Методы оптимизации

1 Теорема Вейерштрасса (метрический вариант).

Задача: Минимизировать функционал $J(u)$ по множеству $U \subset X$, X – метрическое.
$J_* = \inf\limits_{u \in U} J(u)$, $J(u_*) = J_*$, $U_* = \{u \in U | J(u) = J_*\}$

Определение. *Функционал $J(u)$ на U называется полунепревным снизу (сверху), если $\forall u_n \subset U: \ \rho(u_n, u_0) \to 0 \Rightarrow J(u_0) \leqslant \underline{\lim}_{n\to\infty} J(u_n)$ $\left(J(u_0) \geqslant \overline{\lim}_{n\to\infty} J(u_n)\right)$*

Определение. *Последовательность $\{u_n\} \subset U$ называется минимизирующей, если $\exists J(u_n) \to J_*$*

Определение. *Последовательность $\{u_n\} \subset U$ сходится к множеству $K \subset U$, если $\inf\limits_{u \in K} \rho(u_n, u) \to 0$*

Теорема. *(Теорема Вейерштрасса метрический вариант)*
Пусть U - компкатное множество, $J(u)$ - полунепрервный снизу, тогда

1. *$J_* > -\infty$*
2. *U_* - непустое компактное множество*
3. *Любая минимизирующая последовательность $\{u_n\}$ сходится к U_**

2 Слабый вариант теоремы Вейерштрасса. Применение к задаче минимазции квадратичного функционала.

H - гильбертово пространство.

Определение. *Функционал $J(u)$ на U называется слабо полунепревным снизу (сверху), если $\forall u_n \subset U: \ u_n \rightharpoonup u_0 \Rightarrow J(u_0) \leqslant \underline{\lim}_{n\to\infty} J(u_n)$ $\left(J(u_0) \geqslant \overline{\lim}_{n\to\infty} J(u_n)\right)$*

Замечание. *Из слабой полунепрервыности следует сильная полунепрерывность*

Определение. *Последовательность $\{u_n\} \subset H$ слабо сходится к множеству $K \subset H$, если любая слабая предельная точка $\{u_n\}$ принадлежит K.*

Определение. *Множество $U \in H$ называется выпуклым, если $\forall u, v \in U, \ \forall \alpha \in [0,1] : \alpha u + (1-\alpha) v \in U$.*

Определение. *Функционал $J(u)$ называется выпуклым на выпуклом U, если $\forall u, v \in U, \ \forall \alpha \in [0,1]: \ J(\alpha u + (1-\alpha) v) \leqslant \alpha J(u) + (1-\alpha) J(v)$.*

Лемма. *Пусть U - выпуклое замкнутое ограниченное множество, тогда U - слабокомпактное множество.*

Лемма. *Пусть $J(U)$ полунепрерывный снизу и выпуклый на U, тогда он слабо полунепрерывный снизу.*

Теорема. *(Слабый вариант теоремы Вейерштрасса)*
Пусть $U \subset X$ - замкнутое ограниченное выпуклое множество, $J(u)$ - выпуклый и полунепрерывный на U, тогда

1. *$J_* > -\infty$*
2. *U_* - непустое замкнутое ограниченное выпуклое множество*
3. *Любая минимизирующая последовательность $\{u_n\}$ слабо сходится к U_**

Квадратичный функционал

Функионал $J(u) = \|Au - f\|_F^2$, $A : H \to F$ - линейный ограниченный оператор, $f \in F$. Исследуем его свойства:

1. $J(u)$ непрерывный в силу непрерывности А и непрерывности нормы.
2. Проверим выпуклость $J(u)$: $J(\alpha u + (1-\alpha) v) = \|A(\alpha u + (1-\alpha) v) - f\|^2 = \|\alpha(Au - f) + (1-\alpha)(Av - f)\|^2 \leqslant$ $\leqslant (\|\alpha(Au - f)\| + \|(1-\alpha)(Av - f)\|)^2 \leqslant \{x^2 \text{выпукла}\} \leqslant |\alpha| \|(Au - f)\|^2 + |1-\alpha| \|Av - f\|^2$
3. Полунепрерывный снизу
4. В общем случае не является слабо полунепрерывным т.к. если $A = I, f = 0$, то на последовательности ортонормированных векторов слабой сходимости не будет.
5. Слабо полунепрерывен снизу, т.к. полунепрерывен снизу и выпуклый.

3 Существование решения задач минимизации терминального и интегрального квадратичных функционалов на решениях линейной системы ОДУ

Рассматриваем задачу

$$\begin{cases} \dot{x}(t) = D(t)x(t) + B(t)u(t) + y(t), \ t \in (0, T) \\ x(0) = x_0 \end{cases}$$

$u(t)$ - функция управления
Размерности: $x(t) : n \times 1, \ D(t) : n \times n, \ B(t) : n \times r, \ u(t) : r \times 1, \ y(t) : n \times 1$
Базово предполагаем, что $D(t), B(t) \in L_\infty(0, T); \ u \in L_2(0, T); \ x_0 \in \mathbb{R}^n$
Надо найти пару $x(t) \in AC[0, T], u(t) \in L_2(0, T)$. AC - абослютная непревность: 1) п.в. на $[0, T]$ существует производная; 2) верна формула Ньютона-Лейбница

Определение. *Решением Задачи Коши по Каратеодори называется функция $x(t) \in AC[0, 1]$ такая, что уравнение выполняется п.в., а граничное условие выполняется в классическом смысле.*

Определение. *Альтернативным решением называется функция $x(t) \in AC[0, 1]$ такая, что выполняется интегральное соотношение: $x(t) = x_0 + \int\limits_0^t [D(t)x(t) + B(t)u(t) + y(t)]\, dt, \ \forall t \in [0, T]$*

Теорема. *Пусть $D(t), B(t) \in L_\infty(0, T), \ u, y(t) \in L_2(0, T)$ тогда существует и единственно решение Задачи Коши.*

Теорема. *(о существовании решения задачи ОУ линейной системы)*
Пусть $D(t), B(t) \in L_\infty(0, T), \ U$ - слабый компакт, тогда

1. *$J_* > -\infty$*
2. *$U_* \neq \varnothing$*
3. *Любая минимизирующая последовательность $\{u_n\}$ слабо в $L_2(0, T)$ сходится к U_**

4 Существование реения задачи об оптимальном нагреве стержня

Рассматривается задача:

$$\begin{cases} y_t = y_{xx}, x \in (0, l), t \in (0, t) \\ y_x|_{x=0} = 0 \\ y_x + y|_{x=l} = u(t) \\ y|_{t=0} = 0 \end{cases}$$

Решение: $y = y(t; x)$. Рассмотрим функционал $J(u) = \int\limits_0^l (y(T, x, u) - f(x))^2\, dx$. Значение именно в $t = T$, то есть в конце процессса. Минимизируем этот функционал. По сути $y(T, x, u) = Au$, тогда $J(u) = \|Au - f\|^2_{L_2(0,l)}$.
Пусть $y(t, x)$ - дважды гладкая функция, а U - замкнутое ограниченное выпуклое множество.

Теорема. *(о существовании решения обратной задачи)*
Пусть U - слабый компакт в $L_2(0, T)$, тогда

1. *$J_* > \infty$*
2. *$U_* \neq \varnothing$*
3. *Любая минимизирующая последовательность слабо сходится к U_* в $L_2(0, T)$.*

5 Дифференцирование по Фреше. Применение к квадратичному функционалу

Пусть $F : X \to Y, \ X, Y$ - банаховы.

Определение. *Оператор $A : X \to Y$ называется производной по Фреше оператора F в т. $x \in X$, если $F(x + h) - F(x) = Ah + \overline{o}(\|h\|_X)$ при $\|h\|_X \to 0$*

Лемма. *Производная по Фреше определяется единственным образом*

Определение. *Оператор F'' называется второй производной по Фреше от оператора F в т. $x \in X$, если $F'(x + h) - F'(x) = F''h + \overline{o}(\|h\|_X)$ при $\|h\|_x \to 0$.*

Определение. *Оператор F дифференицируем на множестве $U \subset X$, если он определён на множестве $M : U \subset M$ и $\exists F'(u), \forall u \in U$.*

Градиент и Гессиан

Рассматриваем Гильбертово пространство H

Определение. *Функционал F' называется градиентом функционала F, в т. $x \in H$, если $F(x+h) - F(x) = (F'(x), h) + \overline{o}(\|x\|_H)$*

Определение. *Функционал $F''(x)$ называется гессианом функционала F, в т. $x \in H$, если $F'(x+h) - F'(x) = F''(x)h + \overline{o}(\|x\|_H)$*

Найдем градиент и гессиан функционала $J(u) = \|Au - f\|_F^2$:
$J(u+h) - J(u) = \|A(u+h) - f\|^2 - \|Au - f\|^2 = \|(Au - f) + Ah\|^2 - \|Au - f\|^2 = \|Au - f\|^2 + \|Ah\|^2 + 2(Au - f, Ah) - \|Au - f\|^2 = (Au - f, Ah) + \|Ah\|^2$. Покажем, что $\|Ah\|^2 \leqslant \|A\|^2\|h\|^2 = \underline{O}(\|h\|^2) = \overline{o}(\|h\|)$.
В итоге $J'(u) = 2A^*(Au - f)$.
$J'(u+h) - J'(u) = 2A^*Ah \Rightarrow J''(u) = 2A^*A$.

6 Необходимое условие локального минимума

Теорема. *Пусть U - выпуклое множество в H, $u_* \in U$ -локальный минимум $J(u)$ на U и существует $J'(u_*)$ $\Rightarrow (J'(u_*), u - u_*) \geqslant 0,\ \forall u \in U$.*

Теорема. *Пусть U - выпуклое множество в H, $u_* \in intU_*$ и существует $J'(u_*) \Rightarrow (J'(u_*), u - u_*) = 0,\ \forall u \in U$.*

Пример. $J(u) = u,\ u \in [1,2] \subset \mathbb{R}$. *Понятно, что $u_* = 1,\ J_* = 1,\ J'(u) = I, \Rightarrow (J'(u_*), u - u_*) = (1, u-1) \geqslant 0$ т.к. $u \in [1,2]$.*

Пример. *Сложный и непонятный*

7 Градиент терминального граничного функционала

Рассматриваем задачу

$$\begin{cases} \dot{x}(t) = D(t)x(t) + B(t)u(t),\ t \in (0,T) \\ x(0) = 0 \end{cases}$$

Рассматриваем оператор $Au(t) = x(t)$ - сопоставляет решение функции управления. Функционал $J(u) = \|Au - f\|^2$, $J'(u) = 2A^*(Au - f)$. Необходимо найти A^*, т.е. $(Au, v) = (u, A^*v)$
Домножим уравнение скалярно на $\psi(t)$ и проинтегрируем от 0 до T

$$\int_0^T (\psi(t), \dot{x}(t))dt = \int_0^T (\psi(t), D(t)x(t) + B(t)u(t))dt$$

$$(\psi(t), x(t))|_0^T - \int_0^T (\dot{\psi}(t), x(t))dt = \int_0^T (\psi(t), Au(t))dt$$

$$(\psi(T), x(T)) - \int_0^T (\dot{\psi}(t), x(t))dt = \int_0^T (\psi(t), Au(t))dt,\ \text{Потребуем } v(t) = \psi(t)$$

$$(v, Au) = \int_0^T (\dot{\psi}(t), x(t))dt + \int_0^T (\psi(t), D(t)x(t) + B(t)u(t))dt$$

$$(v, Au) = \int_0^T (\dot{\psi}(t), x(t))dt + \int_0^T \left[(D^T\psi(t), x(t)) + (B^T\psi(t), u(t))\right] dt$$

$$(v, Au) = \int_0^T (\dot{\psi}(t) + D^T\psi(t), x(t))dt + \int_0^T (B^T(t)\psi(t), u(t))dt,\ \text{Потребуем } \dot{\psi}(t) + D^T\psi(t) = 0$$

Тогда $(v, Au) = \int_0^T (B^T(t)\psi(t), u(t))dt \Rightarrow (A^*v)(t) = B^T(t)\psi(t)$. $\psi(t)$ определяется из двойственной задачи.

$$\begin{cases} \dot{\psi}(t) = -D^T(t)\psi(t) \\ \psi(t) = v(t) \end{cases}$$

8 Градиент интегрального квадратичного функционала

Рассматриваем задачу

$$\begin{cases} \dot{x}(t) = D(t)x(t) + B(t)u(t),\ t \in (0,T) \\ x(0) = 0 \end{cases}$$

$J_I(u) = \int_0^T |x(t;u) - f(x)|^2 dt$, $Au = x(t;u)$, $J_I(u) = \|Au - f\|_{L_2}^2$, $J_I' = 2A^*(Au - f)$. Надо искать A^*
Нам подойдёт $A^*v = B^T\psi(t;v)$, где

$$\begin{cases} \dot{\psi}(t) = -D^T(t)\psi(t) - v(t),\ t \in (0,T) \\ \psi(T) = 0 \end{cases}$$

Проверим это: $(Au,c)_{L_2} = \int\limits_0^T x(t;u)v(t)dt = \int\limits_0^T x(t;u)\left[-\dot\psi(t) - D^T(t)\psi(t)\right]dt = -\int\limits_0^T x(t;u)\dot\psi(t)dt - \int\limits_0^T x(t;u)D^T(t)\psi(t)dt =$
$- x(t;u)\psi(t)|_{t=0}^{t=T} + \int\limits_0^T \dot x(t;u)\psi(t)dt - \int\limits_0^T x(t;u)D^T(t)\psi(t)dt = \int\limits_0^T [D(t)x(t)+B(t)u(t)]\psi(t)dt - \int\limits_0^T x(t;u)D^T(t)\psi(t)dt =$
$= \int\limits_0^T B(t)\psi(t)u(t)dt = (u(t), B^T\psi)_{L_2}$

9 Градиент функционала в задаче о нагреве стержня

Рассматриваем задачу

$$\begin{cases} y_t = y_{xx}, x\in(0,l), t\in(0,t) \\ y_x|_{x=0} = 0 \\ y_x + y|_{x=l} = u(t) \\ y|_{t=0} = 0 \end{cases}$$

Рассмотрим Функционал $J(u) = \int\limits_0^l |y(T,x;u) - f(x)|^2 dx$, $Au = y(T,x;u)$ тогда $J(u) = \|Au - f\|^2_{L_2}$, $J'(u) = 2A^*(Au - f)$. Нужно найти $A*$.
Умножим уравнение на $\psi(t,x)$ и проинтегрируем по $Q = (0,T)\times(0,l)$.
$\iint\limits_Q [y_{xx} - y_t]\psi dt dx = \int\limits_0^T \left[\int\limits_0^l y_{xx}\psi dx\right]dt - \int\limits_0^l \left[\int\limits_0^T y_t\psi dt\right]dx = \{\text{По частям}\} = \int\limits_0^T \left[y_x\psi|_{x=0}^{x=l} - y\psi_x|_{x=0}^{x=l} + \int\limits_0^l y\psi_{xx}dx\right]dt -$
$\int\limits_0^l \left[y\psi|_{t=0}^{t=T} - \int\limits_0^T y\psi_t dt\right]dx = \int\limits_0^T y_x\psi|_{x=0}^{x=l}dt - \int\limits_0^T y\psi_x|_{x=0}^{x=l}dt - \int\limits_0^l y\psi|_{t=0}^{t=T}dx + \iint\limits_Q y(\psi_{xx} + \psi_t)dt dx =$
$= \{\text{Требуем } \psi_{xx} + \psi_t = 0 \text{ в Q и } \psi_x|_{x=0} = 0, \text{ много что обнуляется из-за граничных условий}\} =$
$= \int\limits_0^T [y_x(t,l)\psi(t,l)]dt - \int\limits_0^T y(t,l)\psi_x(t,l)dt - \int\limits_0^l [y(T,x)\psi(T,x)]dx = \int\limits_0^T [u(t) - y(t,l)]\psi(t,l)dt - \int\limits_0^T y(t,l)\psi_x(t,l)dt - (Au,v)_{L_2} =$
$= (u,\psi|_{x=l})_{L_2} - \int\limits_0^T [\psi(t,l) + \psi_x(t,l)]y(t,l)dt - (Au,v)_{L_2} = \{\text{Потребовали, чтобы } \psi(T,x) = v(x) \text{ и } (\psi_x + \psi)|_{x=l} = 0\} =$
$= (u,\psi|_{x=l})_{L_2} - (Au,v)_{L_2} = 0$
Получили, что $A^*v = \psi|_{x=l}$, где

$$\begin{cases} \psi_t = -\psi_{xx}, x\in(0,l), t\in(0,t) \\ \psi_x|_{x=0} = 0 \\ \psi_x + \psi|_{x=l} = 0 \\ \psi|_{t=T} = v \end{cases}$$

10 Выпуклые функции и функционалы. Теоремы о локальном минимимуме, о множестве Лебега, о касательной плоскости. Критерий оптимальности. Примеры

Определение. *Множество $U \in H$ называется выпуклым, если $\forall u,v\in U,\ \forall\alpha\in[0,1] : \alpha u + (1-\alpha)v \in U$.*

Определение. *Функционал $J(u)$ называется строго выпуклым на выпуклом U, если $\forall u,v\in U,\ \forall\alpha\in[0,1]$: $J(\alpha u + (1-\alpha)v) < \alpha J(u) + (1-\alpha)J(v)$.*

Определение. *Функционал $J(u)$ называется сильно выпуклым на выпуклом U с константой $æ > 0$, если $\forall u,v\in U,\ \forall\alpha\in[0,1]:\ J(\alpha u + (1-\alpha)v) \leqslant \alpha J(u) + (1-\alpha)J(v) - \frac{æ}{2}\alpha(1-\alpha)\|u-v\|^2$.*

Пример. *Функционал $J(u) = \|Au - f\|$ является выпуклым.*

Свойства строгой выпуклости:

1. $J_1(u), J_2(u)$ - строго выпуклы на U и $\alpha_1,\alpha_2 > 0 \Rightarrow \alpha_1 J_1(u) + \alpha_2 J_2(u)$ строго выпулый на U.
2. $J_1(u)$ - строго выпуклый, $J_2(u)$ - выпуклый на $U \Rightarrow J_1(u) + J_2(u)$ - строго выпуклый на U.

Теорема. *(о локальном минимуме)*
$J(u)$ - выпуклый на $U \Rightarrow$ точка локального минимума - точка глобального минимума.

Теорема. *(о множестве Лебега)*
$J(u)$ - выпуклый на $U \Rightarrow$ множество $L_c = \{u\in U | J(u)\leqslant c\}$ - замкнуто $\forall c\in\mathbb{R}$.

Обратное неверно т.к. $J(u) = u^3,\ u\in\mathbb{R},\ L_c = (-\infty, \sqrt[3]{c}]$

Лемма. *Пусть U - выпуклое, $J(u)$ - выпуклый на U, тогда U_* - замкнуто.*

Лемма. *Пусть U - выпуклое, $J(u)$ - строго выпуклый на U, тогда U_* содержит одну точку или $U_* = \varnothing$.*

Пример. *$U_* = \varnothing$:*

1. *$J(u) = u,\ u \in \mathbb{R} \Rightarrow J_* = -\infty, U_* = \varnothing$.*

2. *$J(u) = e^{-u},\ u \in \mathbb{R} \Rightarrow J_* = 0, U_* = \varnothing$.*

Теорема. *(о касательной плоскости)*
Пусть U - выпуклое, $J(u)$ сильно выпуклый на U с $æ > 0$ и в точке v $\exists J'(v) \Rightarrow J(u) \geqslant J(v) + (J'(v), u - v) + \frac{æ}{2}\|u - v\|^2, \forall u \in H$.

Теорема. *(критерий оптимальности)*
Пусть U - выпуклое, $J(u)$ выпуклый на U и $\exists J'(u_*) \Rightarrow u_* \in U_* \Leftrightarrow$ выполнено $(J'(u_*), u - u_*) \geqslant 0, \forall u \in U$.

Пример. *Решить уравнение $Au = f$, $A \in L(H \to H), A = A^*$. Эквивалентна задаче минизации функционала $J(u) = (Au, u) - 2(u, f) \to min$. $J'(u_*) = 0$, $(A + A^*)u_* = 2f \Rightarrow Au_* = f$*

11 Критерий выпуклости функций и функционалов. Выпуклость квадртичного функционала

Теорема. *(критерий выпуклости)*
Пусть U - выпуклое, $J(u) \in C^1(U)$, тогда следующие утверждения эквивалентны

1. *$J(u)$ выпуклый*

2. *$J(u) > J(v) + (J'(v), u - v), \forall u, v \in U$*

3. *$(J'(u) - J'(v), u - v) \geqslant 0, \forall u, v \in U$*

Выпуклость квадратичного доказана в первом билете.

12 Сильно выпуклые функции и функционалы, их свойства. Критерии сильной выпуклости функций и функционалов

Определение. *Функционал $J(u)$ называется сильно выпуклым на выпуклом U с константой $æ > 0$, если $\forall u, v \in U,\ \forall \alpha \in [0, 1]:\ J(\alpha u + (1 - \alpha)v) \leqslant \alpha J(u) + (1 - \alpha)J(v) - \frac{æ}{2}\alpha(1 - \alpha)\|u - v\|^2$.*

Свойства сильно выпуклости:

1. $J_1(u), J_2(u)$ - сильно выпуклы на U и $\alpha_1, \alpha_2 > 0 \Rightarrow \alpha_1 J_1(u) + \alpha_2 J_2(u)$ сильно выпулый на U.

2. $J_1(u)$ - сильно выпуклый, $J_2(u)$ - выпуклый на $U \Rightarrow J_1(u) + J_2(u)$ - сильно выпуклый на U.

Теорема. *(критерий сильной выпуклости)*
Пусть U - выпукло, $J \in C^1(U)$, тогда $J(u)$ сильно выпуклый на U с константой $æ > 0 \Leftrightarrow (J'(u) - J'(v), u - v) \geqslant æ\|u - v\|^2, \forall u, v \in U$.

Теорема. *(второй критерий сильной выпуклости)*
Пусть U - выпукло, $J \in C^2(U)$, $intU \neq \varnothing$, тогда $J(u)$ сильно выпуклый на U с константой $æ > 0 \Leftrightarrow (J''(u)h, h) \geqslant æ\|h\|, \forall u \in U, h \in H$.

Пример. $J(u) = \|u\|^2$, $J'(u) = 2u$, $J''(u) = 2I$, $(J''(u)h, h) = 2\|h\|^2 \geqslant æ\|h\|^2 \Rightarrow æ = 2$

Пример. $J(u) = x^2 + 2xy + y^2 + z^2$, $u \in \mathbb{R}^3$.

$$J''(u) = \begin{pmatrix} 2,2,0 \\ 2,2,0 \\ 0,0,2 \end{pmatrix}$$

Найдем собственные значения: $\lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 4$. По критерию положительно определенных матриц, все с.з. неотрицательны, значит матрица положительна определена: $(J''(u)h, h) \geqslant 0$. Но $J(u)$ не является сильно выпуклым т.к. при λ_1 значение ровно 0.

13 Теорема Вейерштрасса для сильно выпуклых функционалов

Теорема. *Пусть U - выпуклое, замкнутое, $J(u)$ сильно выпуклый на U с $æ > 0$ и полунепрерывный снизу на U, тогда*

1. *$J_* > -\infty$*

2. *$U_* = \{u_*\} \neq \varnothing$*

3. *$\forall u \in U:\ \frac{æ}{2}\|u - u_*\|^2 \leqslant J(u) - J(u_*)$*

14 Метрическая проекция точки на выпуклое замкнутое множество в гильбертовом пространстве, её свойства. Примеры

Определение. *Пусть $U \subset H$. Проекцией элемента $u \in H$ на множество U называется $w \in U$: $\|w - u\| = \inf\limits_{v \in U} \|v - u\|$*

Теорема. *(существование и единственность и свойства проекции)*
Пусть U - выпуклое и замкнутое, тогда

1. $\forall u \in H \ \exists! w = P_u u$.
2. $w = P_u u \Leftrightarrow (w - u, v - w) \geqslant 0, \forall v \in U$.

Теорема. *(о нестрогой сжимаемости)*
Пусть U - выпуклое и замкнутое мноежство $\Rightarrow \forall u, v \in H \quad \|P_u u - P_u v\| \leqslant \|u - v\|$.

Пример. $U = B_R(0)$, $u \in H, w = P_u u$, $w = \begin{cases} u, u \in U \\ \dfrac{u}{\|u\|} R, u \notin U \end{cases}$. *Проверим свойство: cсли $u \in U$ то очевидно, пусть $u \notin U$, тогда $(\dfrac{u}{\|u\|}R - u, v - \dfrac{u}{\|u\|}R) = \left(\dfrac{R}{\|u\|} - 1\right)(u, v - \dfrac{u}{\|u\|}R)$. Первое слагаемое неположительно, второе тоже т.к. $\|v\| \leqslant R$. Поэтому условие выполняется.*

Пример. $U = \{u \in L_2(a,b) | \alpha(t) \leqslant u(t) \leqslant \beta(t), \alpha(t), \beta(t) \in L_2(a,b)\}$, $\|u(t) - h(t)\|^2_{L_2} = \int\limits_a^b |u(t) - h(t)|^2 dt \to \inf$

$$P_U h = \begin{cases} h(t), \alpha(t) \leqslant h(t) \leqslant \beta(t) \\ \beta(t), \beta(t) \leqslant h(t) \\ \alpha(t), h(t) \leqslant \alpha(t) \end{cases}$$

15 Градиентный метод. Метод проекции градиента. Их сходимость

Решаем задачу $J(u) \to \inf$ в гильбертовом пространстве. Многие методы решения укладываются в итерационную схему:

$$u_{k+1} = u_k + \alpha_k p_k \tag{1}$$

u_0 - задано, α_k - шаг, p_k - поправление шага.

1. $p_k = -J'(u_k)$ - наискорейшее локальное убывание
2. α_k можно выбирать например так: $\alpha_k \in (\mathrm{Argmin}_{\alpha \in \mathbb{R}} J(u_k + \alpha p_k))$
3. u_0 хочется выбрать как можно ближе к u_*.
4. Правило останова:
   (a) малость градиента $\|J'(u_k)\| \leqslant \varepsilon$ -строгий
   (b) $\dfrac{\|u_{k+1} - u_k\|}{\|u_k\|} \leqslant \varepsilon$ -слабый
   (c) $\dfrac{|J(u_{k+1}) - J(u_k)|}{|J(u_k)|} \leqslant \varepsilon$ - самый слабый

Градиентный метод

Его имеет смысл применять к задачам вида $J(u) \to \inf, \ u \in U \subset H, \ J(u) \in C^1(H)$
Представляет из себя итерационный процесс $u_{k+1} = u_k - \alpha_k J'(u_k), \ \alpha_k > 0, \ k = 0, 1, 2, \dots$ Выбор длины шага можно делать по-разному:

1. константный шаг (проблемы: зацикливание, перескок)
2. метод дробления: сначала задаем α_*б потом $\alpha_k = \dfrac{\alpha_*}{2^m}$, $m = 0, 1, 2, \dots$. На каждом шаге проверяется будет ли $J(u_k - \dfrac{\alpha_*}{2^m} J'(u_k)) < J(u_k)$ и в качестве m берется первый, при котором выполняется это неравенство.
3. метод скорейшего спуска: выбираем α для оптимального убывания.

Пример. $J(u) = \dfrac{1}{2}x^2 + \dfrac{1}{4}y^4 - \dfrac{1}{2}y^2 \to \inf, \ u \in \mathbb{R}^2$. *Очевидно, что $J_* = -\dfrac{1}{4}$, $U_* = \{(0,1), (0,-1)\}$. Решим с помощью градиентного метода: $J'(u) = (x, y^3 - y)$*
$u_0 = (1, -1)$, $\alpha_k = \dfrac{1}{2}$. $J'(u_0) = (1, 0)$ $u_1 = (\dfrac{1}{2}, -1)$, $u_2 = (\dfrac{1}{4}, -1), \dots, u_k = (2^{-k}, -1)$.

Метод проекции градиента

Отличается тем, что теперь ищем оптимум не во всем пространстве, а на $U \neq H$. Тогда в какой-то момент значение функционала в точке, не принадлежащей множеству U неопределено. Исправляем так:
$u_{k+1} = P_u(u_k - \alpha_k J'(u_k)),\ \alpha_k > 0, k = 0, 1, 2, \dots$.

Теорема. *(о сходимост МПГ)*
Пусть U - выпуклое, замкнутое множество из H и $J(u) \in C^1(U)$ и градиент $J(u)$ удовлетворяет условию Липшица с константой $L > 0$ на U. Пусть $J(u)$ сильно выпуклый на U с константой $æ > 0$ и коэффициенты $\alpha_k = \alpha \in (0, \dfrac{2æ}{L^2})$. Тогда при $\forall$ начальном условии последовательность u_k сходится к решению u_* и справедлива оценка: $\|u_k - u_*\| \leqslant q^k \|u_0 - u_*\|,\ q = \sqrt{1 - 2æ\alpha + \alpha^2 L^2}$

Метод скорейшего спуска

$\alpha_k = \mathrm{Argmin}_{\alpha \geqslant 0} J(u_k - \alpha J'(u_k))$.

Пример. *$J(u) = \|u\|^2,\ J'(u) = 2u$. Пусть $u_0 \in H$, тогда $u_1 = u_0 - \alpha J'(u_0) = u_0 - 2\alpha u_0 = (1 - 2\alpha)u_0$. $J(u_0 - \alpha J'(u_0)) = \|(1 - 2\alpha)u_0)\|^2 = (1 - 2\alpha)^2 \|u_0\|^2 \Rightarrow \alpha_0 = \dfrac{1}{2}$*

16 Метод Ньютона. Его сходимость

Решение задачи условной минимизации: $J(u) \to \inf_{u \in U},\ U \neq H$.
Идея: Пусть уже известно k-ое приближение $u_k \in U$. Берем квадратичную часть приращения $J(u) - J(u_k) = J_k(u) + \overline{o}(\|u - u_k\|^2)$, где $J_k(u) = (J'(u_k), u - u_k) + \dfrac{1}{2}(J''(u_k)(u - u_k), u - u_k),\ u \in U$. И вычисляем $u_{k+1} = \mathrm{Argmin}_{u \in U} J_k(u)$

Теорема. *(О сходимости метода Ньютона)*
Пусть U - выпуклое замкнутое множество, $intU \neq \varnothing$, $J(u)$ сильно выпукла с константой $æ > 0$ на U, $J(u) \in C^2(U)$, $J''(u)$ удовлетворяет на U условию Липшица с константой $L > 0$. Пусть начальное приближение удовлетворяет условию $\|u_0 - u_*\| < \dfrac{2æ}{L}$, u_* -решение задачи и $(J''(u)\xi, \xi) \geqslant æ\|\xi\|^2$. Тогда метода Ньютона порождает последовательность $\{u_k\}:\ \|u_k - u_*\| \leqslant \dfrac{2æ}{L} q^{2^k}$, $q = \dfrac{L\|u_0 - u_*\|}{2æ} < 1$

17 Метод покоординатного спуска

В предыдущих методов нам требовалось вычислять градиент и гессиан функционала, но зачастую функционал не обладает нужной гладкостью. Рассмотрим этот метод для задаче минимизации без ограничений в конечномерном пространстве $J(u) \to \inf_{u \in \mathbb{R}^n}$. $\{e_i\}_{i=1}^{i=n}$ - базис. В дальнейшем будет использоваться бесконечный базис, поэтом доопределим $p_0 = e_1, p_1 = e_2, \dots, p_{n-1} = e_n, p_n = e_1, \dots, p_{2n-1} = e_n, \dots$ (циклически повторяются).
Перед запуском метода выбираем $u_0 \in \mathbb{R}^n$, стартовый шаг $\alpha_0 > 0$ и коэффициент дробления шага $\lambda \in (0, 1)$. Пусть найдено k-е приближение u_k и текущее значение шага $\alpha_k > 0$. Найдем следующее приближение. Вычислим $u = u_k + \alpha_k p_k$

1. Если $J(u_k + \alpha_k p_k) < J(u_k)$, то $u_{k+1} = u_k + \alpha_k p_k$, $\alpha_{k+1} = \alpha_k$ и процесс продолжается со следующим по порядку базисным направлением p_{k+1}.

2. Если $J(u_k - \alpha_k p_k) < J(u_k)$, то $u_{k+1} = u_k - \alpha_k p_k$, $\alpha_{k+1} = \alpha_k$ и процесс продолжается со следующим по порядку базисным направлением p_{k+1}. Будем называть $k + 1$-ую итерацию удачной, если переход от u_k к u_{k+1} произошёл по этому или предыдущему пункту.

3. Если $J(u_k + \alpha_k p_k) \geqslant J(u_k)$, то итерация неудачная. В процессе вычислений ведётся подсчёт числа неудачных итераций, случившихся подряд. Если их число вместе с текущей не достигло n, то полагают $u_{k+1} = u_k$, $\alpha_{k+1} = \alpha_k$ и переходят к следующему базисному направлению. Иначе происходит дробление шага α_k с коэффициентом $\lambda : \alpha_{k+1} = \lambda \alpha_k$.

Теорема. *(о сходимости МПС)*
Пусть $J(u)$ выпуклый, $J(u) \in C^1(\mathbb{R}^n)$, u_0 -начальное приближение, множество Лебега $M_{J(u_0)} = \{u \in \mathbb{R}^n | J(u) \leqslant J(u_0)\}$ ограниченно. Тогда $\lim\limits_{k \to \infty} J(u_k) = J_*$, $\lim\limits_{k \to \infty} \rho(u_k, u_*) = 0$.

Пример. *$J \in C^1(\mathbb{R}^n)$ - существенно. $J(u) = (x-1)^2 + (y-1)^2 + 2|x-y| - 2$. $J_* = -2$, $U_* = \{(1, 1)\}$. Функционал не дифференцируем в $x = y$. Пусть $u_0 = (0, 0)$, тогда все $u_k = (0, 0)$. Все шаги неудачные.*

18 Метод штрафных функций и его сходимость

Позволяет решать задачи с большим количеством ограничений. Нарушая эти ограничения получаем "штрафы".
H - гильбретово, $J(u) \to \inf_{u \in U}, U \subset H$, $U = \{u \in U_0 | g_1(u) \leqslant 0, \dots, g_m(u) \leqslant 0, g_{m+1}(u) = 0, \dots g_{m+s}(u) = 0\}$
Функции g_j как раз задают ограничения. Неструктурированные ограничения, задаваемые множеством U_0

считаем "терпимыми"и обязуемся их не нарушать.
Будем использовать одни из самых распространённых штрафов: за нарушение неравенств будем применять индивидуальные штрафы типа срезки: $g_i^+(u) = max g_i(u), 0$. За нарушения равенств будетм испльзовать модули $g_i^+ = |g_i(u)|$. Из индивидуальных штрафов собирается общий $P(u) = \sum\limits_{i=1}^{m+s}(g_i^+(u))^{P_i},\ P_i \geqslant 1,\ i = \overline{1, m+s}$
Свойства штрафов:

1. $P(u) \geqslant 0$
2. $u \in U \Leftrightarrow \begin{cases} u \in U_0 \\ P(u) = 0 \end{cases} \quad \Leftrightarrow \begin{cases} u \in U_0 \\ g_i^+(u) = 0, i = \overline{1, m+s} \end{cases}$

Общий штраф добавляется к исходному функционалу $J(u)$ и получаем следующую задачу:

$$\Phi_k(u) = J(u) + A_k P(u) \to \inf_{u \in U_0}$$

Задача на терпимом множестве. $\Phi_{k*} = \inf\limits_{u \in U_0} \Phi_k(u) \leqslant \Phi_k(u_k) \leqslant \Phi_{k*} + \varepsilon_k,\ \varepsilon_k > 0.$

Определение. $\{P_k(u)\}$ *- штрафная функция множества U на множестве U_0, если*

1. *$P_k(u)$ определена на U_0 и неотрицательна на U_0*
2. $\lim\limits_{k \to \infty} P_k(u) = \begin{cases} 0, u \in U \\ +\infty, u \in U_0 \backslash U \end{cases}$

Теорема. *(Теорема о сходимости МШФ)*
Пусть H-гильбертово пространство, множество $U_0 \subset H$ слабо замкнут в H, исходная функция $J(u)$ и все индивидуальные штрафы $g_i^+(u)$ слабо полунепрерывны снизу на U_0. Пусть также нижняя грань $J(u)$ на U_0 конечна, а δ - расширение $U(\delta) = \{u \in U_0 | g_i^+(u) \leqslant \delta, i = \overline{1, m+s}\}$ допустимого множества U при некотором δ. Тогда если $A_k \to +\infty$ и $\varepsilon_k \to 0$, то для элементов u_k имеет место сходимость по функционалу $J(u_k) \to J_*$, а у самой последовательности $\{u_k\}_{k=1}^{\infty}$ имеется слабый в H предльные точки, причем каждая из них принадлежит U_*.

Пример. $J(u) = x^2 + xy + y^2 \to \inf$, $U = \{u \in \mathbb{R}^2 | x + y = 2\}$, $J_* = 3$, $u* = (1, 1)$.
$\Phi_k(u) = J(u) + A_k P(u) = x^2 + xy + y^2 + k(|x + y - 2)^2, k \in \mathbb{N}$.
$\Phi_k'(u) = 0 \Rightarrow \begin{pmatrix} 2x + y + 2k(x + y - 2) \\ x + 2y + 2k(x + y - 2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
$u_k = (x_k, y_k):\ x_k = y_k = \dfrac{4k}{3 + 4k} \to 1\ u_* = (1, 1)$.

19 Правило множителей Лагранжа

Та же задача, что и в предыдущем пункте.
Введём функцию Лагранжа $L(u, \lambda) = \lambda_0 J(u) + \sum\limits_{i=1}^{s} \lambda_j g_j(u),\ u \in U_0, \lambda = (\lambda_0, \ldots, \lambda_s), \lambda_j \geqslant 0.$

Теорема. *(Правило множителей Лагранжа)*
Пусть U_0 - выпуклое замкнутое множество, $u_* \in U$ - точка локального минимума $J(u)$. $g_i(u)$ - непрерывно дифференцируемы в окрестности u_*. Тогда $\exists \overline{\lambda^*} = (\lambda_0^*, \ldots, \lambda_s^*) :, \overline{\lambda^*} \neq 0$. $\lambda_j \geqslant 0, j = \overline{0, m}$. $(\dfrac{dL}{du}(u_*, \overline{\lambda^*}), u - u_*) \geqslant 0, \forall u \in U_0,\ \lambda_j^* g_j(u_*) = 0, j = \overline{1, m}$. L - выпукла по u. Тогда $u_* \in Agrmin_{\ u \in U_0} L(u, \overline{\lambda^*})$

20 Теорема Куна-Такера

Задача как в предыдущем билете.
Введём функцию Лагранжа $L(u, \lambda) = \lambda_0 J(u) + \sum\limits_{i=1}^{s} \lambda_j g_j(u),\ u \in U_0, \lambda = (\lambda_0, \ldots, \lambda_s), \lambda_j \geqslant 0$. И возьмём $\lambda_0 = 1$.
$\Lambda_0 = \{\lambda = (\lambda_1, \ldots, \lambda_s) | \lambda_1 \geqslant 0, \ldots, \lambda_m \geqslant 0\}$

Определение. *Точку $(u_*, \lambda^*) \in U_0 \times \Lambda_0$ называют седловой точкой функции Лагранжа, если $L(u_*, \lambda) \leqslant L(u_*, \lambda^*) \leqslant L(u, \lambda^*),\ \forall u \in U_0, \forall \lambda \in \Lambda_0$.*

Теорема. *(достаточное услоиве оптимальности)*
Пусть функции $J(u), g_i(u), i = \overline{1, s}$ определены и конечны на U_0. Пусть (u_*, λ^*) - седловая точка функции $L(u, \lambda)$. Тогда $u_* \in U_*$, $J(u_*) = J_* = L(u_*, \lambda^*)$.