Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 6, n. 2, 2018.

Trabalho apresentado no XXXVIII CNMAC, Campinas - SP, 2018.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Métodos de Pontos Interiores com a Aplicação do Precondicionador de Elman

Ingrid Araújo Sampaio[1]
DECOM - Departamento de Comunicações, FEEC, UNICAMP, Campinas, SP
Yuzo Iano[2]
DECOM - Departamento de Comunicações, FEEC, UNICAMP, Campinas, SP
Aurelio Ribeiro Leite de Oliveira[3]
Instituto de Matemática, Estatística e Computação Científica-IMECC, UNICAMP, Campinas, SP
Lino Marcos da Silva[4]
Colegiado de Engenharia Elétrica, UNIVASF, Juazeiro, BA

Resumo: Neste trabalho, apresentamos um método de pontos interiores com a resolução do sistema linear necessário em cada iteração por métodos iterativos. Vamos utilizar o método dos gradientes conjugados precondicionado para resolução desse sistema linear. Em particular, precondicionadores adaptativos a este sistema linear foi aplicado em conjunto com o método dos gradientes conjugados, obtendo assim bons resultados computacionais. O precondicionador denominado fatoração controlada de Cholesky(FCC) é utilizado nas iterações iniciais e o precondicionador separador especialmente desenvolvido para iterações finais é então utilizado. No entanto, esta abordagem ainda não é robusta porque para muitos problemas existe uma faixa no espectro das iterações onde a fatoração controlada de Cholesky já não é eficiente e o precondicionador separador ainda não obtém resultados satisfatórios. Portanto, nossa proposta é combinar estes precondicionadores com o precondicionador de Elman, visando obter um desempenho computacional ainda superior tanto no aspecto de robustez como no tempo total de processamento. Desta forma, as qualidades dos precondicionadores utilizados seriam combinadas melhorando a eficiência da abordagem.

Palavras-chave: Otimização, Programação Linear, Precondicionadores, Gradientes Conjugados Precondicionados, Métodos de Pontos Interiores.

1 Introdução

Desde o surgimento dos métodos de pontos interiores para programação linear, códigos computacionais sofisticados baseados nessas ideias vêm se firmando como alternativas eficientes para resolução de problemas de grande porte com estruturas genéricas [4, 6–9]. Cada iteração de um método de pontos interiores envolve a solução de um ou mais sistemas lineares [6, 7]. Este é o passo mais caro deste método do ponto de vista computacional.

[1]ingrid_sampaio@yahoo.com.br
[2]yuzo@decom.fee.unicamp.br
[3]aurelio@ime.unicamp.br
[4]lino.silva@univasf.edu.br

Em aplicações reais esses sistemas quase sempre possuem dimensões elevadas e alto grau de esparsidade.

Usualmente métodos diretos são utilizados para resolver esses sistemas. A abordagem mais usada nas implementações existentes é a fatoração de Cholesky nas equações normais [4, 6, 7]. No entanto, por limitações de tempo e memória seu uso torna-se inadequado em muitos problemas de grande porte, fazendo com que abordagens iterativas sejam mais adequadas [1, 8, 9].

O precondicionador fatoração controlada de Cholesky, proposto em 1995 [3], pode ser visto como uma variação da fatoração incompleta de Cholesky, onde o objetivo principal é construir uma matriz precondicionada que possua os autovalores agrupados e próximos à unidade, de forma a acelerar a convergência do método dos gradientes conjugados, através do controle do número de elementos não nulos na fatoração.

O precondicionador separador foi proposto para sistemas aumentados originados de métodos de pontos interiores em [8] para problemas de programação linear, onde a principal característica dessa classe de precondicionadores é que ela apresenta melhores resultados nas proximidades de uma solução ótima, quando os sistemas lineares já são muito mal condicionados.

2 Problemas de Programação Linear

Consideremos o seguinte problema de programação linear na forma padrão:

$$\begin{array}{lrl} \text{minimizar} & c^t x & \\ \text{s.a} & Ax & = b, \\ & x & \geq 0, \end{array} \tag{1}$$

onde $A \in \Re^{m \times n}$, $c, x \in \Re^n$, e $b \in \Re^m$. O problema (1) é chamado primal e associado a este problema tem-se o problema dual, que é dado por

$$\begin{array}{lrl} \text{maximizar} & b^t y & \\ \text{s.a} & A^t y + z & = c, \\ & z & \geq 0, \end{array} \tag{2}$$

onde $y \in \Re^m$ e $z \in \Re^n$.

Os problemas (1) e (2) juntos são chamados de par primal-dual. As condições de otimalidade de primeira ordem (Karush-Kuhn-Tucker) dos problemas (1) e (2) são dadas por:

$$\left\{ \begin{array}{r} Ax - b = 0, \\ A^t y + z - c = 0, \\ XZe = 0, \\ (x, z) \geq 0. \end{array} \right. \tag{3}$$

onde $X = diag(x)$, $Z = diag(z)$ e $e \in \Re^n$ é o vetor com todas as coordenadas iguais a um.

Uma solução deste problema pode ser obtida resolvendo o sistema de equações não lineares (3). Se (x, y, z) for uma solução do problema (3), então x e (y, z) são soluções ótimas de (1) e (2), respectivamente. Um ponto (x, y, z) é dito ser factível se ele satisfaz o conjunto de restrições dos problemas primal e dual. O ponto é dito ser interior se $(x, z) > 0$.

3 Métodos Primais-Duais de Pontos Interiores

Estes métodos, como o próprio nome sugere, buscam uma solução ótima do problema de programação linear percorrendo o interior do ortante positivo. Os métodos de pontos interiores podem ser classificados como primal, dual e primal-dual dependendo do espaço em que estão sendo realizadas as iterações. Alternativamente, eles podem ser classificados em três categorias: métodos afim escala, métodos de redução de potencial e métodos de trajetória central. A classe de métodos de pontos interiores que possui as melhores propriedades práticas e teóricas são os chamados métodos primais-duais pertencentes a categoria trajetória central. Por isso, em nosso trabalho nos restringiremos ao estudo de um método primal-dual preditor corretor pertencente a esta classe.

Os métodos primais-duais de pontos interiores podem ser vistos como aplicações do método de Newton para calcular aproximações da solução de uma sequência de sistemas não lineares (3) perturbados por um parâmetro μ.

$$F_\mu(x, y, z) = \begin{bmatrix} Ax - b \\ A^t y + z - c \\ -XZe + \mu e \end{bmatrix} = 0, \ \ (x, z) \geq 0. \tag{4}$$

Notemos que se $\mu = 0$, então os problemas (3) e (4) são equivalentes. Assim a solução do problema (4) se aproxima da solução do problema (3) conforme $\mu \to 0$. Um método primal-dual obtém uma solução aproximada para o problema gerando uma sequência de pontos (x^k, y^k, z^k) e de parâmetros μ^k. A cada iteração do método é aplicado um passo do método de Newton para resolver o sistema (4), com um dado parâmetro μ^k. A direção de Newton $\triangle = (\triangle x, \triangle y, \triangle z)$ é obtida da solução do seguinte sistema de equações lineares:

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{pmatrix} \begin{bmatrix} \triangle x^k \\ \triangle y^k \\ \triangle z^k \end{bmatrix} = \begin{bmatrix} b - Ax^k \\ c - z^k - A^t y^k \\ \mu^k e - X^k Z^k e \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_{\mu^k} \end{bmatrix}. \tag{5}$$

O novo ponto é dado por

$$(x^{k+1}, y^{k+1}, z^{k+1}) = (x^k, y^k, z^k) + (\alpha_p \triangle x^k, \alpha_d \triangle y^k, \alpha_d \triangle z^k),$$

onde α_p e α_d são os tamanhos do passo primal e dual, respectivamente e preservam a não negatividade das variáveis x e z.

4 Método Preditor-Corretor

O método preditor-corretor difere do método de pontos interiores primal-dual, no cálculo da direção de Newton. A direção de Newton $\triangle = (\triangle x, \triangle y, \triangle z)$ é decomposta em duas partes, $\triangle = \triangle_a + \triangle_c$.

O termo $\triangle_a$ é obtido resolvendo o sistema (5) para $\mu = 0$. A componente $\triangle_a$, conhecida como direção afim, é responsável por determinar uma melhor predição para o parâmetro da barreira. Este parâmetro é escolhido usando a seguinte heurística

$$\mu^k = \left(\frac{(x^k + \alpha_{pa} \triangle_a x^k)^t (z^k + \alpha_{da} \triangle_a z^k)}{(x^k)^t z^k} \right)^p \frac{(x^k + \alpha_{pa} \triangle_a x^k)^t (z^k + \alpha_{da} \triangle_a z^k)}{n}, \qquad (6)$$

onde α_{pa} e α_{da} são os passos que preservam a não negatividade das variáveis x e z. [7] sugere que o valor do expoente seja $p = 2$ ou $p = 3$.

A direção corretora, $\triangle_c$ é dada por,

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{pmatrix} \begin{bmatrix} \triangle_c x^k \\ \triangle_c y^k \\ \triangle_c z^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ r_m \end{bmatrix}. \qquad (7)$$

onde,

$$\triangle_a X^k = diag(\triangle_a x^k), \quad \triangle_a Z^k = diag(\triangle_a z^k) \text{ e } r_m = \mu^k e - \triangle_a X^k \triangle_a Z^k e. \qquad (8)$$

O termo r_m tem por objetivo fazer com que a nova iteração seja a mais central possível e fazer a correção do termo não linear.

Eliminando $\triangle x$ e $\triangle z$ em todos os sistemas teremos $ADA^t \triangle y = r$.

5 Descrição do Problema

O precondicionamento é uma estratégia que deve ser aplicada aos métodos iterativos para melhorar as características de convergência de sistemas que possuem a matriz de coeficientes com autovalores muito dispersos. Precondicionar um sistema linear baseia-se em tornar a matriz dos coeficientes com condições desejadas para que o método que está sendo aplicado na resolução do sistema seja eficiente.

Considerando o sistema linear $Ax = b$, sendo A uma matriz simétrica definida positiva $n \times n$, o objetivo do gradiente conjugado preecondicionado é aplicar o método do gradiente conjugado no sistema transformado

$$\tilde{A}\tilde{x} = \tilde{b},$$

com

$$\tilde{A} = D^{-1} A D^{-1}, \quad \tilde{x} = Dx, \quad \tilde{b} = D^{-1} b,$$

e D simétrica definida positiva, devemos escolher D para que $\tilde{A}$ seja bem condicionada ou uma matriz com autovalores agrupados.

5.1 Fatoração Controlada de Cholesky

Considere a matriz $ADA^t \in \Re^{m \times m}$ fatorada em

$$ADA^t = LL^t = \tilde{L}\tilde{L}^t + R, \qquad (9)$$

sendo

L: o fator obtido pela fatoração completa de Cholesky;
$\tilde{L}$: o fator obtido pela fatoração incompleta de Cholesky;
R: matriz dos resíduos.
Usando $\tilde{L}$ como uma matriz precondicionadora para ADA^t, obtém-se

$$\tilde{L}^{-1}ADA^t\tilde{L}^{-t} = (\tilde{L}^{-1}L)(L^t\tilde{L}^{-t}) = (\tilde{L}^{-1}L)(\tilde{L}^{-1}L)^t.$$

Definindo $E = L - \tilde{L}$ e substituindo na igualdade acima temos

$$\tilde{L}^{-1}ADA^t\tilde{L}^{-t} = (I + \tilde{L}^{-1}E)(I + \tilde{L}^{-1}E)^t.$$

Observemos que, quando $\tilde{L} \to L$ então $E \to 0$ e portanto $\tilde{L}ADA^t\tilde{L}^{-t} \to I$.

A fatoração controlada de Cholesky é construída com base na minimização da norma de Frobenius de E, visto que, quando $||E|| \to 0$ implica $||R|| \to 0$.

5.2 Precondicionador Separador

Consideremos o seguinte precondicionador simétrico em blocos para o sistema aumentado S
$M^{-1} = \begin{pmatrix} F & G \\ H & J \end{pmatrix}$ onde $S = \begin{pmatrix} -D & A^t \\ A & 0 \end{pmatrix}$.
Com ajustes adequados, a matriz precondicionada por M^{-1} é dada por,

$$M^{-1}\begin{pmatrix} D^{-1} & A^t \\ A & 0 \end{pmatrix}M^{-t} = \begin{pmatrix} -I_n + D^{\frac{1}{2}}A^tG^t + GAD^{\frac{1}{2}} & 0 \\ 0 & -D_B^{-1} \end{pmatrix} \tag{10}$$

sendo,
$M^{-1} = \begin{pmatrix} D^{\frac{1}{2}} & G \\ H & 0 \end{pmatrix}, G = H^tD^{-\frac{1}{2}}B^{-1}, HP = [I\,0]$ e $AP^t = [BN]$.

O sistema aumentado tem dimensão $n + m$, onde n é o número de variáveis e m o número de restrições do problema de programação linear. Então podemos considerar que o precondicionador reduz a dimensão do sistema aumentado para n, visto que a parte inferior é formada apenas pela matriz diagonal D_B^{-1} e, portanto de fácil resolução.

A redução a um sistema de equações normais resulta na matriz,

$$I_m + D_B^{-\frac{1}{2}}B^{-1}ND_NN^tB^{-t}D_B^{-\frac{1}{2}}. \tag{11}$$

Observe que a matriz precondicionada é definida positiva e tem autovalores maiores ou iguais a um, isto é, não tem autovalores próximos de zero.

6 Abordagem Adotada

Abordagens de precondicionamento híbrida, usando a fatoração controlada de Cholesky e o precondicinador Separador, para o problema de programação linear têm sido

DOI: 10.5540/03.2018.006.02.0292

adotadas com sucesso [2, 9]. No entanto, o estudo de novos precondicionadores, como o precondicionador de Elman, podem melhorar consideravelmente a técnica. Neste trabalho, propomos uma estratégia de resolução do problema de programação linear por meio de um método de pontos interiores, usando o método dos gradientes conjugados precondicionado com o precondicionador de Elman nas primeiras iterações e usando o método direto usual, a fatoração de Cholesky, nas demais iterações, para resolver os sistemas necessários.

Para a matriz do sistema de equações normais ADA^t o precondicionador de Elman [5] fica da forma:

$$(AA^t)^{-1}(ADA^t)(AA^t)^{-1}.$$

Utilizando o Matlab, realizamos testes com os problemas da netlib apresentados na Tabela 1. Nos testes consideramos o uso do precondicionador de Elman apenas nas iterações de 1 até 4. Todos os problemas testados foram resolvidos até a otimalidade, no entanto, os melhores resultados, com relação ao número de iterações do método de pontos interiores, foram obtidos quando o precondicionador foi utilizado na primeira iteração e, também, estão apresentados na Tabela 1. Os resultados encontrados são promissores uma vez que o número de iterações do método de pontos interiores, para os problemas testados, está compatível com o esperado para o método. Dessa forma, vislumbra-se que em abordagens de precondicionamento híbrido, o uso do precondicionador de Elman juntamente com os precondicionadores Controlada de Cholesky e Separador possam proporcionar uma redução no tempo computacional total do método de pontos interiores.

Os testes foram realizados utilizando Mac OS X, com processador 1.8 GHz Intel Core i5, memória 4gb, 1600 MHz DDR3, através do software Matlab com precisão de 10^{-8}.

Tabela 1: Resultados obtidos

Problemas	Dimensão ($m \times n$)	Iterações GC	Iterações PI	Tempo(segundos)
scsd8	397×2750	169 / 222	20	12.32
czprob	929×3562	214 / 420	85	224.16
nesm	662×3105	310 / 363	18	64.94
truss	1000×8806	122 / 129	33	273.87
adlittle	56×138	69 / 112	20	0.05
scsd1	77×760	30 / 52	14	0.13
scsd6	147×1350	44 / 73	19	0.47
sc50b	50×78	52 / 86	13	0.04
sc105	105×163	74 / 140	15	0.24
scagr25	471×671	146 / 468	29	19.95

7 Conclusão

Neste trabalho, problemas de programação linear foram resolvidos por meio de um método de pontos interiores no qual, nas iterações iniciais, foi utilizado o método dos gradientes conjugados precondicionado com o precondicionador de Elman. A fatoração de Cholesky foi utilizada nas demais. Os resultados obtidos sugerem que o precondicionador de Elman tem potencial de ser utilizado em abordagens de precondicionamento híbrido combinado com a controlada de Cholesky e o precondicinador separador, o que será realizado em trabalhos futuros quando uma implementação em C do precondiciondor for efetivada.

Referências

[1] L. Bergamaschi, J. Gondzio, M. Venturin, and G. Zilli. Inexact constraint preconditioners for linear systems arising in interior point methods, *Comp. Opt. and Appl.*, 36:137-147, 2007. DOI: 10.1007/s10589-006-9001-0.

[2] 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. *Comp. Opt. and Appl.*, 36:149-164, 2007. DOI: 10.1007/s10589-006-9009-5.

[3] F. F. Campos and N. R. C. Birkett. An efficient solver for multi-right hand side linear systems based on the CCCG(η) method with applications to implicit time-dependent partial differential equations, *SIAM J. Sci. Comput.*, 19(1):126-138, 1998. DOI: 10.1137/S106482759630382X.

[4] J. Czyzyk, S. Mehrotra, M. Wagner, and S. J. Wright. PCx an interior point code for linear programming, *Opt. Meth. & Software*, 11-2:397-430, 1999. DOI: 10.1080/10556789908805757.

[5] H. C. Elman, Preconditioning for the steady-state Navier-Stokes equations with low viscosity, *SIAM Journal on Sci. Comp.*, 20:2991-1316, 1999. DOI: 10.1137/S1064827596312547.

[6] J. Gondzio. Multiple centrality corrections in a primal-dual method for linear programming, *Comp. Opt. and Appl.*, 6:137-156, 1996. DOI:10.1007/BF00249643.

[7] S. Mehrotra, On the implementation of a primal-dual interior point method, SIAM Journal on Opt., 2:575-601, 1992. DOI: 10.1137/0802028.

[8] 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 Appl.*, 394:1-24, 2005. DOI: 10.1016/j.laa.2004.08.019.

[9] M. I. Velazco, A. R. L. Oliveira, and F. F. Campos. A note on hybrid preconditions for large scale normal equations arising from interior-point methods, *Opt. Meth. and Software*, 25:321-332, 2010. DOI: 10.1080/10556780902992829.

