8 Метод на спрегнатия градиент. Оценка на скоростта на сходимост.

8.1 Общи теореми за сходимост

Теорема 8.1 *При реализация на метода на спрегнатия градиент съществува $m \leq N$, такова че $x^m = \hat{x}$, където N е размерността на матрицата H.*

Доказателство. Нека допуснем противното. Тогава, $g^k \neq 0$ за $k = 0, 1, \ldots, N$, и следователно в N-мерното Евклидово пространство съществуват $N + 1$ взаимно ортогонални вектора, което е невъзможно. С това теоремата е доказана. ■

Следователно, в метода на спрегнатия градиент след краен брой итерации получаваме точното решение. В този смисъл, методът може да се разглежда, като точен. Тук е необходимо да направим две важни бележки. При компютърната реализация на метода, в резултат на грешките от закръгляне, може да не получим точно решение и за $k > m$. В същото време, особено важен е въпросът, каква е точността на метода (в частност при прилагане на преобуславяне) след изпълнение на итерации, които са значително по-малко от N.

Конструкцията на метода е такава, че резидуалите g^k удовлетворяват условието

$$||g^k||_{H^{-1}} = \min_{g \in T_k} ||g||_{H^{-1}},$$

където елементите на T_k имат вида

$$g = g^0 + \sum_{l=1}^{k} \alpha_l H^l g^0.$$

Нека означим с Π_k^1 полиномите P_k от степен k със свободен член единица, т.е. за които $P_k(0) = 1$. Ще разглеждаме както полиноми, в които независимата променлива е скалар, така и полиноми от матрици с размерност N. Лесно се вижда, че множеството T_k, което използвахме при конструиране на метода на спрегнатия градиент, може да се представи във вида

$$T_k = \{g \in \mathbf{R}^N : g = P_k(H)g^0, P_k \in \Pi_k^1\}.$$

Тогава

$$\begin{aligned} ||g^k||_{H^{-1}} &= \min_{g \in T_k} ||g||_{H^{-1}} = \min_{P_k \in \Pi_k^1} ||P_k(H)g^0||_{H^{-1}} \\ &= \min_{P_k \in \Pi_k^1} [g^{0^T} P_k(H) H^{-1} P_k(H) g^0]^{1/2} \\ &= \min_{P_k \in \Pi_k^1} [g^{0^T} H^{-1} P_k^2(H) g^0]^{1/2}. \end{aligned} \tag{8.1}$$

Тук е използвано равенството $P_k(H)^T H^{-1} P_k(H) = H^{-1} P_k(H)^2$, което се проверява директно.

Следващата теорема има фундаментално значение при анализ на сходимостта на метода на спрегнатия градиент.

Теорема 8.2 *Нека S е множество, което съдържа всички собствени числа на H и нека $M \geq 0$ е такова, че за подходящо избран полином $\tilde{P}_k(\lambda) \in \Pi_k^1$ е в сила неравенството*

$$\max_{\lambda \in S} |\tilde{P}_k(\lambda)| \leq M.$$

Тогава,

$$||x^k - \hat{x}||_H \leq M ||x^0 - \hat{x}||_H. \tag{8.2}$$

Доказателство. Нека собствените числа и собствените вектори на H са $\{\lambda_i, v_i\}_{i=1}^N$, като $0 < \lambda_1 \leq \lambda_2 \leq \cdots \lambda_N$. Без ограничения върху общността на доказателството можем да предположим, че ортогоналните собствени вектори са и ортонормирани, т.е. $v_i^T v_j = \delta_{ij}$. Развиваме началния резидуал по базиса от собствените функции. Така получаваме

$$g^0 = \sum_{i=1}^N \alpha_i v_i, \qquad \alpha_i = v_i^T g^0,$$

от където непосредствено получаваме, че

$$g^{0^T} H^{-1} P_k^2(H) g^0 = \sum_{i=1}^N \alpha_i^2 \lambda_i^{-1} P_k(\lambda_i)^2.$$

Следователно (виж (8.1))

$$||g^k||^2_{H^{-1}} = \min_{P_k \in \Pi^1_k} \sum_{i=1}^{N} \alpha_i^2 \lambda_i^{-1} P_k(\lambda_i)^2.$$

От условието на теоремата следва, че $|\tilde{P}_k(\lambda_i)| \leq M$. Така от последното равенство получаваме оценката

$$||g^k||^2_{H^{-1}} \leq \sum_{i=1}^{N} \alpha_i^2 \lambda_i^{-1} \tilde{P}_k(\lambda_i)^2 \leq M^2 \sum_{i=1}^{N} \alpha_i^2 \lambda_i^{-1},$$

от където и от равенството $||x^k - \hat{x}||_H = ||g^k||_{H^{-1}}$ следва твърдение (8.2) на теоремата. ∎

За да приложим резултата от доказаната теорема е необходимо по подходящ начин да определим множество S съдържащо собствените числа, както и да подберем полином $\tilde{P}_k$, за който константата M е колкото може по-малка.

8.2 Оценка на скоростта на сходимост с помощта на спектралното число на обусловеност.

Нека допуснем, че не ни е известна допълнителна информация за разпределението на собствените числа на матрицата H. В този случай най-естествено е да изберем $S = [\lambda_1, \lambda_N]$. Тогава, оптималният полином $\tilde{P}_k$ удовлетворява условието:

$$\max_{\lambda_1 \leq \lambda \leq \lambda_N} |\tilde{P}_k(\lambda)| = \min_{P_k \in \Pi^1_k} \max_{\lambda_1 \leq \lambda \leq \lambda_N} |P_k(\lambda)|. \tag{8.3}$$

Решението на тази оптимизационна задача е известно и то се получава с помощта на полиномите на Чебишев $T_k(x) \in P_k$, които имат вида

$$T_k(x) = \frac{1}{2}\left[\left(x + \sqrt{x^2-1}\right)^k + \left(x - \sqrt{x^2-1}\right)^k\right].$$

Лема 8.1

$$\max_{-1 \leq x \leq 1} |T_k(x)| = 1 \tag{8.4}$$

Д о к а з а т е л с т в о. Твърдението следва пряко от равенството

$$T_k(x) = \cos[k \cos^{-1}(x)], \qquad -1 \leq x \leq 1.$$

∎

Нека изберем

$$\tilde{P}_k(\lambda) = \frac{T_k[(\lambda_N + \lambda_1 - 2\lambda)/(\lambda_N - \lambda_1)]}{T_k[(\lambda_N + \lambda_1)/(\lambda_N - \lambda_1)]}. \tag{8.5}$$

Тогава, в сила е следната теорема.

Теорема 8.3 *В сила е оценката*

$$||x^k - \hat{x}||_H \leq \frac{1}{T_k[(\lambda_N + \lambda_1)/(\lambda_N - \lambda_1)]} ||x^0 - \hat{x}||_H. \tag{8.6}$$

Д о к а з а т е л с т в о. Твърдението следва директно от (8.2) и (8.4). ∎

Основният резултат в този въпрос се съдържа в следващата теорема.

Теорема 8.4

$$p(\epsilon) \leq \frac{1}{2}\sqrt{\kappa(H)} \ln(2/\epsilon) + 1, \tag{8.7}$$

където с $p(\epsilon)$ е означено най-малкото цяло положително число k, за което

$$||x^k - \hat{x}||_H \leq \epsilon ||x^0 - \hat{x}||_H \qquad \forall x^0 \in \mathbf{R}^N$$

Доказателство. Нека означим с $\alpha = \dfrac{\lambda_N}{\lambda_1} = \kappa(H) > 1$. Тогава

$$T_k\left(\frac{\lambda_N+\lambda_1}{\lambda_N-\lambda_1}\right) = T_k\left(\frac{\alpha+1}{\alpha-1}\right) =$$

$$\frac{1}{2}\left[\left(\frac{\alpha+1}{\alpha-1}+\sqrt{\left(\frac{\alpha+1}{\alpha-1}\right)^2-1}\right)^k+\left(\frac{\alpha+1}{\alpha-1}-\sqrt{\left(\frac{\alpha+1}{\alpha-1}\right)^2-1}\right)^k\right] =$$

$$\frac{1}{2}\left[\left(\frac{\alpha+1+2\sqrt{\alpha}}{\alpha-1}\right)^k+\left(\frac{\alpha+1-2\sqrt{\alpha}}{\alpha-1}\right)^k\right] =$$

$$\frac{1}{2}\left[\left(\frac{(\sqrt{\alpha}+1)^2}{\alpha-1}\right)^k+\left(\frac{(\sqrt{\alpha}-1)^2}{\alpha-1}\right)^k\right] =$$

$$\frac{1}{2}\left[\left(\frac{\sqrt{\alpha}-1}{\sqrt{\alpha}+1}\right)^k+\left(\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}\right)^k\right] >$$

$$\frac{1}{2}\left(\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}\right)^k,$$

където използвахме, че

$$\left(\frac{\alpha+1}{\alpha-1}\right)^2-1 = \frac{4\alpha}{(\alpha-1)^2}.$$

От горната оценка следва, че

$$\frac{||x^k-\hat{x}||_H}{||x^0-\hat{x}||_H} < \frac{2}{\left(\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}\right)^k} \tag{8.8}$$

За да докажем теоремата е необходимо да отговорим на въпроса, колко голямо трябва да е k, за да е в сила неравенството

$$\frac{2}{\left(\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}\right)^k} < \epsilon \tag{8.9}$$

Логаритмуваме (8.9) и получаваме условието

$$k > \ln\frac{2}{\epsilon}\left[\ln\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}\right]^{-1}. \tag{8.10}$$

За да оценим втория логаритмичен множител в дясната страна на (8.10) ще разгледаме спомагателната функция

$$\varphi = \ln \frac{x+1}{x-1} - \frac{2}{x}, \qquad x > 1.$$

Тъй като

$$\varphi'(x) = -\frac{2}{x^2(x^2-1)} < 0,$$

то функцията е строго намаляваща и следователно

$$\varphi(x) > \lim_{x\to\infty} \varphi(x) = 0,$$

от където получаваме неравенството

$$\ln \frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1} > \frac{2}{\sqrt{\alpha}}. \tag{8.11}$$

От (8.10) и (8.11) следва, че ако

$$k > \frac{1}{2}\sqrt{\alpha} \ln \frac{2}{\epsilon} > \ln \frac{2}{\epsilon} \left[\ln \frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1} \right]^{-1}$$

то

$$\frac{2}{\left(\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}\right)^k} < \epsilon$$

с което теоремата е доказана. ■

Забележка. От доказаната теорема следва, че броят на итерациите в метода на спрегнатия градиент за решаване на линейни системи от вида $A\hat{x} = b$, получени при дискретизация на елиптични гранични задачи в двумерни области е $n_{it}^{CG} = O\left(\sqrt{\kappa(A)}\right) = O\left(N^{1/2}\right)$ и следователно общата изчислителна сложност за решаване на системата е $\mathcal{N}^{CG}(A^{-1}b) = O\left(N^{3/2}\right)$. Нека си припомним, че точно такава е асимптотичната сложност на метода на вложените сечения, което е най-добрият известен резултат за пряк метод.

В заключение ще разгледаме една модификация на оценката за скоростта на сходимост на метода на спрегнатия градиент, която подобрява резултата в случая, когато няколко от максималните собствени числа на матрицата H са добре изолирани от останалите.

Теорема 8.5 *Нека m и b са такива, че $1 < m < N$ и $\lambda_{N-m} \leq b < \lambda_{N-m+1}$. Тогава,*

$$p(\epsilon) \leq \frac{1}{2}\sqrt{b/\lambda_1}\ln(2/\epsilon) + m + 1. \tag{8.12}$$

Тази оценка има важно значение, когато m и $(b-\lambda_1)$ са „малки“. При доказателството се избират $S = S_1 \bigcup S_2$, където $S_1 = [\lambda_1, b]$ и $S_2 = \bigcup_{i=N-m+1}^{N} \lambda_i$, както и полинома

$$\tilde{P}_k(\lambda) = \left[\prod_{i=N-m+1}^{N}\left(1 - \frac{\lambda}{\lambda_i}\right)\right] \frac{T_{k-m}\left[(b+\lambda_1-2\lambda)/(b-\lambda_1)\right]}{T_{k-m}\left[(b+\lambda_1)/(b-\lambda_1)\right]}$$

който очевидно е от Π_k^1.