Метод сопряженных градиентов (CG)

(конспект Алексея Черненко, МФТИ)

1 Введение

Решаем систему вида
$$Ax = f, \tag{1}$$
где $A \in \mathcal{R}^{n\times n}$ матрица, при этом предполагается, что $A = A^T > 0$, $x, f \in \mathcal{R}^n$ вектора.

Метод сопряженных градиентов является итерационным алгоритмом Крыловского типа.
Процесс решения методом CG можно рассматривать как минимизацию функционала:
$$J = \frac{1}{2}(Ax, x) - (f, x) \to min \tag{2}$$
используя подпространства Крылова.

2 CG как проекционный метод

Метод CG можно рассматривать как проекционный метод.
Пусть x_0 - начальное приближение к решению системы (1), тогда $r_0 = f - Ax_0$ - начальная невязка. Вводим подпространства Крылова: $\mathcal{K}^m = \{r_0; Ar_0; A^2r_0, ..., A^{m-1}r_0\} \subset \mathcal{R}^n$.
Ищем m-ое приближение решения системы (1) из следующей системы:
$$\begin{cases} x_m = x_0 + \sum\limits_{i=1}^{m} c_i A^{i-1} r_0, \\ r_m = f - Ax_m \perp \mathcal{K}^m; \end{cases} \tag{3}$$

Пусть $x_m = x_0 + \sum\limits_{i=1}^{m} \alpha_i K_i = x_m + \alpha_m K_m$, $\mathcal{K}^m = \{K_1, K_2, ...K_m\}$ - A-ортогональный базис; тогда невязка: $r_m = r_{m-1} - \alpha_m AK_m$. При этом: $r_m \perp \mathcal{K}^m$, отсюда получаем:
$$\alpha_m = \frac{(r_{m-1}, K_m)}{(AK_m, K_m)}; \tag{4}$$

Пусть $K_{m+1} = A^m r_0 + \sum\limits_{i=1}^{m} \gamma_i K_i$, $\mathcal{K}^{m+1} = \{\{r_0, ..., A^{m-1}r_0\},\ A^m r_0\} = \{\{K_1, ..., K_m\},\ K_{m+1}\}$.
$$\begin{cases} r_m = f - Ax_m = f - A(x_0 + \sum\limits_{i=1}^{m} c_i A^{i-1} r_0) = r_0 - \sum\limits_{i=1}^{m} c_i A^i r_0,\ r_m \neq 0,\ c_m \neq 0, \\ K_{m+1} = r_m + \sum\limits_{i=1}^{m} \beta_i K_i; \end{cases} \tag{5}$$
$\{\beta_i\}$ ищем из условия A-ортогональности (или сопряженности):
$$K_{m+1} \perp_A \mathcal{K}^m \longleftrightarrow (K_{m+1}, K_i)_A = 0, \overline{1, m} \tag{6}$$
Отсюда получаем, что $\beta_i = 0, i < m$ и

$$\begin{cases} K_{m+1} = r_m + \beta_m K_m, \\ \beta_m = -\dfrac{(r_m, AK_m)}{(AK_m, K_m)}; \end{cases} \tag{7}$$

Подставляя (7) в (4), получаем итоговое выражение для α_m:

$$\alpha_m = \frac{(r_{m-1}, K_m)}{(AK_m, K_m)} = \frac{(r_{m-1}, r_{m-1})}{(AK_m, K_m)} \tag{8}$$

Подставляя (8) в (7) и используя, что $(r_m, AK_m) = (r_m, \dfrac{r_{m-1} - r_m}{\alpha_m}) = -\dfrac{1}{\alpha_m}(r_m, r_m)$, получаем итоговое выражение для β_m:

$$\beta_m = \frac{(r_m, r_m)}{(r_{m-1}, r_{m-1})} \tag{9}$$

3 Алгоритм CG

В предыдущем пункте получены все необходимые формулы для того, чтобы записать метод CG.

Алгоритм 1 Алгоритм CG

```
задаем начальное приближение x_0
r_0 = f − A x_0
K_1 = r_0
m = 1
цикл пока «не выполнен критерий остановки» выполним
    α_m = (r_{m−1}, r_{m−1}) / (A K_m, K_m)
    x_m = x_{m−1} + α_m K_m
    r_m = r_{m−1} − α_m A K_m
    β_m = (r_m, r_m) / (r_{m−1}, r_{m−1})
    K_{m+1} = r_m + β_m K_m
    m = m + 1
завершим цикл пока
```

Данный итерационный процесс должен сходиться за n итераций, но их может потребоваться больше из-за погрешности в представлении чисел. Выход: в качестве критерия остановки использовать условие малости относительной невязки: $\dfrac{||r_m||}{||f||} < \varepsilon$, где величина ε задается на входе в алгоритм.

4 Алгоритм CG с предоубославливателем

Пусть $Sp(A) \in [\lambda,\ \Lambda]$. Если матрица A плохо обусленная: $\kappa(A)_2 = \dfrac{\Lambda}{\lambda} \gg 1$, то метод CG сходится плохо, поэтому вводится предобуславливатель B, такой, чтобы спектр матрицы $(B^{-1}A)$ был достаточно узким. В этом случае алгорим принимает следующий вид:

Алгоритм 2 Алгоритм CG с предобуславливателем

задаем начальное приближение x_0
$r_0 = f - Ax_0$
$K_1 = B^{-1} r_0$
m = 1
цикл пока «не выполнен критерий остановки» **выполним**
$$\alpha_m = \frac{(r_{m-1}, B^{-1} r_{m-1})}{(AK_m, K_m)}$$
$x_m = x_{m-1} + \alpha_m K_m$
$r_m = r_{m-1} - \alpha_m A K_m$
$$\beta_m = \frac{(r_m, B^{-1} r_m)}{(r_{m-1}, B^{-1} r_{m-1})}$$
$K_{m+1} = B^{-1} r_m + \beta_m K_m$
m = m + 1
завершим цикл пока