Вариационное исчисление. Неофициальный конспект

Лектор: Роман Владимирович Романов
Конспектировал Леонид Данилевич
Редактировал Максим Лаунер

IV семестр, весна 2024 г.

Содержание

Лекция I

15 февраля 2024 г.

1 Что мы будем изучать

Вариационное исчисление занимается поиском экстремумов в задаче, где число переменных бесконечно.

Рассмотрим конечномерную ситуацию. Пусть имеется $f : M \to \mathbb{R}$, где M — какое-то многообразие.

При поиске экстремумов формируются следующие направления:

1. Необходимое условие: $(\nabla f)(x) = 0$.
2. Достаточное: форма $(D^2 f)(x)$ знакоопределён ($>< 0$).
3. Поиск экстремумов сужения $f\big|_N$ на подмногообразие (метод множителей Лагранжа).

В случае вариационного исчисления вместо M стоит некоторое бесконечномерное пространство, например, пространство функций. В основном мы будем заниматься аналогами 1 и 3 пунктов.

Функция, которая в свою очередь задана на пространстве функций часто называется *функционал*. Чтобы визуально различать «обычные» функции, и функционалы, образ точки f под действием функционала J будем обозначать $J[f]$.

Пускай X — (пока произвольное) метрическое пространство, $J : X \to \mathbb{R}$ — функция.

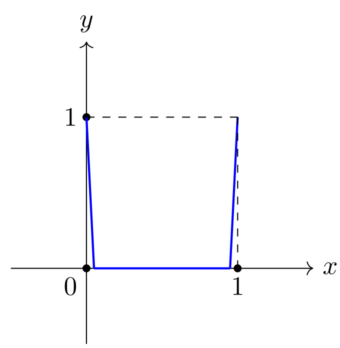
Определение 1.1 ($x \in X$ — строгий локальный минимум)**.** $\exists \delta > 0 : \forall y \in U_\delta(x) : J[y] > J[x]$.

Аналогично определяются нестрогий минимум и максимумы. Также стоит вспомнить про существование глобальных строгих и нестрогих минимумов и максимумов.

Пример (Чего такого особенного в бесконечномерии?). Пусть $X = \{f \in C[0,1] | f(0) = f(1) = 1\}$, норма на $C[0,1]$ определена формулой $\|f\| = \max\limits_{x \in [0,1]} |f(x)|$.

Пусть $J[f] \coloneqq \int\limits_0^1 f^2(x)\, \mathrm{d}x$. Очевидно, J непрерывен.

Ясно, что $\forall f \in X : J[f] > 0$. С другой стороны, $\inf\limits_{f \in X} J[f] = 0$ — можно рассматривать функции вида



С третьей стороны, X замкнуто: равномерный предел равномерных непрерывен, и условия на значения на концах уважают предел. Получается, в данном случае теорема Кантора не работает. В чём дело?

Оказывается, проблема в том, что нет компактности: в бесконечномерном пространстве замкнутое ограниченное множество необязательно компактно.

1.1 Интегральные функционалы

В дальнейшем мы будем рассматривать не произвольные функционалы, а ограничимся некоторым их подмножеством.

Пусть задано непрерывное $L : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, положим $J[u] := \int\limits_a^b L(t, u(t), \dot{u}(t))\,\mathrm{d}t$. Мы будем заниматься множеством $X = C^1[a,b] = C^1([a,b] \to \mathbb{R}^n)$ (далее не будем указывать область значений, ясно из контекста) и его замкнутыми подмножествами.

Такие J называются *интегральные функционалы*. Мы их изучаем, так как на них возможна богатая теория, и вместе с тем, интегральные функционалы часто встречаются в приложениях.

Примеры.

- $X = \left\{u \in C^1[a,b] \middle| u(a) = u_a, u(b) = u_b\right\}$, $J[u] = \int\limits_a^b \sqrt{1 + (u')^2}\,\mathrm{d}x$ — функционал длин графиков кривых, соединяющих две данные точки.
- $J = \int\limits_a^b (\frac{\dot{u}^2}{2} - V(u))\,\mathrm{d}x$, где V — заданная функция. В механике называется *действием*.

Сначала убедимся, что они непрерывны.

Замечание (О норме). Для $f \in C^1[a,b]$: $\|f\| = \max\limits_{x\in[a,b]} |f(x)| + \max\limits_{x\in[a,b]} |f'(x)|$ — очевидно норма. В дальнейшем мы всегда будем использовать такую норму для C^1.

Предложение 1.1. *Пусть $X = C^1[a,b], L \in C([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$. Тогда интегральный функционал J непрерывен на X.*

Доказательство. Пусть $u, \widetilde{u} \in X, \|u - \widetilde{u}\| < \delta < 1$.

$$|J[u] - J[\widetilde{u}]| = \left| \int\limits_a^b L(x, \widetilde{u}(x), \dot{\widetilde{u}}(x)) - L(x, u(x), \dot{u}(x))\,\mathrm{d}x \right| \leqslant$$

Заметим, что $\|(x, \widetilde{u}(x), \dot{\widetilde{u}}(x)) - (x, u(x), \dot{u}(x))\|_{\mathbb{R}^{2n+1}} < \delta$

Рассмотрим $K = [a,b] \times \overline{B_{\|u\|_X+1}} \times \overline{B_{\|u\|_X+1}}$ — компакт в $\mathbb{R}^{2n+1}$.

$$\leqslant \int\limits_a^b \omega_{L\big|_K}(\delta)\,\mathrm{d}x = (b-a)\omega_{L\big|_K}(\delta) \xrightarrow[\delta\to 0]{} 0$$

где ω — модуль непрерывности. Он определён, так как $L\big|_K$ непрерывна на компакте. ☐

Пусть X — нормированное пространство (необязательно замкнутое), $J : X \to \mathbb{R}$.

Определение 1.2 (Производная функционала J в точке x по направлению $h \in X$)**.**

$$\delta J[x,h] = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} J[x + th]$$

Иначе эту штуку называют *вариация* J по направлению h.

Свойства (Вариация).

- Однородность: $\delta J[x, ch] = c \cdot \delta J[x,h]$.
- Не следует ожидать аддитивность. Так, $\exists \delta J[x, h_1], \delta J[x, h_2]$ не влечёт существование $\delta J[x, h_1 + h_2]$, а если последнее и существует, то не обязано быть суммой.

  Примеры этого были в анализе, здесь бесконечномерной специфики нет.

- Как и в конечномерном анализе, в критической (экстремальной) точке вариация (коли существует) должна обращаться в нуль.

  А именно, $x \in X$ — локальный экстремум J, тогда $\forall h : \exists \delta J[x,h] \Rightarrow \delta J[x,h] = 0$.

  Доказательство. Сужение $\alpha(t) = J[x+th]$ тоже имеет локальный экстремум, значит, если производная в $t=0$ есть, то она равна нулю. $\square$

2 Формула первой вариации. Уравнение Эйлера — Лагранжа

2.1 Лемма Дюбуа-Реймона

Лемма 2.1 (Дюбуа-Реймон)**.** *Пускай $f \in C[a,b]$, и для всех $\omega \in C^1[a,b]$, таких, что $\omega(a) = \omega(b) = 0$, известно, что $\int\limits_a^b f\omega' = 0$.*

Тогда $f \equiv \mathrm{const}$.

Доказательство. Если бы f сама была гладкой, то можно было бы интегрировать по частям. $\int f'\omega = 0 \Rightarrow f' \equiv 0$ — можно взять ω, сосредоточенную там, где f' одного знака.

Мы надеемся, что f — константа, то есть равна своему среднему $\overline{f} \overset{def}{=} \frac{1}{b-a}\int\limits_a^b f$.

Определим ω, интегрируя $f - \overline{f}$: $\omega(x) \coloneqq \int\limits_a^x \left(f(x') - \overline{f}\right) \mathrm{d}x'$. Понятно, что $\omega \in C^1$. Более того, несложно видеть, что $\omega(a) = \omega(b) = 0$.

Подставим данную ω в посылку теоремы.

$$0 = \int\limits_a^b f\omega' = \int\limits_a^b (f - \overline{f})\omega' = \int\limits_a^b (f - \overline{f})^2 \,\mathrm{d}x$$

Так как интеграл нуль, то получаем $f \equiv \overline{f}$. $\square$

2.2 Формула первой вариации

Опять $X = C^1[a,b]$, и функционал того же самого вида $J[u] = \int\limits_a^b L(t, u(t), \dot{u}(t))\,\mathrm{d}t$.

Лемма 2.2 (Формула первой вариации)**.** *Пусть $L \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$. Градиент L по второму и третьему аргументам будем обозначать $\nabla_u L$ и $\nabla_{\dot{u}} L$ соответственно, это векторы из $\mathbb{R}^n$.*

Тогда производная J в точке u по направлению h существует, и равна

$$\int\limits_a^b \left[\Big\langle (\nabla_u L)(t, u(t), \dot{u}(t)), h(t) \Big\rangle + \Big\langle (\nabla_{\dot{u}} L)(t, u(t), \dot{u}(t)), \dot{h}(t) \Big\rangle\right] \mathrm{d}t$$

Доказательство. $J[u+\tau h] - J[u] = \int\limits_a^b \left[L\Big(t, u(t) + \tau h(t), \dot{u}(t) + \tau\dot{h}(t)\Big) - L\Big(t, u(t), \dot{u}(t)\Big)\right] \mathrm{d}t$.

Применяя формулу Лагранжа, получаем для некой $\tau_* = \tau_*(t) \in [0, \tau]$:

$$J[u+\tau h] - J[u] = \tau \int\limits_a^b \Big[\Big\langle (\nabla_u L)(t, u(t) + \tau_* h(t), \dot{u}(t) + \tau_* \dot{h}(t)), h(t) \Big\rangle + \\ + \Big\langle (\nabla_{\dot{u}} L)(t, u(t) + \tau_* h(t), \dot{u}(t) + \tau_* \dot{h}(t)), \dot{h}(t) \Big\rangle\Big] \mathrm{d}t$$

Поделив на τ, получаем $\frac{J[u+\tau h]-J[u]}{\tau} = \int\limits_a^b \ldots$ — вот тот, что выше.

Сперва разберёмся с первым слагаемым. Покажем, что

$$\underbrace{\int\limits_a^b \left\langle (\nabla_u L)(t, u(t) + \tau_* h(t), \dot{u}(t) + \tau_* \dot{h}(t)), h(t) \right\rangle \mathrm{d}t}_{I} \xrightarrow[\tau \to 0]{} \underbrace{\int\limits_a^b \langle (\nabla_u L)(t, u(t), \dot{u}(t)), h(t) \rangle \mathrm{d}t}_{I\!I}$$

Модуль разности аргументов не превосходит $\tau_* \|h\|_X$. Отсюда $\|\nabla_u L(\ldots) - \nabla_u L(\ldots)\|_{\mathbb{R}^n} \leqslant \omega_{L\big|_K}(\tau_* \|h\|_X)$, здесь $K \coloneqq [a, b] \times \overline{B_{\|u\|+\|h\|}} \times \overline{B_{\|u\|+\|h\|}}$ (мы считаем, что $\tau \leqslant 1$, откуда $\tau_* \leqslant 1$).

Значит, $|(I) - (I\!I)| \leqslant \int\limits_a^b \omega_{L\big|_K}(\tau_* \|h\|)\,\mathrm{d}t \leqslant (b-a)\omega_{L\big|_K}(\tau \|h\|)\,\mathrm{d}t \xrightarrow[\tau \to 0]{} 0$.

Таким образом, у первого слагаемого под интегралом — естественный предел. Аналогично со вторым слагаемым, получаем утверждение леммы. $\square$

2.3 Уравнение Эйлера — Лагранжа

Пусть $u \in X$ — экстремум. Тогда $\forall h \in X : \delta J[u, h] = 0$.

Условие обнуления градиента — некое уравнение на точку. Мы хотим уравнение на $u(t)$, избавимся от h. Подгоним под лемму Дюбуа-Реймона (лемма 2.1).

Введём $R(x) \coloneqq \int\limits_a^x (\nabla_u L)(t, u(t), \dot{u}(t))\,\mathrm{d}t$. Согласно (лемма 2.2)

$$\delta J[x, h] = \int\limits_a^b \left\langle \dot{R}(t), h(t) \right\rangle + \left\langle (\nabla_{\dot{u}} L)(t, u(t), \dot{u}(t)), \dot{h}(t) \right\rangle \mathrm{d}t$$

Поскольку $R(a) = 0$, интегрируя по частям, получим $\langle R(b), h(b) \rangle + \int\limits_a^b \Big\langle \underbrace{(\nabla_{\dot{u}} L)(t, u(t), \dot{u}(t)) - R(t)}_{\xi(t)}, \dot{h}(t) \Big\rangle \mathrm{d}t$

И это равно нулю $\forall h \in C^1[a, b]$. Рассмотрим h, обращающийся на концах в ноль: $h(a) = h(b) = 0$. Теперь $\int\limits_a^b \left\langle \xi(t), \dot{h}(t) \right\rangle \mathrm{d}t = 0$, и мы покомпонентно можем применить лемму Дюбуа-Реймона, получая $\xi(t) = C \equiv \mathrm{const}$. Но $R(t) \in C^1$, значит, $\nabla_{\dot{u}} L(t, u(t), \dot{u}(t)) \in C^1$ тоже.

Дифференцируя ξ, получаем уравнение: $\frac{\mathrm{d}}{\mathrm{d}t}(\nabla_{\dot{u}} L)(t, u(t), \dot{u}(t)) - (\nabla_u L)(t, u(t), \dot{u}(t)) = 0$. Оно называется *уравнение Эйлера — Лагранжа*, это основное уравнение вариационного исчисления.

Замечание. В случае общего положения уравнение Эйлера — Лагранжа — дифференциальное второго порядка, что соответствует $u \in C^2$: при вычислении $\frac{\mathrm{d}}{\mathrm{d}t}(\nabla_{\dot{u}} L)(t, u(t), \dot{u}(t))$ появится в общем случае вторая производная u. Такая ситуация, на самом деле, довольно общая: экстремаль «регулярнее», чем произвольный элемент своего пространства (например, (предложение 2.1))).

2.4 Случай свободных концов

Теперь рассмотрим совсем произвольную $h \in C^1$, получим уравнение ($C = \xi$ — определена выше):

$$0 = \delta J[u, h] = \langle R(b), h(b) \rangle + \int\limits_a^b \left\langle C, \dot{h}(t) \right\rangle \mathrm{d}t = \langle R(b), h(b) \rangle + \langle C, h(b) \rangle - \langle C, h(a) \rangle$$

1. Рассмотрим такую h, что $h(b) = 0, h(a) = C$. Для неё $\delta J[u, h] = -\|C\|^2$, значит, $\xi = C = 0$.

   Подставляя в определение ξ, получаем $R(a) = 0$, то есть $(\nabla_{\dot{u}} L)(a, u(a), \dot{u}(a)) = 0$.

2. Теперь рассмотрим такую h, что $h(b) = R(b)$. В этом случае $\delta J[u, h] = \|R(b)\|^2 \Rightarrow R(b) = 0$. Получили $(\nabla_{\dot{u}} L)(b, u(b), \dot{u}(b)) = 0$.

Итак, помимо уравнения Эйлера — Лагранжа, мы получили два условия (но в разных точках) на уравнение второго порядка, можно надеяться, что хватит, чтобы найти решения (но это совсем не факт — так, может существовать одно решение, а может их вовсе не быть, или быть бесконечно много).

Подытожим в теорему.

Теорема 2.1 (Задача со свободными концами)**.** Пусть $L \in C^1([a, b] \times \mathbb{R}^n \times \mathbb{R}^n)$, пусть $X = C^1[a, b]$, пусть u — локальный экстремум J.

Тогда

1. $\left(\nabla_{\dot{u}} L\right)(t, u(t), \dot{u}(t)) \in C^1[a, b]$.
2. $\frac{\mathrm{d}}{\mathrm{d}t} \nabla_{\dot{u}} L = \nabla_u L$ — уравнение Эйлера — Лагранжа.
3. $(\nabla_{\dot{u}} L)(a, u(a), \dot{u}(a)) = 0$
4. $(\nabla_{\dot{u}} L)(b, u(b), \dot{u}(b)) = 0$

2.5 Случай фиксированных концов

Теперь обсудим, что происходит, если имеются граничные условия на функции из множества, по котором ищутся экстремумы функционалов.

Рассмотрим $X = \left\{ f \in C^1[a, b] \middle| f(a) = f_a, f(b) = f_b \right\}$. Это не подпространство (не имеет линейной структуры), нельзя определить производную по направлению.

Функционал $J : X \to \mathbb{R}$ задан той же формулой.

Какая здесь характеризация локальных экстремумов?

Рассмотрим $\widetilde{J} : C^1[a, b] \to \mathbb{R}$ — с той же формулой, что и J. Тогда $\forall u, h : \exists \delta \widetilde{J}[u, h]$.

С другой стороны, если $h \in C^1[a, b], h(a) = h(b) = 0$, то $\forall u \in X, t \in \mathbb{R} : u + th \in X$. Имеем право рассмотреть $J[u + th]$. Если u — локальный экстремум, то $\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} J[u + th] = 0$. Она существует, так как это $\frac{\mathrm{d}}{\mathrm{d}t} \widetilde{J}[u + th]$.

Тем самым, такие функции h прибавлять можно, будем это тоже называть вариацией: $\delta J[u, h]$ задаётся той же формулой. Дальше работает то же самое рассуждение, все действия те же самые, только при интегрировании по частям внеинтегральный член занулится, никаких дополнительных соотношений не возникнет.

Теорема 2.2 (Задача с фиксированными концами)**.** Пусть $L \in C^1([a, b] \times \mathbb{R}^n \times \mathbb{R}^n)$, пусть $X = \left\{ f \in C^1[a, b] \middle| f(a) = f_a, f(b) = f_b \right\}$, пусть u — локальный экстремум J. Тогда

1. $(\nabla_{\dot{u}} L)(t, u(t), \dot{u}(t)) \in C^1[a, b]$.
2. $\frac{\mathrm{d}}{\mathrm{d}t} \nabla_{\dot{u}} L = \nabla_u L$ — уравнение Эйлера — Лагранжа.

Заметим, что у нас по-прежнему два условия (теперь уже данные в самой задаче) и уравнение второго порядка, значит, по-прежнему, данных для решения задачи как раз столько, что стоит надеяться на получение решения.

Лекция II

29 февраля 2024 г.

Распишем чуть подробнее уравнение Эйлера — Лагранжа, пусть для определённости функции скалярны ($d = 1$). Если все производные определены, то оно имеет вид

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{u}} = \frac{\partial^2 L}{\partial t \partial \dot{u}} + \frac{\partial^2 L}{\partial u \partial \dot{u}} \dot{u} + \frac{\partial^2 L}{\partial \dot{u}^2} \ddot{u} \tag{$*$}$$

Общая теорема (теорема 2.1) говорит, что $\nabla_{\dot u} L$ имеет C^1 гладкость, однако совсем не утверждается, что при разложении $(*)$ каждое слагаемое будет гладким, или даже просто будет существовать. И правда, такого и не наблюдается.

Контрпример. Рассмотрим функционал $J[u] = \int\limits_{-1}^{1} u^2(\dot u - 2x)^2 \,\mathrm{d}x$, где $X = \left\{ u \in C^1[-1,1] \,\middle|\, \begin{matrix} u(-1) = 0 \\ u(1) = 1 \end{matrix} \right\}$ и функцию $u \in X, u(t) = \begin{cases} 0, & x \in [-1,0] \\ x^2, & x \in [0,1] \end{cases}$. u — экстремаль, например, потому что это глобальный минимум. При этом $u \notin C^2$, хотя $\frac{\partial L}{\partial \dot u} = 2u^2(\dot u - 2x) \equiv 0$ — бесконечно гладкая.

Что нужно потребовать, чтобы все слагаемые $(*)$ существовали?

В примере сам лагранжиан $L(x, u, \dot u) = u^2(\dot u - 2x^2)$ — бесконечно гладкий. Но $\ddot u$ можно выразить из $(*)$ только если $\frac{\partial^2}{\partial \dot u^2} L \neq 0$.

Следующее предложение формулируется в случае, когда L задан на $[a,b] \times \mathbb{R}^d \times \mathbb{R}^d$; в общем случае сужения L на некоторое подмножество принципиально ничего не поменяется.

Предложение 2.1. *Пусть $L \in C^2(\Omega)$, где $\Omega = [a,b] \times \mathbb{R}^d \times \mathbb{R}^d$, пусть $\det\left(\frac{\partial^2 L}{\partial \dot u^2}\right) \neq 0$ везде в Ω.*

Пусть u — локальный экстремум функционала J. Утверждается, что $u \in C^2[a,b]$.

Доказательство. Введём функцию

$$\begin{aligned} \xi : [a,b] \times \mathbb{R}^d &\to \mathbb{R}^d \\ (t, v) &\mapsto (\nabla_{\dot u} L)(t, u(t), \dot u(t)) - (\nabla_{\dot u} L)(t, u(t), v) \end{aligned}$$

Согласно посылке теоремы, $\frac{\partial}{\partial v}\xi \neq 0$ для всех t, v. Так как u — экстремум, то $\xi \in C^1$.

По теореме о неявной функции $\forall t_0 \in (a,b) : \exists \delta > 0 : \{(t,v) | \xi(t,v) = 0, |t - t_0| < \delta\}$ — график некоторой функции $v \in C^1\left((t_0 - \delta, t_0 + \delta) \to \mathbb{R}^d\right)$. Но $v \equiv \dot u\big|_{(t_0-\delta, t_0+\delta)}$. Значит, $u \in C^2(a,b)$.

Случай концов ($t_0 = a, b$)— упражнение. $\square$

3 Условные экстремумы

Согласно полуисторической, полулегендарный справке, некогда Дидона прибыла на берег некоего африканского государства, и потребовала, на основании своего высокого происхождения, выделить ей столько земли, сколько можно опоясать ремешком из шкуры одного быка...

Напоминание конечномерного случая: пусть $\Omega \subset \mathbb{R}^d$ — область, $f, g \in C^1(\Omega)$, $\mathcal{M} = \{x \in \Omega | g(x) = 0\}$.

Заинтересуемся экстремумами сужения $f\big|_{\mathcal{M}}$. Пусть $x_0 \in \Omega$ – экстремум. Построим кривую $x : (-\varepsilon, \varepsilon) \to \Omega$ так, что $x(0) = x_0$. Условие $g(x(t)) \equiv 0$ влечёт, что $f(x(t))$ имеет локальный экстремум в нуле. Другими словами, $\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x(t)) = \langle (\nabla f)(x_0), \dot x(t) \rangle = 0$.

Поскольку кривую можно выбрать с любым вектором скорости, то $(\nabla f)(x_0) \perp T_{x_0}\mathcal{M}$. Если $(\nabla g)(x_0) \neq 0$ в x_0, то $T_{x_0}\mathcal{M}$ — пространство коразмерности 1. Найдём какой-нибудь вектор, перпендикулярный $\mathcal{M}$. Это как раз градиент: $g(x(t)) = 0 \Rightarrow \langle (\nabla g)(x_0), \dot x \rangle = 0$.

Иными словами $\exists \lambda \in \mathbb{R} : \nabla(f - \lambda g)(x_0) = 0$. Далее для поиска экстремумов ищут критические точки $f - \lambda g$, выделяют те, которые в $\mathcal{M}$, а с обнулениями градиента g разбираются отдельно.

Пускай X — нормированное замкнутое пространство, $G \in C^1(X)$ — задающий условие функционал. Расшифруем условие $G \in C^1(X)$:

- $\forall x \in X : \exists G'(x) \in X^* : |G(x+s) - G(x) - G'(x)s| = o(\|s\|)$ — *сильная дифференцируемость* в точке x.

- Определённое в предыдущем пункте отображение $G' : X \to X^*$ непрерывно.

Для применения метода множителей Лагранжа нам понадобится лемма, утверждающая, что в направлении всякого вектора из $\operatorname{Ker} G'(x_0)$ можно пустить путь — как в конечномерном случае.

Лемма 3.1. *Пусть* $x_0 \in \mathcal{M} \coloneqq \{x \in X | G(x) = 0\}$. *Пусть* $G'(x_0) \neq 0$.

Тогда $\forall h \in \operatorname{Ker} G'(x_0) : \exists x \in C^1((-\delta, \delta) \to \mathcal{M}) : x(0) = x_0, \dot{x}(0) = h$.

Доказательство. Зафиксируем произвольный $\xi \notin \operatorname{Ker} G'(x_0)$. Определим $r(t, \tau) \coloneqq G[x_0 + t\xi + \tau h]$. Ясно, что $r \in C^1\left([-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]\right)$.

$r(0, 0) = 0$, $\frac{\partial r}{\partial t}(0, 0) = G'[x_0]\xi \neq 0$. Применяя теорему о неявной функции, получаем $\exists \delta > 0$: $\{(t, \tau) | \tau \in (-\delta, \delta), r(t, \tau) = 0\}$ — график C^1 функции $t = t(\tau)$, где $t : (-\delta, \delta) \to \mathbb{R}$.

Проверим, что $x : \tau \mapsto x_0 + t(\tau)\xi + \tau h$ — искомая кривая:

1. По построению $x : (-\delta, \delta) \to \mathcal{M}$ — класса C^1.
2. Дифференцируя тождество $G[x(t)] = 0$ в точке $t = 0$, получаем $G'[x(0)] \cdot \dot{x}(0) = 0$, значит, $\dot{x}(0) \in \operatorname{Ker} G'[x_0]$. С другой стороны, $\dot{x}(0) = \dot{t}(0)\xi + h$, откуда $\dot{t}(0) = 0$ (ведь $\xi \notin \operatorname{Ker} G'[x_0]$). Тем самым, $\dot{x}(0) = h$. □

Пускай $F \in C^1(X), x_0 \in \mathcal{M}$ — точка локального экстремума сужения $F\big|_{\mathcal{M}}$.

Рассмотрим только что построенную кривую $x(\tau)$ ($x'(0) = h$). Так как x_0 — экстремаль, то в частности должно быть $\frac{\mathrm{d}}{\mathrm{d}\tau}\Big|_{\tau=0} F[x(\tau)] = 0$. С другой стороны, это равно $F'[x_0] \cdot h$.

Значит, $\operatorname{Ker} G'[x_0] \subset \operatorname{Ker} F'[x_0]$, причём $\exists \lambda \in \mathbb{R} : (F'[x_0] - \lambda G'[x_0]) = 0$ — и F', и G' обнуляются на пространстве коразмерности 1. Формальнее, выберем $\eta \notin \operatorname{Ker} G'[x_0]$, и разложим

$$\forall h \in X : h = \underbrace{\left(h - \frac{G'[x_0]h}{G'[x_0]\eta}\eta\right)}_{\in \operatorname{Ker} G'[x_0]} + \frac{G'[x_0]h}{G'[x_0]\eta}\eta$$

Значит, $(F'[x_0] - \lambda G'[x_0])(h) = \frac{G'[x_0]h}{G'[x_0]\eta} \cdot (F'[x_0]\eta - \lambda G'[x_0]\eta)$. Видно, что подойдёт $\lambda = \frac{F'[x_0]\eta}{G'[x_0]\eta}$.

Получилась теорема:

Теорема 3.1. Пускай $F, G \in C^1(X)$, пускай x_0 — точка локального экстремума F на $\mathcal{M} \coloneqq \{x \in X | G[x] = 0\}$, пусть $G'[x_0] \neq 0$.

Тогда $\exists \lambda \in \mathbb{R} : \forall h \in X : \delta(F - \lambda G)[x_0, h] = 0$ (отметим, что так как $F, G \in C^1$, то $\exists \delta(F - \lambda G)$.)

Упражнение 3.1. *Задача с фиксированными концами*

3.1 Случай нескольких условий

Даны $F, G_1, \ldots, G_n \in C^1(X), \mathcal{M} \coloneqq \{x \in X | G_1[x] = \cdots = G_n[x] = 0\}$.

Образуем линейный оператор $\mathbb{G}'[x_0] = \begin{pmatrix} G_1'[x_0] \\ \vdots \\ G_n'[x_0] \end{pmatrix} : X \to \mathbb{R}^n$ по правилу $\mathbb{G}'[x_0]h = \begin{pmatrix} G_1'[x_0]h \\ \vdots \\ G_n'[x_0]h \end{pmatrix}$.

Теорема 3.2. Пусть x_0 — точка локального экстремума F на $\mathcal{M}$, пусть $\operatorname{Ran} \mathbb{G}'[x_0] = \mathbb{R}^n$ (иными словами $\sum\limits_{j=1}^{n} c_j G_j'[x_0] = 0 \Rightarrow \forall j : c_j = 0$) (Ran (от англ. range) — образ).

Тогда $\exists \lambda_1, \ldots, \lambda_n \in \mathbb{R} : \forall h \in X : \delta(F - \sum \lambda_j G_j)[x_0, h] = 0$

Доказательство.

Лемма 3.2. *В тех же предположениях невырожденности* $\operatorname{Ran}\mathbb{G}'(x) = \mathbb{R}^n$. *Пусть* $h \in \bigcap\limits_{j=1}^{n} \operatorname{Ker} G'_j[x_0]$. *Тогда* $\exists x : (-\delta, \delta) \to \mathcal{M}, x \in C^1, x(0) = x_0, \dot{x}(0) = h$.

Доказательство леммы.

Аналогично (лемма 3.1), тоже теорема о неявной функции. □

Полностью аналогично скалярному случаю $n = 1$. □

Замечание. Можно попробовать применить скалярную теорему с $n = 1$ к функционалу $G[x] = \sum\limits_{i=1}^{n} G_i^2[x]$. Однако это ничего не даст, так как $G'[x] = 0$ везде на $\mathcal{M}$.

Упражнение 3.2. *Доказать аналогичное утверждение для задачи с фиксированными концами.*

Теперь применим (теорема 3.2) к случаю интегральных функционалов.

Теорема 3.3. Пускай $L, r_1, \ldots, r_n \in C^1([a,b] \times \mathbb{R}^d \times \mathbb{R}^d)$, и определены $J[u] \coloneqq \int\limits_a^b L(t, u(t), \dot{u}(t))\,\mathrm{d}t$, $R_j[u] \coloneqq \int\limits_a^b r_j(t, u(t), \dot{u}(t))\,\mathrm{d}t$.

Пусть u_0 — точка локального экстремума $J\big|_{\bigcap \operatorname{Ker} R_j}$; оператор $\mathbb{R}(u_0) \coloneqq \begin{pmatrix} R'_1(u_0) \\ \vdots \\ R'_n(u_0) \end{pmatrix}$ имеет полный ранг (можно явно выразить производные при помощи (лемма 2.2))). Вообще, если я правильно понимаю, лемма даёт производные по Гато, что не гарантирует существование производных по Фреше. Тем не менее, несложно видеть, что доказательство не меняется, если считать производные по Фреше — все оценки прячутся в $o(\|h\|)$.

Тогда

1. $\exists \lambda_1, \ldots, \lambda_n : \nabla_{\dot{u}}(L - \sum \lambda_j r_j)(t, u_0(t), \dot{u}_0(t)) \in C^1[a,b]$.
2. Выполнено уравнение Эйлера — Лагранжа: $\frac{\mathrm{d}}{\mathrm{d}t}\nabla_{\dot{u}}(L - \sum \lambda_j R_j) = \nabla_u(L - \sum \lambda_j R_j)$
3. $\nabla_{\dot{u}}(L - \sum \lambda_j r_j)\Big|_{t=a} = \nabla_{\dot{u}}(L - \sum \lambda_j r_j)\Big|_{t=b} = 0$.

Доказательство. Вытекает из доказательства (теорема 2.1) (там использовалось только то, что вариация обращается в нуль, а не то, что u_0 — экстремаль) и (теорема 3.1). □

Пример. Пускай $\Omega \subset \mathbb{R}^3$ — ограниченная область, граница которой — поверхность класса C^1. Введём пространство непрерывных на этой границе функций $X \coloneqq C(\partial\Omega)$.

Заведём $J[\sigma] \coloneqq \iint\limits_{\partial\Omega\,\partial\Omega} \frac{\sigma(x)\sigma(y)\,\mathrm{d}S(x)\,\mathrm{d}S(y)}{|x-y|}$.

J непрерывен на X, поскольку $\xi : y \mapsto \int\limits_{\partial\Omega} \frac{\sigma(x)}{|x-y|}\,\mathrm{d}S(x)$ непрерывно.

$$J[\sigma + s] - J[\sigma] = 2\iint\limits_{\partial\Omega\,\partial\Omega} \frac{s(x)\sigma(y)}{|x-y|}\,\mathrm{d}S(x)\,\mathrm{d}S(y) + \mathcal{O}(\|s\|_C^2)$$

Также $s \mapsto \int\limits_{\partial\Omega} s(x)\xi(x)\,\mathrm{d}x$ непрерывен, откуда J — даже функционал класса $C^1(X)$.

Лекция III

14 марта 2024 г.

Заинтересуемся экстремумами с постоянным значением $G[\sigma] = \int\limits_{\partial\Omega} \sigma(x)\,\mathrm{d}x$. Решение будет отвечать распределению зарядов на поверхности, минимизирующее энергию системы — физический принцип говорит, что конечное положение экстремально.

Уже проверили, что $J, G \in C^1(X)$.

Пусть σ — экстремаль $J\big|_{\{\sigma\in X\,|\,G(\sigma)=Q\}}$. Тогда $\forall h \in X$: $\delta(J-\lambda G)[\sigma,h] = 0$. Посчитаем

$$\delta(J-\lambda G)[\sigma,h] = (J-\lambda G)[\sigma+h] - (J-\lambda G)[\sigma] = \\ = 2\int\limits_{\partial\Omega}\int\limits_{\partial\Omega} \frac{\sigma(x)h(y)}{|x-y|}\,\mathrm{d}S(x)\,\mathrm{d}S(y) - \lambda\int\limits_{\partial\Omega} h(y)\,\mathrm{d}S(y) + \int\limits_{\partial\Omega}\int\limits_{\partial\Omega} \frac{h(x)h(y)}{|x-y|}\,\mathrm{d}S(x)\,\mathrm{d}S(y)$$

Третье слагаемое $\mathcal{O}(\|h\|_X^2)$: $\iint \frac{1}{|x-y|}$ сходится. Заметим, что остальная часть — линейный функционал от h, где коэффициент непрерывен от σ. Это в точности значит, что $J \in C^1$.

$$J[\sigma+h] - J[\sigma] = l_\sigma(h) + o(\|h\|), \text{ где } l_\sigma : h \mapsto \iint\limits_{\partial\Omega\times\partial\Omega} \frac{\sigma(x)h(y)}{|x-y|}\,\mathrm{d}S(y) = \int\limits_{\partial\Omega} \xi(y)h(y)\,\mathrm{d}S(y)$$

Запишем условие экстремальности:

$$\forall h : \delta(J-\lambda G)[\sigma,h] = 2\int h(y)\,\mathrm{d}S(y)\left(2\int \frac{\sigma(x)\,\mathrm{d}S(x)}{|x-y|} - \lambda\right) = 0$$

По «нулевой лемме Дюбуа-Реймона», выражение в скобочках должен быть всегда нулём.

Получили

$$\boxed{\lambda = 2\int \frac{\sigma(x)\,\mathrm{d}S(x)}{|x-y|}}$$

Экстремаль σ ищется, как решение «интегрального уравнения»: заведём $K : f \mapsto \int \frac{f(x)\,\mathrm{d}S(x)}{|x-y|}$, это ограниченный непрерывный интегральный оператор. Таким образом, σ — решение $K\sigma = \frac{\lambda}{2}\mathbb{1}$.

Иными словами, потенциал, создаваемый распределением заряда на самой поверхности постоянен. Такая постановка задачи не очень естественна — например, бывают точечные заряды. Естественнее было бы рассматривать задачи вида σ — борелевская мера на $\partial\Omega$, $J[\sigma] = \int \frac{\mathrm{d}\sigma(x)\,\mathrm{d}\sigma(y)}{|x-y|}$. Тут уже уместно задавать вопросы о существовании интеграла, сходимости, и прочем, мы не будем это выяснять по причине нехватки аппарата.

4 Функционалы на кривых

В задаче Дидоны ответом является некоторая кривая, а кривая, в зависимости от того, как выбрать параметр, может не реализовываться, как график функции. С другой стороны, хотим независимость от параметризации, потому что зачем.

Определение 4.1 (Кривая $\gamma \in C$)**.** Непрерывное отображение $\gamma : [a,b] \to \mathbb{R}^n$.

Определение 4.2 (Параметризованная кривая γ)**.** $\forall x \in [a,b] : \gamma'(x) \neq 0$.

Определение 4.3 (Кривая γ класса C^j)**.** Кривая $\gamma \in C^j$.

Пусть $\gamma_1 : [a_1,b_1] \to \mathbb{R}^n, \gamma_2 : [a_2,b_2] \to \mathbb{R}^n$ — две параметризованные кривые.

Определение 4.4 (Эквивалентность кривых γ_1 и γ_2)**.** Диффеоморфизм $\kappa \in C^j([a_1,b_1] \to [a_2,b_2])$, такой, что $\gamma_1 = \gamma_2 \circ \kappa$ и $\forall x : \kappa'(x) > 0$. Класс эквивалентности относительно данного отношения зовётся *ориентированная кривая*.

За Γ^j обозначим множество ориентированных кривых, представители которых — кривые класса C^j. Ещё используют $\Gamma^j[a,b]$.

Пускай $\mathcal{F} \in C(\mathbb{R}^n \times \mathbb{R}^n)$, пусть она однородная порядка 1 по второму аргументу:

$$\forall \lambda > 0 : \mathcal{F}(z, \lambda w) = \lambda \mathcal{F}(z, w)$$

Пусть $\gamma : [a_\gamma, b_\gamma] \to \mathbb{R}^d$ — кривая класса C^1. Определим $J[\gamma] \coloneqq \int\limits_{a_\gamma}^{b_\gamma} \mathcal{F}[\gamma(t), \dot\gamma(t)]\,\mathrm{d}t$.

Предложение 4.1. *В этой ситуации J задаёт функционал на Γ^1.*

Доказательство. Пусть $\gamma_1 : [a_1, b_1] \to \mathbb{R}^d, \gamma_2 : [a_2, b_2] \to \mathbb{R}^d$ — два эквивалентных представителя.

$$J[\gamma_1] = \int\limits_{a_1}^{b_1} \mathcal{F}[\gamma_1(t), \dot\gamma_1(t)]\,\mathrm{d}t = \int\limits_{a_1}^{b_1} \mathcal{F}[\gamma_2(\kappa(t)), \dot\kappa(t) \cdot \dot\gamma_2(\kappa(t))]\,\mathrm{d}t =$$

$$= \int\limits_{a_1}^{b_1} \mathcal{F}[\gamma_2(\kappa(t)), \dot\gamma_2(\kappa(t))]\dot\kappa(t)\,\mathrm{d}t = \left\| \begin{matrix} \tau = \kappa(t) \\ \mathrm{d}\tau = \dot\kappa(t)\,\mathrm{d}t \end{matrix} \right\| = \int\limits_{a_2}^{b_2} \mathcal{F}[\gamma_2(\tau), \dot\gamma_2(\tau)]\,\mathrm{d}\tau \quad \square$$

Примеры.

- $\mathcal{F}(z, w) = |w|$. Функционал J — длина кривой
- $\mathcal{F}(z, w) = |w| \cdot f(z)$, где, например, $d = 2$ (кривая плоская), $f(z) = z_2^\alpha$ (здесь $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$).
  - При $\alpha = 0$ это предыдущий случай.
  - При $\alpha = -1$ это длина в гиперболической плоскости (в модули Пуанкаре в верхней полуплоскости).
  - При $\alpha = 1$ это координата центра масс кривой, а ещё — площадь поверхности вращения.
  - При $\alpha = -\frac{1}{2}$ это время, требуемое шарику, чтобы скатиться по жёлобу данной формы.

Пусть $L \in C([a, b] \times \mathbb{R}^n \times \mathbb{R}^n)$, $X = C^1[a, b]$, $J[u] = \int\limits_a^b L(t, u(t), \dot u(t))\,\mathrm{d}t$. Превратим последний в функционал на кривой. Заведём $\mathcal{F} : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}, \mathcal{F}(z, w) = L(z_1, z_2, \ldots, z_{n+1}, \frac{w_2}{|w_1|}, \ldots, \frac{w_{n+1}}{|w_1|})|w_1|$. Он имеет требуемую однородность. Типа сопоставим функции $u(t)$ кривую $\gamma_u : t \mapsto (t, u(t))$.

Рассмотрим $\widetilde{J}[\gamma] \coloneqq \int \mathcal{F}(\gamma(t), \dot\gamma(t))\,\mathrm{d}t$, утверждается, что если L «разумная», то $\widetilde{J}$ — функционал на кривых.

Предложение 4.2. $\widetilde{J}[\gamma_u] = J[u]$.

Доказательство. $\dot\gamma_u(t) = (1, \dot u(t))$. $\square$

Утверждение 4.1. *Пусть $\mathcal{F} \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})) \cap C(\mathbb{R}^n \times \mathbb{R}^n)$.*

Пусть $\forall \lambda > 0 : \mathcal{F}(z, \lambda w) = \lambda \mathcal{F}(z, w)$.

Пусть $\gamma \in \Gamma^2$, $\gamma : [a_\gamma, b_\gamma] \to \mathbb{R}^d$. $J[\gamma] = \int\limits_{a_\gamma}^{b_\gamma} \mathcal{F}[\gamma, \dot\gamma]\,\mathrm{d}t$, $E\{\gamma\} \coloneqq (\nabla_z \mathcal{F})(\gamma, \dot\gamma) - \frac{\mathrm{d}}{\mathrm{d}t}(\nabla_w \mathcal{F})(\gamma, \dot\gamma)$. Определение осмысленно, так как кривая параметризована, и $\dot\gamma \neq 0$, а $\mathcal{F} \in C^2(\ldots)$.

Теперь пусть $s \in C^2\left([a, b] \times [-\varepsilon, \varepsilon] \to \mathbb{R}^n\right)$, и пусть $s(\_, \tau)$ — параметризованная кривая. Тогда $\frac{\mathrm{d}}{\mathrm{d}\tau} J[s(\_, \tau)] = \int\limits_a^b \left\langle E\{s(x, \tau)\}, \frac{\partial s}{\partial \tau}(x, \tau) \right\rangle \mathrm{d}x + \left\langle (\nabla_w \mathcal{F})(s(x, \tau), \frac{\partial s}{\partial x}(x, \tau)), \frac{\partial s}{\partial \tau}(x, \tau) \right\rangle \Big|_{x=a}^{x=b}$. Наверно, так.

Доказательство. Упражнение. $\square$

Лемма 4.1. *Пусть $\gamma_1 \sim \gamma_2$ — представители кривой $\gamma \in \Gamma^2$ ($\gamma_1 = \gamma_2 \circ \kappa$). Тогда $(E\{\gamma_1\})(x) = \kappa'(x)\,(E\{\gamma_2\})\,(\kappa(x))$.*

Доказательство.

$$E\{\gamma_1\}(x) = (\nabla_z \mathcal{F})(\gamma_1(x), \gamma_1'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}(\nabla_w \mathcal{F})(\gamma_1(x), \dot\gamma_1(x)) =$$
$$= \kappa'(x)(\nabla_z \mathcal{F})\,(\gamma_2(\kappa(x)), \gamma_2'(\kappa(x))) - \frac{\mathrm{d}}{\mathrm{d}x}(\nabla_w \mathcal{F})(\gamma_1(x), \dot\gamma_1(x)) =$$

Дифференцируя по w равенство $\mathcal{F}(z, \lambda w) = \lambda \mathcal{F}(z, w)$, получаем $\lambda(\nabla_w \mathcal{F})(z, \lambda w) = \lambda(\nabla_w \mathcal{F})(z, w)$

$$= \kappa'(x)(\nabla_z \mathcal{F})\,(\gamma_2(\kappa(x)), \gamma_2'(\kappa(x))) - \kappa'(x)\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=\kappa(x)}(\nabla_w \mathcal{F})(\gamma_2(s), \gamma_2'(s)) = \kappa'(x)\cdot(E\{\gamma_2\})(\kappa(x))$$

□

Следствие 4.1. $E\{\gamma_1\} \equiv 0 \iff E\{\gamma_2\} \equiv 0$ *при* $\gamma_1 \sim \gamma_2$.

Заведём метрику на Γ^2, чтобы определить экстремумы. Пусть $\xi : [a_\xi, b_\xi] \to \mathbb{R}^d$ и $\nu : [a_\nu, b_\nu] \to \mathbb{R}^d$ — представители $\gamma_\xi, \gamma_\nu \in \Gamma^2$.

Перепараметризуем ξ и ν так, что они определены на $[0, 1]$, и их скорости постоянны: $|\dot\xi| \equiv \mathrm{const}, |\dot\nu| \equiv \mathrm{const}$.

Положим $\|\gamma_\xi - \gamma_\nu\| = \|\xi - \nu\|_{C^2[0,1]}$.

Упражнение 4.1. *Проверить, что это метрика на* Γ^2.

С метрикой также пришли всевозможные локальные, глобальные, строгие, нестрогие, минимумы и максимумы.

Теорема 4.1. Пусть $\gamma \in \Gamma^2$ — локальный максимум J, непрерывная $\mathcal{F} \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$, $\forall \lambda > 0 : \mathcal{F}(z, \lambda w) = \lambda \mathcal{F}(z, w)$.

Пусть $\gamma_a, \gamma_b \in \mathbb{R}^n, \mathcal{D} = \left\{\gamma \in \Gamma^2 \middle| \gamma(a_\gamma) = \gamma_a, \gamma(b_\gamma) = \gamma_b\right\}$ (J задаётся на пространстве $\mathcal{D}$).

Тогда $E\{\gamma\} = 0$.

Доказательство. Рассмотрим $\gamma + \tau h$, где $h(a_\gamma) = h(b_\gamma) = 0$, $h : [a_\gamma, b_\gamma] \to \mathbb{R}^d, h \in C^2$

Так как $\dot\gamma(s) \neq 0$, то $\|\dot\gamma(s)\| \neq 0$, и при достаточно малых τ : $\min\|\dot\gamma + \tau\dot h\| > \varepsilon$. Значит, при подстановке мы попадём в область, где $\mathcal{F} \in C^2$, и существует $\frac{\mathrm{d}}{\mathrm{d}\tau}\big|_{\tau=0} J[\gamma + \tau h] = 0$,

Раз γ — экстремум, то производная равна нулю.

$$J[\gamma + \tau h] - J[\gamma] = \tau \int\limits_{a_\gamma}^{b_\gamma} [\langle(\nabla_z \mathcal{F})(\gamma, \dot\gamma), h\rangle + \langle(\nabla_w \mathcal{F})(\gamma, \dot\gamma), h'\rangle]\,\mathrm{d}t + \mathcal{O}(\tau^2)$$

Интегрируя по частям, получаем $\int\limits_a^b \langle E\{\gamma\}, h\rangle\,\mathrm{d}t + 0 + \mathcal{O}(\tau^2)$. Применяя «нулевую лемму Дюбуа-Реймона», получаем $E\{\gamma\} = 0$. □

Таким образом, мы доказали, что чтобы показать экстремальность кривой γ, достаточно проверять экстремальность любой её перепараметризации, а экстремальность проверяется при помощи уравнения, аналогичного уравнению Эйлера — Лагранжа: $E\{\gamma\} = 0$.

Лекция IV

28 марта 2024 г.

5 Условия трансверсальности. Задача Лагранжа

Пусть J — функционал на кривых, концы которых должны находиться на двух заданных многообразиях $M_1, M_2 \subset \mathbb{R}^n$. Считаем, что $J, M_1, M_2 \in C^1$. Функционал рассматривается на пространстве $X = \left\{\gamma \in \Gamma^2 \middle| \gamma(a_\gamma) \in M_1, \gamma(b_\gamma) \in M_2\right\}$ и задан обычной интегральной формулой

$$J[\gamma] = \int \mathcal{F}[\gamma, \dot\gamma]\,\mathrm{d}t, \text{ где } \mathcal{F} \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})) \cap C(\mathbb{R}^n \times \mathbb{R}^n), \forall \lambda > 0 : \mathcal{F}(z, \lambda w) = \lambda \mathcal{F}(z, w)$$

Пусть γ_0 — локальный экстремум J на X. Тогда γ_0 — экстремум на $\{\gamma \in X | \gamma(a_\gamma) = \gamma_0(a_{\gamma_0}), \gamma(b_\gamma) = \gamma_0(b_{\gamma_0})\}$, следовательно, $E\{\gamma_0\} = 0$.

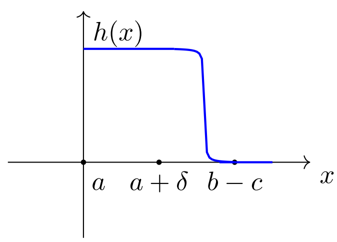
Изучим граничные условия. Теперь кривая вида $\gamma_0 + \tau h$ не лежит в X, поэтому просто изучить вариацию не получится.

Попробуем подвигать один из концов кривой так, чтобы он оставался на многообразии, и через некоторое расстояние подвинутая кривая сливалась с изначальной.

Запишем это формулами. Пусть $a \coloneqq a_{\gamma_0}, b \coloneqq b_{\gamma_0}$, пусть $r : (-\varepsilon, \varepsilon) \to M_1, r \in C^2, r(0) = \gamma_0(a)$. Пусть $\delta > 0$ мало, $c \in (0, b-a]$. Рассмотрим s вида $s(t, \tau) = \begin{cases} \gamma_0(t) + (r(\tau) - \gamma_0(a)), & t \in [a, a+\delta) \\ \gamma_0(t), & t \in [b-c, b] \end{cases}$.

Потребуем $s(\_, 0) = \gamma_0$ и $s \in C^2([a, b] \times (-\varepsilon, \varepsilon))$, $s(t, \_) \in X$. Также потребуем, чтобы $\forall t, \tau : \frac{\partial s}{\partial t} \neq 0$, то есть кривая $s(\_, \tau)$ должна быть регулярной.

Пример. Пусть h — шапочка, как ниже, тогда подойдёт $s(t, \tau) = \gamma_0(t) + h(t) \cdot (r(\tau) - r(0))$.



Так как γ_0 — экстремум, то $J[s(\_, \tau)]$, как функция от τ, имеет экстремум при $\tau = 0$. Положим $f(\tau) = J[s(\_, \tau)]$, и посчитаем $f'(\tau)$, приблизив разность первым членом ряда Тейлора:

$$f(\tau) - f(0) = \int \langle \nabla_u \mathcal{F}, s - \gamma_0 \rangle + \langle \nabla_{\dot u} \mathcal{F}, \dot s - \dot\gamma_0 \rangle + \mathcal{O}(\|s(\_, \tau) - \gamma_0\|^2_{\Gamma^2}) =$$

Формула применима, так как $\mathcal{F} \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$, и $\dot u$ отделён от нуля.

$\frac{\partial s}{\partial t} \neq 0$, равномерно отделена от нуля, а отрезок фиксирован, откуда $\mathcal{O}\left(\|s(\_, \tau) - \gamma_0\|_{\Gamma^2}\right) \equiv \mathcal{O}\left(\|s(\_, \tau) - \gamma_0\|_{C^2}\right)$.

Теперь применим формулу Тейлора по τ:

$$= \tau \int_a^b \left[\left\langle \nabla_u \mathcal{F}, \frac{\partial s}{\partial \tau}(\_, 0) \right\rangle + \left\langle \nabla_{\dot u} \mathcal{F}, \frac{\partial \dot s}{\partial \tau}(\_, 0) \right\rangle \right] \mathrm{d}\tau + \mathcal{O}(\tau^2) = \quad \text{(интегрируем по частям)}$$

$$= \tau \left[\int_a^b \left\langle E\{\gamma_0\}, \frac{\partial s}{\partial \tau}(\_, 0) \right\rangle \mathrm{d}\tau - \left\langle \nabla_{\dot u} \mathcal{F}, \frac{\partial s}{\partial \tau}(a, 0) \right\rangle \right] + \mathcal{O}(\tau^2) =$$

$$= -\tau \left\langle \nabla_{\dot u} \mathcal{F}(\gamma_0)(a), \frac{\partial s}{\partial \tau}(a, 0) \right\rangle + \mathcal{O}(\tau^2)$$

Так как по построению $\frac{\partial s}{\partial \tau}(b, 0) = 0$, то внеинтегральный член в b равен нулю. При этом $\frac{\partial s}{\partial \tau}(a, 0) = \frac{\partial r}{\partial \tau}(0)$. Итак,

$$f'(0) = 0 \Rightarrow \left\langle (\nabla_{\dot u} \mathcal{F})(\gamma_0)(a), \frac{\partial r}{\partial \tau}(0) \right\rangle = 0$$

Так как r — любая кривая через $\gamma_0(a)$, то $(\nabla_{\dot u}\mathcal{F})(\gamma_0)(a) \perp T_{\gamma_0(a)}M$.

Аналогично со вторым концом.

Запишем всё это в теорему:

Теорема 5.1. Пускай J — функционал на кривой ($\mathcal{F} \in C(\mathbb{R}^n \times \mathbb{R}^n) \cap C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$), как водится, $\forall \lambda > 0 : \mathcal{F}(z, \lambda w) = \lambda \mathcal{F}(z, w)$, $M_1, M_2 \subset \mathbb{R}^n$ — многообразия класса C^1, пусть γ_0 — локальный экстремум J на X.

Тогда

1. $E\{\gamma_0\} = 0$.
2. $\begin{cases}(\nabla_{\dot u}\mathcal{F})(\gamma_0)(a) \perp T_{\gamma_0(a)}M_1 \\ (\nabla_{\dot u}\mathcal{F})(\gamma_0)(b) \perp T_{\gamma_0(b)}M_2\end{cases}$ — условия трансверсальности.

Примеры.

- $\mathcal{F}(z,w) = |w|$. Минимум этого функционала — расстояние от M_1 до M_2.

  Условия из теоремы означают, что $\dot\gamma_0(a) \perp T_{\gamma(a)}M_1, \dot\gamma_0(b) \perp T_{\gamma(b)}M_2$, а также $\nabla_w \mathcal{F} = \frac{w}{|w|} \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\dot\gamma_0}{|\dot\gamma_0|}) = 0$ (уравнение Эйлера — Лагранжа). Экстремаль — отрезок, соединяющий два многообразия, и перпендикулярный обоим многообразиям.

- $\mathcal{F}(z,w) = g(z)|w|$. $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$; мы видели, что тут масса полезностей при разных g. Так как $\nabla_w \mathcal{F} = g(z)\frac{w}{|w|} \parallel w$, то в этом случае условия трансверсальности тоже сводятся к условиям ортогональности: $\dot\gamma_0(a) \perp M_1, \dot\gamma_0(b) \perp M_2$.

Рассмотрим частный случай: $n = 2, J[y] = \int\limits_{a_y}^{b_y} L(x, y(x), y'(x))\,\mathrm{d}x$, L — гладкая (везде, где нужно) на $\mathbb{R}^3$; $\phi, \psi \in C^2$ таковы, что $y(a_y) = \phi(a_y), y(b_y) = \psi(b_y)$. Пусть $y \in C^2$. Таким образом, многообразия $M_1 = \{(x, \phi(x)) | x \in \mathbb{R}\}, M_2 = \{(x, \psi(x)) | x \in \mathbb{R}\}$ — графики функций, и концы y лежат на этих графиках.

Сведёмся к уже доказанной теореме. Пусть $\mathcal{F}(z,w) \in C(\mathbb{R}^2 \times \mathbb{R}^2)$ такова: $\mathcal{F}(z,w) = L\left(z_1, z_2, \frac{w_2}{|w_1|}\right)|w_1|$.

Будем рассматривать область $w_1 > 0$ (на интересующей нас кривой-графике $w_1 \equiv 1$).

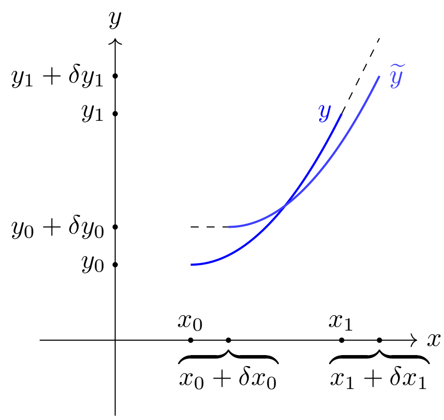
$$\nabla_w \mathcal{F} = \begin{pmatrix} L\left(z_1, z_2, \frac{w_2}{w_1}\right) - \frac{w_2 w_1}{w_1^2}\frac{\partial L}{\partial \dot u}\left(z_1, z_2, \frac{w_2}{w_1}\right) \\ \frac{\partial L}{\partial \dot u}\left(z_1, z_2, \frac{w_2}{w_1}\right)\end{pmatrix}$$

$\gamma(t) = (t, y(t))$, $\dot\gamma(t) = (1, \dot y(t))$. Запишем условия трансверсальности $(\nabla_w \mathcal{F})(\gamma)(a) \perp \begin{pmatrix} 1 \\ \dot\phi(a)\end{pmatrix}$ и $(\nabla_w \mathcal{F})(\gamma)(b) \perp \begin{pmatrix} 1 \\ \dot\psi(b)\end{pmatrix}$ через скалярное произведение: $(L - \dot y \frac{\partial L}{\partial \dot y})(a) \cdot 1 + \frac{\partial L}{\partial \dot y}\dot\phi(a) = 0$, аналогично со вторым концом. Это часто записывают в виде

$$\begin{cases} L + \frac{\partial L}{\partial \dot y} \cdot (\dot\phi(a) - \dot y(a)) = 0 \\ L + \frac{\partial L}{\partial \dot y} \cdot (\dot\psi(b) - \dot y(b)) = 0 \end{cases} \qquad (\triangle)$$

Продемонстрируем элементарный вывод этого факта. Видимо, ему будет недоставать некоторой строгости, так что доказательством считаться метод не будет, но так должны были думать те, кто впервые эти уравнения вывели.

Пусть $y, \widetilde{y}$ — две кривые, продлим их касательными так, чтобы они было определены на одном большем отрезке.



Пусть $h \coloneqq \widetilde{y} - y$, запишем вариацию:

$$J[\widetilde{y}] - J[y] = \int\limits_{x_0+\delta x_0}^{x_1+\delta x_1} L(x, \widetilde{y}, \dot{\widetilde{y}}) - \int\limits_{x_0}^{x_1} L(x, y, \dot{y}) =$$

$$= \int\limits_{x_0}^{x_1} (L(x, \widetilde{y}, \dot{\widetilde{y}}) - L(x, y, \dot{y})) - \int\limits_{x_0}^{x_0+\delta x_0} L(x, \widetilde{y}, \dot{\widetilde{y}}) + \int\limits_{x_1}^{x_1+\delta x_1} L(x, \widetilde{y}, \dot{\widetilde{y}}) =$$

Далее к первому слагаемому применим формулу Тейлора по второму и третьему аргументу, и проинтегрируем по частям, чтобы получить общий множитель h:

$$= \int\limits_{x_0}^{x_1} \left(\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial \dot{y}} \right) h \,\mathrm{d}x + \frac{\partial L}{\partial \dot{y}} h \Big|_{x_0}^{x_1} + L(x_1, y(x_1), \dot{y}(x_1))\delta x_1 - L(x_0, y(x_0), \dot{y}(x_0))\delta x_0 +$$

$$+ \mathcal{O}(\|h\|^2_{C^1}) + o(\delta x_1) + o(\delta x_2) =$$

Рассматривая финитные h, обращающиеся в нуль на концах, получаем, что первый член $\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial \dot{y}} = 0$ (это, кстати, уравнение Эйлера — Лагранжа).

Так как мы продолжили y и $\widetilde{y}$ касательными, то $h(x_0) = \delta y_0 - \dot{y}(x_0)\delta x_0$ и $h(x_0) = \delta y_1 - \dot{y}(x_1)\delta x_1$.

$$= L(x_1, y(x_1), \dot{y}(x_1))\delta x_1 - L(x_0, y(x_0), \dot{y}(x_0))\delta x_0 +$$

$$+ \frac{\partial L}{\partial \dot{y}}(x_1, y(x_1), \dot{y}(x_1)) \cdot [\delta y_1 - \dot{y}(x_1)\delta x_1] - \frac{\partial L}{\partial \dot{y}}(x_0, y(x_0), \dot{y}(x_0)) \cdot [\delta y_0 - \dot{y}(x_0)\delta x_0] =$$

Теперь вспомним, что концы y и $\widetilde{y}$ лежат на графиках ϕ и ψ, откуда с точностью до некоторых малых поправок, $\delta y_1 = \dot{\psi}(x_1)\delta x_1$ и $\delta y_0 = \dot{\phi}(x_0)\delta x_0$.

$$= \delta x_1 \left(L(x_1, y(x_1), \dot{y}(x_1)) + \frac{\partial L}{\partial \dot{y}}(x_1, y(x_1), \dot{y}(x_1))(-\dot{y}(x_1) + \dot{\psi}(x_1)) \right) +$$

$$+ \delta x_0 \cdot \left(L(x_0, y(x_0), \dot{y}(x_0)) + \frac{\partial L}{\partial \dot{y}}(x_0, y(x_0), \dot{y}(x_0))(-\dot{y}(x_0) + \dot{\psi}(x_0)) \right)$$

Рассматривая такие $\widetilde{y}$, что по очереди $\delta x_0 = 0$ и $\delta x_1 = 0$, получаем ($\triangle$).

6 Инвариантность уравнения Эйлера — Лагранжа

Если сделать замену переменных в решении уравнения Эйлера — Лагранжа, то получится решение задачи, в которой так же заменили переменные.

Пусть $T:\mathbb{R}^n\to\mathbb{R}^n$ — диффеоморфизм, $J[\gamma]=\int\mathcal{F}(\gamma,\dot\gamma)\,\mathrm{d}t, \mathcal{F}(z,\lambda w)=\lambda\mathcal{F}(z,w), \lambda>0$, и как и раньше, $\mathcal{F}\in C(\mathbb{R}^n\times\mathbb{R}^n)\cap C^2(\mathbb{R}^n\times(\mathbb{R}^n\setminus\{0\}))$.

Пусть γ — ориентированная кривая, тогда $T\circ\gamma$ — также ориентированная кривая того же класса гладкости. Определим функционал $J_T[\gamma]\coloneqq J[T\circ\gamma]$, то есть $J_T[\gamma]=\int\mathcal{F}[T(\gamma(t)),T'(\gamma(t))\cdot\dot\gamma(t)]\,\mathrm{d}t$.

Функция $\mathcal{F}_T(z,w)\coloneqq\mathcal{F}[T(z),T'(z)w]$ имеет ту же однородность. Пусть E_T — функция E, построенная по $\mathcal{F}_T$, то есть $E_T\{\gamma\}=(\nabla_z\mathcal{F}_T)(\gamma,\dot\gamma)-\frac{\mathrm{d}}{\mathrm{d}t}(\nabla_w\mathcal{F}_T)(\gamma,\dot\gamma)$.

Пусть поставлена некоторая задача с фиксированными концами. Согласно (лемма 2.2) после интегрирования по частям получаем: $\delta J_T[\gamma,h]=\int\langle E_T\{\gamma\},h\rangle\,\mathrm{d}t+o(\|h\|)$. Пусть $T(\gamma+h)=T(\gamma)+T'(\gamma)h+r$, где $\|r\|_{C^1}$ конечно же мала (порядка $\|h\|_{C^1}$).

$$\begin{aligned}J_T[\gamma+h]-J_T[\gamma]&=J[T(\gamma)+T'(\gamma)h+r]-J[T(\gamma)]=\\&=\int\left[\left\langle\frac{\partial\mathcal{F}}{\partial\gamma},T'(\gamma)h+r\right\rangle+\left\langle\frac{\partial\mathcal{F}}{\partial\dot\gamma},(T'(\gamma)h+r)'\right\rangle\right]\mathrm{d}t+o(\|h\|_{C^1})=\\&=\int\left[\left\langle\frac{\partial\mathcal{F}}{\partial\gamma},T'(\gamma)h\right\rangle-\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial\dot\gamma},T'(\gamma)h\right\rangle\right]\mathrm{d}t+o(\|h\|_{C^1})=\int\left\langle T'(\gamma)^tE\{T(\gamma)\},h\right\rangle+o(\|h\|_{C^1})\end{aligned}$$

Итак, $E_T\{\gamma\}=T'(\gamma)^tE\{T(\gamma)\}$ — заявленная инвариантность уравнения Эйлера — Лагранжа.

В частности, $E_T\{\gamma\}\equiv 0\Leftrightarrow E\{T(\gamma)\}\equiv 0$.

Лекция V

11 апреля 2024 г.

7 Прямые методы вариационного исчисления

7.1 Поиск решения задачи Штурма — Лиувилля

У уравнения Эйлера — Лагранжа есть некоторые недостатки — так, выявление характера экстремума является отдельной, зачастую весьма сложной, задачей.

Здесь пойдёт речь о методах, пытающихся построить точки максимума или минимума непосредственно. Платой за такое удобство будет общность.

Здесь всё будет одномерно и скалярно: пусть $p\in C^1[a,b], q\in C[a,b]$. Рассмотрим задачу с фиксированными концами для функционала следующего вида: $J[u]=\int\limits_a^b(pu'^2+qu^2)\,\mathrm{d}x$. Пусть $X_0\coloneqq\left\{u\in C^1[a,b]\middle|u(a)=u(b)=0\right\}$, однако в X_0 (ввиду однородности по u) инфимум J всегда либо 0, либо $-\infty$. Поэтому добавим нормировочное условие: $X\coloneqq\left\{u\in X_0\middle|\int\limits_a^b u^2=1\right\}$.

В рамках ранее рассмотренной теории это является задачей на условный экстремум (при $G[u]=\int\limits_a^b u^2=1$). Уравнение Эйлера — Лагранжа для $J-\lambda G$ получится

$$-(pu')'+qu=\lambda u \tag{$*$}$$

Из общей теории (и C^1-гладкости p) следует, что для экстремума u: $pu'\in C^1$, то есть $u\in C^2$ вне окрестности тех точек, где p обращается в 0. Потребуем, чтобы этих точек не было: $\forall x\in[a,b]: p(x)>0$.

Задача поиска решения уравнения ($*$) при условиях вида $\begin{cases}\alpha_au(a)+\beta_au'(a)=0\\\alpha_bu(b)+\beta_bu'(b)=0\end{cases}$ $\left(\text{где считается, что }\begin{cases}\alpha_u^2+\beta_a^2\neq0\\\alpha_b^2+\beta_b^2\neq0\end{cases}\right)$ называется задачей Штурма — Лиувилля. Её можно рассматривать, как задачу поиска собственных векторов оператора $\mathscr{L}:X_0\to X_0$, $\mathscr{L}:u\mapsto-(pu')'+qu$.

Тем самым, λ, при которых $(*)$ имеет решение, называются *собственными числами*, и соответствующие функции u — *собственные функции*.

Положим $\lambda_J := \inf\limits_X J$. Очевидно, что

1. $\lambda_J \geqslant \min\limits_{x\in[a,b]} q(x)$

2. Из однородности J по u: $\forall u \in X_0 : J[u] \geqslant \lambda_J \int\limits_a^b u^2\,\mathrm{d}x$.

Пускай $u_n \in X$ — минимизирующая последовательность, такая, что $J[u_n] \searrow \lambda_J$. Мы докажем, что из соображений компактности можно выбрать равномерно сходящуюся подпоследовательность, и что предел обладает свойствами, которые от него ожидаются.

Итак, имеется последовательность $u_n \in X$, такая, что $\int\limits_a^b pu_n'^2 + qu_n^2 \underset{n\to\infty}{\longrightarrow} \lambda_J$. Оценим

$$J[f] \geqslant \underbrace{\min_{[a,b]} p}_{>0} \cdot \int\limits_a^b f'^2 + \min_{[a,b]} q \cdot \underbrace{\int\limits_a^b f^2}_{1}$$

Из ограниченности $J[u_n]$ получаем, что $\sup\limits_n \int\limits_a^b u_n'^2 < \infty$.

Отсюда сразу следует равностепенная непрерывность:

$$|u_n(x_1) - u_n(x_2)| = \left| \int\limits_{x_1}^{x_2} u_n' \right| \leqslant \sqrt{|x_1 - x_2|} \cdot \left(\int\limits_{x_1}^{x_2} u_n'^2 \right)^{1/2}$$

Эта же оценка показывает равномерную ограниченность: принимая $x_2 = b$ и $x = x_1$, получаем $|u_n(x)| \leqslant \sqrt{b-a} \cdot C$.

Тем самым по теореме Арцела — Асколи из последовательности $\{u_n\}$ можно выбрать сходящуюся в C подпоследовательность; без потери общности, эта последовательность совпадает с исходной: $\exists u = \lim\limits_{n\to\infty} u_n$, где предел берётся в $C[a,b]$.

Эта предельная u — кандидат на минимизирующую функцию. Но пока u даже в интеграл не подставить: про гладкость ничего не известно. Тем не менее, конечно, $\int\limits_a^b u^2 = 1$.

Теорема 7.1. Так построенное $u \in C^2$ (в том числе $u \in X$), выполнено $(*)$, и $J[u] = \lambda_J$.

Доказательство. Сначала докажем $(*)$ в слабом смысле: убедимся, что

$$\forall h \in C^2[a,b] : h(a) = h(b) = 0 \Rightarrow \int\limits_a^b (-(ph')' + qh)u = \lambda_J \int\limits_a^b hu \qquad (**)$$

Это равенство можно было бы получить, умножив $(*)$ на h, и дважды проинтегрировав первое слагаемое по частям, если бы $(*)$ и C^2-гладкость u были уже даны. Нам же придётся пойти в обратном направлении.

Попробуем поварьировать J, добавляя $h \in C^2[a,b]$, $h(a) = h(b) = 0$. Функция вида $u_n + \varepsilon h$ совсем необязательно лежит в X, подставим эту функцию в $\widetilde{J}$, определённый на X_0, и заданный той же формулой, что и J.

$$\widetilde{J}[u_n + \varepsilon h] = J[u_n] + 2\varepsilon \int\limits_a^b (pu_n'h' + qu_nh) + \varepsilon^2 \widetilde{J}[h]$$

Интегрируя по частям $pu_n'h'$, получаем $2\varepsilon\int\limits_a^b u_n(-(ph')'+qh)$, внеинтегральные члены обнулились.

С другой стороны, $\widetilde{J}[u_n+\varepsilon h]\geqslant\lambda_J\int\limits_a^b(u_n+\varepsilon h)^2=\lambda_J+2\varepsilon\lambda_J\int\limits_a^b u_nh+\varepsilon^2\lambda_J\int\limits_a^b h^2$. Переходя к пределу в неравенствах, и сокращая λ_J, получаем

$$2\varepsilon\int\limits_a^b u(-(ph')'+qh)+\varepsilon^2\widetilde{J}[h]\geqslant 2\varepsilon\lambda_J\int\limits_a^b uh+\varepsilon^2\lambda_J\int\limits_a^b h^2$$

Так как можно выбирать ε разных знаков, то предельный переход показывает равенство линейных по ε членов. То есть $(**)$ выполнено.

Рассмотрим $\xi\in C[a;b]$ и подберём h такое, что $\xi=-(ph')'$.

А именно, $h(x)\coloneqq-\int\limits_a^x\frac{1}{p(t)}\left(\int\limits_a^t\xi(s)\,\mathrm{d}s+C\right)\mathrm{d}t$, где C выбрана так, что $h(b)=0$. Если посчитать, то

$$C=-\frac{\int\limits_a^b\frac{1}{p(t)}\int\limits_a^t\xi(s)\,\mathrm{d}s\,\mathrm{d}t}{\int\limits_a^b\frac{\mathrm{d}t}{p(t)}}$$

По построению $h\in C^2, h(a)=h(b)=0$, значит, можно подставить h в $(**)$:

$$0=\int\limits_a^b\xi u-\int\limits_a^b(q(x)-\lambda_J)u(x)\int\limits_a^x\frac{1}{p(t)}\left(\int\limits_a^t\xi(s)\,\mathrm{d}s+C\right)\mathrm{d}t\,\mathrm{d}x=$$

переставим интегралы так, чтобы интеграл по s был внешним

$$=\int\limits_a^b\xi u-\int\limits_a^b\xi(s)\int\limits_s^b\frac{1}{p(t)}\int\limits_t^b(q(x)-\lambda_J)u(x)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}s+$$
$$+\underbrace{\frac{1}{\int\limits_a^b\frac{1}{p}}\int\limits_a^b\xi(s)\int\limits_s^b\frac{\mathrm{d}t}{p(t)}\,\mathrm{d}s}_{C}\cdot\int\limits_a^b(q(x)-\lambda)u(x)\int\limits_a^x\frac{\mathrm{d}\tau}{p(\tau)}\,\mathrm{d}x$$

Обозначив $K\coloneqq\left(\int\limits_a^b\frac{1}{p}\right)^{-1}\int\limits_a^b(q(x)-\lambda)u(x)\int\limits_a^x\frac{\mathrm{d}\tau}{p}\,\mathrm{d}x$ (константа, не зависящая от ξ), получаем

$$0=\int\limits_a^b\xi(s)\left[u(s)-\int\limits_s^b\frac{1}{p(t)}\int\limits_t^b(q(x)-\lambda_J)u(x)\,\mathrm{d}x\,\mathrm{d}t+K\int\limits_s^b\frac{\mathrm{d}t}{p(t)}\right]\mathrm{d}s$$

Итак, $0=\int\limits_a^b\xi(s)\left[\cdots\right]\mathrm{d}s$, откуда «по нулевой лемме Дюбуа-Реймона» выражение в скобках равно нулю везде: $u(s)=\int\limits_s^b\frac{1}{p(t)}\int\limits_t^b(q(x)-\lambda_J)u(x)\,\mathrm{d}x\,\mathrm{d}t-K\int\limits_s^b\frac{\mathrm{d}t}{p(t)}$, в частности сразу $u\in C^1$.

Значит, u можно продифференцировать, получим

$$u'(s)=-\frac{1}{p(s)}\int\limits_s^b(q(x)-\lambda_J)u(x)\,\mathrm{d}x+\frac{K}{p(s)},$$

откуда $u \in C^2$. Тем самым, $(p(s)u'(s))' = (q(s) - \lambda_J)u(s)$, значит, $u \in X$, $J[u] = \int\limits_a^b pu'^2 + qu^2$. Интегрируя по частям, получаем ровно $\int\limits_a^b ((-pu')' + qu)u \, \mathrm{d}x = \lambda_J \int\limits_a^b u^2 = \lambda_J$. $\square$

Из данной теоремы получаются следующие выводы: u — нестрогий глобальный минимум, причём если $J[u] = \lambda_J = J[\widetilde{u}] \Rightarrow u = \pm\widetilde{u}$. Это следует из того, что u — решение соответствующего диффура, то есть лежит в одномерном $\mathbb{R}$-пространстве, а нормировка фиксирует точку с точностью до знака.

Пусть u_n — минимизирующая последовательность. Соображения компактности имеют тот недостаток, что они не предоставляют алгоритма выбора сходящейся подпоследовательности; но в данном случае это можно сделать постфактум, зная результат теоремы.

Выберем $c \in [a, b]$ так, что $|u|(c) \neq 0$ (между прочим, конечный перебор — у $|u|$ не более, чем конечное число нулей). ($|u|$ находим, как предел $|u_n|$ — он существует, никаких подпоследовательностей выбирать не надо)

Подправим u_n так, что $u_n(c) \geqslant 0$. Теперь $u_n \underset{n\to\infty}{\longrightarrow} u$, и опять мы справились найти u без выбора сходящихся подпоследовательностей.

7.2 Построение ортонормированного базиса

Сейчас мы поймём, что множество собственных чисел $\{\lambda \in \mathbb{R} | (*) \text{ имеет решение}\}$ не очень велико, и решения обладают всякими чудесными свойствами.

Пусть $u \in X$ — минимизирующая $J[u] = \lambda_J =: \lambda_1$.

Положим $X^{(1)} = \left\{ f \in X \middle| \int\limits_a^b fu = 0 \right\}$. Обозначим $\lambda_2 := \inf\limits_{X^{(1)}} J$. Ясно, что $\lambda_2 \geqslant \lambda_1$.

Выберем минимизирующую последовательность из $X^{(1)}$, на которой J сходится к λ_2, и пусть u_2 — предел сходящейся подпоследовательности из данной минимизирующей. Далее по сути повторим доказательство (теорема 7.1), показав, что u_2 — глобальный минимум J на $X^{(1)}$.

Так как интеграл терпит поточечные предельные переходы, то $\int\limits_a^b u_2^2 = 1, u_2(a) = u_2(b) = 0, \int\limits_a^b u_2 u = 0$.

Аналогично доказательству (теорема 7.1), для u_2 верно: $\forall h \in C^2 : h(a) = h(b) = 0, \int\limits_a^b hu = 0 \Rightarrow$

$$\lambda_2 \int\limits_a^b uh = \int\limits_a^b u_2(-(ph')' + qh) \qquad (**)$$

Условие $\int\limits_a^b hu = 0$ должно быть выполнено, так как в J в какой-то момент придётся подставить отнормированное $u_2 + \varepsilon h$, то есть должно быть выполнено включение $\frac{u_2 + \varepsilon h}{\|u_2 + \varepsilon h\|_{L^2}} \in X^{(1)}$. Отсюда следует, что $u_2 \in C^2[a, b]$, и $(*)$ выполнено для λ_2.

Здесь есть некоторая тонкость: при решении используется лемма Дюбуа-Реймона, и для её применения хотелось бы избавиться от условия ортогональности $h \perp u$. На самом деле, оно и правда лишнее: достаточно убедиться, что $(**)$ выполнено для $h = u$, так как любую функцию можно разложить в ортогональную часть и часть, пропорциональную u. А при $h = u$ оно выполнено тривиальным образом: обе части нули, так как $\int\limits_a^b u_2 u = 0$.

Отметим, что мы сразу получили, что $\lambda_2 > \lambda_1$, так как при равенстве u_2 было бы решением того же дифференциального уравнения $(*)$, и должно было бы быть пропорционально u.

Лекция VI

25 апреля 2024 г.

Рассуждая тем же образом, можно рассмотреть $X^{(n)} \coloneqq \left\{ f \in X \middle| \int\limits_a^b fu_1 = \dots = \int\limits_a^b fu_n = 0 \right\}$, где $u_1 = u$. Пусть $\lambda_n \coloneqq \inf\limits_{X^{(n-1)}} J$, тогда по тем же причинам $\lambda_n \geqslant \lambda_{n-1}$, инфимум достигается на некотором $u_n \in C^2[a,b]$ — равенство $(**)$ верно и для h, ортогональных $u_1, \dots, u_{n-1}$, и для тех, что лежат в их линейной оболочке тоже верно, поэтому мы можем сформулировать

Теорема 7.2.

1. $\forall n \geqslant 1 : \lambda_n = \inf\limits_{X^{(n-1)}} J$ достигается: $\lambda_n = J[u_n]$ для некоторого $u_n \in C^2[a,b] \cap X^{(n-1)}$, и для него выполнено $(*)$.
2. Такое u_n единственно с точностью до домножения на ± 1, и $\lambda_j \nearrow +\infty$.

Доказательство. Из ещё не обсуждённых вещей в формулировке появилось только условие $\lim\limits_{j\to\infty} \lambda_j = \infty$.

Понятно, что $\lambda_n > \lambda_{n-1}$, так как в случае равенства u_n и u_{n-1} были бы решениями дифференциального уравнения $(*)$, и, следовательно, были бы пропорциональны, однако они ортогональны (и не нули).

Предположим, что $\lambda_n \nearrow \lambda_* < \infty$. Тем самым, $\sup\limits_n J[u_n] =: S < \infty$. Но, расписав J, получаем $\int\limits_a^b pu_n'^2 + qu_n^2 \geqslant \min q + \min p \int\limits_a^b u_n'^2$, то есть в последовательности u_n функции равномерно ограничены и равностепенно непрерывны, а значит, имеется подпоследовательность u_{n_k}, сходящаяся в $C[a,b]$ к $u_@$. Это противоречие к попарной ортогональности: $\int\limits_a^b u_{n_k} u_{n_{k-1}} = 0$, но левая часть при устремлении $k \to \infty$ даёт норму $u_@$, то есть 1. Противоречие. $\square$

Предложение 7.1. *Построенная в предыдущей теореме последовательность $\{u_n\}$ — ортонормированный базис в $L^2[a,b]$.*

Доказательство. Достаточно показать, что это выполнено в некотором плотном множестве, скажем, в $C_0^\infty[a,b]$.

Для $f \in C_0^\infty[a,b]$ надо проверить, что $f_N \coloneqq \sum\limits_{n=1}^N (f, u_n)u_n \underset{N\to\infty}{\longrightarrow} f$. По построению $\frac{f-f_N}{\|f-f_N\|_{L^2}} \in X^{(N)}$.

$$J\left(\frac{f-f_N}{\|f-f_N\|_{L^2}}\right) \geqslant \lambda_{N+1} \iff \underbrace{J[f-f_N]}_{\int\limits_a^b p(f'-f_N')^2 + q(f-f_N)^2} \geqslant \lambda_{N+1}\|f-f_N\|_{L^2}^2$$

Интегрируя по частям, получаем

$$\int\limits_a^b (-(p(f'-f_N'))' + q(f-f_N)) \cdot (f-f_N) = (\mathscr{L}(f-f_N), f-f_N)_{L^2[a,b]}$$

Здесь $\mathscr{L} : g \mapsto -(pg')' + qg$. Так как $f - f_N \perp X^{(N)}$ по построению, а $\mathscr{L}(f_N) \in X^{(N)}$, то $(\mathscr{L}(f-f_N), f-f_N)_{L^2[a,b]} = (\mathscr{L}(f), f-f_N)_{L^2[a,b]}$. Оценим $(\mathscr{L}(f), f-f_N)_{L^2[a,b]} \leqslant \|\mathscr{L}f\|_{L^2} \cdot \|f-f_N\|_{L^2}$.

Итак, $\|\mathscr{L}f\|_{L^2} \cdot \|f-f_N\|_{L^2} \geqslant \lambda_{n+1}\|f-f_N\|_{L^2}^2$

Сокращая на $\|f-f_N\|$ (если это нуль для некоторого N, то f уже разложилась в конечную сумму), получаем $\|f-f_N\| \leqslant \frac{1}{\lambda_{N+1}}\|\mathscr{L}f\|_{L^2} \underset{N\to\infty}{\longrightarrow} 0$. $\square$

Следствие 7.1. $\{\lambda_j\}_{j=1}^\infty = \left\{\lambda \in \mathbb{C} \middle| (*) \text{ *имеет решение для* } u \in C^2([a,b] \to \mathbb{C}) \text{ *где* } u(a) = u(b) = 0, u \not\equiv 0\right\}$

Доказательство. $\subseteq$ уже доказали выше. Обратное включение докажем от противного.

Пусть $\lambda_* \notin \{\lambda_j\}_{j=1}^{\infty}$, и она отвечает собственной функции u_*: $\mathscr{L}u_* = \lambda_* u_*$.

$$\lambda_*(u_*, u_n)_{L^2} = (\mathscr{L}u_*, u_n)_{L^2} = \int\limits_a^b (-(pu_*')' + qu_*)u_n \underset{\text{по частям}}{=} \int\limits_a^b u_*((-pu_n')' + qu_n) = \lambda_n(u_*, u_n)_{L^2}$$
$$\Rightarrow \forall n\ (u_*, u_n) = 0 \Rightarrow u_* = 0$$

Ещё можно было заметить, что при рассмотрении $\mathscr{L}$, как оператора $X_0 \to L^2$, верно $\forall u, v \in X_0$: $(\mathscr{L}u, v) = (u, \mathscr{L}v)_{L^2}$ — выкладка выше. $\square$

7.3 Нули собственных функций

Каждая из функций u_n имеет разве что конечное число нулей на (a, b), как нетривиальное решение дифференциального уравнения второго порядка.

Введём для фиксированного λ функция $u(x, \lambda)$ — решение задачи Коши $\begin{cases} -(pu')' + qu = \lambda u \\ u(a, \lambda) = 0 \\ u'(a, \lambda) = 1 \end{cases}$ (все производные по x). Тогда $\{\lambda_j\}_{j=1}^{\infty} = \{\lambda | u(b, \lambda) = 0\}$.

Теорема 7.3. u_n имеет в точности $n - 1$ нуль на (a, b).

Доказательство.

- $u(x, \lambda)$ не имеет нулей на $(a, b]$ при $\lambda < \min q$: $(*)$ переписывается в виде $(pu')' = (q - \lambda)u$, то есть $u' > 0$ везде.

- Запишем $(*)$ в виде $(pu')' + (\lambda - q)u = 0$. При $\lambda < \lambda'$: $\lambda' - q > \lambda - q$, то есть (как обсуждалось на диффурах, теорема Штурма) между соседними нулями $u(x, \lambda)$ есть хотя бы один нуль $u(x, \lambda')$. Тем самым, u_n имеет хотя бы $n - 1$ нуль, и осталось доказать, что их не больше.

- Обозначим $\lambda_* = \inf\{\lambda | u(\_, \lambda)$ имеет нуль на $(a, b]\}$. Утверждается, что $\lambda_* = \lambda_1$. В самом деле, предположим на минуточку, что $\lambda_* < \lambda_1$. Тогда $u(\_, \lambda_*)$ (если имеет нуль на $(a; b]$) имеет нуль на (a, b) ($u(b, \lambda_*) = 0$ противоречит определению λ_1).

  Сначала покажем, что инфимум достигается: $u(\_, \lambda_*)$ имеет нуль на (a, b). Устремим некоторую последовательность к λ_* сверху и выберем точку сгущения на $(a; b)$: $\lambda_n \to \lambda_*, x_{\lambda_n} \to x_*$.

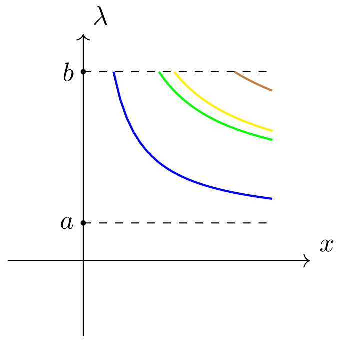
  $u(x_*, \lambda_*) = 0$. Рассмотрим $u(x_* - \varepsilon, \lambda)$ и $u(x_* + \varepsilon, \lambda_*)$ (где $\varepsilon > 0$ — произвольный, такой, что $x_* \pm \varepsilon \in (a, b)$, и $\forall x : 0 < |x - x_*| < \varepsilon \Rightarrow u(x, \lambda_*) \neq 0$). Если они разных знаков, то мы получаем противоречие с определением λ_* — есть и меньше, согласно непрерывной зависимости решений от параметров.

  Тем самым, $u(x_n - \varepsilon, \lambda_*)$ и $u(x_n + \varepsilon, \lambda_*)$ одного знака при всех $\varepsilon > 0$. Но тогда $u(\_, \lambda_*)$ имеет в x_n нуль минимум второй кратности, и значит, $u(\_, \lambda_*)$, как решение диффура, равно тождественному нулю.

- И так далее. Теперь $\lambda_* = \inf\{\lambda | u(\_, \lambda)$ имеет хотя бы два нуля на $(a, b]\}$. Пусть $u(x_{1,\lambda}, \lambda) = u(x_{2,\lambda}, \lambda) = 0$. Выберем подпоследовательность так, чтобы $x_{1,\lambda} \longrightarrow x_1$ и $x_{2,\lambda} \longrightarrow x_2$. Если $x_1 \neq x_2$, то всё аналогично.

  Если же $x_1 = x_2 = \widetilde{x}$, то по теореме Лагранжа между $x_{1,\lambda}$ и $x_{2,\lambda}$ имеется точка s_λ, в которой $u'(s_\lambda, \lambda) = 0$. По принципу двух полицейских $s_\lambda \underset{\lambda \to \lambda_*}{\longrightarrow} \widetilde{x}$, а так как производная u тоже непрерывно зависит от параметров, то в $\widetilde{x}$ — нуль кратности хотя бы 2. $\square$

Тем самым, нули появляются в точке b, и монотонно едут к точке a:



Непрерывная зависимость нулей от λ следует из непрерывной зависимости решений от параметров и того, что все нули — простые (кратности 1).