

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Sistemas Lineares Aproximados em Métodos de Pontos Interiores

Luciana Yoshie Tsuchiya¹

Instituto Federal do Paraná, IFPR, Paranavaí, PR

Aurelio Ribeiro Leite de Oliveira²

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

Resumo. Uma das abordagens utilizadas para resolver o sistema linear que surge a cada iteração nos métodos de pontos interiores primal-dual é reduzi-lo a um sistema linear equivalente simétrico definido positivo, conhecido como sistema de equações normais, e aplicar a fatoração de Cholesky na matriz do sistema. A grande desvantagem desta abordagem é o preenchimento gerado durante a fatoração, o que pode tornar seu uso inviável, por limitação de tempo e memória. Com o intuito de contornar o problema de preenchimento gerado na fatoração de Cholesky, neste trabalho, estamos propondo uma abordagem que resolve de forma direta sistemas lineares aproximados do sistema de equações normais e que exerce um certo controle sobre o preenchimento.

Palavras-chave. Programação linear, Método de pontos interiores primal-dual, Fatoração de Cholesky, Fatoração controlada de Cholesky, Sistema de equações normais.

1 Introdução

Um problema de programação linear na forma padrão e o seu problema dual associado são dados respectivamente por

$$\begin{array}{ll} \min c^t x & \max b^T y \\ Ax = b & A^T y + z = c \\ x \geq 0 & y \in \mathbb{R}^m, z \geq 0, \end{array}$$

em que $A^{m \times n}$ é uma matriz de posto completo e c, x, b, y e z são vetores colunas de dimensões apropriadas.

Os métodos de pontos interiores (MPI) primal-dual, encontram uma solução ótima (x^*, y^*, z^*) para o par de problemas acima, aplicando variações do método de Newton às três igualdades das condições de otimalidade: $Ax = b$, $A^T y + z = c$, $XZe = 0$ e $(x, z) > 0$ [1].

¹luciana.tsuchiya@ifpr.edu.br

²aurelio@ime.unicamp.br

A cada iteração do método deve-se então resolver um sistema de Newton para encontrar a direção de busca (dx, dy, dz) . Em aplicações reais esse sistema linear quase sempre possui dimensões elevadas e um alto grau de esparsidade e sua resolução é o passo mais caro do método.

Realizando operações algébricas neste sistema, pode-se reduzi-lo a um sistema equivalente simétrico definido positivo dado por $ADA^tdy = h$, em que $D = XZ^{-1}$, com $X = \text{diag}(x)$, $Z = \text{diag}(z)$, e h é calculado de maneira apropriada. Este sistema é conhecido como sistema de equações normais.

Uma das abordagens utilizadas para resolver o sistema de equações normais é a aplicação da fatoração de Cholesky na matriz do sistema, já que ela é simétrica definida positiva. Assim, a solução se dá por meio da resolução de dois sistemas triangulares. Esta abordagem é uma forma bastante estável para calcular soluções de sistemas lineares com a matriz dos coeficientes simétrica definida positiva, por isso, ela é utilizada nos MPI primal-dual sempre que a decomposição não é extremamente cara.

A grande desvantagem da fatoração de Cholesky é o preenchimento gerado durante a decomposição, o que leva à necessidade de mais espaço de armazenamento e maior tempo computacional para a resolução dos sistemas lineares. Logo, seu uso pode se tornar inviável por limitação de tempo e memória, fazendo com que a utilização de métodos iterativos se torne mais apropriada. Neste contexto, a escolha mais natural de um método para resolver o sistema de equações normais é o método dos gradientes conjugados preconditionado (GCP).

Em geral, a matriz do sistema de equações normais sofre mudanças significativas entre as iterações dos MPI primal-dual e se torna extremamente mal condicionada nas iterações finais, por isso, junto ao método GCP nos MPI é mais adequado a utilização de preconditionadores híbridos compostos por mais de um tipo de preconditionador, em que cada um é mais apropriado para uma determinada etapa da resolução do problema [2, 3].

Um preconditionador que se mostra adequado para as iterações iniciais dos MPI é construído pela fatoração controlada de Cholesky(FCC) [4], que é um tipo de fatoração incompleta de Cholesky.

Neste tipo de fatoração, ao fatorar uma matriz quadrada, obtemos uma matriz triangular inferior \tilde{L} , cujos elementos são calculados como na fatoração de Cholesky, porém devido a algum tipo de controle de preenchimento, \tilde{L} é mais esparsa que o fator de Cholesky.

A consequência de se resolver o sistema de Newton que surge nos MPI por um método iterativo é o surgimento de um erro residual do lado direito da equação, mas se este erro satisfizer certas condições em cada iteração, ainda é possível calcular uma boa direção no sentido de que a convergência do método é atingida [5, 6].

Com o intuito de contornar o problema de preenchimento gerado na fatoração de Cholesky e motivados pelo fato de que uma direção aproximada com boas propriedades pode ser calculada no método de pontos interiores, neste trabalho, estamos propondo uma abordagem que resolve de forma direta, sistemas aproximados do sistema de equações normais e que exerce um certo controle sobre o preenchimento. Isto é feito substituindo a fatoração de Cholesky da matriz do sistema de equações normais, pela fatoração controlada de Cholesky.

2 Fatoração Controlada de Cholesky

Na fatoração incompleta de Cholesky de uma matriz quadrada \mathcal{A} de ordem $m \times m$, uma vez calculada uma coluna de \tilde{L} , o controle sobre o preenchimento é feito descartando por algum critério um certo número de elementos da coluna. Após este descarte, passa-se para o cálculo da próxima coluna, realizando o mesmo procedimento. Desta forma, obtemos uma matriz triangular inferior \tilde{L} , tal que $\mathcal{A} = \tilde{L}\tilde{L}^T - R$, onde R é uma matriz chamada de matriz resto.

Em geral, as fatorações incompletas de Cholesky são utilizadas como preconditionadores para os métodos de gradientes conjugados preconditionados. Em [7] é mostrado que o número de iterações do gradiente conjugado está quase que diretamente relacionado com a norma da matriz R . Isto motivou a construção da FCC baseada na minimização da norma de Frobenius da matriz $E = \tilde{L} - L$. De fato, reescrevendo R da forma

$$\begin{aligned} R &= LL^T - \tilde{L}\tilde{L}^T \\ &= LL^T - L\tilde{L}^T + L\tilde{L}^T - \tilde{L}\tilde{L}^T \\ &= L(L - \tilde{L})^T + (L - \tilde{L})\tilde{L}^T, \end{aligned}$$

podemos observar que, $\|R\| \rightarrow 0$ quando $\|E\| \rightarrow 0$.

No problema

$$\min \|E\|_F^2 = \sum_{j=1}^m c_j, \tag{1}$$

com $c_j = \sum_{i=1}^m |l_{ij} - \tilde{l}_{ij}|^2$, onde l_{ij} e \tilde{l}_{ij} , $i, j = 1, \dots, m$ são os elementos de L e \tilde{L} respectivamente, os elementos c_j podem ser reescritos como soma de dois somatórios

$$c_j = \sum_{k=1}^{n_j+\eta} |l_{ikj} - \tilde{l}_{ikj}|^2 + \sum_{k=n_j+\eta+1}^m |l_{ikj}|^2, \tag{2}$$

onde n_j é o número de elementos não nulos da coluna j de \mathcal{A} e η é um parâmetro que determina a quantidade de preenchimentos permitidos em relação a n_j , por coluna na fatoração.

A primeira parcela de (2) contém todos os elementos não nulos da coluna j de \tilde{L} e a segunda contém apenas os elementos restantes do fator completo L , cujo correspondente em \tilde{L} é nulo. Então, se consideramos que $\tilde{l}_{ij} \approx l_{ij}$, uma redução da norma de Frobenius de E pode ser obtida por meio da seguinte heurística:

- Aumentar o valor de η permitindo mais preenchimento. Isto faz com que c_j , $j = 1, \dots, m$ e conseqüente $\|E\|_F$ decresça, pois o primeiro somatório em (2) conterá mais elementos e o segundo menos elementos.
- Para um η fixo, escolher os $n_j+\eta$ maiores elementos da coluna j de L , para $j = 1, \dots, m$, em valor absoluto. Assim, para uma determinada quantidade de armazenamento, um fator incompleto \tilde{L} ótimo é obtido, já que os maiores elementos estarão no primeiro somatório e os menores no segundo.

A FCC combina os dois itens acima. Uma vez calculada a coluna j como na fatoração de Cholesky, ela propõe utilizar um parâmetro η tomando o número máximo de elementos não-nulos desta coluna como sendo $k_j = \eta + n_j$. Os k_j elementos maiores em valor absoluto calculados, serão mantidos na coluna j de \tilde{L} e os demais descartados, isto é, serão nulos.

Esta heurística permite prever o armazenamento necessário na fatoração e exercer um controle de preenchimento de acordo com a memória disponível. Como para um dado problema, o número de elementos não nulos de \mathcal{A} e m são fixos, a dimensão de \tilde{L} depende apenas do parâmetro η , que por sua vez, pode assumir qualquer valor de $-m$ a m . Se $\eta = -m$, a matriz \tilde{L} será diagonal, se $\eta = 0$, ela requer o mesmo espaço de armazenamento da matriz \mathcal{A} e se $\eta = m$, a matriz \tilde{L} é o fator de Cholesky.

3 Nova abordagem na solução direta do sistema de equações normais

A nova abordagem propõe que, ao invés de resolver em cada iteração o sistema linear $LL^T dy = h$, onde L é o fator de Cholesky da matriz ADA^T , resolvamos, ainda diretamente, um sistema da forma

$$\tilde{L}\tilde{L}^T \tilde{d}y = h, \tag{3}$$

em que \tilde{L} é o fator incompleto da matriz ADA^T , calculado por meio da fatoração controlada de Cholesky. Assim, a cada iteração é encontrada uma solução $\tilde{d}y$ que é uma aproximação da solução original. Observamos que se $\tilde{d}y$ satisfaz (3), então

$$ADA^T \tilde{d}y = h + \tilde{r}, \tag{4}$$

onde $\tilde{r} = ADA^T \tilde{d}y - h$.

Tomando a norma dois de \tilde{r} , temos que

$$\begin{aligned} \|\tilde{r}\|_2 &= \|ADA^T \tilde{d}y - h\|_2 \\ &= \|LL^T \tilde{d}y - \tilde{L}\tilde{L}^T \tilde{d}y\|_2 \\ &\leq \|LL^T - \tilde{L}\tilde{L}^T\|_F \|\tilde{d}y\|_2 \\ &= \|L(L - \tilde{L})^T + (L - \tilde{L})\tilde{L}^T\|_F \|\tilde{d}y\|_2 \\ &\leq (\|L\|_F \|(L - \tilde{L})^T\|_F + \|L - \tilde{L}\|_F \|\tilde{L}^T\|_F) \|\tilde{d}y\|_2 \\ &= \|L - \tilde{L}\|_F (\|L\|_F + \|\tilde{L}^T\|_F) \|\tilde{d}y\|_2. \end{aligned}$$

Assim, se $\|L - \tilde{L}\|_F$ e $\|\tilde{d}y\|_2$ forem limitadas, temos que $\|\tilde{r}\|_2 \rightarrow 0$, quando $\|L - \tilde{L}\|_F \rightarrow 0$. Logo, desejamos minimizar a norma de Frobenius de $E = L - \tilde{L}$, o que justifica a escolha da FCC para a construção do fator incompleto na abordagem proposta.

O parâmetro de preenchimento η da FCC pode ser controlado convenientemente, de forma que comecemos com um valor que permita na primeira iteração, uma matriz \tilde{L} o mais esparsa possível. À medida que o MPI avança nas iterações, o valor de η aumenta progressivamente, permitindo um aumento gradual da densidade de \tilde{L} ao longo das iterações, até atingir a fatoração de Cholesky (completa). Ao mesmo tempo espera-se que a convergência do método seja alcançada. Desta forma, a resolução do sistema se torna mais rápida nas iterações iniciais, já que haverá uma economia de tempo tanto no cálculo dos elementos do fator incompleto, quanto na resolução dos dois sistemas triangulares.

4 Experimentos Computacionais

A abordagem proposta foi incorporada ao código PCx [8] que é uma implementação do método primal-dual preditor-corretor. Todas as rotinas deste código são implementadas em Linguagem C, exceto a parte responsável pela solução dos sistemas lineares que é feita aplicando a fatoração de Cholesky no sistema de equações normais, utilizando o código da fatoração de Cholesky esparsa projetada por Ng e Peyton [9], em FORTRAN77. Na implementação do código com a nova abordagem, que chamaremos de código PCx modificado (PCx_{mod}), foi admitida a existência de duas fases. Na primeira fase são calculados fatores de Cholesky incompletos, utilizando um código da FCC implementado em FORTRAN77 que foi fornecido pelo seu autor [4]. Na segunda fase, são calculadas fatorações de Cholesky (completas) utilizando o código de Ng e Peyton. Isto porque, o código da FCC não é eficiente para o cálculo do fator completo, pois necessita de operações adicionais e não explora o fato de que a fatoração final é conhecida.

A mudança para a segunda fase é feita quando o número de elementos não nulos do fator incompleto calculado é 95% do número de elementos não nulos do fator completo ou quando não está sendo obtido progresso com a fatoração controlada de Cholesky. O progresso de uma iteração para outra é medido através do quociente $\rho = \frac{\mu^{k-1}}{\mu^k}$, em que $\mu^k = (x^k)^t z^k / n$ e μ^{k-1} são os gap de dualidade da iteração atual e anterior respectivamente. Em nossa implementação, consideramos que se $\rho \geq 0,99$, nenhum progresso está sendo obtido, assim a mudança para a segunda fase é feita na iteração imediatamente seguinte.

Ainda na primeira fase, uma vez determinado um valor de η inicial, a partir da segunda iteração, seu valor é atualizado ou não, em função de ρ , da seguinte forma:

$$\eta = \begin{cases} \eta, & \text{se } \rho < 0,3 \\ \eta + a\rho, & \text{se } 0,3 \leq \rho \leq 0,7 \\ \eta + b\rho, & \text{se } \rho \geq 0,7 \end{cases}$$

onde a e b são constantes, com $a < b$.

Foram realizados testes computacionais com alguns problemas. A tabela 1 exibe em suas colunas o nome dos problemas testados, o número de linhas, o número de colunas, a densidade do fator de Cholesky da matriz do sistema de equações normais e a Biblioteca a qual pertence cada problema, respectivamente.

O valor inicial de η e o valor dos parâmetros a e b utilizados no processamento de cada problema foram os melhores obtidos após vários testes.

Pela tabela 2, pode-se comparar o número de iterações dos códigos PCx e PCx_{mod} através da segunda e terceira coluna da tabela, respectivamente. Na quarta coluna temos o número de iterações da primeira fase, ou seja, o número de iterações em que se utilizou a FCC. Através das duas últimas colunas da tabela, podemos comparar o tempo de processamento para a resolução dos problemas pelos dois códigos. A última linha da tabela exibe a soma total de cada coluna da tabela.

Na última linha da tabela 2, podemos observar que a redução no tempo total de processamento dos problemas pelo PCx_{mod} compensou o aumento no número de iterações necessárias para resolver os problemas.

Tabela 1: Problemas

Problema	Linhas	Colunas	Densidade Chol	Biblioteca
baxter	27441	30733	0,020	Meszaros
karted	46502	133115	0,035	Meszaros
seymour	4944	6316	0,387	Meszaros
tp-6	142752	1014301	0,002	Meszaros
pds-30	49944	158489	0,013	Kennington
pds-50	83060	275814	0,013	Kennington
pds-70	114944	390005	0,013	Kennington
pds-90	142823	475448	0,011	Kennington
kra30a	18059	85725	0,517	Qaplib
scr20	5079	15980	0,517	Qaplib
128-64-10	9138	47833	0,110	Mnetgen
128-64-12	9088	48389	0,105	Mnetgen
128-128-11	17355	98316	0,039	Mnetgen
512-64-8	34684	189375	0,057	Mnetgen

Tabela 2: Comparação do número de iterações e do tempo

Problema	Iteração			Tempo	
	PCx	PCx _{mod}	FCC	PCx	PCx _{mod}
baxter	32	31	22	47,09	86,58
karted	20	25	11	1087,31	894,07
seymour	19	22	10	48,21	46,79
tp-6	30	35	28	1137,89	2270,30
pds-30	62	66	26	563,14	383,26
pds-50	68	72	30	2760,82	1770,20
pds-70	72	82	39	7425,77	4685,90
pds-90	75	83	45	13348,96	7095,25
kra30a	26	29	9	4420,87	4987,00
scr20	20	22	9	114,52	121,06
128-64-10	26	27	27	51,79	47,15
128-64-12	22	22	22	40,43	33,67
128-128-11	30	29	29	75,58	95,32
512-64-8	25	27	9	1026,90	886,78
Total	527	572	316	32149,28	23355,33

5 Conclusões

Apresentamos neste trabalho uma nova abordagem para a resolução de sistemas lineares oriundos de métodos de pontos interiores. Os resultados computacionais mostraram

que o método apresentado é promissor. Como um trabalho futuro, vemos a necessidade de uma melhora do código implementado, buscando heurísticas que padronizem e determinem os melhores valores de η em cada iteração.

Agradecimentos

Os autores deste trabalho agradecem a Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) e ao CNPq pelo apoio recebido.

Referências

- [1] S. J. Wright, *Primal–Dual Interior–Point Methods*. Philadelphia, PA, USA: SIAM Publications, SIAM, 1996.
- [2] C. T. L. S. Ghidini, A. R. L. Oliveira, and D. C. Sorensen, “Computing a hybrid preconditioner approach to solve the linear systems arising from interior point methods for linear programming using the gradient conjugate method,” *Annals of Management Science*, 2013.
- [3] M. I. Velazco, A. R. L. Oliveira, and F. F. Campos, “Heuristics for implementation of a hybrid preconditioner for interior-point methods,” *Pesquisa Operacional*, vol. 34, no. 9, pp. 2553–2561, 2011.
- [4] F. F. Campos, *Analysis of Conjugate Gradients - type methods for solving linear equations*. PhD thesis, Oxford University Computing Laboratory, Oxford, 1995.
- [5] G. Al-Jeiroudi and J. Gondzio, “Convergence analysis of the inexact infeasible interior-point method for linear optimization,” *Journal of Optimization Theory and Applications*, vol. 141, pp. 231–247, 2009.
- [6] V. Baryamureeba and T. Steihaug, *On the convergence of an inexact primal-dual interior point method for linear programming*. Berlin/Heidelberg: in Lecture Notes in Computer Science, Springer, 2006.
- [7] I. S. Duff and G. A. Meurant, “The effect of ordering on preconditioned conjugate gradients,” *BIT*, vol. 29, no. 4, pp. 635–657, 1989.
- [8] J. Czyzyk, S. Mehrotra, S. J. Wright, and M. Wagner, “Pcx: An interior-point code for linear programming,” *Optimization Methods and Software*, vol. 11(1), pp. 397–430, 1999.
- [9] E. Ng and B. P. Peyton, “Block sparse Cholesky algorithms on advanced uniprocessors computers,” *SIAM Journal on Scientific Stat. Computing*, vol. 14, pp. 1034–1056, 1993.