aqui denominada PCx_F3) gerou uma matriz pré-condicionadora dada por $\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{\Theta} \\ \mathbf{H} & \mathbf{0} \end{bmatrix}$, onde $\mathbf{\Theta} = \mathbf{H}^T \mathbf{D}_B \mathbf{B}^{-1}$, $\mathbf{H} \mathbf{P}^T = [\mathbf{I}_m \ \mathbf{0}]$, $\mathbf{A} \mathbf{P}^T = [\mathbf{B} \ \mathbf{N}]$, e uma matriz pré-condicionada dada por $\begin{bmatrix} -\mathbf{D} + \mathbf{A}^T \mathbf{\Theta}^T + \mathbf{\Theta} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D}_B \end{bmatrix}$.

Uma segunda abordagem (denominada PCx_F4), foi elaborada objetivando construir a matriz

$$\mathbf{D}_N^{-1} + \mathbf{D}_N^{-1} \mathbf{N}^T \mathbf{B}^{-T} \mathbf{D}_B \mathbf{B}^{-1} \mathbf{N} \mathbf{D}_N^{-1} \quad (8)$$

multiplicando à esquerda e à direita da matriz (6b) por $\mathbf{D}_N^{-\frac{1}{2}}$. Obteve-se a matriz pré-condicionadora $\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{D}_N^{-1} & \mathbf{\Gamma} \\ \mathbf{H} & \mathbf{0} \end{bmatrix}$, onde $\mathbf{\Gamma} = \mathbf{H}^T \mathbf{B}^{-1}$, $\mathbf{H} \mathbf{P}^T = [\mathbf{I}_m \ \mathbf{0}]$ e $\mathbf{A} \mathbf{P}^T = [\mathbf{B} \ \mathbf{N}]$, e a matriz pré-condicionada $\begin{bmatrix} -\mathbf{D}_N^{-1} + \mathbf{\Gamma} \mathbf{A} \mathbf{D}_N^{-1} + \mathbf{D}_N^{-1} \mathbf{A}^T \mathbf{\Gamma}^T & \mathbf{0} \\ \mathbf{0} & -\mathbf{D}_B \end{bmatrix}$.

Estas duas estratégias foram incorporadas à versão híbrida, de modo que a fase 2 utilize as abordagens propostas em lugar do pré-condicionamento das equações normais, já implementada no PCx_MOD. Os sistemas lineares das matrizes (7) e (8) foram resolvidos com o método de Gradientes Conjugados pré-condicionado.

5 Experimentos numéricos

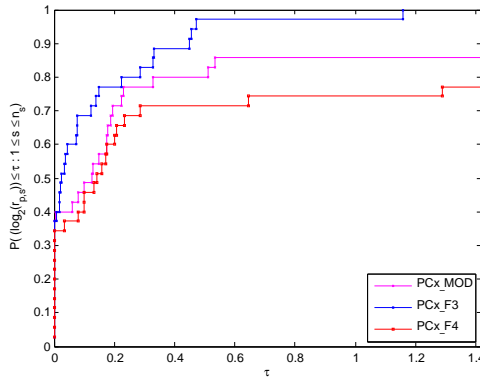
Os testes foram realizados em um computador equipado com processador Intel® Core™ i7 920 de 2,66GHz, além de 8GB de memória RAM. O sistema operacional da máquina é Li-

Tabela 1: Dados gerais dos problemas abordados.

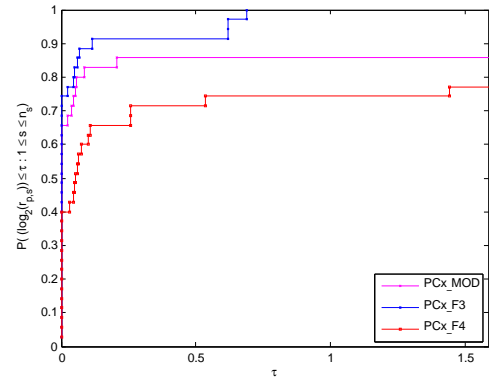
Coleção	Problema	Linhas (m)	Colunas (n)	$n - m$	$nmz(\mathbf{A})$	$\bar{\eta}$	η_0
KENN	cre-b	5.328	36.382	31.054	112.233	10	-38
KENN	cre-d	4.094	28.601	24.507	86.704	10	-38
KENN	ken11	9.964	16.740	6.776	38.157	10	23
KENN	ken13	22.365	36.561	14.196	82.191	50	23
KENN	ken18	78.538	128.434	49.896	297.886	10	23
MESZ	aa03	690	8.572	7.882	60.898	-20	-82
MESZ	air06	690	8.572	7.882	60.898	-20	-82
MESZ	aircraft	3.754	7.517	3.763	20.267	-140	-398
MESZ	bas1lp	5.410	9.824	4.414	587.771	-150	-260
MESZ	baxter_mat	23.871	30.122	6.251	106.586	-50	-42
MESZ	dbic1	34.205	174.457	140.252	819.746	-20	-51
MESZ	lp11	34.037	89.383	55.346	263.081	90	25
MESZ	nemsemm1	3.230	41.048	37.818	584.436	0	-73
MESZ	nemsemm2	4.526	26.754	22.228	91.127	100	28
MESZ	nemspmm1	2.227	7.145	4.918	43.606	420	-42
MESZ	nemspmm2.in	2.081	7.944	5.863	60.235	300	-43
MESZ	nsir2	4.450	10.054	5.604	154.935	-70	-73
MESZ	nsir1	4.406	10.010	5.604	143.247	-50	-69
MESZ	nw14	73	123.409	123.336	904.910	50	-46
MESZ	pcb3000	3.852	7.532	3.680	56.422	70	-31
MESZ	seymour	4.827	6.082	1.255	38.259	-100	-146
MESZ	slp-tsk	2.861	3.347	486	72.465	-1.140	-1.267
NETL	pilot	1.368	4.543	3.175	41.879	200	-63
NETL	pilots	1.368	4.543	3.175	41.879	200	-63
NETL	pilot87	1.971	6.373	4.402	72.163	90	-79
NETL	stocfor3	15.362	22.228	6.866	62.960	20	26
QAPL	chr22b	5.587	10.417	4.830	36.520	-40	-47
QAPL	chr25a	8.149	15.325	7.176	53.725	-30	-50
QAPL	els19	4.350	13.186	8.836	50.882	-30	-51
QAPL	scr20	5.079	15.980	10.901	61.780	-50	-52
QAPL	scr15	2.234	6.210	3.976	24.060	-10	-46
QAPL	rou20	7.359	37.640	30.281	152.980	0	-68
RAIL	rail507	507	62.689	62.182	404.727	0	-69
RAIL	rail516	516	47.827	47.311	315.412	0	-75
RAIL	rail582	582	54.906	54.324	394.113	0	-74

nux Kubuntu 15.10 e foi utilizado compilador GCC. Consideramos 35 problemas selecionados das bibliotecas KENNINGTON (<http://www.netlib.org/lp/data/kennington/>), MESZAROS (<http://www.sztaki.hu/~meszaros/publicftp/lptestset/misc/>), NETLIB (<http://www.netlib.org/lp/data/>), QAPLIB (<http://anjos.mgi.polymtl.ca/qaplib/>) e RAIL (<http://plato.asu.edu/ftp/lptestset/rail/>).

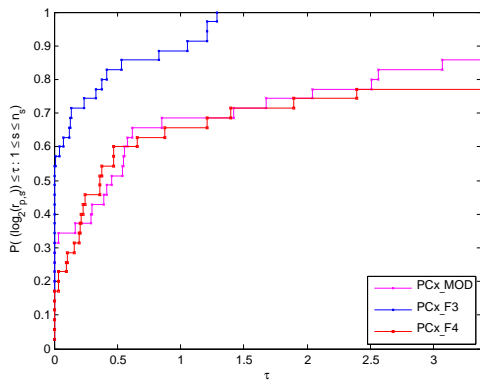
A Tabela 1 apresenta os dados básicos dos problemas abordados. A quantidade de linhas e colunas apresentada se refere à formulação resultante da etapa de pré-processamento. A sigla nmz indica o número de elementos não nulos da matriz após o pré-processamento. O parâmetro $\bar{\eta}$, fornecido como dado de entrada, foi escolhido a partir de testes prévios com alguns valores pré-estabelecidos e usado aquele que apresentou o melhor resultado



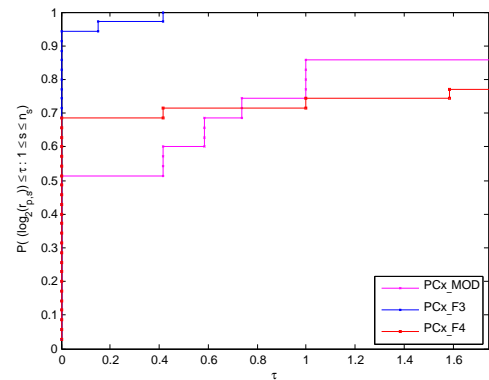
(a) Tempo total de solução



(b) Iterações de pontos interiores



(c) Iterações de Gradientes Conjugados



(d) Número de cálculos de uma nova **B**

Figura 1: Perfis de desempenho para as métricas consideradas

dentre os avaliados, para as três versões consideradas neste trabalho.

Os resultados foram resumidos em perfis de desempenho. Dos 35 problemas teste selecionados, a versão PCx_MOD não conseguiu resolver 5 e a versão PCx_F4 não conseguiu resolver 8, enquanto que a versão PCx_F3 resolveu toda a lista de problemas. As métricas consideradas foram tempo total de solução, iterações de pontos interiores e de gradientes conjugados, além da quantidade de vezes que a fase 2 necessitou uma nova matriz **B**.

A figura 1 apresenta os perfis de desempenho para estas quatro métricas. Em apenas um problema, as versões PCx_F3 e PCx_F4 fizeram mais cálculos de uma nova **B** que a versão PCx_MOD. Para os demais problemas, menos trocas foram observadas e, nos problemas onde houve mesma quantidade de cálculos, estes foram feitos mais tardiamente para as versões PCx_F3 e PCx_F4. Para as demais métricas, de modo geral, podemos ver que a versão PCx_F3 se destaca, necessitando de menos iterações de gradientes conjugados e de pontos interiores, além de resolver em menor tempo a lista de problemas propostos.

A versão PCx_F4 tem ligeira vantagem no comportamento computacional em relação à versão PCx_MOD, tendo em vista que dos 8 problemas não resolvidos, todos terminaram com status “desconhecido”, sendo 6 deles por convergência lenta. A versão PCx_MOD deixou de resolver 5 problemas, com 1 status “infectível” e 4 como “desconhecido”, dos quais apenas 1 foi por convergência lenta.

6 Considerações Finais

Desenvolvemos duas abordagens para recuperação da solução do sistema aumentado pré-condicionado pelo pré-condicionador separador, no contexto da solução de problemas de programação linear usando o método de pontos interiores de Mehrotra com pré-condicionamento híbrido. As abordagens apresentam, de maneira geral, melhor comportamento computacional que a versão híbrida proposta em trabalhos anteriores, para a maioria dos problemas testados. Elas necessitam de menor tempo computacional e iterações de pontos interiores e gradientes conjugados, além de realizarem poucos cálculos de uma nova matriz não singular para o pré-condicionador separador.

Agradecimentos

Agradecemos ao PICME–OBMEP, CNPq e FAPESP, pelo apoio financeiro.

Referências

- [1] S. Mehrotra. On the Implementation of a Primal-Dual Interior Point Method. *SIAM Journal on Optimization*, 2(4):575–601, 1992. DOI: 10.1137/0802028.
- [2] A. R. L. Oliveira and 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*, 394:1–24, 2005. DOI: 10.1016/j.laa.2004.08.019.
- [3] S. Bocanegra, F. F. Campos, and A. R. L. Oliveira. Using a hybrid preconditioner for solving large-scale linear systems arising from interior point methods. *Computational Optimization and Applications*, 36(2):149–164, 2007. DOI: 10.1007/s10589-006-9009-5.
- [4] M. I. Velazco, A. R. L. Oliveira, and F. F. Campos. A Note on Hybrid Preconditioners for Large-scale Normal Equations Arising from Interior-point Methods. *Optimization Methods Software*, 25(2):321–332, 2010. DOI: 10.1080/10556780902992829.
- [5] S. J. Wright. *Primal-Dual Interior-Point Methods*. Society for Industrial and Applied Mathematics, Philadelphia, 1997. ISBN: 978-0-89871-382-4.