Лекція 13. Метод найшвидшого спуску

Складність, пов‘язана із методом послідовної верхньої релаксації, чебишовськими процедурами прискорення і такого є типу методами, полягає в тому, що вони залежать від параметрів, правильний вибір яких іноді буває важким. Наприклад, для того щоб чебишовські процедури прискорення були вдалими, потрібні добрі оцінки для найбільшого і найменшого власних значень відповідної матриці переходу G. Якщо ця матриця не улаштована особливим чином, то отриманнЯ їх в аналітичному вигляді, найімовірніше, неможливе, а обчислення – дорого.

На щастя, існує метод, в якому немає цієї трудності, – добре відомий метод спряжених градієнтів Хестенса-Штифеля (Hestenes M.R., Stiefel E. (1952)). Метод спряжених гражієнтів (CG) хоча і є ітераційним, але збігається до істинного розв‘язку лінійної системи за скінченне число ітерацій за відсутності похибок округлення. Вивід цього методу почнемо з методу найшвидшого спуску.

Найшвидший спуск

Почнемо з того, що обговоримо, як може відбуватися рискання при мінімізації функціонала $\varphi(x)$, що має вигляд

$$\varphi(x) = \frac{1}{2}(Ax, x) - (b, x),$$

де $b \in R^n$ і матриця $A \in R^{n\times n}$ вважається додатно визначеною і симетричною.

Лема 1. Мінімальне значення φ дорівнює $-\frac{1}{2}(A^{-1}b, b)$ і досягається при $x = A^{-1}b$. Таким чином, мінімізація φ і розв‘язування системи $Ax = b$ – еквівалентні задачі.

Доведення. Запишемо функціонал $\varphi(x)$ як квадратичну форму

$$\varphi(x) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j - \sum_{i=1}^{n} b_i x_i .$$

Необхідною і достатньою умовою мінімума цієї квадратичної форми є рівність її градієнта нулю:

$$grad\,\varphi(x) = 0.$$

Отже, маємо

$$grad\,\varphi(x) = \left\{\frac{\partial\varphi}{\partial x_i}\right\}_{i=1}^{n} = \left\{\frac{1}{2}\frac{\partial}{\partial x_i}\left(\sum_{i=1}^{n}\sum_{j\neq i}^{n} a_{ij}x_i x_j + \sum_{i=1}^{n} a_{ii}x_i^2\right) - b_i\right\}_{i=1}^{n} =$$

$$= \left|\begin{matrix} a_{ij} = a_{ji} \\ a_{ij}x_j + a_{ji}x_j = 2a_{ij}x_j \end{matrix}\right| = \left\{\sum_{j\neq i}^{n} a_{ij}x_j + \sum_{i=1}^{n} a_{ii}x_i - b_i\right\}_{i=1}^{n} = \left\{\sum_{i=1}^{n} a_{ij}x_j - b_i\right\}_{i=1}^{n} = Ax - b = 0$$

Таким чином, мінімізація функціонала еквівалентна розв‘язуванню СЛАУ $Ax = b$, і досягається мінімум в точці $x = A^{-1}b$.

Однією із найпростіших стратегій мінімізації φ є метод найшвидшого спуску. В поточній точці x_c функція φ убуває найшвидше в напрямку антіградієнта:

$$-grad\,\varphi(x_c) = b - Ax_c\,.$$

Означення 1. Назвемо $r_c = b - Ax_c$ *відхилом* вектора x_c.

Якщо відхил ненульовий, то $\varphi(x_c + \alpha r_c) < \varphi(x_c)$ для деякого додатного α.

Лема 2. В методі найшвидшого спуску (з точною мінімізацією на прямій) вибираємо

$$\alpha = \frac{(r_c, r_c)}{(Ar_c, r_c)},$$

яке лоставляє мінімум для $\varphi(x_c + \alpha r_c)$.

Доведення. Запишемо

$$\varphi(x_c + \alpha r_c) = \frac{1}{2}\left(A(x_c + \alpha r_c), x_c + \alpha r_c\right) - (b, x_c + \alpha r_c) =$$

$$= \frac{1}{2}(Ax_c, x_c) + \alpha(Ax_c, r_c) + \frac{1}{2}\alpha^2(Ar_c, r_c) - \alpha(b, r_c) - (b, x_c) =$$

$$= \frac{1}{2}\alpha^2(Ar_c, r_c) + \alpha(Ax_c - b, r_c) + \frac{1}{2}(Ax_c - 2b, x_c).$$

Оскільки припускається, що матриця A є додатно визначеною, то $(Ar_c, r_c) > 0$ і квадратний трьохчлен $q(\alpha) = \varphi(x_c + \alpha r_c)$ досягає мінімума при $q'(\alpha) = 0$, тобто при

$$\alpha = \frac{(b - Ax_c, r_c)}{(Ar_c, r_c)} = \frac{(r_c, r_c)}{(Ar_c, r_c)}.$$

Отже отримуємо алгоритм

$k = 0;\ x_0 = 0;\ r_0 = b;$

while $r_k \neq 0$

$k = k + 1;$

$$\alpha_k = \frac{(r_{k-1}, r_{k-1})}{(Ar_k, r_k)}; \qquad (1)$$

$x_k = x_{k-1} + \alpha r_{k-1};$

$r_k = b - Ax_k;$

end

Вибір початкового вектора у вигляді $x_0 = 0$ не є обмеженням. Якщо $\tilde{x}$ вважається кращим посатковим наближенням, то ми просто застосуємо (1), замінивши b на $b - A\tilde{x}$. В такому випадку шуканий розв‘язок апроксимується векторами $x_k + \tilde{x}$.

Теорема (про збіжність методу найшвидщого спуску). Наближення методу найшвидшого спуску задовольняють співвідношення

$$\varphi_0(x_k) \le \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^{2k} \varphi_0(x_0),$$

де $\varphi_0(x) = (A(x - \bar{x}), x - \bar{x})$, $\lambda_{\min} \le \lambda_i(A) \le \lambda_{\max}$ – мінімальне і максимальне власне число матриці A, відповідно.

Доведення. Покладемо $y_k = x_k$ і зробимо одну ітерацію допоміжного оптимального однокрокового процесу

$$y_{k+1} = y_k - \frac{2}{\lambda_{\max} + \lambda_{\min}}(Ay_k - b).$$

Похибки ітерацій $r_k = y_k - \bar{x}$ пов‘язані співвідношенням

$$r_{k+1} = \left(I - \frac{2}{\lambda_{\max} + \lambda_{\min}} A\right) r_k .$$

Нехай $\{e_i\}_{i=1}^n$ – ортонормована система власних векторів матриці A: $Ae_i = \lambda_i e_i$,

$$(e_i, e_j) = \begin{cases} 1, & \textit{якщо } i = j, \\ 0, & \textit{якщо } i \ne j. \end{cases}$$

Оскільки $\lambda_{\min} \le \lambda_i(A) \le \lambda_{\max}$, то для всіх індексів i виконуються співвідношення

$$-\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \le 1 - \frac{2}{\lambda_{\max} + \lambda_{\min}} \lambda_i \le \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} .$$

Це випливає із таких співвідношень:

$$\lambda_{\min} \le \lambda_i \le \lambda_{\max} \Rightarrow -\lambda_{\max} \le -\lambda_i \le -\lambda_{\min} \Rightarrow$$

$$\Rightarrow -\frac{2}{\lambda_{\max} + \lambda_{\min}} \le -\frac{2}{\lambda_{\max} + \lambda_{\min}} \lambda_i \le -\frac{2}{\lambda_{\max} + \lambda_{\min}} \Rightarrow$$

$$\Rightarrow 1 - \frac{2}{\lambda_{\max} + \lambda_{\min}} \le 1 - \frac{2}{\lambda_{\max} + \lambda_{\min}} \lambda_i \le 1 - \frac{2}{\lambda_{\max} + \lambda_{\min}} \Rightarrow$$

$$\Rightarrow -\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \le 1 - \frac{2}{\lambda_{\max} + \lambda_{\min}} \lambda_i \le \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} .$$

Отже,

$$\left| 1 - \frac{2}{\lambda_{\max} + \lambda_{\min}} \lambda_i \right| \le \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} . \tag{2}$$

Нехай $r_k = \sum_{i=1}^n c_i e_i$. Тоді мають місце співвідношення

$$(Ar_k, r_k) = \left(\sum_{i=1}^n c_i \lambda_i e_i, \sum_{i=1}^n c_i e_i\right) = \sum_{i=1}^n \lambda_i c_i^2,$$

$$r_{k+1} = \sum_{i=1}^n d_i e_i \text{ , де } d_i = \left(1 - \frac{2}{\lambda_{\max} + \lambda_{\min}}\right) c_i,$$

$$\left(Ar_{k+1}, r_{k+1}\right) \le \sum_{i=1}^{n} \lambda_i d_i^2 = \sum_{i=1}^{n} \lambda_i \left(1 - \frac{2\lambda_i}{\lambda_{\max} + \lambda_{\min}}\right)^2 c_i^2 .$$

Із урахуванням (2) отримуємо

$$\left(Ar_{k+1}, r_{k+1}\right) \le \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^2 \left(Ar_k, r_k\right).$$

Оскільки $\varphi_0\left(y_k\right) = \left(Ar_k, r_k\right)$, то це означає, що

$$\varphi_0\left(y_{k+1}\right) \le \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^2 \varphi_0\left(y_k\right) = \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^2 \varphi_0\left(x_k\right).$$

Тепер наближення y_{k+1} запишемо в формі, подібній до методу найближчого спуску

$$y_{k+1} = x_k + \alpha\, grad\, \varphi\left(x_k\right), \quad \alpha = \frac{1}{\lambda_{\max} + \lambda_{\min}}, \tag{3}$$

де $\varphi\left(x_k\right) = \left(Ax_k, x_k\right) - 2\left(b, x_k\right)$.

Оскільки на x_{k+1}, що отримується за методом найшвидшого спуску, досягається мінімум $\varphi(x)$ серед усіх наближень вигляду (3), то $\varphi\left(x_{k+1}\right) \le \varphi\left(y_{k+1}\right)$. Звідси випливає оцінка

$$\varphi_0\left(x_{k+1}\right) \le \varphi_0\left(y_{k+1}\right) \le \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^2 \varphi_0\left(x_k\right) \le \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^{2k} \varphi_0\left(x_0\right).$$

Зауваження 1. Із теореми і оцінок

$$\lambda_{\min} \left\| y - \bar{x} \right\|_2^2 \le \varphi_0\left(y\right) \le \lambda_{\max} \left\| y - \bar{x} \right\|_2^2$$

випливає, що

$$\left\| x_k - \bar{x} \right\|_2 \le \sqrt{\frac{\varphi_0\left(x_k\right)}{\lambda_{\min}}} \le \left(\sqrt{\varphi_{0,\max}}\right)^k \sqrt{\frac{\varphi_0\left(x_0\right)}{\lambda_{\min}}} \le \left(\sqrt{\varphi_{0,\max}}\right)^k \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \left\| x_0 - \bar{x} \right\|_2$$

$$\left\| x_k - \bar{x} \right\|_2 \le \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^k \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \left\| x_0 - \bar{x} \right\|_2 .$$

Зауваження 2. Наближення методу найшвидшого спуску задовольняють співвідношення

$$\left\| x_k - \bar{x} \right\|_2 \le \left(1 - \frac{\lambda_{\min}}{\lambda_{\max}}\right)^k \left\| x_0 - \bar{x} \right\|_2 .$$