

Методы оптимизации

Сюй Минчуань

27 января 2021 г.

Содержание

1	Теоремы Вейерштрасса	4
1.1	Метрическая теорема Вейерштрасса	4
1.1.1	Классическая теорема Вейерштрасса	4
1.1.2	Метрические пространства	4
1.1.3	Компактное множество	5
1.1.4	Метрическая теорема Вейерштрасса	5
1.2	Слабая теорема Вейерштрасса	5
1.2.1	Нормированное пространство	5
1.2.2	Евклидово пространство	5
1.2.3	Слабые сходимость и компактность	5
1.2.4	Слабые непрерывность и полунепрерывность снизу	6
1.2.5	Слабая теорема Вейерштрасса	6
1.3	Свойства простейших функционалов	6
1.3.1	Линейный функционал	7
1.3.2	Квадратичный функционал типа невязки	7
1.3.3	Квадратичный функционал общего вида	7
1.4	Свойства простейших множеств	7
1.4.1	Невырожденный эллипсоид в \mathbb{H}	7
1.4.2	Параллелепипед в $L^2(a, b)$	7
2	Элементы дифференциального исчисления	7
2.1	Дифференцируемость по Фреше	7
2.1.1	Производные по Фреше	7
2.1.2	Свойства производной	8
2.1.3	Градиент и гессиан	8
2.2	Важные формулы дифференцирования	9
2.2.1	Формула Тейлора для функционала	9
2.2.2	Формула конечных приращений первого порядка	9
2.2.3	Формула конечных приращений второго порядка	9
2.3	Примеры дифференциальных вычислений	9
2.3.1	Линейный функционал	9
2.3.2	Квадратичный функционал типа невязки	9
2.3.3	Квадратичный функционал общего вида	9
2.3.4	$J(u) = \ u\ _{\mathbb{H}}^p$	9
2.3.5	Функционал вида интеграла	10
2.4	Необходимые условия экстремума	10
3	Управление линейными динамическими системами	10
3.1	Постановка задачи	10
3.2	Упрощение исходной постановки	10
3.2.1	Упрощение	10
3.2.2	Операторная постановка	11
3.2.3	Теорема о решениях задачи	11

3.3	Вычисление градиентов $J_1(u)$	11
3.3.1	Вычисление $J'_1(u)$	11
3.3.2	Вычисление $J'_2(u)$	11
4	Элементы выпуклого анализа	11
4.1	Основные понятия	11
4.2	Критерии сильной выпуклости	12
4.2.1	Критерий сильной выпуклости 1-го порядка	12
4.2.2	Критерий сильной выпуклости 2-го порядка	12
4.2.3	Критерии выпуклости	12
4.2.4	Выпуклость и сильная выпуклость простейших функционалов	12
4.3	Свойства и теоремы для выпуклых функционалов	13
4.3.1	Свойство точек минимума выпуклых функционалов	13
4.3.2	Множество Лебега	13
4.3.3	Сильно выпуклая теорема Вейерштрасса	13
4.3.4	Критерий оптимальности для выпуклых задач	13
4.4	Метрическая проекция	13
4.4.1	Свойства (критерий) метрической проекции	13
4.4.2	Проектирование на гиперплоскость	13
4.4.3	Проектирование на шар	14
4.4.4	Проектирование на замкнутые подпространства	14
4.4.5	Проектирование на параллелепипед в $L^2(a, b)$	14
4.4.6	Проектирование на параболоид в l^2	14
4.4.7	Проекционная форма критерия оптимальности	14
5	Итерационные методы минимизации	15
5.1	Градиентный метод	15
5.1.1	Метод скорейшего спуска и его сходимость	15
5.1.2	Геометрический смысл метода скорейшего спуска	15
5.1.3	Непрерывный аналог градиентного метода	15
5.2	Метод проекции градиента	16
5.3	Метод условного градиента	16
5.4	Метод Ньютона	17
5.4.1	Классический метод и его сходимость	17
5.4.2	Метод Ньютона с выбором шага по Армихо	17
5.5	Квазиньютоновские методы	17
5.5.1	Семейство методов ранга один	18
5.5.2	Семейство методов ранга два	18
5.6	Линейное программирование	19
5.6.1	Постановка задачи	19
5.6.2	Каноническая задача	19
5.6.3	Угловые точки	19
5.6.4	Симплекс-метод	19
5.7	Метод сопряженных градиентов	20
5.7.1	Сходимость	21
5.7.2	Случай бесконечномерного пространства	21
5.8	Метод покоординатного спуска	21
5.8.1	Схема метода	22
5.8.2	Сходимость метода	22
6	Методы снятия ограничений	22
6.1	Метод штрафных функций	22
6.2	Правило множителей Лагранжа	23
6.2.1	Теорема Куна-Таккера и условие Слейтера	23
6.2.2	Правило множителей Лагранжа для гладких задач	23
6.3	Двойственные задачи	24

7	Принцип максимума Л.С.Понтрягина	25
7.1	Постановка задачи оптимального управления	25
7.2	Принцип максимума	25
7.3	Схема применения	25
8	Регуляризация по А.Н.Тихонову	26
8.1	Некорректно поставленные и неустойчивые экстремальные задачи	26
8.2	Регуляризация Тихонова	26

1 Теоремы Вейерштрасса

Постановка задачи $J(u) \rightarrow \inf, \quad u \in \mathbf{U} \in \mathbb{M}$. Искомая величина - $J_* = \inf_{u \in \mathbf{U}} J(u)$. Множество оптимальных решений - $\mathbf{U}_* = \operatorname{Argmin}_{u \in \mathbf{U}} J(u) = \{u \in \mathbf{U} : J(u) = J_*\}$. Оптимальный элемент - $u_* \in \mathbf{U}_*$.

1.1 Метрическая теорема Вейерштрасса

1.1.1 Классическая теорема Вейерштрасса

Теорема Вейерштрасса в \mathbb{R}^n Пусть \mathbf{U} - замкнутое ограниченное множество пространства \mathbb{R}^n , функция $J(u)$ непрерывна на \mathbf{U} . Тогда $J_* > -\infty, \mathbf{U}_* \neq \emptyset$.

Определение Последовательность $\{u_n\} \in \mathbf{U}$ называется минимизирующей в задаче минимизации, если

$$\lim_{n \rightarrow \infty} J(u_n) = J_*$$

Определение Функция $J(u)$, определенная на множестве \mathbf{U} пространства \mathbb{R}^n , называется полунепрерывной снизу в точке u_0 этого множества, если любой сходящейся к u_0 последовательности $\{u_n\}$ элементов множества выполнено предельное соотношение

$$\liminf_{n \rightarrow \infty} J(u_n) \geq J(u_0)$$

Теорема Пусть \mathbf{U} - замкнутое ограниченное множество пространства \mathbb{R}^n , функция $J(u)$ полунепрерывна снизу во всех точках множества \mathbf{U} . Тогда $J_* > -\infty, \mathbf{U}_* \neq \emptyset$, и кроме того, все минимизирующие последовательности сходятся к множеству \mathbf{U}_* , то есть

$$\lim_{n \rightarrow \infty} \inf_{u \in \mathbf{U}_*} \|u_n - u\| = 0$$

1.1.2 Метрические пространства

Определение Пространство \mathbb{M} называется метрическим, если введен функционал $\rho : \mathbb{M} \times \mathbb{M} \rightarrow \mathbb{R}^1$, обладающий следующими свойствами:

1. $\rho(u, v) = 0 \Leftrightarrow u = v$
2. $\rho(u, v) = \rho(v, u) \geq 0 \quad \forall u, v \in \mathbb{M}$
3. $\rho(u, v) \leq \rho(u, w) + \rho(w, v) \quad \forall u, v, w \in \mathbb{M}$

Этот функционал ρ называется метрикой или расстоянием в пространстве \mathbb{M} .

Определение Последовательность $\{u_n\}$ элементов метрического пространства \mathbb{M} с введенной на нем метрикой ρ называется сильно сходящейся к элементу $u \in \mathbb{M}$, если $\lim_{n \rightarrow \infty} \rho(u_n, u) = 0$.

Определение Последовательность $\{u_n\}$ элементов метрического пространства \mathbb{M} называется фундаментальной, если

$$\lim_{\substack{m \rightarrow \infty \\ n \rightarrow \infty}} \rho(u_m, u_n) = 0$$

Определение Метрическое пространство \mathbb{M} называется полным, если для любой фундаментальной последовательности $\{u_n\}$ его элементов существует элемент $u \in \mathbb{M}$, к которому она сильно сходится.

Определение Функционал $J(u) : \mathbb{M} \rightarrow \mathbb{R}^1$ называется непрерывным в точке u_0 , если для любой сильно сходящейся к u_0 последовательности выполнено

$$\lim_{n \rightarrow \infty} J(u_n) = J(u_0)$$

Определение Функционал $J(u) : \mathbb{M} \rightarrow \mathbb{R}^1$ называется полунепрерывным снизу в точке u_0 , если для любой сильно сходящейся к u_0 последовательности выполнено

$$\liminf_{n \rightarrow \infty} J(u_n) \geq J(u_0)$$

1.1.3 Компактное множество

Определение Множество \mathbf{U} из метрического пространства \mathbb{M} с введенной метрикой ρ называется компактным, если из любой последовательности $\{u_n\} \subset \mathbf{U}$ можно выделить подпоследовательность, сильно сходящуюся в метрике ρ к некоторому элементу $u \in \mathbf{U}$.

В конечномерном случае компактность эквивалентно его замкнутости и ограниченности. В бесконечномерном случае - единичный шар - некомпактное множество. Из компактности следует его замкнутость и ограниченность.

1.1.4 Метрическая теорема Вейштрасса

Теорема Пусть \mathbf{U} - компактное множество из метрического пространства \mathbb{M} с метрикой ρ , функционал $J(u)$ определен и полунепрерывен снизу на \mathbf{U} . Тогда

1. $J_* > -\infty$
2. Множество \mathbf{U}_* непусто и компактно.
3. Любая минимизирующая последовательность $\{u_n\}$ сильно сходится к множеству \mathbf{U}_* , т.е.

$$\lim_{n \rightarrow \infty} \inf_{u \in \mathbf{U}_*} \rho(u_n, u) = 0$$

Если выполнено $J_* > -\infty$, $\mathbf{U}_* \neq \emptyset$, и любая минимизирующая последовательность сходится к множеству \mathbf{U}_* , то говорят, что задача оптимизации корректно поставлена.

1.2 Слабая теорема Вейерштрасса

1.2.1 Нормированное пространство

Определение Линейное пространство \mathbb{L} называется нормированным, если введен функционал $\|\cdot\| : \mathbb{L} \rightarrow \mathbb{R}^1$, обладающий следующим свойством:

1. $\|u\| = 0 \Leftrightarrow u = \Theta$;
2. $\|u + v\| \leq \|u\| + \|v\| \quad \forall u, v \in \mathbb{L}$;
3. $\|\lambda u\| = |\lambda| \cdot \|u\| \quad \forall u \in \mathbb{L}, \lambda \in \mathbb{R}^1$.

Этот функционал называется нормой пространства \mathbb{L} .

Определение $\rho(u, v) = \|u - v\|$ - метрика, порожденная нормой.

Определение Последовательность $\{u_n\}$ сильно сходится по норме к элементу u , если $\lim_{n \rightarrow \infty} \|u_n - u\| = 0$.

Определение Нормированное пространство, полное относительно метрики, порожденной введенной на нем нормой, называется банаховым.

1.2.2 Евклидово пространство

Определение Линейное пространство \mathbb{H} называется евклидовым, если введен функционал $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}$, обладающий следующим свойством:

1. $\langle h, h \rangle \geq 0 \quad \forall h \in \mathbb{H}; \langle h, h \rangle = 0 \Leftrightarrow h = \Theta$;
2. $\langle h_1, h_2 \rangle = \langle h_2, h_1 \rangle \quad \forall h_1, h_2 \in \mathbb{H}$;
3. $\langle h_1 + h_2, h_3 \rangle = \langle h_1, h_3 \rangle + \langle h_2, h_3 \rangle \quad \forall h_1, h_2, h_3 \in \mathbb{H}$;
4. $\langle \lambda h_1, h_2 \rangle = \lambda \langle h_1, h_2 \rangle \quad \forall h_1, h_2 \in \mathbb{H}, \forall \lambda \in \mathbb{R}$.

Этот функционал называется скалярным произведением в пространстве \mathbb{H} .

Определение $\|h\| = \sqrt{\langle h, h \rangle}$ - норма порожденная скалярным произведением. $\rho(u, v) = \sqrt{\langle u - v, u - v \rangle}$ - метрика, порожденная скалярным произведением.

Определение Евклидово пространство, полное относительно метрики, порожденной введенной на нем скалярным произведением, называется гильбертовым.

1.2.3 Слабые сходимости и компактность

Определение Последовательность $\{u_n\}$ элементов евклидова пространства \mathbb{H} с введенным на нем скалярным произведением $\langle \cdot, \cdot \rangle$ называется слабо сходящейся к элементу u_0 этого пространства, если

$$\lim_{n \rightarrow \infty} \langle u_n, h \rangle = \langle u_0, h \rangle$$

В бесконечномерном случае из сильно сходимости вытекает слабую сходимость.

Определение Множество \mathbf{U} из евклидова пространства \mathbb{H} называется слабо компактным, если из любой последовательности $\{u_n\} \subset \mathbf{U}$ можно выделить подпоследовательность, слабо сходящуюся в метрике ρ к некоторому элементу $u \in \mathbf{U}$.

В бесконечномерном случае из компактности вытекает слабую компактность.

Определение Множество \mathbf{U} из линейного пространства \mathbb{L} называется выпуклым, если для любых двух точек u и v из множества \mathbf{U} и любого $\alpha \in [0, 1]$ точка $\alpha u + (1 - \alpha)v$ также лежит в множестве \mathbf{U} .

Определение Множество \mathbf{U} из нормированного пространства \mathbb{L} с введенной нормой $\|\cdot\|$ называется ограниченным, если существует такое число $R > 0$, что $\|u\| \leq R \quad \forall u \in \mathbf{U}$.

Определение Множество \mathbf{U} из метрического пространства \mathbb{M} с введенной метрикой ρ называется замкнутым, если \mathbf{U} содержит все свои предельные точки, т.е. из $\{u_n\} \subset \mathbf{U}, \{u_n\} \xrightarrow{\rho} u$ следует $u \in \mathbf{U}$.

Теорема - достаточное условия слабо компактности Если множество \mathbf{U} из евклидова пространства \mathbb{H} с введенным на нем скалярным произведением $\langle \cdot, \cdot \rangle$ выпукло, замкнуто и ограничено, то оно слабо компактно.

Обратно: если слабо компактно то замкнуто и ограничено, но не следует выпуклость.

В любом \mathbb{H} единичный шар - слабо компактно.

1.2.4 Слабые непрерывность и полунепрерывность снизу

Определение Функционал $J(u) : \mathbb{H} \rightarrow \mathbb{R}^1$ называется слабо непрерывным в точке u_0 , если для любой слабо сходящейся к u_0 последовательности $\{u_n\}$ выполнено

$$\lim_{n \rightarrow \infty} J(u_n) = J(u_0)$$

Из слабой непрерывности следует его непрерывность.

Определение Функционал $J(u) : \mathbb{H} \rightarrow \mathbb{R}^1$ называется слабо полунепрерывным снизу в точке u_0 , если для любой слабо сходящейся к u_0 последовательности $\{u_n\}$ выполнено

$$\liminf_{n \rightarrow \infty} J(u_n) \geq J(u_0)$$

Из слабой полунепрерывности снизу следует его полунепрерывности снизу.

Определение Функционал $J(u)$ называется выпуклым на выпуклом множестве \mathbf{U} , если выполнено $J(\alpha u + (1 - \alpha)v) \leq \alpha J(u) + (1 - \alpha)J(v) \quad \forall u, v \in \mathbf{U}, \forall \alpha \in [0, 1]$.

Теорема - достаточное условия слабо полунепрерывности снизу Пусть \mathbf{U} - выпуклое множество из евклидова пространства \mathbb{H} с введенным на нем скалярным произведением $\langle \cdot, \cdot \rangle$. Если функционал $J(u)$ полунепрерывен снизу в каждой точке множества \mathbf{U} и является выпуклым на множестве \mathbf{U} , то он слабо полунепрерывен снизу в каждой точке множества \mathbf{U} .

1.2.5 Слабая теорема Вейштрасса

Теорема Пусть \mathbf{U} - слабо компактное множество из евклидова пространства \mathbb{H} со скалярным произведением $\langle \cdot, \cdot \rangle$, функционал $J(u)$ определен и слабо полунепрерывен снизу на \mathbf{U} . Тогда

1. $J_* > -\infty$
2. Множество U_* непусто и компактно.
3. Любая минимизирующая последовательность $\{u_n\}$ слабо сходится к множеству U_* , т.е. все ее слабые предельные точки принадлежат множеству U_* .

$$\lim_{n \rightarrow \infty} \inf_{u \in U_*} \rho(u_n, u) = 0$$

1.3 Свойства простейших функционалов

Пусть \mathbb{H}, \mathbb{F} - евклидова пространства с введенным на нем скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}, \langle \cdot, \cdot \rangle_{\mathbb{F}}$.

1.3.1 Линейный функционал

$$J(u) = \langle c, u \rangle, \quad c \in \mathbb{H}$$

Он слабо непрерывен в любой точке, слабо полунепрерывен снизу, непрерывен и полунепрерывен снизу на всем пространстве.

1.3.2 Квадратичный функционал типа невязки

$$J(u) = \|Au - f\|_{\mathbb{F}}^2, \quad f \in \mathbb{F}$$

где $A \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{F})$.

Он непрерывен, полунепрерывен снизу, слабо полунепрерывен снизу, нет слабой непрерывности.

1.3.3 Квадратичный функционал общего вида

$$J(u) = \langle Au, u \rangle_{\mathbb{H}} + \langle b, u \rangle_{\mathbb{H}} + c, \quad b \in \mathbb{H}, c \in \mathbb{R}$$

где $A \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{H})$.

Он непрерывен, полунепрерывен снизу, может не быть слабой непрерывности и слабой полунепрерывности снизу. Но если A - неотрицательно определена, то слабая полунепрерывность снизу есть.

1.4 Свойства простейших множеств

Пусть \mathbb{H}, \mathbb{F} - евклидова пространства с введенным на нем скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}, \langle \cdot, \cdot \rangle_{\mathbb{F}}$. $A \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{F})$ и $\exists A^{-1} \in \mathcal{L}(\mathbb{F} \rightarrow \mathbb{H})$

1.4.1 Невырожденный эллипсоид в \mathbb{H}

$$U = \{u \in \mathbb{H} : \|Au - f\| \leq R\}, \quad R > 0, f \in \mathbb{F}$$

Множество U - выпукло, замкнуто и ограничено. \Rightarrow слабо компактно.

1.4.2 Параллелепипед в $L^2(a, b)$

$$U = \{u = u(t) \in L^2(a, b) : f(t) \stackrel{\text{п.в.}}{\leq} u(t) \stackrel{\text{п.в.}}{\leq} g(t), t \in (a, b)\}, \quad f = f(t), g = g(t) \in L^2(a, b)$$

Множество U - выпукло, замкнуто и ограничено. \Rightarrow слабо компактно. Но он не компактен.

2 Элементы дифференциального исчисления

2.1 Дифференцируемость по Фреше

2.1.1 Производные по Фреше

Определение Пусть \mathbb{X} и \mathbb{Y} - нормированные пространства с нормами $\|\cdot\|_{\mathbb{X}}, \|\cdot\|_{\mathbb{Y}}$, а $F : \mathbb{X} \rightarrow \mathbb{Y}$ - отображение, определенное в некоторой окрестности $O(x_0, \varepsilon) = \{x \in \mathbb{X} : \|x - x_0\|_{\mathbb{X}} \leq \varepsilon\}$ точки x_0 . Отображение F называется дифференцируемым по Фреше в точке x_0 , если справедливо равенство

$$F(x_0 + h) - F(x_0) = Ah + \alpha(h; x_0) \quad \forall h : \|h\|_{\mathbb{X}} \leq \varepsilon$$

В котором $A \in \mathcal{L}(\mathbb{X} \rightarrow \mathbb{Y})$, а остаточный член $\alpha(h; x_0)$ имеет по отношению к приращению h более высокий порядок малости, т.е.

$$\lim_{\|h\|_{\mathbb{X}} \rightarrow 0} \frac{\|\alpha(h; x_0)\|_{\mathbb{Y}}}{\|h\|_{\mathbb{X}}} = 0$$

При этом оператор A - производная Фреше отображения F в точке x_0 и обозначается $F'(x_0)$.

Определение Пусть \mathbb{X} и \mathbb{Y} - нормированные пространства с нормами $\|\cdot\|_{\mathbb{X}}, \|\cdot\|_{\mathbb{Y}}$, а $F : \mathbb{X} \rightarrow \mathbb{Y}$ - отображение, определенное в некоторой окрестности $O(x_0, \varepsilon) = \{x \in \mathbb{X} : \|x - x_0\|_{\mathbb{X}} \leq \varepsilon\}$ точки x_0 .

Отображение F называется дважды дифференцируемым по Фреше в точке x_0 , если отображение $F'(x) : \mathbb{X} \rightarrow \mathcal{L}(\mathbb{X} \rightarrow \mathbb{Y})$ дифференцируемо по Фреше в точке x_0 , то есть

$$F'(x_0 + h) - F'(x_0) = \mathcal{B}h + \beta(h; x_0) \quad \forall h : \|h\|_{\mathbb{X}} \leq \varepsilon$$

В котором $\mathcal{B} \in \mathcal{L}(\mathbb{X} \rightarrow \mathcal{L}(\mathbb{X} \rightarrow \mathbb{Y}))$, а остаточный член $\beta(h; x_0) \in \mathcal{L}(\mathbb{X} \rightarrow \mathbb{Y})$ имеет по отношению к приращению h более высокий порядок малости, т.е.

$$\lim_{\|h\|_{\mathbb{X}} \rightarrow 0} \frac{\|\beta(h; x_0)\|_{\mathcal{L}(\mathbb{X} \rightarrow \mathbb{Y})}}{\|h\|_{\mathbb{X}}} = 0$$

При этом оператор \mathcal{B} - вторая производная Фреше отображения F в точке x_0 и обозначается $F''(x_0)$.

2.1.2 Свойства производной

Теорема - о дифференцировании сложного отображения Пусть $\mathbb{X}, \mathbb{Y}, \mathbb{Z}$ - нормированные пространства, отображение $F : \mathbb{X} \rightarrow \mathbb{Y}$ определено в окрестности $O(x_0, \gamma)$ точки $x_0 \in \mathbb{X}$ и дифференцируемо в ней по Фреше, отображение $G : \mathbb{Y} \rightarrow \mathbb{Z}$ определено в окрестности $O(y_0, \delta)$ точки $y_0 = F(x_0)$ и дифференцируемо в ней по Фреше. Тогда отображение $GF : \mathbb{X} \rightarrow \mathbb{Z}$ дифференцируемо по Фреше в точке x_0 , причем $(GF)'(x_0) = G'(y_0)F'(x_0)$.

Свойство 1: Если отображение $F(x)$ дифференцируемо по Фреше в точке x_0 , то его производная по Фреше $F'(x_0)$ единственна.

Свойство 2: Если $F_1(x), F_2(x)$ - два дифференцируемых по Фреше в точке x_0 отображения, действующих из \mathbb{X} в \mathbb{Y} , то отображение $F(x) = \alpha F_1(x) + \beta F_2(x)$ ($\alpha, \beta \in \mathbb{R}$) также дифференцируемо по Фреше в точке x_0 , причем $F'(x_0) = \alpha F_1'(x_0) + \beta F_2'(x_0)$.

2.1.3 Градиент и гессиан

Определение Пусть \mathbb{H} - гильбертово пространство со скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$. Функционал $J(u)$, определенный в некоторой окрестности $O(u_0, \varepsilon) = \{x \in \mathbb{H} : \|u - u_0\|_{\mathbb{H}} \leq \varepsilon\}$ точки $u_0 \in \mathbb{H}$. Функционал $J(u)$ называется дифференцируемым по Фреше в точке u_0 , если справедливо равенство

$$J(u_0 + h) - J(u_0) = \langle c, h \rangle_{\mathbb{H}} + \alpha(h; u_0) \quad \forall h : \|h\|_{\mathbb{H}} \leq \varepsilon$$

В котором c - не зависящий от h элемент из \mathbb{H} , а остаточный член $\alpha(h; u_0)$ имеет по отношению к приращению h более высокий порядок малости, т.е.

$$\lim_{\|h\|_{\mathbb{H}} \rightarrow 0} \frac{\|\alpha(h; u_0)\|}{\|h\|_{\mathbb{H}}} = 0$$

Элемент c - *градиент* функционала $J(u)$ в точке u_0 и обозначается $c = J'(u_0)$.

Определение Пусть \mathbb{H} - гильбертово пространство со скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$. Функционал $J(u)$, определенный в некоторой окрестности $O(u_0, \varepsilon) = \{x \in \mathbb{H} : \|u - u_0\|_{\mathbb{H}} \leq \varepsilon\}$ точки $u_0 \in \mathbb{H}$. Функционал $J(u)$ называется дважды дифференцируемым по Фреше в точке u_0 , если отображение-градиент $J'(u) : \mathbb{H} \rightarrow \mathbb{H}$ дифференцируемо по Фреше в точке u_0 , то есть справедливо

$$J'(u_0 + h) - J'(u_0) = \mathcal{A}h + \beta(h; u_0) \quad \forall h : \|h\|_{\mathbb{H}} \leq \varepsilon$$

В котором $\mathcal{A} \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{H})$ - не зависящий от h , а остаточный член $\beta(h; u_0) \in \mathbb{H}$ имеет по отношению к приращению h более высокий порядок малости, т.е.

$$\lim_{\|h\|_{\mathbb{H}} \rightarrow 0} \frac{\|\beta(h; u_0)\|_{\mathbb{H}}}{\|h\|_{\mathbb{H}}} = 0$$

Оператор \mathcal{A} - *гессиан* функционала $J(u)$ в точке u_0 и обозначается $\mathcal{A} = J''(u_0)$.

Замечание Вторая производная по Фреше и гессиан - симметричные операторы.

2.2 Важные формулы дифференцирования

2.2.1 Формула Тейлора для функционала

Теорема Пусть \mathbb{H} - гильбертово пространство со скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$. Функционал $J(u) : \mathbb{H} \rightarrow \mathbb{R}$, определен и дифференцируем по Фреше в шаре $O(u_0, \varepsilon) = \{x \in \mathbb{H} : \|u - u_0\|_{\mathbb{H}} \leq \varepsilon\}$, и кроме того, существует $J''(u_0)$. Тогда справедлива формула Тейлора:

$$J(u) = J(u_0) + \langle J'(u_0), u - u_0 \rangle_{\mathbb{H}} + \frac{1}{2} \langle J''(u_0)(u - u_0), u - u_0 \rangle_{\mathbb{H}} + o(\|u - u_0\|_{\mathbb{H}}^2) \quad u \in O(u_0, \gamma)$$

2.2.2 Формула конечных приращений первого порядка

Теорема Пусть \mathbb{H} - гильбертово пространство со скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $\mathbf{U} \subseteq \mathbb{H}$ - выпуклое множество, $J(u) \in \mathbf{C}^1(\mathbf{U})$. Тогда для любых $u, u + h \in \mathbf{U}$ верно

$$J(u + h) - J(u) = \int_0^1 \langle J'(u + th), h \rangle_{\mathbb{H}} dt = \langle J'(u + \theta h), h \rangle_{\mathbb{H}}, \quad \theta \in [0, 1]$$

2.2.3 Формула конечных приращений второго порядка

Теорема Пусть \mathbb{H} - гильбертово пространство со скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $\mathbf{U} \subseteq \mathbb{H}$ - выпуклое множество, $J(u) \in \mathbf{C}^2(\mathbf{U})$. Тогда для любых $u, u + h \in \mathbf{U}, v \in \mathbb{H}$ верно

$$\langle J'(u + h) - J'(u), v \rangle_{\mathbb{H}} = \int_0^1 \langle J''(u + th)h, v \rangle_{\mathbb{H}} dt = \langle J''(u + \theta h)h, v \rangle_{\mathbb{H}}, \quad \theta \in [0, 1]$$

2.3 Примеры дифференциальных вычислений

2.3.1 Линейный функционал

$$\begin{aligned} J(u) &= \langle c, u \rangle_{\mathbb{H}}, \quad c \in \mathbb{H} \\ J'(u) &\equiv c, \quad J''(u) = \mathcal{O} \end{aligned}$$

2.3.2 Квадратичный функционал типа невязки

$$J(u) = \|Au - f\|_{\mathbb{F}}^2, \quad f \in \mathbb{F}$$

Где $\mathcal{A} \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{F})$.

$$J'(u) = 2\mathcal{A}^*(Au - f), \quad J''(u) \equiv 2\mathcal{A}^*\mathcal{A}$$

2.3.3 Квадратичный функционал общего вида

$$J(u) = \langle \mathcal{A}u, u \rangle_{\mathbb{H}}$$

Где $\mathcal{A} \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{H})$.

$$J'(u) = (\mathcal{A} + \mathcal{A}^*)u, \quad J''(u) \equiv \mathcal{A} + \mathcal{A}^*$$

2.3.4 $J(u) = \|u\|_{\mathbb{H}}^p$

$$\begin{aligned} J'(u) &= \begin{cases} p\|u\|_{\mathbb{H}}^{p-2}u, & u \neq \Theta_{\mathbb{H}} \\ \Theta_{\mathbb{H}}, & u = \Theta_{\mathbb{H}}, p > 1 \\ \nexists, & u = \Theta_{\mathbb{H}}, p \leq 1 \end{cases} \\ J''(u) &= \begin{cases} p(p-2)\|u\|_{\mathbb{H}}^{p-4}\langle u, h \rangle_{\mathbb{H}}u + p\|u\|_{\mathbb{H}}^{p-2}\mathcal{E}, & u \neq \Theta_{\mathbb{H}} \\ \Theta_{\mathbb{H}}, & u = \Theta_{\mathbb{H}}, p > 2 \\ 2\mathcal{E}, & p = 2 \\ \nexists, & u = \Theta_{\mathbb{H}}, p < 2 \end{cases} \end{aligned}$$

2.3.5 Функционал вида интеграла

$$J(u) = \|Au\|_{L^2(0,1)}^2 = \int_0^1 \left(\int_0^{\sqrt{t}} u(s) ds \right)^2 dt, \quad u(t) \in L^2(0,1)$$

где

$$\begin{aligned} \mathcal{A}u(t) &= \int_0^{\sqrt{t}} u(s) ds, \quad \mathcal{A}^*v(t) = \int_{t^2}^1 v(s) ds \\ J'(u) &= 2 \int_{t^2}^1 \int_0^{\sqrt{\tau}} u(s) ds d\tau, \quad J''(u)[h(t)] = 2 \int_{t^2}^1 \int_0^{\sqrt{\tau}} h(s) ds d\tau \end{aligned}$$

2.4 Необходимые условия экстремума

Теорема Пусть функционал $J(u)$ определен на гильбертовом пространстве \mathbb{H} , пусть $J_* = \inf_{u \in \mathbb{H}} J(u) = J(u_*)$. Если $J(u)$ дифференцируема по Фреше в точке u_* , то необходимо выполняется равенство $J'(u_*) = \Theta_{\mathbb{H}}$. Если кроме того, $J(u)$ дважды дифференцируем по Фреше в точке u_* , то необходимо выполняется ещё и условие $\langle J''(u_*)h, h \rangle_{\mathbb{H}} > 0 \quad h \in \mathbb{H}$.

3 Управление линейными динамическими системами

3.1 Постановка задачи

Линейная задача оптимального управления

$$\begin{cases} \dot{x}(t) \stackrel{\text{п.б.}}{=} A(t)x(t) + B(t)u(t) + f(t), \\ x(0) = x_0 \end{cases}$$

$$0 \leq t \leq T, A(t) \in L_{n \times n}^\infty(0, T), B(t) \in L_{n \times m}^\infty(0, T), f(t) = L_n^\infty(0, T), x_0 \in \mathbb{R}^n,$$

$$x(t) = (x_1(t), \dots, x_n(t))^T, u(t) = (u_1(t), \dots, u_m(t))^T \in \mathbf{U} \subseteq L_m^2(0, T)$$

Интегральный функционал:

$$J_1(u) = \int_0^T \|x(t; u) - y(t)\|_{\mathbb{R}^n}^2 dt \rightarrow \inf_{u \in \mathbf{U}}$$

Терминальный функционал:

$$J_2(u) = \|x(T; u) - y\|_{\mathbb{R}^n}^2 \rightarrow \inf_{u \in \mathbf{U}}$$

где y - заданный вектор из \mathbb{R}^n , $y(t)$ - заданная вектор-функция из $L_n^2(0, T)$.

$L^\infty(0, T)$ - пространство существенно ограниченных измеримых функций. Оно является банаховым с нормой:

$$\|u\|_{L^\infty(0, T)} = \operatorname{esssup}_{(0, T)} |u(t)| = \inf_{M > 0: |u(t)| \leq M} M = \lim_{p \rightarrow \infty} \left((L) \int_0^T |u(t)|^p dt \right)^{1/p}$$

Можно доказать существование и единственность решения задачи Коши, используя принцип сжимающих отображений после приведения задачи на интегральный вид.

3.2 Упрощение исходной постановки

3.2.1 Упрощение

$$\begin{cases} \dot{x}(t) \stackrel{\text{п.б.}}{=} A(t)x(t) + B(t)u(t), \\ x(0) = 0 \end{cases}$$

$$0 \leq t \leq T, A(t) \in L_{n \times n}^\infty(0, T), B(t) \in L_{n \times m}^\infty(0, T),$$

$$x(t) = (x_1(t), \dots, x_n(t))^T, u(t) = (u_1(t), \dots, u_m(t))^T \in \mathbf{U} \subseteq L_m^2(0, T)$$

$$J_1(u) = \int_0^T \|x(t; u) - z(t)\|_{\mathbb{R}^n}^2 dt \rightarrow \inf_{u \in \mathbf{U}} \quad \text{или} \quad J_2(u) = \|x(T; u) - z\|_{\mathbb{R}^n}^2 \rightarrow \inf_{u \in \mathbf{U}}$$

где z - заданный вектор из \mathbb{R}^n , $z(t)$ - заданная вектор-функция из $L_n^2(0, T)$.

3.2.2 Операторная постановка

Можно записать функционалы в операторном виде, и оказываются что, эти операторы линейные и ограниченные.

$$\begin{aligned} J_1(u) &= \|\mathcal{I}u - z\|_{L_n^2(0, T)}^2, \quad \mathcal{I} : L_m^2(0, T) \rightarrow L_n^2(0, T), \mathcal{I}u = x(t; u), \\ J_2(u) &= \|\mathcal{T}u - z\|_{\mathbb{R}^n}^2, \quad \mathcal{T} : L_m^2(0, T) \rightarrow \mathbb{R}^n, \mathcal{T}u = x(T; u). \end{aligned}$$

3.2.3 Теорема о решениях задачи

Теорема Пусть в задачах оптимального управления имеется постановка как начало параграфа, множество \mathbf{U} слабо компактно. Тогда $J_{1*} = \inf_{u \in \mathbf{U}} J_1(u) > -\infty$, $J_{2*} = \inf_{u \in \mathbf{U}} J_2(u) > -\infty$, $\mathbf{U}_{1*} \neq \emptyset$, $\mathbf{U}_{2*} \neq \emptyset$

3.3 Вычисление градиентов $J_1(u)$, $J_2(u)$

3.3.1 Вычисление $J_1'(u)$

$$J_1'(u) = 2\mathcal{I}^*(\mathcal{I}u - z)$$

Сопряженная задача:

$$\begin{cases} \dot{\psi}(t) + A^T(t)\psi(t) + v(t) \stackrel{\text{п.б.}}{=} 0, \\ \psi(T) = 0. \end{cases}$$

Схема вычисления $J_1'(u)$:

1. $u(t)$ подставляется в исходную задачу Коши, находим решение $\mathcal{I}u = x(t; u)$.
2. $v(t) = \mathcal{I}u - z = x(t; u) - z(t)$ подставляется в сопряженную задачу, находим $\psi(t) = \psi(t; u)$.
3. $J_1'(u) = 2B^T(t)\psi(t; u)$.

3.3.2 Вычисление $J_2'(u)$

$$J_2'(u) = 2\mathcal{T}^*(\mathcal{T}u - z)$$

Сопряженная задача:

$$\begin{cases} \dot{\psi}(t) + A^T(t)\psi(t) \stackrel{\text{п.б.}}{=} 0, \\ \psi(T) = v. \end{cases}$$

Схема вычисления $J_2'(u)$:

1. $u(t)$ подставляется в исходную задачу Коши, находим решение $\mathcal{T}u = x(T; u)$.
2. $v(t) = \mathcal{T}u - z = x(T; u) - z(t)$ подставляется в сопряженную задачу, находим $\psi(t) = \psi(t; u)$.
3. $J_2'(u) = 2B^T(t)\psi(t; u)$.

4 Элементы выпуклого анализа

4.1 Основные понятия

Определение Множество \mathbf{U} из линейного пространства \mathbb{L} называется выпуклым, если выполнено условие

$$\alpha u_1 + (1 - \alpha)u_2 \in \mathbf{U} \quad \forall u_1, u_2 \in \mathbf{U}, \alpha \in [0; 1]$$

Определение Функционал $J(u) : \mathbf{U} \rightarrow \mathbb{R}^1$ называется выпуклым на выпуклом множестве \mathbf{U} из линейного пространства \mathbb{L} , если выполнено условие

$$J(\alpha u_1 + (1 - \alpha)u_2) \leq \alpha J(u_1) + (1 - \alpha)J(u_2) \quad \forall u_1, u_2 \in \mathbf{U}, \alpha \in [0; 1]$$

Определение Функционал $J(u) : \mathbf{U} \rightarrow \mathbb{R}^1$ называется строго выпуклым на выпуклом множестве \mathbf{U} из линейного пространства \mathbb{L} , если выполнено условие

$$J(\alpha u_1 + (1 - \alpha)u_2) < \alpha J(u_1) + (1 - \alpha)J(u_2) \quad \forall u_1, u_2 \in \mathbf{U}, \alpha \in (0; 1)$$

Определение Функционал $J(u) : \mathbf{U} \rightarrow \mathbb{R}^1$ называется сильно выпуклым на выпуклом множестве \mathbf{U} из нормированного пространства \mathbb{H} с введенным на нем нормой $\|\cdot\|_{\mathbb{H}}$, если существует такое число $\mu > 0$, что выполнено условие

$$J(\alpha u_1 + (1 - \alpha)u_2) \leq \alpha J(u_1) + (1 - \alpha)J(u_2) - \frac{\mu}{2}\alpha(1 - \alpha)\|u_1 - u_2\|_{\mathbb{H}}^2 \quad \forall u_1, u_2 \in \mathbf{U}, \alpha \in [0; 1]$$

Максимальное из чисел μ , при которых это условие верно, называют константой сильной выпуклости функционала $J(u)$.

4.2 Критерии сильной выпуклости

4.2.1 Критерий сильной выпуклости 1-го порядка

Пусть \mathbf{U} - непустое выпуклое множество из евклидова пространства \mathbb{H} с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $J(u) \in C^1(\mathbf{U})$. Тогда функционал $J(u)$ является сильно выпуклым на множестве \mathbf{U} тогда и только тогда, когда при некотором $\mu > 0$ справедливо

$$J(u) \geq J(v) + \langle J'(v), u - v \rangle_{\mathbb{H}} + \frac{\mu}{2}\|u - v\|_{\mathbb{H}}^2 \quad \forall u, v \in \mathbf{U}$$

или справедливо неравенство

$$\langle J'(u) - J'(v), u - v \rangle_{\mathbb{H}} \geq \mu\|u - v\|_{\mathbb{H}}^2 \quad u, v \in \mathbf{U}$$

4.2.2 Критерий сильной выпуклости 2-го порядка

Лемма Пусть \mathbf{U} - выпуклое множество из нормированного пространства \mathbb{H} с введенной на нем нормой $\|\cdot\|_{\mathbb{H}}$, $\text{int}\mathbf{U} \neq \emptyset$, $u \in \text{int}\mathbf{U}$, $v \in \text{Gr}\mathbf{U}$. Тогда $\forall \alpha \in (0, 1] : v_{\alpha} = v + \alpha(u - v) \in \text{int}\mathbf{U}$.

Теорема Пусть \mathbf{U} - непустое выпуклое множество из евклидова пространства \mathbb{H} с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $\text{int}\mathbf{U} \neq \emptyset$, $J(u) \in C^2(\mathbf{U})$. Тогда функционал $J(u)$ является сильно выпуклым на множестве \mathbf{U} тогда и только тогда, когда при некотором $\mu > 0$ справедливо

$$\langle J''(u)h, h \rangle_{\mathbb{H}} \geq \mu\|h\|_{\mathbb{H}}^2 \quad \forall u \in \mathbf{U}, \forall h \in \mathbb{H}$$

4.2.3 Критерии выпуклости

Теорема Пусть \mathbf{U} - непустое выпуклое множество из евклидова пространства \mathbb{H} с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $J(u) \in C^1(\mathbf{U})$. Тогда функционал $J(u)$ является выпуклым на множестве \mathbf{U} тогда и только тогда, когда при некотором $\mu > 0$ справедливо

$$J(u) \geq J(v) + \langle J'(v), u - v \rangle_{\mathbb{H}} \quad \forall u, v \in \mathbf{U}$$

или справедливо неравенство

$$\langle J'(u) - J'(v), u - v \rangle_{\mathbb{H}} \geq 0 \quad u, v \in \mathbf{U}$$

Теорема Пусть \mathbf{U} - выпуклое множество из евклидова пространства \mathbb{H} с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $\text{int}\mathbf{U} \neq \emptyset$, $J(u) \in C^2(\mathbf{U})$. Тогда функционал $J(u)$ является сильно выпуклым на множестве \mathbf{U} тогда и только тогда, когда при некотором $\mu > 0$ справедливо

$$\langle J''(u)h, h \rangle_{\mathbb{H}} \geq 0 \quad \forall u \in \mathbf{U}, \forall h \in \mathbb{H}$$

4.2.4 Выпуклость и сильная выпуклость простейших функционалов

1. $J(u) = \langle c, u \rangle_{\mathbb{H}}$, $c \in \mathbb{H}$ - выпуклый.
2. $J(u) = \|\mathcal{A}u - f\|_{\mathbb{F}}^2$, $\mathcal{A} \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{F})$, $f \in \mathbb{F}$ - выпуклый. сильно выпуклый при существовании обратного оператора $\mathcal{A}^{-1} \in \mathcal{L}(\mathbb{F} \rightarrow \mathbb{H})$. (или $2\|\mathcal{A}(u - v)\|_{\mathbb{F}}^2 \geq \mu\|u - v\|_{\mathbb{H}}^2$ при некотором $\mu > 0$.)
3. $J(u) = \langle \mathcal{A}u, u \rangle_{\mathbb{H}}$, $\mathcal{A} \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{H})$ - выпуклый. сильно выпуклый при неотрицательной определенности оператора \mathcal{A} .

4.3 Свойства и теоремы для выпуклых функционалов

4.3.1 Свойство точек минимума выпуклых функционалов

Теорема Пусть \mathbf{U} - непустое выпуклое множество из нормированного пространства \mathbb{H} с введенной на нем нормой $\|\cdot\|_{\mathbb{H}}$, функционал $J(u)$ является выпуклым на \mathbf{U} . Тогда

1. Любая точка локального минимума $J(u)$ на множестве \mathbf{U} является точкой его глобального минимума.
2. Если множество \mathbf{U}_* непусто, то оно выпукло.
3. Если функционал $J(u)$ является строго выпуклым на \mathbf{U} , а множество \mathbf{U}_* непусто, то оно состоит из единственной точки u_* .

4.3.2 Множество Лебега

$$\mathbf{U}(v) = \{u \in \mathbf{U} : J(u) \leq J(v)\}.$$

Лемма Пусть \mathbf{U} - выпуклое замкнутое множество из нормированного пространства \mathbb{H} с введенной на нем нормой $\|\cdot\|_{\mathbb{H}}$, функционал $J(u)$ является сильным выпуклым и полунепрерывным снизу на \mathbf{U} . Тогда множество $\mathbf{U}(v) = \{u \in \mathbf{U} : J(u) \leq J(v)\}$ непусто, выпукло, замкнуто и ограничено (как следствие, слабо компактно) при $\forall v \in \mathbf{U}$.

4.3.3 Сильно выпуклая теорема Вейерштрасса

Пусть \mathbf{U} - выпуклое замкнутое множество из нормированного пространства \mathbb{H} с введенной на нем нормой $\|\cdot\|_{\mathbb{H}}$, функционал $J(u)$ является сильным выпуклым с константой μ и полунепрерывен снизу на \mathbf{U} . Тогда

1. $J_* > -\infty$, множество \mathbf{U}_* непусто и состоит из единственного элемента u_* .
2. Справедливо $\frac{\mu}{2}\|u - u_*\|_{\mathbb{H}}^2 \leq J(u) - J(u_*) \quad \forall u \in \mathbf{U}$.

4.3.4 Критерий оптимальности для выпуклых задач

Пусть \mathbf{U} - выпуклое множество из евклидова пространства \mathbb{H} с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $J(u) \in C^1(\mathbf{U})$. Тогда

1. Если $u_* \in \mathbf{U}_*$, то верно неравенство $\langle J'(u_*), u - u_* \rangle_{\mathbb{H}} \geq 0 \quad \forall u \in \mathbf{U}$,
2. Если $u_* \in \mathbf{U}_* \cup \text{int} \mathbf{U}$, то $J'(u_*) = \Theta_{\mathbb{H}}$.
3. Если функционал является выпуклым на \mathbf{U} , то неравенство в пункте 1. и является достаточным.

4.4 Метрическая проекция

Пусть \mathbf{U} - множество из метрического пространства \mathbb{M} с введенной на нем метрикой $\rho(\cdot, \cdot)$. Метрической проекцией (или просто *проекцией*) элемента h на множество \mathbf{U} называется такой элемент $p \in \mathbf{U}$, что $\rho(p, h) = \inf_{u \in \mathbf{U}} \rho(u, h)$. Обозначение: $\mathcal{P}_{\mathbf{U}}(h)$.

4.4.1 Свойства (критерий) метрической проекции

Пусть \mathbb{H} - гильбертово пространство с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, $\mathbf{U} \subseteq \mathbb{H}$ - выпуклое замкнутое множество. Тогда справедливы:

1. $\forall h \in \mathbb{H} \quad \exists! \mathcal{P}_{\mathbf{U}}(h)$.
2. $p = \mathcal{P}_{\mathbf{U}}(h) \leftrightarrow \langle p - h, u - p \rangle_{\mathbb{H}} \geq 0 \quad \forall u \in \mathbf{U}$. - *характеристическое свойство проекции*.
3. $\|\mathcal{P}_{\mathbf{U}}(h_1) - \mathcal{P}_{\mathbf{U}}(h_2)\|_{\mathbb{H}} \leq \|h_1 - h_2\|_{\mathbb{H}} \quad \forall h_1, h_2 \in \mathbb{H}$ - свойство нестрогой сжимаемости.

4.4.2 Проектирование на гиперплоскость

Гиперплоскость в евклидовом пространстве \mathbb{H} - это множество

$$\mathbf{U} = \{u \in \mathbb{H} : \langle c, u \rangle_{\mathbb{H}} = \alpha\}, \quad c \in \mathbb{H}, c \neq \Theta_{\mathbb{H}}, \alpha \in \mathbb{R}^1$$

Вектор c ортогонален любому вектору, лежащему в гиперплоскости.

$$p = \mathcal{P}_{\mathbf{U}}(h) = h + \frac{\alpha - \langle c, h \rangle_{\mathbb{H}}}{\|c\|_{\mathbb{H}}^2} c$$

4.4.3 Проектирование на шар

Шар с центром u_0 в точке радиуса R в евклидовом пространстве \mathbb{H} :

$$B = \{u \in \mathbb{H} : \|u - u_0\|_{\mathbb{H}} \leq R\}, \quad u_0 \in \mathbb{H}, R > 0$$

$$\mathcal{P}_U(h) = \begin{cases} u_0 + \frac{R}{\|h - u_0\|_{\mathbb{H}}} (h - u_0), & \|h - u_0\|_{\mathbb{H}} > R \\ h, & \|h - u_0\|_{\mathbb{H}} \leq R \end{cases}$$

4.4.4 Проектирование на замкнутые подпространства

Замкнутое линейное подпространство $\mathbf{L} = \mathcal{L}(a_1, a_2, \dots, a_n)$, a_1, a_2, \dots, a_n - ЛНЗ элементы. Пусть построена ортогональная система ЛНЗ $e_1, e_2, \dots, e_n \in \mathbf{L}$.

$$p = \mathcal{P}_L(h) = \sum_{i=1}^n \frac{\langle h, e_i \rangle_{\mathbb{H}}}{\|e_i\|_{\mathbb{H}}^2} e_i$$

4.4.5 Проектирование на параллелепипед в $L^2(a, b)$

Параллелепипед в $L^2(a, b)$:

$$\mathbf{U} = \{u = u(t) \in L^2(a, b) : f(t) \overset{\text{п.в.}}{\leq} u(t) \overset{\text{п.в.}}{\leq} g(t), t \in (a, b)\}, \quad f(t), g(t) \in L^2(a, b)$$

Пусть $(a, b) = \mathbf{T}_1 \cap \mathbf{T}_2 \cap \mathbf{T}_3$

$$h(t) \overset{\text{п.в.}}{\geq} g(t), t \in \mathbf{T}_1, \quad f(t) \overset{\text{п.в.}}{\leq} h(t) \overset{\text{п.в.}}{\leq} g(t), t \in \mathbf{T}_2, \quad h(t) \overset{\text{п.в.}}{\leq} f(t), t \in \mathbf{T}_3,$$

$$p = p(t)\mathcal{P}_U(h) = \begin{cases} g(t), & t \in \mathbf{T}_1 \\ h(t), & t \in \mathbf{T}_2 \\ f(t), & t \in \mathbf{T}_3 \end{cases}$$

4.4.6 Проектирование на параболоид в l^2

Параболоид в l^2 :

$$\mathbf{U} = \{x = (x_1, x_2, \dots, x_n, \dots) \in l^2 : x_1 \geq \sum_{n=2}^{\infty} x_n^2\}$$

$$\mathcal{P}_U(h) = \begin{cases} h, & \text{если } x_1 \geq \sum_{n=2}^{\infty} x_n^2 \\ (x_1 + \frac{1-\lambda}{2\lambda}, \lambda x_2, \lambda x_3, \dots, \lambda x_n, \dots), & \text{если } x_1 < \sum_{n=2}^{\infty} x_n^2, \end{cases}$$

где λ - корень уравнения $2 \left(\sum_{n=2}^{\infty} x_n^2 \right) \lambda^3 = (2x_1 - 1) \lambda + 1$, лежащий на интервале $(0, 1)$

4.4.7 Проекционная форма критерия оптимальности

Пусть \mathbf{U} - выпуклое замкнутое множество из евклидова пространства \mathbb{H} , а функционал $J(u)$ непрерывно дифференцируем на \mathbf{U} . Тогда если множество \mathbf{U}_* непусто, то

$$u_* \in \mathbf{U}_* \quad \Rightarrow \quad u_* = \mathcal{P}_U(u_* - \alpha J'(u_*)) \quad \forall \alpha > 0$$

Если функционал $J(u)$ ещё и выпуклый на \mathbf{U} , то равенство является критерием для $u_* \in \mathbf{U}_*$

5 Итерационные методы минимизации

5.1 Градиентный метод

В евклидовом пространстве \mathbb{H} рассматривается задача минимизации без ограничений, и пусть $J(u) \in C^1(\mathbb{H})$.

$$J(u) \rightarrow \inf_{u \in \mathbb{H}}$$

при $J'(u) \neq \Theta_{\mathbb{H}}$ направление самого быстрого убывания функционала $J(u)$ в точке u совпадает с направлением градиента $-J'(u)$.

Пусть u_0 задана. Итерация строится по правилу

$$u_{k+1} = u_k - \alpha_k J'(u_k), \quad k = 0, 1, \dots$$

α_k - шаг метода. Всегда можно выбрать шаг метода так, чтобы выполнялось неравенство $J(u_{k+1}) < J(u_k)$. Процесс прекращается при $J'(u_k) = \Theta_{\mathbb{H}}$.

5.1.1 Метод скорейшего спуска и его сходимость

α_k находят из условия

$$J_k(\alpha_k) = \inf_{\alpha \geq 0} J_k(\alpha), \quad J_k(\alpha) = J(u_k - \alpha J'(u_k))$$

Этот метод имеет линейную скорость сходимости, то есть $\|u_{k+1} - u_*\|_{\mathbb{H}} \leq C \cdot \|u_k - u_*\|_{\mathbb{H}}$. В связи с трудностью нахождения можно, например выбрать так на практике:

$$\inf_{\alpha \geq 0} J_k(\alpha) \leq J_k(\alpha_k) \leq \inf_{\alpha \geq 0} J_k(\alpha) + \delta_k, \quad \delta_k \geq 0, k = 0, 1, \dots \quad \sum_{k=0}^{\infty} \delta_k = \delta < \infty$$

Теорема о сходимости метода скорейшего спуска Пусть \mathbb{H} - евклидово пространство с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, функционал $J(u)$ является сильно выпуклым на \mathbb{H} с константой сильной выпуклости $\mu > 0$ и $J(u) \in C^{1,1}(\mathbb{H})$, то есть градиент $J'(u)$ удовлетворяет условию Липшица с константой $L > 0$. Тогда для любого начального приближения $u_0 \in \mathbb{H}$ приближения u_k , полученные методом скорейшего спуска таковы, что выполняются следующие неравенства:

$$0 \leq J(u_k) - J_* \leq q^k (J(u_0) - J_*), \quad \|u_k - u_*\|_{\mathbb{H}} \leq \sqrt{\frac{2}{\mu}} q^k (J(u_0) - J_*), \quad k = 0, 1, \dots$$

где u_* - единственная точка минимума $J(u)$ на \mathbb{H} , $q = 1 - \frac{\mu}{L}$, $0 \leq q < 1$.

5.1.2 Геометрический смысл метода скорейшего спуска

Метод скорейшего спуска имеет простой геометрический смысл: точка u_{k+1} , получаемая с его помощью, лежит на луче $L_k = \{u \in \mathbb{H} : u = u_k - \alpha J'(u_k), \alpha \geq 0\}$ в точке его касания поверхности уровня $\Gamma_{k+1} = \{u \in \mathbb{H} : J(u) = J(u_{k+1})\}$, а сам луч L_k перпендикулярен к поверхности уровня $\Gamma_k = \{u \in \mathbb{H} : J(u) = J(u_k)\}$.

5.1.3 Непрерывный аналог градиентного метода

Вместо итерации $u_{k+1} = u_k - \alpha_k J'(u_k)$ за основу берется задача Коши

$$= \begin{cases} u'(t) = -\alpha(t) J'(u(t)), \\ u(0) = u_0, \quad t > 0 \end{cases}$$

Теорема о сходимости непрерывного аналога градиентного метода Пусть \mathbb{H} - евклидово пространство с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, функционал $J(u)$ является сильно выпуклым на \mathbb{H} с константой сильной выпуклости $\mu > 0$ и $J(u) \in C^{1,1}(\mathbb{H})$, то есть градиент $J'(u)$ удовлетворяет условию Липшица с константой $L > 0$. Кроме этого, пусть

$$\alpha(t) \in \mathcal{C}[0, +\infty), \quad \int_0^{+\infty} \alpha(t) dt = +\infty.$$

Тогда для любого начального приближения $u_0 \in \mathbb{H}$ траектория $u(t)$ дифференциального уравнения сходится к единственной точке минимума u_* функционала $J(u)$ на \mathbb{H} , причем верно неравенство

$$\|u(t) - u_*\|_{\mathbb{H}} \leq \|u_0 - u_*\|_{\mathbb{H}} \exp \left\{ -\mu \int_0^t \alpha(\tau) d\tau \right\} \quad \forall t \geq 0$$

5.2 Метод проекции градиента

В евклидовом пространстве \mathbb{H} рассматривается задача условной минимизации, и пусть $J(u) \in C^1(\mathbb{H})$.

$$J(u) \rightarrow \inf_{u \in \mathbb{H} \subseteq \mathbb{H}}$$

Пусть $u_0 \in \mathbf{U}$, строить последовательность $\{u_k\}$ по правилу

$$u_{k+1} = \mathcal{P}_{\mathbf{U}}(u_k - \alpha_k J'(u_k)), \quad \forall k = 0, 1, \dots$$

где $\alpha_k > 0, k = 0, 1, \dots$ шаг метода.

Теорема о сходимости метода проекции градиента с постоянным шагом Пусть \mathbb{H} - гильбертово пространство с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, множество \mathbf{U} - выпукло и замкнуто, функционал $J(u)$ является сильно выпуклым на \mathbb{H} с константой сильной выпуклости $\mu > 0$ и $J(u) \in C^{1,1}(\mathbb{H})$, то есть градиент $J'(u)$ удовлетворяет условию Липшица с константой $L > 0$. Тогда для любого начального приближения $u_0 \in \mathbb{H}$ приближения u_k , полученные методом проекции градиента с постоянным шагом $\alpha_k \equiv \alpha, 0 < \alpha < 2\mu L^{-2}$, сходятся к единственному решению u_* задачи $J(u) \rightarrow \inf, u \in \mathbf{U}$, причем справедлива оценка

$$\|u_k - u_*\|_{\mathbb{H}} \leq q^k(\alpha) \|u_0 - u_*\|_{\mathbb{H}}, \quad k = 0, 1, \dots,$$

в которой $q(\alpha) = \sqrt{1 - 2\mu\alpha + L^2\alpha^2}, 0 < q(\alpha) < 1$.

5.3 Метод условного градиента

В евклидовом пространстве \mathbb{H} рассматривается задача условной минимизации, и пусть $J(u) \in C^1(\mathbb{H})$.

$$J(u) \rightarrow \inf_{u \in \mathbb{H} \subseteq \mathbb{H}}$$

\mathbf{U} - выпукло, замкнуто, ограничено. Пусть $u_0 \in \mathbf{U}$ - начальное приближение. Если приближение $u_k \in \mathbf{U}$ найдено, вместо функционала $J(u)$ рассматривается линейное приближение и решается задача минимизации $J_k(u)$ на множестве \mathbf{U} .

$$J_k(u) = J(u_k) + \langle J'(u_k), u - u_k \rangle_{\mathbb{H}} \rightarrow \inf, \quad u \in \mathbf{U}.$$

Пусть \bar{u}_k - одно из решений. Следующее приближение вычисляется по правилу

$$u_{k+1} = u_k + \alpha_k(\bar{u}_k - u_k), \quad 0 \leq \alpha_k \leq 1.$$

Величина α_k может выбираться из условий

$$f_k(\alpha_k) = \min_{0 \leq \alpha \leq 1} f_k(\alpha), \quad f_k(\alpha) = J(u_k + \alpha(\bar{u}_k - u_k)).$$

Для квадратичного и сильно выпуклого функционала эта задача минимизации имеет аналитическое решение. Рассмотрим

$$J(u) = \frac{1}{2} \langle \mathcal{A}u, u \rangle_{\mathbb{H}} - \langle b, u \rangle_{\mathbb{H}}, \quad \mathcal{A} = \mathcal{A}^* \in \mathcal{L}(\mathbb{H} \rightarrow \mathbb{H}), b \in \mathbb{H} \quad \exists \mu > 0 : \langle \mathcal{A}h, h \rangle_{\mathbb{H}} \geq \mu \|h\|_{\mathbb{H}}^2 \quad \forall h \in \mathbb{H}.$$

$$\alpha_k = \min\{\alpha_{k*}, 1\}, \quad \alpha_{k*} = - \frac{\langle J'(u_k), \bar{u}_k - u_k \rangle_{\mathbb{H}}}{\langle \mathcal{A}(\bar{u}_k - u_k), \bar{u}_k - u_k \rangle_{\mathbb{H}}}$$

Теорема о сходимости метода условного градиента Пусть $\mathbf{U} \subset \mathbb{H}$ - выпукло, замкнуто и ограничено, функционал $J(u) \in C^{1,1}(\mathbf{U})$ и выпуклый на \mathbf{U} . Тогда для любого начального приближения $u_0 \in \mathbb{H}$ приближения u_k , полученные методом условного градиента с выбором шага по формуле $\alpha_k = \operatorname{argmin}_{0 \leq \alpha \leq 1} J(u_k + \alpha(\bar{u}_k - u_k))$, минимизирует $J(u)$ на \mathbf{U} причем справедлива оценка

$$0 \leq J(u_k) - J_* \leq \frac{C}{k}, \quad k = 1, 2, \dots, \quad C = \operatorname{const} > 0$$

Если, кроме того, функционал $J(u)$ является сильно выпуклым на \mathbf{U} с константой μ , то

$$\|u_k - u_*\| \leq \sqrt{\frac{2C}{\mu k}}, \quad k = 1, 2, \dots$$

5.4 Метод Ньютона

В методе условного градиента вместо линейной части приращения берется квадратичная часть. Сначала находим точку $\bar{u}_k \in \mathbf{U}$ из условия.

$$J_k(\bar{u}_k) = \inf_{u \in \mathbf{U}} J_k(u), \quad J_k(u) = \langle J'(u_k), u - u_k \rangle_{\mathbb{H}} + \frac{1}{2} \langle J''(u_k)(u - u_k), u - u_k \rangle_{\mathbb{H}}$$

потом вычисляем $u_{k+1} = u_k + \alpha_k(\bar{u}_k - u_k)$, $\alpha_k \in [0, 1]$. α_k - шаг метода Ньютона. В случае $\mathbf{U} = \mathbb{H}$ и $\langle J''(u_k)h, h \rangle_{\mathbb{H}} \geq \mu \|h\|_{\mathbb{H}}^2$ $h \in \mathbb{H}$, $\mu \in \mathbb{R}^1$, точка \bar{u}_k может вычислена по формуле $\bar{u}_k = u_k - (J''(u_k))^{-1} J'(u_k)$.

5.4.1 Классический метод и его сходимость

Если взять $\alpha_k \equiv 1$, то $u_{k+1} = \bar{u}_k$. Получаем классический метод Ньютона.

Он имеет квадратичную скорость сходимости, (то есть $\|u_{k+1} - u_*\|_{\mathbb{H}} \leq C \cdot \|u_k - u_*\|_{\mathbb{H}}^2$) но сходится только локально.

Теорема о сходимости классического метода Ньютона Пусть \mathbb{H} - евклидово пространство с введенным скалярным произведением $\langle \cdot, \cdot \rangle_{\mathbb{H}}$, функционал $J(u)$ является сильно выпуклым на \mathbf{U} с константой сильной выпуклости $\mu > 0$ и $J(u) \in C^2(\mathbf{U})$, $J''(u)$ удовлетворяет условию Липшица на \mathbf{U} с константой $L > 0$, множество \mathbf{U} - выпукло и замкнуто, $\text{int} \mathbf{U} \neq \emptyset$. $u_0 \in \mathbf{U}$ - начальное приближение. Тогда если выполнено соотношение

$$q = \frac{L}{2\mu} \|u_0 - u_*\|_{\mathbb{H}} < 1, \quad u_* \in \mathbf{U}_*$$

то последовательность $\{u_k\}$ сильно сходится к u_* , причем

$$\|u_k - u_*\|_{\mathbb{H}} \leq \frac{2\mu}{L} \cdot q^{2^k}, \quad k = 0, 1, \dots$$

5.4.2 Метод Ньютона с выбором шага по Армихо

Теорема о сходимости метода Ньютона с выбором шага по Армихо Пусть \mathbb{H} - евклидово пространство с введенным скалярным произведением $\langle \cdot, \cdot \rangle$, множество $\mathbf{U} \subseteq \mathbb{H}$ выпукло, замкнуто, $\text{int} \mathbf{U} \neq \emptyset$. Пусть функционал $J(u)$ является сильно выпуклым на \mathbf{U} с константой сильной выпуклости μ , $J(u) \in C^2(\mathbf{U})$, $J''(u)$ удовлетворяет на \mathbf{U} условию Липшица с константой L , существует такая константа $M > 0$, что $\langle J''(u)h, h \rangle_{\mathbb{H}} \leq M \|h\|_{\mathbb{H}}^2 \quad \forall u \in \mathbf{U}, h \in \mathbb{H}$.

Тогда последовательность $\{u_k\}$, полученная методом Ньютона с выбором шага α_k по способу $\alpha_k = \lambda^m$, m - минимальное целое неотрицательное число, при котором верно неравенство

$$J(u_k) - J(u_k + \lambda^m(\bar{u}_k - u_k)) \geq \frac{1}{2} \lambda^m |J_k(\bar{u}_k)|$$

при любом начальном приближении $u_0 \in \mathbf{U}$ существует и сходится к точке u_* - единственному решению исходной задачи, причем найдутся число $q \in (0, 1)$ и номер k_0 такой, что при всех $k \geq k_0$ будет справедливо $\alpha_k = 1$ и, кроме того,

$$\|u_k - u_*\|_{\mathbb{H}} \leq \frac{2\mu}{L} q^{2^k}, \quad k = k_0, k_0 + 1, \dots$$

5.5 Квазиньютоновские методы

Будем рассматривать задачу безусловной минимизации в \mathbb{R}^n , то есть

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}^n$$

пусть $f(x) \in C^2(\mathbb{R}^n)$. Общая итерационная схема такая:

1. Выбираются начальное приближение $x_0 \in \mathbb{R}^n$ и матрица $H_0 \in \mathbb{R}^{n \times n}$.
2. Далее вычисляются по формуле

$$x_{k+1} = x_k - t_k H_k f'(x_k), \quad H_{k+1} = U(x_k, H_k), \quad k = 0, 1, 2, \dots$$

где $t_k \in \mathbb{R}^1$, $U(x_k, H_k)$ такова, что $\{H_k\}$ удовлетворяет квазиньютоновскому условию.

$$H_{k+1}(f'(x_{k+1}) - f'(x_k)) = x_{k+1} - x_k, \quad k = 0, 1, 2, \dots$$

5.5.1 Семейство методов ранга один

Переход от H_k к H_{k+1} делается с помощью формулы

$$H_{k+1} = H_k + \frac{1}{d_k^T(f'(x_{k+1}) - f'(x_k))} ((x_{k+1} - x_k) - H_k(f'(x_{k+1}) - f'(x_k))) d_k^T$$

где вектор столбец d_k выбирается так, чтобы $d_k^T(f'(x_{k+1}) - f'(x_k)) \neq 0$. При таком способе H_{k+1} квазиньютоновское условие выполнено. Эта формула задает семейство квазиньютоновских методов ранга один, так как матрица поправки $\Delta H_{k+1} = H_{k+1} - H_k$ как произведение вектор-столбца на вектор строку, имеет ранг равный единице. d_k - параметризация семейства.

Метод Бройдена: $t_k \equiv 1, d_k = f'(x_{k+1}) - f'(x_k), k = 0, 1, \dots$

Метод МакКормика: $t_k \equiv 1, d_k = x_{k+1} - x_k, k = 0, 1, \dots$

Теорема о сходимости методов Бройдена и МакКормика Пусть функция $f(x)$ дважды непрерывно дифференцируема и сильно выпукла с константой сильной выпуклости μ в некоторой окрестности $X = \{x \in \mathbb{R}^n : \|x - x_*\| \leq C\}$ точки $x_*, f'(x_*) = 0$, и существует число $K > 0$ такое, что выполнено неравенство

$$\|f''(x) - f''(x_*)\| \leq K \|x - x_*\|, \quad \forall x \in X.$$

Тогда методы Бройдена и МакКормика локально и сверхлинейно сходятся к x_* , то есть существуют такие $\varepsilon, \delta > 0$, что если верно

$$\|x_0 - x_*\| \leq \varepsilon, \quad \left\| H_0 - (f''(x_*))^{-1} \right\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n \left(H_0 - (f''(x_*))^{-1} \right)_{ij}^2 \leq \delta,$$

то

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

Что означает $\|x_{k+1} - x_*\| \leq q \|x_k - x_*\|$ - q уменьшается при каждой итерации.

5.5.2 Семейство методов ранга два

Имеется единственный способ получения квазиньютоновского метода ранга один с симметричными матрицами H_k : H_0 выбирается симметричной, затем на каждом шаге брать $d_k = z_k - H_k y_k$. Но такой метод не локально сходится.

Можно преодолеть эту трудность, рассматривая квазиньютоновские методы ранга два. Пусть $x_0 \in \mathbb{R}^n$ и $H_0 \in \mathbb{R}^{n \times n}$ заданы. Далее

$$x_{k+1} = x_k - H_k f'(x_k)$$

$$H_{k+1} = H_k + \frac{(z_k - H_k y_k) d_k^T + d_k (z_k - H_k y_k)^T}{d_k^T y_k} - \frac{y_k^T (z_k - H_k y_k) d_k d_k^T}{(d_k^T y_k)^2}$$

где $z_k = x_{k+1} - x_k, y_k = f'(x_{k+1}) - f'(x_k)$. ΔH_k имеет ранг равный двум. Если H_k - симметричная, то H_{k+1} - симметричная.

Метод Гринстада: $d_k = y_k = f'(x_{k+1}) - f'(x_k), k = 0, 1, \dots$

Метод BFGS: $d_k = z_k = x_{k+1} - x_k, k = 0, 1, \dots$

Теорема о сходимости методов Гринстада и BFGS Пусть функция и точка удовлетворяют условиям предыдущей теоремы. Тогда методы Гринстада и BFGS локально и сверхлинейно сходятся к x_* , то есть существуют такие $\varepsilon, \delta > 0$, что если верно

$$\|x_0 - x_*\| \leq \varepsilon, \quad \left\| H_0 - (f''(x_*))^{-1} \right\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n \left(H_0 - (f''(x_*))^{-1} \right)_{ij}^2 \leq \delta,$$

то

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

5.6 Линейное программирование

5.6.1 Постановка задачи

Задача минимизации или максимизации линейной функции

$$f(x) = \langle c, x \rangle \rightarrow \inf, \quad x \in \mathbf{X}$$

$$\mathbf{X} = \{x = (x_1, \dots, x_n)^T \in \mathbb{R}^n : x_k \geq 0, k \in I_+, \langle a_i, x \rangle \leq b_i, i = 1, \dots, m; \langle a_i, x \rangle = b_i, i = m+1, \dots, m+s\}$$

Матричная форма

$$f(x) = \langle c^1, x^1 \rangle + \langle c^2, x^2 \rangle \rightarrow \inf, \quad x \in \mathbf{X}$$

$$\mathbf{X} = \{x = (x^1, x^2)^T : x^1 \in \mathbb{R}^{n_1}, x^2 \in \mathbb{R}^{n_2}, A_{11}x^1 + A_{12}x^2 \leq b^1, A_{21}x^1 + A_{22}x^2 = b^2; x^1 \succcurlyeq \Theta\}$$

5.6.2 Каноническая задача

$$g(u) = \langle c, u \rangle \rightarrow \inf, \quad u \in \mathbf{U} = \{u \in \mathbb{R}^{n_1+n_2+m_1} : u \succcurlyeq \Theta, Au = b\}$$

где

$$c = \begin{bmatrix} c^1 \\ c^2 \\ -c^2 \\ \Theta \end{bmatrix}, \quad u = \begin{bmatrix} x^1 \\ v \\ w \\ z \end{bmatrix}, \quad A = \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & -A_{12} & I \\ \hline A_{21} & A_{22} & -A_{22} & O \end{array} \right], \quad b = \begin{bmatrix} b^1 \\ b^2 \end{bmatrix}.$$

5.6.3 Угловые точки

Определение Точка $v \in \mathbf{X}$ называется *угловой точкой* множества \mathbf{X} , если равенство

$$v = \alpha v^1 + (1 - \alpha)v^2, \quad v^1, v^2 \in \mathbf{X}, 0 < \alpha < 1.$$

верно, только если $v^1 = v^2$.

Теорема о алгебраическом критерии угловой точки Пусть $A \neq O$ ранг матрицы A равен r . Для того, чтобы точка $v = (v_1, \dots, v_n) \in \mathbf{U}$ была угловой точкой множества $\mathbf{U} = \{x_1, \dots, x_n \in \mathbb{R}^n : x \succcurlyeq \Theta, Ax = b\}$, необходимо и достаточно, чтобы существовали номера $j_1, \dots, j_r, 1 \leq j_i \leq n, i = 1, \dots, r$, такие что

$$A_{j_1}v_{j_1} + \dots + A_{j_r}v_{j_r} = b; \quad v_j = 0, j \neq j_1, \dots, j_r$$

причем столбцы A_{j_1}, \dots, A_{j_r} линейно независимы.

Определение Систему векторов A_{j_1}, \dots, A_{j_r} , входящих в критерий угловой точки, называют *базисом угловой точки* v , а соответствующие им переменные v_{j_1}, \dots, v_{j_r} - *базисными координатами угловой точки* v .

Определение Если v_{j_1}, \dots, v_{j_r} все положительны, то такую угловую точку называют *невыврожденной точкой*. Если среди из них есть нулевая координата, то называют *вырожденной точкой*. При фиксированном базисе A_{j_1}, \dots, A_{j_r} переменные x_{j_1}, \dots, x_{j_r} - *базисные переменные* угловой точки, а остальные переменные - *свободные переменные*.

5.6.4 Симплекс-метод

Рассмотрим каноническую задачу линейного программирования

$$f(x) = \langle c, x \rangle \rightarrow \inf, \quad x \in \mathbf{U} = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x \succcurlyeq \Theta, Ax = b\}$$

Пусть $J_b = \{j_1, \dots, j_r\}$ - базисные номера, $J_f = \{i_1, \dots, i_{n-r}\} = \{1, 2, \dots, n\} \setminus J_b$, $x_b = (x_{j_1}, \dots, x_{j_r})^T$ - базисные переменные, $x_f = (x_{i_1}, \dots, x_{i_{n-r}})$ - свободные переменные.

Пусть $B = (A_{j_1} | \dots | A_{j_r})$. В силу ограничения-равенства $Ax = b$ имеем

$$b = \sum_{j \in J_b} A_j x_j + \sum_{j \in J_f} A_j x_j = Bx_b + \sum_{j \in J_f} A_j x_j \quad (1)$$

Так как A_{j_1}, \dots, A_{j_r} - ЛНЗ, то существует B^{-1} . Пусть $v_b = (v_{j_1}, \dots, v_{j_r})^T, c_b = (c_{j_1}, \dots, c_{j_r})^T, F = (A_{i_1} | \dots | A_{i_{n-r}})$. Поскольку $v_f = (v_{i_1}, \dots, v_{i_{n-r}})^T = \Theta_{n-r}^T$, подставляя в его и получаем $Bv_b = b, v_b = B^{-1}b$. Домножив B^{-1} слева на (1), получаем

$$\begin{aligned}\Theta = v_b = B^{-1}b &= x_b + \sum_{j \in J_f} B^{-1}A_j x_j \Rightarrow x_b = v_b - \sum_{j \in J_f} B^{-1}A_j x_j = v_b - B^{-1}F x_f \\ f(x) &= \langle c_b, x_b \rangle + \sum_{k \in J_f} c_k x_k = \langle c_b, x_b \rangle - \sum_{k \in J_f} (\langle c_b, B^{-1}A_k \rangle - c_k) x_k = f(v) - \sum_{k \in J_f} \Delta_k x_k \\ j(x_f) &= f(v) - \sum_{k \in J_f} \Delta_k x_k \rightarrow \inf, \quad x_f \succcurlyeq \Theta, x_b \succcurlyeq \Theta\end{aligned}$$

где $\Delta = (\Delta_{i_1}, \dots, \Delta_{i_{n-r}})^T = \langle c_b, B^{-1}F \rangle - c_f = (B^{-1}F)^T c_b - c_f$.

Схема метода: На каждом шаге метода

1. Обрабатывается угловая точка $v \in \mathbf{U}$ и составляется функцию $j(x_f)$.

2. Выбираются номера из J_f , изменения которых может уменьшить значение функции $j(x_f)$: $J_f^+ = \{k \in J_f : \Delta_k > 0\}$.

2.1. $J_f^+ = \emptyset$, то есть нельзя уменьшить значение функции \Rightarrow процесс останавливается и $v \in \mathbf{U}_*, f_* = f(v)$.

2.2. $J_f^+ \neq \emptyset$ и $\exists k \in J_f^+$: для x_k нет ограничений, тогда процесс останавливается и $f_* = -\infty$. Обычно смотрим все столбцы $\gamma_k, k \in J_f^+$ матрицы $B^{-1}F$. Если в каждом столбце есть положительные числа, то переходим к следующему случаю, если в каком-то столбце нет положительных чисел, то останавливается.

2.3. $J_f^+ \neq \emptyset$ и $\forall k \in J_f^+$: для x_k есть ограничений. Тогда выбираем какую-то одну свободную переменную, и присваиваем ей максимальное возможное значение, и остальные равны нулю. В результате получаем новую угловую точку w , причем $f(w) \leq f(v)$. - Это действие называется *шагом симплекс-метода*. Если v - невырожденная, то $f(w)$ точно меньше; если v - вырожденная, то возможно $f(w) = f(v)$ - закликивание. Возможный подход - *правило Блэнда*.

Процесс перехода от точки v к точке w

1) Сначала выбираем какое-то $\Delta_{i_s} > 0$. Его номер i_s будет номером свободной переменной, *вводимой в базис*. Затем рассматривается столбец γ_{i_s} , числам из него сверху вниз присваиваются номера j_1, j_2, \dots, j_r . Из этих номеров мы выбираем только те номера j_k , для которых $\gamma_{i_s, j_k} > 0$, после чего находятся величины θ_{j_k} по правилу

$$\theta_{j_k} = \frac{v_{j_k}}{\gamma_{i_s, j_k}}, \text{ где } v_{j_k} - \text{координаты угловой точки } v$$

2) Затем находим минимальное из этих чисел (если их несколько, берем какое-то одно). Его номер j_q является номером базисной переменной, *выводимой из базиса*. Его значение (обозначим его θ используется для перехода к новой угловой точке w по правилу

$$w_b = v_b - \theta \cdot \gamma_{i_s}$$

(при этом координата w_{j_q} обязательно станет равна нулю), в качестве значения координаты w_{i_s} берется число θ , остальные координаты точки w берутся равными нулю. Из списка J_b базисных координат выкидывается номер j_q , вместо него вводим номер i_s .

5.7 Метод сопряженных градиентов

Рассмотрим квадратичный функционал

$$J(u) = \frac{1}{2} \langle Au, u \rangle - \langle f, u \rangle, \quad A = A^T > 0, f \in \mathbb{R}^n$$

Задача минимизации на \mathbb{R}^n эквивалентна задаче решения СЛАУ

$$J'(u) = \frac{1}{2}(A + A^T)u - f = Au - f = \Theta \Leftrightarrow Au = f$$

Пусть $\{p_0, p_1, \dots, p_n\}$ - базис в \mathbb{R}^n , $u_0 \in \mathbb{R}^n$ - начальное приближение, u_* - точка минимума функционала $J(u)$.

Схема метода:

Итерация 0. Пусть u_0 - любое и $p_0 = -J'(u_0)$. Если $p_0 = \Theta$, то заканчивается работа.

Итерация 1.

$$\alpha_0 = -\frac{\langle Au_0 - f, p_0 \rangle}{\langle Ap_0, p_0 \rangle}, \quad u_1 = u_0 + \alpha_0 p_0, \quad p_1 = -J'(u_1) + \beta_0 p_0, \quad \beta_0 = \frac{\langle J'(u_1), Ap_0 \rangle}{\langle Ap_0, p_0 \rangle}$$

Итерация $k \geq 2$.

$$\alpha_{k-1} = -\frac{\langle J'(u_{k-1}), p_{k-1} \rangle}{\langle Ap_{k-1}, p_{k-1} \rangle}, \quad u_k = u_{k-1} + \alpha_{k-1} p_{k-1}, \quad \beta_{k-1} = \frac{\langle J'(u_k), Ap_{k-1} \rangle}{\langle Ap_{k-1}, p_{k-1} \rangle}, \quad p_k = -J'(u_k) + \beta_{k-1} p_0.$$

На последней итерации n найдено $u_n = u_*$.

5.7.1 Сходимость

Лемма о свойствах коэффициентов α_k Пусть $J(u) = \frac{1}{2} \langle Au, u \rangle - \langle f, u \rangle$, $A = A^T > 0$; p_0, p_1, \dots, p_n - базис $\mathbb{H} = \mathbb{R}^n$ из сопряженных относительно матрицы A векторов, то есть $\langle Ap_i, p_j \rangle = 0, \forall i \neq j$, u_* - точка минимума на всем пространстве \mathbb{R}^n функционала $J(u)$, α_k - коэффициенты из разложения

$$u_* - u_0 = \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{n-1} p_{n-1}$$

Тогда

$$\alpha_k = \frac{\langle f - Au_0, p_k \rangle}{\langle Ap_k, p_k \rangle} = \operatorname{argmin}_{\alpha \in \mathbb{R}} J(u_k + \alpha p_k) = -\frac{\langle J'(u_k), p_k \rangle}{\langle Ap_k, p_k \rangle}; \quad \langle J'(u_{k+1}), p_k \rangle = 0, \quad k = 0, 1, \dots, n-1$$

Теорема о сходимости метода сопряженных градиентов Пусть $\mathbb{H} = \mathbb{R}^n$, $J(u) = \frac{1}{2} \langle Au, u \rangle - \langle f, u \rangle$, $A = A^T > 0$. Тогда итерационный процесс метода сопряженных градиентов при любом начальном приближении $u_0 \in \mathbb{H}$ не более, чем за $n = \dim \mathbb{H}$ находит точку u_* - решение задачи $J(u) \rightarrow \inf_{\mathbb{H}}$, причем верны формулы

$$\begin{aligned} \langle Ap_k, p_m \rangle &= 0, \quad \langle J'(u_k), J'(u_m) \rangle = 0, \quad \forall k, m \in \{0, 1, \dots, n-1\}, k \neq m \\ \langle J'(u_k), p_m \rangle &= 0, \quad \forall k, m \in \{0, 1, \dots, n-1\}, m < k \end{aligned}$$

5.7.2 Случай бесконечномерного пространства

Выбираем произвольные начальное приближение $u_0 \in \mathbb{H}$, берем $p_0 = -J'(u_0)$, после чего вычисляются по формулам:

$$\begin{aligned} u_{k+1} &= u_k + \alpha_k p_k, \quad p_{k+1} = -J'(u_{k+1}) + \beta_k p_k, \quad k = 0, 1, 2, \dots \\ \alpha_k &= \operatorname{argmin}_{\alpha \in \mathbb{R}} J(u_k + \alpha p_k), \quad \beta_k = \frac{\langle J'(u_{k+1}), J'(u_{k+1}) - J'(u_k) \rangle}{\|J'(u_k)\|^2} \end{aligned}$$

Обычно метод сопряженных градиентов используется в сочетании с методом Ньютона, в результате получаем такую оценку сходимости.

$$\|u_m - u_*\| \leq \frac{2\mu}{L} \cdot q^{(\sqrt{2})^m}$$

5.8 Метод покоординатного спуска

Этот метод позволяет решить задачу минимизации, не вычисляя значение первой и второй производных. Рассмотрим задачу минимизации функции $f(x) \in \mathbb{R}^n$.

$$f(x) \rightarrow \inf, \quad x \in \mathbb{R}^n$$

Выберем некоторый базис $\{e_i\}_{i=1}^n$ из \mathbb{H} , например стандартный базис. Производится циклический перебор этих векторов. Для удобства описания итерации запишем

$$p_0 = e_1, p_1 = e_2, \dots, p_{n-1} = e_n, \quad p_n = e_1, p_{n+1} = e_2, \dots, p_{2n-1} = e_n, \quad p_{2n} = e_1$$

5.8.1 Схема метода

Сначала выбирается начальное приближение $x_0 \in \mathbb{R}^n$, стартовый шаг $\alpha_0 > 0$, коэффициент дробления $\lambda \in (0, 1)$. Пусть найдено k -ое приближение x_k и текущее значение шага $\alpha_k > 0$. Для x_{k+1} имеем:

1. Если $f(x_k + \alpha_k p_k) < f(x_k)$, то $x_{k+1} = x_k + \alpha_k p_k$, $\alpha_{k+1} = \alpha_k$
2. Если $f(x_k + \alpha_k p_k) \geq f(x_k)$ и $f(x_k - \alpha_k p_k) < f(x_k)$, то $x_{k+1} = x_k - \alpha_k p_k$, $\alpha_{k+1} = \alpha_k$.
3. Если $f(x_k \pm \alpha_k p_k) \geq f(x_k)$ - неудачная итерация - и притом количество неудачных итераций, случившихся подряд, меньше n , то $x_{k+1} = x_k$, $\alpha_{k+1} = \alpha_k$.
4. Если Если $f(x_k \pm \alpha_k p_k) \geq f(x_k)$ и притом количество неудачных итераций, случившихся подряд, равно n , то $x_{k+1} = x_k$, $\alpha_{k+1} = \lambda \alpha_k$, количество неудач становится 0.

5.8.2 Сходимость метода

Теорема Пусть функция $f(x)$ выпукла на \mathbb{R}^n и принадлежит классу $C^1(\mathbb{R}^n)$, а начальное приближение x_0 таково, что множество Лебега $M(x_0) = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ ограничено. Тогда последовательность x_k , сходится и по функции, и по аргументу.

$$\lim_{k \rightarrow \infty} f(x_k) = f_*, \quad \lim_{k \rightarrow \infty} \rho(x_k, X_*) = 0$$

6 Методы снятия ограничений

6.1 Метод штрафных функций

Будем рассматривать задачу:

$$J(u) \rightarrow \inf, \quad u \in \mathbf{U} = \{u \in \mathbf{U}_0 \subseteq \mathbb{H} : g_1(u) \leq 0, \dots, g_m(u) \leq 0; g_{m+1}(u) = 0, \dots, g_{m+s}(u) = 0\}.$$

где \mathbb{H} - евклидово пространство, \mathbf{U}_0 множество простого вида.

Введем *штрафную функцию*:

$$P(u) = \sum_{i=1}^{m+s} (g_i^+(u))^{p_i}, \quad p_i \geq 1, i = 1, \dots, m+s$$

числа p_i мы выбираем сами, функции

$$g_i^+(u) = \max\{g_i(u), 0\}, i = 1, \dots, m; \quad g_i^+(u) = |g_i(u)|, i = m+1, \dots, m+s$$

Эти функции называют *индивидуальными штрафами*. Ясно, что $P(u) \geq 0, \forall u \in \mathbb{H}$ и

$$u \in \mathbf{U} \Leftrightarrow \begin{cases} u \in \mathbf{U}_0, \\ P(u) = 0 \end{cases} \Leftrightarrow \begin{cases} u \in \mathbf{U}_0 \\ g_i^+(u) = 0, i = 1, \dots, m+s \end{cases}$$

Рассмотрим последовательность задач минимизации

$$\Phi_k(u) = J(u) + A_k P(u) \rightarrow \inf, \quad u \in \mathbf{U}_0, k = 1, 2, \dots, \lim_{k \rightarrow \infty} A_k = +\infty$$

Пусть $u_k \in \mathbf{U}_0$ - приближенные решения этих задач, то есть

$$\Phi_{k_*} = \inf_{\mathbf{U}_0} \Phi_k(u) \leq \Phi_k(u_k) \leq \Phi_{k_*} + \varepsilon_k, \quad k = 1, 2, \dots, \lim_{k \rightarrow \infty} \varepsilon_k = 0$$

Теорема о сходимости метода штрафных функций Пусть \mathbb{H} - гильбертово пространство, множество $\mathbf{U}_0 \subseteq \mathbb{H}$ слабо замкнуто, функционалы $J(u)$ и $g_i^+(u), i = 1, \dots, m+s$ слабо полунепрерывны снизу на \mathbf{U}_0 , множество

$$\mathbf{U}(\delta) = \{u \in \mathbf{U}_0 : g_i^+ \leq \delta, i = 1, \dots, m+s\}$$

ограничено при некотором $\delta > 0$,

$$J = \inf_{\mathbf{U}_0} J(u) > -\infty, \quad \lim_{k \rightarrow \infty} A_k = +\infty, \quad \lim_{k \rightarrow \infty} \varepsilon_k = 0$$

Тогда последовательность $\{u_k\}$, полученная методом штрафных функций такова, что $\lim_{k \rightarrow \infty} J(u_k) = J_*$, и все ее слабые предельные точки принадлежат множеству \mathbf{U}_* .

6.2 Правило множителей Лагранжа

Будем рассматривать задачу:

$$J(u) \rightarrow \inf, \quad u \in \mathbf{U} = \{u \in \mathbf{U}_0 \subseteq \mathbb{H} : g_1(u) \leq 0, \dots, g_m(u) \leq 0\}.$$

где \mathbb{H} - линейное пространство, \mathbf{U}_0 - выпуклое множество простого вида, $J(u), g_i(u)$ - выпуклые. Такие задачи называют *выпуклыми*, так как минимизируется выпуклое на выпуклом множестве.

Введем функцию Лагранжа

$$\mathcal{L}(u, \lambda) = \lambda_0 J(u) + \sum_{i=1}^m \lambda_i g_i(u), \quad u \in \mathbf{U}_0, \lambda_i \in \mathbb{R}, i = 0, 1, \dots, m$$

6.2.1 Теорема Куна-Таккера и условие Слейтера

Теорема Куна-Таккера Пусть в рассматриваемой задаче множество \mathbf{U}_0 выпуклое, $J(u), g_i(u), i = 1, \dots, m$ выпуклы на \mathbf{U}_0 . Тогда если $u_* \in \mathbf{U}_*$, то существуют множители Лагранжа $\lambda^* = (\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*) \neq \Theta$ такие, что выполнены условия

$$\begin{aligned} \mathcal{L}(u_*, \lambda^*) &\leq \mathcal{L}(u, \lambda^*) \quad \forall u \in \mathbf{U}_0, & (\text{принцип минимума}) \\ \lambda_0^* &\geq 0, \lambda_1^* \geq 0, \dots, \lambda_m^* \geq 0, & (\text{неотрицательность множителей Лагранжа}) \\ \lambda_1^* g_1(u_*) &= 0, \dots, \lambda_m^* g_m(u_*) = 0. & (\text{условия дополняющей нежесткости}) \end{aligned}$$

Кроме того, если для пары (u_*, λ^*) выполнены эти три условия и $u_* \in \mathbf{U}, \lambda_0^* > 0$, то $u_* \in \mathbf{U}_*$.

$\lambda_0^* > 0$ - условие регулярности. Достаточным его условием является *условие Слейтера*:

$$\exists u_0 \in \mathbf{U}_0 : g_i(u_0) < 0, \quad \forall i = 1, \dots, m$$

Определение Точку (x_*, y^*) называют *седловой точкой* функции $f(x, y)$ на $\mathbf{X} \times \mathbf{Y}$, если

$$(x_*, y^*) \in \mathbf{X} \times \mathbf{Y}, \quad f(x_*, y) \leq f(x_*, y^*) \leq f(x, y^*), \quad \forall x \in \mathbf{X}, \forall y \in \mathbf{Y}$$

Теорема о седловой форме теоремы Куна-Таккера Пусть выполнены все условия исходной теоремы и условие регулярности Слейтера. Тогда для того, чтобы $u_* \in \mathbf{U}_*$, необходимо и достаточно, чтобы у классической функции Лагранжа

$$\mathcal{L}(u, \lambda) = J(u) + \sum_{i=1}^m \lambda_i g_i(u), \quad u \in \mathbf{U}_0, \lambda_i \in \mathbb{R}, i = 1, \dots, m$$

существовала седловая точка вида (u_*, λ^*) на $\mathbf{U}_0 \times \mathbb{R}_+^m$.

6.2.2 Правило множителей Лагранжа для гладких задач

Будем рассматривать задачу:

$$J(u) \rightarrow \inf, \quad u \in \mathbf{U} = \{u \in \mathbb{H} \subseteq \mathbb{H} : g_1(u) \leq 0, \dots, g_m(u) \leq 0, g_{m+1}(u) = 0, \dots, g_{m+s}(u) = 0\}.$$

где \mathbb{H} - гильбертово пространство, $\mathbf{U}_0 = \mathbb{H}$, $J(u), g_i(u), i = 1, \dots, m+s$ - гладкие функционалы. Рассмотрим необходимые условия локального минимума для таких задач. Введем функцию Лагранжа

$$\mathcal{L}(u, \lambda) = \lambda_0 J(u) + \sum_{i=1}^{m+s} \lambda_i g_i(u), \quad u \in \mathbb{H}, \lambda_i \in \mathbb{R}, i = 0, 1, \dots, m+s$$

Теорема о необходимых условиях оптимальности Пусть в рассматриваемой задаче u_* - точка локального минимума, $J(u), g_i(u), i = 1, \dots, m$ непрерывно дифференцируемы по Фреше в окрестности

\mathbf{U}_ε точки u_* . Тогда существуют множители Лагранжа $\lambda^* = (\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*, \lambda_{m+1}^*, \dots, \lambda_{m+s}^*) \neq \Theta$ такие, что выполнены условия

$$\begin{aligned}\mathcal{L}'_u(u_*, \lambda^*) &= \lambda_0^* J'(u_*) + \sum_{i=1}^{m+s} \lambda_i^* g'_i(u_*) = \Theta, & (\text{условие стационарности}) \\ \lambda_0^* &\geq 0, \lambda_1^* \geq 0, \dots, \lambda_m^* \geq 0, & (\text{неотрицательность множителей для неравенств и } J(u)) \\ \lambda_1^* g_1(u_*) &= 0, \dots, \lambda_m^* g_m(u_*) = 0. & (\text{условия дополняющей нежесткости})\end{aligned}$$

Достаточное условие регулярности задачи: Рассмотрим оператор $G'(u_*) : \mathbb{H} \rightarrow \mathbb{R}^s$:

$$G'(u_*)h = (\langle g'_{m+1}(u_*), h \rangle, \langle g'_{m+2}(u_*), h \rangle, \dots, \langle g'_{m+s}(u_*), h \rangle)$$

Если выполнены все условия теоремы о необходимых условиях оптимальности, и верно

$$\text{Im} G'(u_*) = \mathbb{R}^s, \quad \exists h \in \text{Ker} G'(u_*) : \langle g'_i(u_*), h \rangle < 0, i = 1, 2, \dots, m$$

Тогда в любом наборе множителей Лагранжа, соответствующем u_* , обязательно $\lambda_0^* > 0$. Если нет неравенств, то $-\text{Im} G'(u_*) = \mathbb{R}^s$; если нет равенств, то $\exists h \in \mathbb{H} : \langle g'_i(u_*), h \rangle < 0, i = 1, 2, \dots, m$

6.3 Двойственные задачи

Будем рассматривать задачу

$$J(u) \rightarrow \inf, \quad u \in \mathbf{U} = \{u \in \mathbf{U}_0 \mathbb{H} : g_i(u) \leq 0, i = 1, \dots, m, g_i(u) = 0, i = m+1, \dots, m+s\}$$

где \mathbb{H} - линейное пространство, \mathbf{U}_0 - множество простого вида. Запишем классическую функцию Лагранжа:

$$\mathcal{L}(u, \lambda) = J(u) + \sum_{i=1}^{m+s} \lambda_i g_i(u), \quad u \in \mathbf{U}_0, \lambda \in \Lambda^0 = \{\lambda = (\lambda_1, \dots, \lambda_{m+s}), \lambda_1 \geq 0, \dots, \lambda_m \geq 0\}$$

Рассмотрим функцию и задачу

$$\varphi(u) = \sup_{\lambda \in \Lambda^0} \mathcal{L}(u, \lambda) = \begin{cases} J(u), & u \in \mathbf{U}, \\ +\infty, & u \in \mathbf{U}_0 \setminus \mathbf{U} \end{cases} \rightarrow \inf_{u \in \mathbf{U}_0}$$

Эта задача эквивалентна исходной задаче. Формально меняя порядок взятия максимума и минимума, получаем

$$\Psi(\lambda) = \inf_{u \in \mathbf{U}_0} \mathcal{L}(u, \lambda) \rightarrow \sup_{\lambda \in \Lambda^0}$$

Это двойственная к исходной задаче. При этом мы максимизируем $\Psi(\lambda)$ лишь на множестве $\Lambda = \{\lambda \in \Lambda^0 : \Psi(\lambda) > -\infty\}$. Введем обозначения

$$\Psi^* = \sup_{\lambda \in \Lambda^0} \Psi(\lambda) = \sup_{\lambda \in \Lambda} \Psi(\lambda), \quad \Lambda^* = \{\lambda \in \Lambda^0 : \Psi(\lambda) = \Psi^*\}$$

Теорема о свойствах решений двойственных задач Всегда имеют место неравенства

$$\Psi(\lambda) \leq \Psi^* \leq \varphi_* \leq \varphi(u) \quad \forall \lambda \in \Lambda^0, u \in \mathbf{U}_0$$

Для того, чтобы выполнялось $\Psi^* = \varphi_* = J_*$, $\mathbf{U}_* \neq \emptyset$, $\Lambda^* \neq \emptyset$, необходимо и достаточно, чтобы классическая функция Лагранжа $\mathcal{L}(u, \lambda)$ имела седловую точку на множестве $\mathbf{U}_0 \times \Lambda^0$. При этом множество всех её седловых точек совпадает с множеством $\mathbf{U}_* \times \Lambda^*$.

Двойственная задача хороша тем, что

1. если ее удастся выписать, то как минимум можно оценить снизу искомую точную нижнюю грань J_* , а в хорошем случае удастся и найти решение исходной задачи.
2. Двойственная задача конечномерная, что проще исходной задачи.
3. Она всегда является выпуклой задачей.

7 Принцип максимума Л.С.Понтрягина

7.1 Постановка задачи оптимального управления

Рассмотрим задачу Коши

$$\begin{cases} \dot{x}(t) \stackrel{\text{п.в.}}{=} f(x(t), t, u(t)), & t_0 < t < T \\ x(t_0) = x^0 \end{cases}$$

где $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ - фазовая переменная, $u(t) = (u_1(t), u_2(t), \dots, u_m(t))$ - управление. Функция правой части $f: \mathbb{R}^n \times \mathbb{R}^1 \times \mathbb{R}^m \rightarrow \mathbb{R}^n$. Начальное состояние $x^0 = (x_1^0, x_2^0, \dots, x_n^0)$. t_0, T - известны. $u \in \mathbf{U} = \{u = u(t) \mathbb{H} : u(t) \stackrel{\text{п.в.}}{\in} \mathbf{V}, t \in (t_0, T)\}$, где \mathbf{V} - заданное множество из \mathbb{R}^n .

Задача оптимального управления заключается в нахождении точек минимума $u = u(t)$ - оптимальных управлений - на множестве \mathbf{U} функционала

$$J(u) = \int_{t_0}^T f^0(x(t; u), t, u(t)) dt + \Phi(x(T; u)).$$

где $f^0: \mathbb{R}^n \times \mathbb{R}^1 \times \mathbb{R}^m \rightarrow \mathbb{R}^1$, $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^1$ заданы. Пусть

$$f(x, t, u), \frac{\partial f}{\partial x}(x, t, u), f^0(x, t, u), \frac{\partial f^0}{\partial x}(x, t, u)$$

непрерывны по совокупности переменных (x, t, u) и удовлетворяют условию Липшица по x, u , $\Phi(x), \Phi'(x)$ удовлетворяют условию Липшица по x , множество \mathbf{V} - ограничено.

7.2 Принцип максимума

Функция Гамильтона-Понтрягина

$$H(x, t, u, \psi) = f^0(x, t, u) + \langle \psi, f(x, t, u) \rangle : \mathbb{R}^n \times \mathbb{R}^1 \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^1$$

где $\psi = (\psi_1, \psi_2, \dots, \psi_n)$ - сопряженная переменная

Теорема о принципе максимума Понтрягина Пусть в рассматриваемой задаче оптимального управления выполнены все указанные предположения, $u_* = u_*(t)$ - оптимальное управление, $x_*(t) = x(t; u_*)$ - соответствующая ему оптимальная траектория, $\psi_*(t)$ - соответствующее им решение сопряженной системы

$$\begin{cases} \dot{\psi}(t) \stackrel{\text{п.в.}}{=} -\frac{\partial H}{\partial x}(x_*(t), t, u_*(t), \psi(t)), & t_0 < t < T \\ \psi(T) = \Phi'(x_*(T)) \end{cases}$$

Тогда

$$H(x_*(t), t, u_*(t), \psi_*(t)) \stackrel{\text{п.в.}}{=} \min_{v \in \mathbf{V}} H(x_*(t), t, v, \psi_*(t)), t \in (t_0; T)$$

7.3 Схема применения

Итак, наша задача - найти оптимальное управление $u_*(t)$. В соответствии с принципом максимума, если взять какое-то управление $u_*(t) \in \mathbf{U}$ (сейчас мы только предполагаем, что оно оптимальное), то надо сделать следующее:

- 1) Подставить это управление в исходную задачу Коши, решить ее и найти соответствующую ему фазовую траекторию $x_*(t) = x(t; u_*)$;
- 2) Построить сопряженную задачу, подставить туда управление $u_*(t)$ и найденную траекторию $x_*(t)$, затем найти ее решение $\psi_*(t) = \psi(t; u_*)$;
- 3) Подставить функции $u_*(t), x_*(t), \psi_*(t)$ в функцию Гамильтона-Понтрягина и получить функцию

$$H_*(t) = H(x(t; u_*), t, u_*(t), \psi(t; u_*))$$

- 4) При каждом $t \in [t_0; T]$ построить функцию $H(v) = H(x(t; u_*), t, v, \psi(t; u_*))$ и проверить условие из принципа максимума

$$\min_{v \in \mathbf{V}} H(v) = H_*(t), \text{ минимум достигается при } v = u_*(t)$$

И если при почти всех $t \in (t_0; T)$ это верно, то управление $u_*(t)$ может быть оптимальным. - Надо решить континуальное число задач минимизаций.

На самом деле

- 1) Составляют $H(x, t, u, \psi)$ и рассматривают ее как функцию m переменных $u(u_1, \dots, u_m)$.
- 2) При фиксированном (x, t, ψ) решают задачу минимизации

$$H(x, t, u, \psi) \rightarrow \inf, \quad u \in \mathbf{V}$$

и находят $u = \bar{u}(x, t, \psi) \in \mathbf{V}$ такой что $H(x, t, \bar{u}(x, t, \psi), \psi) = \inf_{v \in \mathbf{V}} H(x, t, v, \psi)$

- 3) После этого рассматривают систему из $2n$ дифференциальных уравнений

$$\begin{cases} \dot{x}(t) \stackrel{\text{п.в.}}{=} f(x(t), t, \bar{u}((x, t), t, \psi(t))), \\ \dot{\psi}(t) \stackrel{\text{п.в.}}{=} -\frac{\partial H}{\partial x}(x(t), t, \bar{u}((x, t), t, \psi(t)), \psi(t)), \quad t_0 < t < T \\ x(t_0) = x^0, \quad \psi(T) = \Phi'(x(T)) \end{cases}$$

относительно $(x(t), \psi(t))$. Такая задача - *краевая задача принципа максимума*. Найдя ее решение $(x_*(t), \psi_*(t))$ можно утверждать, что управление

$$u_*(t) = \bar{u}(x_*(t), t, \psi_*(t))$$

может быть оптимальным. И оптимальное управление обязательно является решением краевой задачи принципа максимума.

8 Регуляризация по А.Н.Тихонову

8.1 Некорректно поставленные и неустойчивые экстремальные задачи

Определение Задача минимизации $J(u) \rightarrow \inf, u \in \mathbf{U}$ называется *корректно поставленной*, если

1. $J_* = \inf_{u \in \mathbf{U}} J(u) > -\infty, \quad \mathbf{U}_* = \{u \in \mathbf{U} : J(u) = J_*\} \neq \emptyset.$
2. Если $\{u_k\} \in \mathbf{U}, J(u_k) \rightarrow J_*$, то $u_k \xrightarrow{P} \mathbf{U}_*$ (то есть $\lim_{k \rightarrow \infty} \inf_{u \in \mathbf{U}} \rho(u_k, u) = 0$)

Определение Задача минимизации $J(u) \rightarrow \inf, u \in \mathbf{U}$ называется *слабо корректно поставленной*, если

1. $J_* = \inf_{u \in \mathbf{U}} J(u) > -\infty, \quad \mathbf{U}_* = \{u \in \mathbf{U} : J(u) = J_*\} \neq \emptyset.$
2. Если $\{u_k\} \in \mathbf{U}, J(u_k) \rightarrow J_*$, то $u_k \xrightarrow{\text{слабо}} \mathbf{U}_*$ (то есть если $u_{m_k} \xrightarrow{\text{слабо}} u_0$, то $u_0 \in \mathbf{U}_*$)

Определение Если малые изменения значений функционала $J(u)$ приводят к большой разнице между решениями исходной и приближенной задач, то такие задачи называют *неустойчивыми*.

8.2 Регуляризация Тихонова

Метод регуляризации Тихонова применяется, когда задача минимизации нарушает одно из условий сильной корректности. Также мы знаем что

$$|\tilde{J}(u) - J(u)| \leq \delta(1 + \|u\|^2) \quad \forall u \in \mathbf{U}$$

Рассматривается семейство задач минимизации функционала Тихонова

$$T_\alpha(u) = \tilde{J}(u) + \alpha\|u\|^2 \rightarrow \inf \quad u \in \mathbf{U}$$

$\alpha > 0$ - *параметр регуляризации*. Ищутся элементы $\tilde{u} = \tilde{u}(\alpha, \delta, \varepsilon)$ из условия

$$\tilde{u} \in \mathbf{U}, \quad T_\alpha(\tilde{u}) \leq \inf_{u \in \mathbf{U}} T_\alpha(u) + \varepsilon, \quad \varepsilon > 0$$

Теорема о сходимости метода Тихонова Пусть в задаче минимизации $J(u) \rightarrow \inf, u \in \mathbf{U}$ множество \mathbf{U} - выпукло и замкнуто, функционал $J(u)$ - выпуклый и полунепрерывный снизу на $\mathbf{U}, J_* > -\infty, \mathbf{U}_* \neq \emptyset$.

Тогда получаемое методом А.Н.Тихонова семейство точек $u_\alpha = u(\alpha(\delta), \delta, \varepsilon(\delta))$ при выполнении условий согласования

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = \lim_{\delta \rightarrow 0} \varepsilon(\delta) = 0, \lim_{\delta \rightarrow 0} \frac{\delta}{\alpha(\delta)} = 0, \lim_{\delta \rightarrow 0} \frac{\varepsilon(\delta)}{\alpha(\delta)} = 0$$

обладает свойствами

$$\lim_{\delta \rightarrow 0} \tilde{J}(u(\alpha(\delta), \delta, \varepsilon(\delta))) = J_*, \quad \lim_{\delta \rightarrow 0} \|u(\alpha(\delta), \delta, \varepsilon(\delta)) - u_*\|_{\mathbb{H}} = 0$$

где u_* - *нормальное решение* исходной задачи (т.е. решение, минимальное по норме).