Лекція №7 – Метод регуляризаії для матричного рівняння

В основі побудови стійких методів розв'язку некоректних задач лежить поняття регуляризуючого алгоритму (РА) і пов'язаного з ним поняття регуляризованого сімейства розв'язків, введеного А.М. Тихоновим. Погано обумовлені СЛАР слід розглядати як некоректно поставлені задачі і при їх наближеному розв'язанні необхідно застосовувати ідеї регуляризації.

Домовимося розрізняти точні дані – пару $\{A, b\}$, котрі формують задачу (1.7) і нам не відомі, і наближені дані $\{A_h, b_\delta\}$, $\|A_h - A\| \le h$, $\|b_\delta - b\| \le \delta$ з рівнем похибок h, δ, якими ми володіємо. Суть методу регуляризації наближеного розв'язання стосується побудови послідовності векторів $x_{h,\delta}$, яка збігається до розв'язку або псевдорозв'язку рівняння (1.7) при $\delta \vee h \to 0$. За наближений розв'язок $x_{h,\delta}$ не можна брати точний розв'язок або квазірозв'язок рівняння з даними

$$A_h \cdot x = b_\delta \tag{1.28}$$

Нехай ми маємо спосіб (правило), який по парі $\{A_h, b_\delta\}$ і додатньому параметру α однозначно будує вектор $x^\alpha(A_h; b_\delta)$. Якщо існує залежність параметра $\alpha(\delta, h)$ від похибок δ, h вихідних даних така, що

$$\lim_{\substack{\delta \to 0 \\ h \to 0}} \left\| x^{\alpha(\delta, h)}(A_h; b_\delta) - \tilde{x} \right\| = 0, \tag{1.29}$$

тоді множину $\left\{ x^{\alpha(\delta, h)}(A_h; b_\delta) \right\}$ називають регуляризованим сімейством наближених розв'язків, а сам спосіб побудови $x^\alpha(A_h; b_\delta)$ – регуляризуючим алгоритмом для задачі (1.7). Тут вектор $\tilde{x}$ – розв'язок, нормальний розв'язок або псевдорозв'язок системи (1.7) залежно від того, чи розв'язується ця система однозначно, чи має множину розв'язків чи немає.

Співвідношення (1.29) засвідчує, що наближений розв'язок $x^\alpha(A_h; b_\delta)$ тим краще апроксимує точний розв'язок $\tilde{x}$, чим менша похибка вихідних даних δ, h. Таким чином, регуляризуючий алгоритм дає теоретичну базу для конструювання стійкого до збурень вихідних даних наближеного розв'язку системи (1.7) загального виду, включаючи погано обумовлені системи.

Важливо розуміти наступну обставину. Якщо A^{-1} існує, тобто A – невироджена матриця, то для достатньо малих h A_h^{-1} теж існує і розв'язок $x_{h,\delta}$ рівняння (1.28) теоретично буде збігатися до розв'язку рівняння (1.7) при $\delta, h \to 0$. Однак якщо A – погано обумовлена матриця ($\mu(A)$ – велике, $\mu(A)$ – число обумовленості [13, с.37]), то величина похибки $\|\tilde{x} - x_{\delta,h}\|$, навіть при малих δ, h, може бути недопустимо великою і задачу слід вважати практично нестійкою (некоректною). Метод регуляризації саме і направлений на те, щоб зменшити вплив похибок (вхідних даних, обчислень) і одержати практично стійкий наближений розв’язок в цих несприятливих обставинах.

Тепер перейдемо до опису конкретних процедур побудови регуляризованих наближених розв’язків $x^{\alpha}\left(A_h;b_{\delta}\right)$. Далі для скорочення запису будемо опускати залежність $x^{\alpha}\left(A_h;b_{\delta}\right)$ від $\left(A_h;b_{\delta}\right)$ і записувати просто $x^{\alpha(\delta,h)}$.

Розглянемо спочатку частинний випадок – схему М.М. Лаврент’єва, коли A – симетрична додатня напіввизначена матриця, для якої система (1.7) при заданому векторі b може бути розв’язаною.

Перейдемо від (1.6) до регуляризованої системи

$$(A+\alpha E)x^{\alpha}=b+\alpha x^{0}, \tag{1.30}$$

де α – додатній параметр, E – одинична матриця, x^0 – пробний розв’язок, тобто деяке наближення до шуканого розв’язку (якщо інформація про розв’язок відсутня, то можна позначити $x^0=0$).

При зроблених припущеннях, СЛАР (1.30) має єдиний розв’язок x^{α}, який збігається при $\alpha\to 0$ до нормального розв’язку $\tilde{x}$.

Твердження 1.1. Нехай $\left\{A_h,b_{\delta}\right\}$, $\|A_h-A\|\le h$, $\|b_{\delta}-b\|\le\delta$ – наближені дані задачі і A_h – симетрична додатна напіввизначена матриця. Тоді СЛАР

$$(A_h+\alpha E)x^{\alpha}=b_{\delta}+\alpha x^{0} \tag{1.31}$$

однозначно розв’язана і при зв’язку параметра α з похибками δ,h такими, що $\alpha(\delta,h)\to 0$, $\dfrac{h+\delta}{\alpha(\delta,h)}\to 0$, коли $\delta\to 0$, $h\to 0$, тобто $x^{\alpha(\delta,h)}$ збігається до нормального розв’язку $\tilde{x}$ рівняння (1.7). Це і є розв’язок, який найменше ухиляється від вектору x^0.

Таким чином, згідно з означенням, наведеним вище, розв’язки СЛАР (1.31) $\left\{x^{\alpha(\delta,h)}\right\}$ утворюють регуляризоване сімейство наближених розв’язків для системи (1.7); причому, вибір параметра за формулою $\alpha=\sqrt[p]{\delta+h}$ $(p>1)$ задовольняє необхідним вимогам, оскільки $\alpha=\sqrt[p]{\delta+h}\to 0$, $\dfrac{\delta+h}{\sqrt[p]{\delta+h}}=(\delta+h)^{1-\frac{1}{p}}$, коли $\delta,h\to 0$.

Зауваження 1.1. Нехай A – додатня напіввизначена вироджена матриця і $\|A\|=1$. Якщо $\mu(A)=\infty$, то $\mu(A+\alpha E)\le\dfrac{1+\alpha}{\alpha}$. З цієї причини при розумному виборі параметра α можна досягти гарної обумовленості систем (1.30), (1.31) і задовільної апроксимації $x^{\alpha}\approx\tilde{x}$, не зважаючи на те, що ці вимоги мають протиріччя.

Роль параметра регуляризації α добре видно, якщо записати розв’язок системи ($\mu(A)$) (при $x^0=0$) у вигляді

$$x^{\alpha}=\sum_{i=1}^{n}\frac{b_i}{\lambda_i+\alpha}\cdot u_i\,,$$

де λ_i власні значення $(\lambda_i \geq 0)$, а u_i ортонормовані власні вектори матриці A. Це зображення показує, що при малих λ_i додавання додатнього параметра суттєво збільшує знаменник і тим самим послаблює вплив можливих похибок у відповідних компонентах b_i $\left(\tilde{b}_i = b_i + \Delta b_i\right)$. Одночасно для $\lambda_i \geq 0$ вплив малого параметра α незначний (і навіть такий, що ним можна знехтувати).

Тепер відмовимося від вимоги симетричності і додатньості матриці A. Нехай матриця B така, що для деякого α_0 $(A + \alpha_0 B)$ є невиродженою матрицею і, значить, існує й обернена матриця. Тоді можлива регуляризація в наступній формі:

$$(A + \alpha B)x^\alpha = b, \tag{1.32}$$

де параметр α довільного знака і $|\alpha| \leq |\alpha_0|$.

Зауваження 1.2. Нехай $\left\|(A + \alpha B)^{-1} A\right\| \leq c < \infty$ (при $\alpha \to 0$), $\|A_h - A\| \leq h$, $\|B_\mu - B\| \leq \mu$, $\|b_\delta - b\| \leq \delta\|b\|$. Тоді при достатньо малих h, μ, δ СЛАР

$$\left(A_h + \alpha B_\mu\right)x^\alpha = b_\delta \tag{1.33}$$

має єдиний розв'язок x^α і справедлива оцінка похибки

$$\left\|x^\alpha - \tilde{x}_B\right\| \leq c \cdot \frac{|\alpha| + \mu + \delta + h}{|\alpha|}, \tag{1.34}$$

де $\tilde{x}_B$ – розв'язок системи (1.7), що задовольняє умові ($\overline{X}$ – множина розв'язків системи (1.7))

$$\|B\tilde{x}\| = \min\left\{\|B_x\| : x \in \overline{X}\right\}.$$

Із оцінки (1.34) одразу випливає, що якщо $\alpha(\delta, h, \mu) \to 0$, $\dfrac{\delta + h}{|\alpha(\delta, h, \mu)|} \to 0$ при $\delta, h, \mu \to 0$, має місце збіжність δ, h, μ

$$\lim_{\delta, h, \mu \to 0} \left\|x^\alpha - \tilde{x}_B\right\| = 0.$$

Найбільш важливим моментом в описаній регуляризації є підбір матриці B, для якої $(A + \alpha_0 B)$ невироджена і $\left\|(A + \alpha B)^{-1} A\right\| < \infty$.

Дослідимо, врешті, загальну ситуацію, коли система (1.7), взагалі кажучи, нерозв'язна. В цьому випадку шуканим є псевдорозв'язок. Розв'язується задача стійкої апроксимації цього псевдорозв'язку в умовах задання вхідних даних з похибкою. В якості регуляризованого наближення розв'язку приймемо вектор x^α, що задовольняє СЛАР

$$\left(A_h^* A_h + \alpha E\right)x^\alpha = A_h^* b_\delta + \alpha x^0. \tag{1.35}$$

Твердження 1.2. Нехай $\|A_h - A\| \leq h$, $\|b_\delta - b\| \leq \delta$, $\alpha > 0$. Тоді СЛАР (1.35) однозначно розв'язана і справедлива оцінка

$$\left\|\tilde{x} - x^\alpha\right\| \leq c_1\alpha + \frac{h}{\alpha}\left(\|A\tilde{x} - b\| + 2c_2^2\alpha^2\right)^{\frac{1}{2}} + \frac{1}{\sqrt{a}}(c_3 h + \delta), \tag{1.36}$$

де c_i $(i = 1,2,3)$ константи, які залежать від норми $\|\tilde{x}\|$ псевдорозв'язку.

Наслідок 1.1. Нехай h, δ величини порядку ε, причому, ε достатньо мале число. Якщо точне рівняння (1.7) має розв'язок, (тобто $\|A\tilde{x} - b\| = 0$), тоді права частина оцінки (1.36) за характером залежності від α та ε, є функція вигляду

$$\varphi(\alpha) = \alpha + \varepsilon + \frac{\varepsilon}{\sqrt{\alpha}}. \tag{1.37}$$

При $\alpha = \varepsilon^{\frac{2}{3}}$ вона набуває значення порядку $\varepsilon^{\frac{2}{3}}$. Якщо СЛАР (1.6) нерозв'язна $\left(\|A\tilde{x} - b\| \neq 0\right)$, то права частина нерівності (1.36) є функція вигляду

$$\psi(\alpha) = \alpha + \frac{\varepsilon}{\alpha} + \frac{\varepsilon}{\sqrt{\alpha}}. \tag{1.38}$$

При $\alpha = \sqrt{\varepsilon}$ вона набуває значення порядку $\varepsilon^{\frac{1}{2}}$. Ці порядки одержуються за допомогою мінімізації функцій $\varphi(\alpha)$, $\psi(\alpha)$ (тобто є розв'язками рівнянь $\varphi'(\alpha) = 0$, $\psi'(\alpha) = 0$).

Таким чином, якщо вхідні дані рівняння (1.7) задані з точністю порядку ε, то псевдорозв'язок може бути визначений з точністю порядку $\varepsilon^{\frac{2}{3}}$, у випадку розв'язності точного рівняння, і з точністю порядку $\varepsilon^{\frac{1}{2}}$ – в оберненому випадку. Помітимо, однак, що більш складний спосіб вибору параметра α дозволяє апроксимувати псевдорозв'язок з точністю порядку апроксимації $h + \delta$.

Задача (1.35) еквівалентна задачі на мінімум

$$\min\left\{\|A_h x - b_\delta\|^2 + \alpha\left\|x - x^0\right\|^2 : x \in R^n\right\}, \tag{1.39}$$

При $\alpha = 0$ (1.39) переходить в метод найменших квадратів (МНК), який нестійкий відносно збурень матриці. Перехід від МНК до його регуляризованого аналога (1.39) відтворює стійкість наближеного розв'язку.