INF1608 – Análise Numérica
Projeto: Resolução de Sistemas Esparsos

Prof. Waldemar Celes
Departamento de Informática, PUC-Rio

Descrição

O método dos Gradientes Conjugados é um método iterativo para resolução de sistemas lineares na forma $A\mathbf{x} = \mathbf{b}$. O método é largamente empregado na resolução de problemas de engenharia. Uma vantagem do método é que, em cada iteração, usa-se praticamente 4 tipos de operações: multiplicação de matriz por vetor, adição de vetores, multiplicação de escalar por vetor e produto escalar entre vetores. Portanto, cada iteração tem um custo computacional de $O(n^2)$, onde n é a dimensão do sistema. Se o método convergir em k iterações, o custo total é $O(kn^2)$. Se $k \ll n$, o método se torna muito atrativo em relação ao uso de Eliminação de Gauss (que tem complexidade $O(n^3)$).

Além disso, para sistemas esparsos, o custo de cada iteração pode ser reduzido, se fizermos a operação de multiplicação de matriz por vetor processar apenas os elementos não nulos da matriz. Para tanto, devemos armazenar apenas os elementos não nulos da matriz (o que por si só já traz um benefício em uso de memória). Uma estratégia simples é representar a matriz esparsa por um vetor de vetores. Cada elemento desse vetor de vetores tem um ponteiro para o vetor linha da matriz, que armazena apenas os elementos não nulos. Assim, além do valor do elemento da matriz, cada entrada desse vetor deve também armazenar a coluna j do elemento. O acesso a um elemento (i, j) qualquer da matriz pode ser feito da seguinte forma: o acesso ao vetor linha i é feito indexando diretamente o vetor de vetores; o acesso ao elemento j da linha é feito realizando, por exemplo, uma busca binária em j no vetor linha. No entanto, numa operação de multiplicação de matriz por vetor, a busca pela coluna j não é necessária: basta processar cada elemento da linha, multiplicando-o pelo elemento do vetor.

O método dos Gradientes Conjugados na sua forma direta, no entanto, apresenta problemas de instabilidade numérica. Para amenizar o problema, usa-se pré-condicionadores M, transformando o sistema:

$$A\mathbf{x} = \mathbf{b} \quad \longrightarrow \quad M^{-1}A\mathbf{x} = M^{-1}\mathbf{b}$$

O pré-condicionador SSOR (*symmetric successive over-relaxation*) é dado por:

$$\begin{aligned} A &= L + D + U \\ M &= (D + wL)D^{-1}(D + wU), \quad w \in [0, 2] \end{aligned}$$

Com $w = 1$, temos o pré-condicionador Gauss-Seidel; com $w > 1$, temos um condicionador com sobre-relaxamento, que tende a aumentar a taxa de convergência do sistema.

Tarefa

O objetivo deste trabalho é testar a eficiência do método dos Gradientes Conjugados com pré-condicionadores para a resolução de sistemas esparsos. Considere o sistema esparso $A\mathbf{x} = \mathbf{b}$, de dimensão $n \times n$, onde $A(i,i) = i+1$, $A(i,i+1) = A(i+1,i) = 0.5$, $A(i,i+2) = A(i+2,i) = 0.5$ e $A(i,2i) = A(2i,i) = 0.5$, para todo i que corresponde a uma entrada válida de A. A solução do sistema é o vetor $\mathbf{x}_e$ com todos os valores iguais a 1. Faça $\mathbf{b} = A\mathbf{x}_e$ e use o método de Gradientes Conjugados para achar uma solução $\bar{\mathbf{x}}$ dentro de uma certa tolerância. Use o vetor com valores todos iguais a zero como estimativa inicial.

Para valores de n variando entre 100, 1000, 10000, 100000, verifique o desempenho em número de iterações e tempo de processamento para o método convergir, considerando:

- O método sem pré-condicionador: $M = I$
- O método com pré-condicionador de Gauss-Seidel: $M = (D + wL)D^{-1}(D + wU)$, com $w = 1.0$.
- O método com pré-condicionador SSOR: $M = (D + wL)D^{-1}(D + wU)$, com $w > 1.0$.

Análise

Ao desenvolver seu trabalho e testá-lo, procure, baseado em experimentos computacionais, responder às seguintes perguntas:

- Como a matriz esparsa está sendo representada? A representação tira proveito do fato da matriz ser simétrica?
- Como as configurações de pré-condicionamento afetam o desempenho (número de iterações e tempo) do método para a matriz em questão? Faça um comparativo considerando os valores de $n = 100, 1000, 10000, 100000$. Variando valores de $w \in [1.0, 2.0]$, avalie o efeito do sobre-relaxamento.
- Baseado no experimento computacional, qual o custo computacional para resolver esse sistema em relação a n? Está de acordo com o custo teórico esperado?