Лекция 12. IV МО

25/03

а) Общие соображения.

Будем к задачу $J(u) \to \inf$, $u \in U \subseteq H$, (1) где H – гильб. пр-во

мотено и в банах.

Многие методы реш-я зад. (1) укладываются в итерац. схему $u_{k+1} = u_k + \alpha_k h_k$, $k = 0, 1, \ldots$, (2), где $h = h_k$ – некот. направление из H, (движемся к реш. зад.) (надеемся получить лучшее приближение) $\alpha = \alpha_k > 0$ – длина шага метода по этому направлению.

$J(u_k) \to J_*$, $u_k \xrightarrow{\|\cdot\|} u_*$

Основные трудности, возникающие при формировании процесса (2).

1. Выбор направления h_k: он зависит от имеющейся инф-ции о ф-ции $J(u)$, о мн-ве U, о т. u_k (зависит от зад., а решаем). Например, в случае, когда $U = H$ или $u_k \in \operatorname{int} U$, а ф-ция $J(u) \in C^1(U)$, то в кач-ве h_k можно взять $h_k = -J'(u_k)$ – антиградиент. Обсудим предпосл. выбор шага: $J(u_k + \alpha h) - J(u_k) =$ выбираем h: чтобы разность была как можно меньше $= \alpha \langle J'(u_k), h \rangle + \bar{o}(\alpha) = \alpha \left(\underbrace{\langle J'(u_k), h \rangle}_{\text{как можно меньше должна быть}} + \frac{\bar{o}(\alpha)}{\alpha} \right)$ (3)

При $J'(u_k) \neq 0$ главн. лин. часть приращения можно оценить, используя нер-во К–Б:

$-\|J'(u_k)\| \|h\| \leqslant \langle J'(u_k), h \rangle \leqslant \|J'(u_k)\| \|h\|$ (4)

← как раз наим. знач.

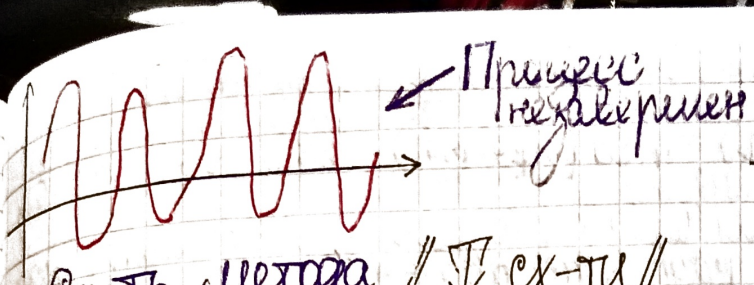
Тогда выбираем $h_k = -J'(u_k)$ //один из выборов// (*)

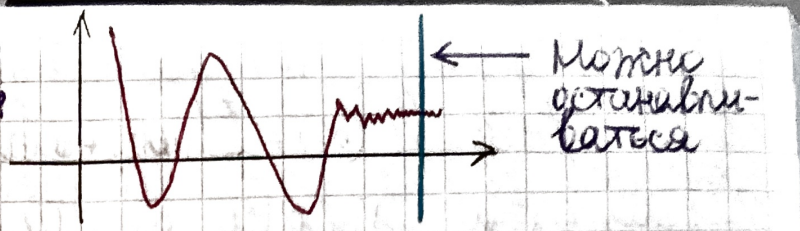
Наилучшее убывание получим, когда лев. нер-во обращ-ся в р-во, т.е. при $h = -\alpha J'(u_k)$, $\alpha > 0$. Отсюда и из (2) видно, что при малых α_k т. $u_{k+1} \in U$ и $J(u_{k+1}) < J(u_k)$. Это справедливо и для (*) в (2). Но, если $U \neq H$ и $u_k \in \Gamma r U$, то т. u_{k+1} может и не $\in U$, в связи с чем возникают осложнения при выборе шага – в разных методах они преодолеваются по-разному.

2. Выбор α_k: (зависит от решаемой задачи //минимизация//) метод скорейшего спуска (лучший, но сложный); простое дробление шага, исходя из требования $J(u_{k+1}) < J(u_k)$ (средний); $\alpha_k = \text{const}$ (худший).

3. Выбор начал. т.: (//как захочется, но иногда из проблемы//) общих правил нет, прим-ся использ-ть эвристич. соображ-я, опыт реш-я подобн. зад.

4. Правило останова процесса (2): Эвристические правила вида $\|u_{k+1} - u_k\| \leqslant \varepsilon$, $\|J'(u_k)\| \leq \varepsilon$, $|J(u_{k+1}) - J(u_k)| \leq \varepsilon$





5. Сх-ть метода // V сх-ти //
6. Влияние погреш-тей на метод → неточн. (либо) входные данные → задачи решаем числ. ⇒ погреш-ти
7. Задачи большой размер-ти // распараллеливание //
8. Вычислительные задачи: проблема перебора, полиномиальные методы // „Неизученный пока раздел" //
9. Общие методы: Методы вида $u_{k+1} = A(u_k)$, где A – нелин. опер-р
10. Многошаг. процессы: $u_{k+1} = A(u_k, u_{k-1}, \dots, u_{k-l})$
11. Непр. методы: в 2 вместо (2) исп-ся ДУ

1) Метод проекции градиента // Базовый метод // $h_k = -J'(u_k)$

1.1) Предметный метод. Его имеет смысл применять к задачам вида: $J(u) \to \inf$, $u \in U = H$; $J(u) \in C^1(H)$ (1)

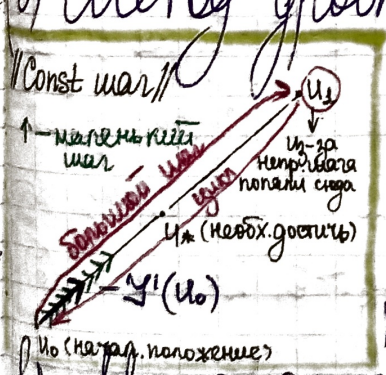
Градиент. метод представляет из себя итерац. процесс:

$u_{k+1} = u_k - \alpha_k J'(u_k)$, $\alpha_k > 0$, $k = 0, 1, \dots$ (2) // В т. движемся по направлению скорейшего убывания //

Выбор длины шага можно производить по-разному. Варианты:

а) Const шаг: $\alpha_k = \alpha = const > 0$ // Проблемы: перескок, зацикливание, slow //

б) Метод дробления: // один из способов // Задаем некот. $\alpha_* > 0$, далее берем $\alpha_k = \frac{\alpha_*}{2^m}$, $m = 0, 1, \dots$, и при каждом m проверяется, будет ли $J(u_k - \frac{\alpha_*}{2^m} J'(u_k)) < J(u_k)$, $m = 0, 1, \dots$ В кач-ве α_k берется тот из номеров m, когда это нер-во выполн-ся // Problem: slow //



в) Метод скорейшего спуска: // выбираем оптимал. α: было бы наим. убывание //

Введем функ-ю $f_k(\alpha) = J(u_k - \alpha J'(u_k))$, $\alpha \geq 0$ и решим зад.: $f_k(\alpha) \to \inf$, $\alpha \geq 0$ (3). Т. α_k: $f_k(\alpha_k) = \inf_{\alpha \geq 0} f_k(\alpha)$, берем эл. длину шага в (2) ☆

квадратич. ф-ции $\mathcal{J}(u) = \|Au - b\|_F^2$, $A \in \mathcal{L}(H \to F)$ зад. (3) мб реш-а явно. Д/нее $f_k(\alpha) = \mathcal{J}(u_k + \alpha h) = \|A(u_k + \alpha h) - b\|_F^2$, $\alpha \geq 0$. В этом случае относительно α получаем параболу // Этот метод быстрый, но самый сложный // // If $\alpha = 0$, то надо остановиться // // Самый применимый метод — метод дробления //

Пример // показывает зачем д-ть сх-ть //

$\mathcal{J}(u) = \frac{1}{2}x^2 + \frac{1}{4}y^4 - \frac{1}{2}y^2 \hookrightarrow \inf$, $u = (x, y) \in E^2$

Обратим внимание, что ответ можно получить сразу $\mathcal{J}(u) = \frac{1}{2}x^2 + \frac{1}{4}y^4 - \frac{1}{2}y^2 = \frac{1}{2}x^2 + \frac{1}{4}(y^2-1)^2 - \frac{1}{4} \Rightarrow \mathcal{J}_* = -\frac{1}{4}$, $U_* = \{u_{*1} = (0,1), u_{*2} = (0,-1)\}$.

Потребуем решить зад., используя град. метод. Д/этого нам понадобится $\mathcal{J}'(u) = (x, y^3 - y)$. // $\mathcal{J}'' = \begin{pmatrix} 1 & 0 \\ 0 & 3y^2-1 \end{pmatrix}$ //

• Возьмем в кач-ве начал. приближ-я $u_0 = (1, -1)$, в кач-ве шага $\alpha_k = \alpha = \frac{1}{2}$. Тогда $\mathcal{J}'(u_0) = (1; 0)$ и $u_1 = (1, -1) - \frac{1}{2}(1, 0) = (\frac{1}{2}, -1)$, $u_2 = (\frac{1}{4}, -1)$, $u_k = (\frac{1}{2^k}, -1) \xrightarrow{k \to \infty} u_* = (0, -1)$ // На практике ∞ не будет //

• Возьмем в кач-ве начал. приближ-я $u_0 = (1, 0)$, в кач-ве шага $\alpha_k = \alpha = \frac{1}{2} \Rightarrow \mathcal{J}'(u_0) = (1, 0)$; $u_1 = (1,0) - \frac{1}{2}(1,0) = (\frac{1}{2}, 0)$; $u_2 = (\frac{1}{4}, 0)$; $u_k = (\frac{1}{2^k}, 0) \xrightarrow{k \to \infty}$ и $(0,0)$ — стационар. т.

1.2) Метод проекции градиента. Его имеет смысл применять к задач. вида $\mathcal{J}(u) \hookrightarrow \inf$, $u \in U \neq H$; $\mathcal{J}(u) \in C^1(H)$ (H). Град. метод в этой зад. не годится, т.к. может оказ-ся, что $u_k \in \Gamma U$ и тогда $u_k - \alpha \mathcal{J}'(u_k)$ может и не $\in U$ $\forall \alpha > 0$, представляя из себя итерац. процесс. Тогда можно применить метод проекции градиента, а получается из град. метода проектир-нием прав. части на мн-во U: $\boxed{u_{k+1} = P_U(u_k - \alpha_k \mathcal{J}'(u_k)), \alpha_k > 0, k = 0, 1, \ldots}$ (5)

Чтобы проекция $\exists$ и однозначно опред-сь, прих-ся счит-ть, что U — вып. и замкн.

Выбор длины шага α_k в (5) можно проводить из тех же

отображ-й, что и в методе (2) (разумеется, счит-ся, что $f_k(\alpha) = f(P_U(u_k - \alpha f'(u_k)))$

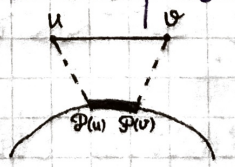
На метод (5) можно посмотреть с друг. стороны. Вспомнить кр. опт-ти (д/вып. ф-ций) в проекц. форме. $u_* \in U_* \Leftrightarrow u_* = P_U(u_* - \alpha f'(u_*)), \forall \alpha > 0$ (6)

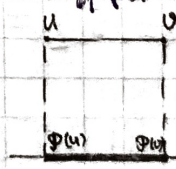
Если теперь ввести отображ. $Au = P_U(u - \alpha f'(u))$, где $\alpha > 0$ – фикс., то соотнош. (6) означает, что $u_* \in U_*$ – неподвижн. т. этого отображ-я $u_* = Au_*$ и тогда метод проекции град. можно записать в виде $u_{k+1} = Au_k$, $\alpha_k = \alpha$.

Нам известно, что если A – сжимающ. оп-р и действ. в себя, то такой итерац. процесс будет сх-ся. Будет ли $Au = P_U(u - \alpha f'(u))$ сжимающим?

Т (о сх-ти метода проекции градиента) Пусть U – вып., замк. мн-во из H, кроме того $f(u) \in C^1(U)$ и $\operatorname{grad} f'(u)$ удовл. усл. Липшица, т.е. $\exists L = \text{const} > 0$: $\|f'(u) - f'(v)\| \le L\|u - v\|$, $\forall u, v \in U$ (7). Ф-ция $f(u)$ сил. вып. на U и, поэтому, $\exists \mu = \text{const} > 0$: $\mu\|u - v\|^2 \le \langle f'(u) - f'(v), u - v\rangle$, $\forall u, v \in U$ (8) (кр. сил. вып.). Пусть в методе (5) $\alpha_k = \alpha$, где $0 < \alpha < \frac{2\mu}{L^2}$. Тогда при $\forall$ выборе нач. приближ-я $u_0 \in U$ посл-ть $\{u_k\}$, порождаемая методом (5), сх-ся к реш. зад. u_* зад. (4), причем справедлива оценка: $\|u_k - u_*\|_H \le q^k \|u_0 - u_*\|_H$, $k = 0, 1, \ldots$, где (лин. ск-ть сх-ти) $\forall \alpha$ (9) $q = q(\alpha) = \sqrt{1 - 2\mu\alpha + \alpha^2 L^2} \in (0, 1)$ (10)

▲ Убедимся, что реш-е $\exists$! Заметим, что по Т Вейер. у/сил. вып. ф-ций зад. (4) имеет притом ! реш. u_*. Убедимся, что оп-р $Au = P_U(u - \alpha f'(u)): U \to U$ явл-ся сжимающ. при $\alpha \in (0; \frac{2\mu}{L^2})$

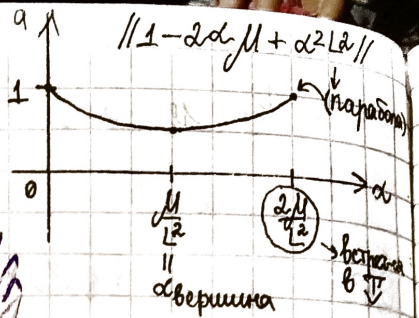




// Нам нужен не расширяющ. оп-р, а сжимающ.!

Имеем: $\|Au - Av\|_H^2 = \|P_U(u - \alpha f'(u)) - P_U(v - \alpha f'(v))\|_H^2 \le \|u - \alpha f'(u) - v + \alpha f'(v)\|^2 = \|u - v\|^2 + \alpha^2 \underbrace{\|f'(u) - f'(v)\|^2}_{\text{по Липшицу}} - 2\alpha \underbrace{\langle f'(u) - f'(v), u - v\rangle}_{\text{по кр. сил. вып-ти}} \le \|u - v\|^2 (1 - 2\mu\alpha +$

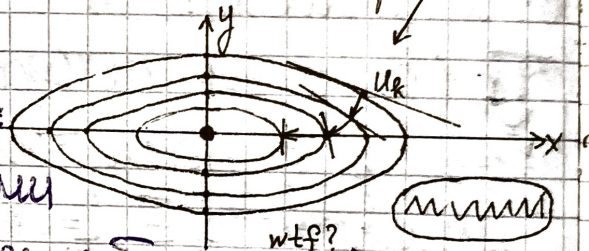
$+\alpha^2L^2) = \|u-v\|_H^2 q^2(\alpha)$



Т.о.: $\|\mathcal{A}u-\mathcal{A}v\|_H \le q_f(\alpha)\|u-v\|_H$, $q_f(\alpha)\in(0,1)$, $\forall\alpha\in(0;\frac{2\mu}{L^2}) \Rightarrow$ $\mathcal{A}$ – сжимающ. оп-р $\Rightarrow \exists!$ неподвиж. т. u_*, являющ-ся реш. зад. (*). Из $\exists!$ неподвижн. т. оп-ра $u_{k+1} = \mathcal{A}u_k$, $\forall u_0\in W$ сх-ся к т. u_*. Далее оценим ск-ть сх-ти $\|u_{k+1}-u_*\|_H = \|\mathcal{A}u_k - \mathcal{A}u_*\|_H \le$
$\le q_f\|u_k-u_*\| = q_f\|\mathcal{A}u_{k-1}-\mathcal{A}u_*\| \le q_f^2\|u_{k-1}-u_*\| \le \dots \le q_f^{k+1}\|u_0-u_*\|$ ■

Достоинства: прозрачность идеи, простота реализации в случае "хороших" ЛН-в (ЛН-ва нас волнуют, т.к. надо делать проекцию), когда известны оценки $\gamma/\mu, L$

Недостатки: плохая сх-ть, особенно, if $\mu \ll L$ (овражность, тестовые с-мы) даже в сил. вып. зад.



Зам. $\star$ подробнее $q_f(\alpha) = \sqrt{1-2\mu\alpha+\alpha^2L^2} \in (0,1)$. Под $\sqrt{\ }$ у нас парабола ветвями вверх с вершиной в т. $\alpha = \mu/L^2$. здесь и будет min знач. $q_f = \sqrt{1-(\frac{\mu}{L})^2} \ge 0$, причем $q_f = 0$ при $\mu = L$, что справедливо, например, д/ ф-ции $\gamma(u) = \|u\|^2$. В общем случае, $\mu \le L$, т.к. $\mu\|u-v\|^2 \le L\|u-v\|^2$ и в случае, когда $\mu \ll L$ q_f оч. близко к 1 и сх-ть метода плохая

Это есть случай, когда велик разброс собств. чисел м-цы γ''

Пример $\gamma(u) = \frac{x^2}{10^n} + y^2$; $\gamma'_{(u)} = (\frac{2x}{10^n}, 2y)$, $\gamma''(u) = 2\begin{pmatrix}\frac{1}{10^n} & 0\\ 0 & 2\end{pmatrix}$
Тогда $\mu = \vartheta_1 = 2/10^n \ll \vartheta_2 = 2 = L$

Зам. $\exists$ непр. аналоги метода проекции grad: $\dot u(t) = \mathcal{P}_U(u(t) - \alpha(t)\gamma'(u(t))) - u(t)$, $t \ge 0$