MAC0300 - Métodos Numéricos para Álgebra Linear - 2015/S2

António Martins Miranda (Nº 7644342) {*amartmiranda@gmail.com*}
António Rui Castro Júnior (Nº 5984327) {*antonio.castro@usp.br*}

EP2 - Métodos iterativos para sistemas lineares: Gradientes Conjugados

1 Método de Gradientes Conjugados

1.1 Descrição

O método dos gradientes conjugados consiste em um algoritmo para solução de sistemas lineares cujas matrizes são simétricas positivas definidas.

Quando A é simétrica positiva definida podemos definir um novo produto interno induzido pela matriz $A : \langle x, y\rangle_A = \langle Ax, y\rangle$, onde $\langle Ax, y\rangle$ denota o produto escalar usual entre os vetores Ax e y. A comutatividade do produto interno induzido por A decorre da simetria da matriz, enquanto que o fato da matriz ser positiva definida garante que $\langle x, x\rangle_A > 0$, se $x \neq 0$.

O método para a solução de uma equação $Ax = b$ parte de uma aproximação inicial qualquer x_0, a partir da qual define-se uma direção inicial:

$d_0 = r_0 = b - Ax_0$

Se denotarmos a solução exata do sistema linear por $\bar{x}$, obtemos a seguinte relação entre o resíduo r_0 e o erro $e_0 = \bar{x} - x_0$:

$r_0 = Ae_0$

Vamos procurar eliminar do erro qualquer componente da direção inicial d_0. Para isso poderíamos em princípio subtrair do erro sua projeção ortogonal na direção d_0 (para o que teríamos que avaliar $\langle e_0, d_0\rangle$). Isto não é viável uma vez que não conhecemos e_0. O problema é no entanto contornável em vez de usarmos o produto interno usual do R^n, optarmos por usar o produto interno induzido pela matriz A, uma vez que $\langle e_0, d_0\rangle_A = \langle Ae_0, d_0\rangle = \langle r_0, d_0\rangle$, que podemos calcular mesmo desconhecendo e_0. Assim obtemos um novo erro e_1, que será ortogonal à direção inicial d_0, dado por

$e_1 = e_0 - \frac{\langle r_0,d_0\rangle}{\langle d_0,d_0\rangle_A} d_0$

Note que a equação anterior equivale a definirmos uma nova aproximação

$x_1 = x_0 + \frac{\langle r_0,d_0\rangle}{\langle d_0,d_0\rangle_A} d_0$

que podemos calcular apenas através da matriz A e da aproximação inicial. A nova aproximação x_1 define um novo resíduo $r_1 = b - Ax_1$, valendo também que $r_1 = Ae_1$. Este novo resíduo forma com a direção inicial d_0 uma base de um plano (espeço de dimensão 2). Vamos agora querer tornar o novo erro e_1 ortogonal a este plano. Para tal vamos primeiramente tornar a base do plano ortogonal (em relação ao produto interno induzido pela matriz A), definindo a direção d_1 ortogonal a d_0, como:

$d_1 = r_1 - \frac{\langle r_1,d_0\rangle_A}{\langle d_0,d_0\rangle_A} d_0$

Para tornarmos o erro e_1 ortogonal ao plano temos apenas que retirar sua projeção ortogonal na direção d_1 (uma vez que ele já é ortogonal a d_0). Assim obteremos:

$e_2 = e_1 - \frac{\langle r_1,d_1\rangle}{\langle d_1,d_1\rangle_A} d_1$

o que equivale a definirmos a nova aproximação

$x_2 = x_1 + \frac{\langle r_1, d_1 \rangle}{\langle d_1, d_1 \rangle_A} d_1$

Esta nova aproximação definirá um novo resíduo que formará junto com as direções anteriores um espaço de dimesão 3. Vamos tornar a base deste espaço ortogonal, retirando do novo resíduo sua projeção ortogonal nas direções anteriores. Aí só teremos que projetar o erro nesta nova direção, somando o resultado à aproximação anterior, obtendo assim um erro ortogonal a este espaço de dimensão 3. O processo continua com a adição de novas direções definidas através dos resíduos(ortogonalizados em relação às direções anteriores), sempre tornando o erro ortogonal a esses espaços de dimensão crescente. O processo terminará no máximo após n etapas, quando teremos o erro ortogonal a um sub-espaço de dimensão n do próprio R^n (ou seja, o sub-espaço será o próprio R^n). Isto só é possível com erro igual a zero, ouseja, se chegarmos à solução exata da equação. Note que todo o processo só se torna possível ao utilizarmos o produto interno induzido pela matriz A, que nos permite calcular a projeção do erro em diversas direções, mesmo sem conhecer este erro.

O algoritmo completo fica então definido como:

- Escolha x_0 e calcule $d_0 = r_0 = b - Ax_0$. Defina $k = 0$.
- Enquanto $k < n$ e $r_k \neq 0$ faça
  - Defina $\alpha_k = \langle r_k, d_k \rangle / \langle d_k, d_k \rangle_A$ ($\langle e_k, d_k \rangle_A / \langle d_k, d_k \rangle_A$)
  - Calcule $x_{k+1} = x_k + \alpha_k d_k$
  - Calcule $r_{k+1} = b - Ax_{k+1} = r_k - \alpha_k A d_k$ (Ad_k já foi calculado)
  - Defina $\beta_k = \langle r_{k+1}, d_k \rangle_A / \langle d_k, d_k \rangle_A$
  - Defina $d_{k+1} = r_{k+1} - \beta_k d_k$
  - Incremente $k = k + 1$
- A solução é dada por x_k!

1.2 Implementação

1.2.1 Método de gradientes conjugados

...

1.2.2 Gerador de matrizes esparsas definidas positivas

...

1.3 Experimentos

...

1.4 Análise de resultados

...