Вариационное исчисление. Неофициальный конспект

Лектор: Роман Владимирович Романов
Конспектировал Леонид Данилевич

IV семестр, весна 2024 г.

Содержание

Лекция I

15 февраля 2023 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: $(\operatorname{grad} f)(x) = 0$.
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.

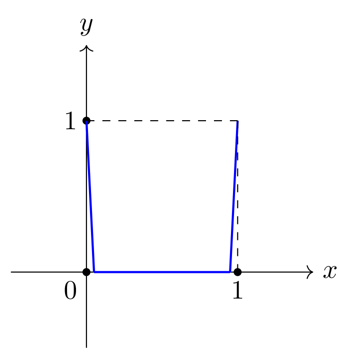
Определение 1.1 ($x \in X$ — строгий локальный минимум)**.** $\exists \delta > 0 : \forall y \in U_\delta(x) : J[y] > J[x]$. Квадратные скобочки — косметическое.

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.
- $J = \int\limits_a^b (\frac{\dot{u}^2}{2} - V(u))\,\mathrm{d}x$, 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. $\square$

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 $\exists$) 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 нуль. □

2 Формула первой вариации. Уравнение Эйлера — Лагранжа

2.1 Лемма Дюбуа-Реймона

Лемма 2.1 (Дюбуа-Реймона)**.** *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 $\overline{f} \overset{def}{=} \frac{1}{b-a} \int\limits_a^b f$.

Проинтегрируем $f - \overline{f}$: $\omega(x) \coloneqq \int\limits_a^x \left(f(x') - \overline{f}\right) \mathrm{d}x'$. 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 $f \equiv \overline{f}$. □

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

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

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

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

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

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 $K \coloneqq [a,b] \times \overline{B_{\|u\|+\|h\|}} \times \overline{B_{\|u\|+\|h\|}}$.

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. □

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

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

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. Это, на самом деле, довольно общая ситуация: экстремаль «регулярнее», чем произвольный элемент своего пространства.

2.4 Случай свободных концов

Теперь рассмотрим совсем произвольную $h \in C^1$, и получим уравнение на вариацию

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

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

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.